Physics 550 Final Project
Time-resolved Shadow Tomography of Open System Dynamics

Onkar Apte, Arman Babakhani, Joseph Barreto

April 2023

Contents

1 Introduction

We show that shadow state tomography is applicable to open quantum systems, and highlight its shortcomings. Shadow state tomography provides an efficient way to estimate the expectation values of operators, requiring exponentially less copies of a state to approximate

the expectation values of a set of operators, with the cost of introducing some error in the certainty of our measurement outcomes. In this work, we investigate the dependence of shadow state tomography using single Pauli unitaries on the specifics of the system Hamiltonian. We highlight the fact that the number of copies required to estimate the measurement outcomes of a set of operators is related to the locality of the operators, and the locality of operators as a function of time could increase depending on the specifics of the Hamiltonians and models in consideration. Because of this fact, the efficiency of shadow state tomography to ascertain expectation values after some time evolution, using only shadows of the initial state, seems to be heavily dependent on the specifics of the dynamics. In general, shadow state tomography might be limited to short-time state estimation. However, in certain settings, shadow state tomography can be used for state estimation for long-time evolutions.

In this work, we show that for a specific open system dephasing dynamics, shadow state tomography could be used to efficiently estimate the expectation value of local operators for times of the order of dephasing time. However, we suspect that this is only specific to the model we have chosen, and in general, one would not be able to efficiently approximate the expectation values of local operators without worrying about increasing the number of copies of initial state required to maintain a reasonable error.

2 Outline

Our goal is to apply shadow tomography to efficiently estimate the measurement outcomes of operators of a state $\rho_S(t)$. Directly measuring the state $\rho_S(t)$ of the system might not be feasible as it will disturb the state, and one has to possibly employ weak measurement methods, that will further complicate the process. If, however, one has access to measurements of the initially prepared state ($\rho_S(0)$), with the knowledge of the noise model that defines the open system time evolution of ρ_S, one can use shadow tomography to estimate the measurement outcomes of local operators at time $t > 0$.

Given a set of operators $\{O_n\}$, the expectation value of the O_i at time t is

$$\langle O_i \rangle(t) = \mathrm{Tr}(O_i \rho_S(t)) = \mathrm{Tr}(O_i \mathcal{L}_t[\rho_S(0)]) = \mathrm{Tr}\Big(\mathcal{L}_t^\dagger[O_i]\rho_S(0)\Big)\,. \tag{1}$$

where $\mathcal{L}_t[\bullet]$ is the Lindbladian describing the open system dynamics of the state ρ_S. If we intend to estimate $\langle O_i \rangle(t)$ from $\rho_S(0)$, we can compute $O_i(t)$ through Heisenberg picture time evolution of open system dynamics, and use shadow state tomography to ascertain $\langle O_i \rangle(t)$. The shadow norm of an operator using random Pauli unitaries is bounded by an exponential function of the locality of the operator. Since the locality of the operator O_i might change in the Heisenberg picture time evolution, increasing the required number of copies of $\rho_S(0)$ to attain measurements within a given error. This is because shadow state tomography requires number of copies of the order that depends on the shadow norm of the operator being measured. More details about the relationship between the shadow norm and the number of copies of the system will be provided in the next section (Sec. 3). In Sec. 4, we will specify the details of our measurement operators and the models considered. In section Sec. 5, we will provide some analysis justifying our hypothesis that in general one cannot efficiently use shadow tomography to estimate the measurement outcomes of local

operators for arbitrarily large times, having access to the initial copies of the system. In sections Sec. 5 and Sec. 6, we summarize our results and outline our conclusions.

3 Background

Shadow tomography generally refers to a measurement protocol which, given some number of identical copies of a state ρ and a list of observables O_i, estimates the expectation values $o_i = \mathrm{Tr}(O_i\rho)\ \forall i$ with far fewer shots (i.e. copies of ρ) compared to a naive protocol which relies on state tomography, and thus an exponential number of preparations of ρ. Extended to processes, one could also seek to predict $o_{ij} = \mathrm{Tr}(\mathcal{E}(\rho_j)O_i)$ given some number of identical applications of a fixed process $\mathcal{E}$. Instead of performing process tomography on $\mathcal{E}$, the shadow tomography technique directly estimates the desired observables for some set of O_i and input states ρ_j with sub-exponential uses of $\mathcal{E}$.

3.1 Shadow state tomography

3.1.1 Aaronson

The paper by Aaronson [1] introduces the problem of shadow tomography, which is as follows:

Given an unknown D-dimensional quantum state ρ, two-outcome measurements $E_1, \cdots, E_M$ which accepts or rejects ρ with probability $\mathrm{Tr}(E_i\rho)$ and $1-\mathrm{Tr}(E_i\rho)$ respectively, output numbers $b_1, b_2, \cdots, b_M \in [0,1]$ such that $|b_i - \mathrm{Tr}(E_i\rho)| \leq \varepsilon\ \forall i$ with success probability at least $1-\delta$. Do this via measuring k copies of ρ, where $k = k(D, M, \varepsilon, \delta)$ is as small as possible. The author shows that this problem is solvable using

$$k = \tilde{O}\left(\frac{\log 1/\delta}{\varepsilon^4} \cdot \log M \cdot \log D\right) \tag{2}$$

copies of ρ where $\tilde{O}$ hides polynomial functions of $\log\log M, \log\log D, \log\frac{1}{\varepsilon}$.

It is also shown that the lower bound for this problem is

$$k = \Omega\left(\frac{\min\{D^2, \log M\}}{\varepsilon^2}\right). \tag{3}$$

The shadow tomography procedure given by the author involves combining two ideas, post-selected learning on quantum states [**aaronson˙2005**] and gentle search procedure [**aaronson˙2006**]. Even though we can estimate the probabilities of exponentially many measurements using only a polynomial copies of ρ, the time complexity of this method is still exponential. Huang et al. [2] gives a time-efficient shadow tomography method which is discussed in Sec. 3.1.2 . We will be mainly using this algorithm.

3.1.2 Huang

The procedure for shadow tomography given by [2] is as follows. Given an n-qubit state ρ, we randomly choose a unitary gate U from a particular ensemble of unitaries $\mathcal{U}$ and apply

it to ρ, and then measure each qubit of the resulting state in the Z-basis, yielding a random bitstring $|\hat{b}\rangle$. This particular bitstring $|\hat{b}\rangle$ occurs with a probability given by the Born rule:

$$\Pr\left[\hat{b} = b\right] = \langle b|U\rho U^\dagger|b\rangle\,. \tag{4}$$

It is important that the ensemble $\mathcal{U}$ be tomographically complete, i.e.

$$\exists U, b, \rho \neq \sigma \quad \text{s.t.} \quad \langle b|U\sigma U^\dagger|b\rangle \neq \langle b|U\rho U^\dagger|b\rangle\,, \tag{5}$$

so it is always possible to distinguish two unequal states from each other given an appropriate choice of U (in the limit of infinite shots). An efficient, classical description of $U^\dagger|\hat{b}\rangle\langle\hat{b}|U$ is stored in classical memory [1]. We refer to this quantity as a *snapshot*. If one were to average these snapshots over all unitaries (classically sampled) and bitstrings (quantumly sampled), the net effect has the form of a quantum channel $\mathcal{M}$:

$$\mathcal{M}(\rho) \equiv \mathbb{E}\left[U^\dagger|\hat{b}\rangle\langle\hat{b}|U\right]\,, \tag{6}$$

$$= \mathbb{E}_{U\sim\mathcal{U}} \sum_{b\in\{0,1\}^n} \langle b|\,U\rho U^\dagger\,|b\rangle\,U^\dagger\,|b\rangle\langle b|\,U\,. \tag{7}$$

Though the inverse of this channel, $\mathcal{M}^{-1}$, may not be a valid quantum operation, it is still numerically possible to compute classically. Moreover, it is Hermitian: $\mathrm{Tr}(X\mathcal{M}^{-1}(Y)) = \mathrm{Tr}(\mathcal{M}^{-1}(X)Y)$. One can then define the *classical shadow* of ρ, denoted as $\hat{\rho}$, via

$$\hat{\rho} = \mathcal{M}^{-1}(U^\dagger|\hat{b}\rangle\langle\hat{b}|U)\,. \tag{8}$$

The classical shadow $\hat{\rho}$ is a matrix with unit trace, but need not be positive semi-definite, and reproduces ρ in expectation:

$$\mathbb{E}(\hat{\rho}) = \mathcal{M}^{-1}\big(\mathbb{E}\left[U^\dagger|\hat{b}\rangle\langle\hat{b}|U\right]\big)\,, \tag{9}$$

$$= \mathcal{M}^{-1}(\mathcal{M}(\rho)) = \rho\,. \tag{10}$$

The classical shadow is the single-shot limit of the linear-inversion estimate of ρ, so snapshots of all possible U in the ensemble could be used to perform state tomography. However, it is usually the case that we seek only some observables of ρ, not the full density matrix, and thus only a few snapshots may be needed to estimate these observables. Specifically, averaging over shadows reproduces an observable in expectation:

$$\hat{o} = \mathrm{Tr}(O\hat{\rho}) \implies \mathbb{E}[\hat{o}] = \mathrm{Tr}(O\rho)\,. \tag{11}$$

The authors of [2] show that for a fixed observable O, the variance in $\hat{o}$ obeys

$$\mathrm{Var}[\hat{o}] \leq \left\| O - \frac{\mathrm{Tr}(O)}{2^n}\mathbb{I} \right\|^2_{\text{shadow}} \tag{12}$$

[1] This is only possible depending on the choice of unitary ensemble, i.e. Clifford gates, which have known efficient representations via the stabilizer formalism.

which is to say, the variance is bounded by the square of the *shadow norm* of the traceless part of O. The shadow norm depends only on the measurement protocol

$$||O||_{\text{shadow}} = \max_{\sigma} \left(\mathbb{E}_{U\sim\mathcal{U}} \sum_{b\in\{0,1\}^n} \langle b| U\sigma U^\dagger |b\rangle\langle b| U\mathcal{M}^{-1}(O)U^\dagger |b\rangle^2 \right)^{1/2} \tag{13}$$

and is non-negative, homogeneous, and obeys the triangle inequality (so it is a valid norm).

With this, the main theorem of [2] asserts that for a chosen ensemble $\mathcal{U}$, a set of observables $O_1, \ldots, O_M$, and parameters $\epsilon, \delta \in [0,1]$, pick two sampling parameters K and N such that

$$K = 2\log(2M/\delta)\,, \tag{14}$$

$$N = \frac{34}{\epsilon^2} \max_{1\le i\le M} \left\| O_i - \frac{\text{Tr}(O_i)}{2^n}\mathbb{I} \right\|^2_{\text{shadow}}\,. \tag{15}$$

A collection of NK classical shadows is sufficient to estimate all $\text{Tr}(O_i\rho)$ to error ϵ via *median-of-means* estimation:

$$|\hat{o}_i(N,K) - \text{Tr}(O_i\rho)| \le \epsilon\,, \quad \forall\, 1\le i\le M \tag{16}$$

with probability at least $1-\delta$. The median-of-means estimator $\hat{o}_i(N,K)$ denotes the procedure of dividing the NK shadows into K sets of N shadows each, and taking the median of the sample means over the K sets; the same NK shadows is used for each observable O_i. The total sample complexity scales with the number of shadows NK and depends crucially on the ensemble $\mathcal{U}$, which determines the shadow norm.

$$N_{tot} = NK = \mathcal{O}\left(\frac{\log M}{\epsilon^2} \max_{1\le i\le M} \left\| O_i - \frac{\text{Tr}(O_i)}{2^n}\mathbb{I} \right\|^2_{\text{shadow}} \right). \tag{17}$$

Notice the logarithmic scaling with the number of observables M. Two noteworthy examples of N_{tot} are given in the appendix of [2], for the cases where $\mathcal{U}$ is either $\text{Cl}(2^n)$ or $\text{Cl}(2)^{\otimes n}$, i.e. Clifford gates which act either globally on all n qubits (case 1), or tensor products of single-qubit Cliffords (case 2). The latter case is equivalent to measuring each qubit in a random Pauli basis. For case 1, it can be shown that

$$\hat{\rho} = (2^n+1)U^\dagger|\hat{b}\rangle\langle\hat{b}|U - \mathbb{I}\,, \tag{18}$$

$$\left\| O - \frac{\text{Tr}(O)}{2^n}\mathbb{I} \right\|^2_{\text{shadow}} \le 3\,\text{Tr}\big(O^2\big)\,, \tag{19}$$

so for non-local Cliffords, the bound scales with the rank of O, thus case 1 is naturally suited to O which are projectors, but requires physically executing highly-entangled n-qubit operations. For case 2, we have

$$\hat{\rho} = \bigotimes_{j=1}^{n} \left(3U_j^\dagger|\hat{b}_j\rangle\langle\hat{b}_j|U_j - \mathbb{I}\right), \tag{20}$$

$$\left\| O - \frac{\text{Tr}(O)}{2^n}\mathbb{I} \right\|^2_{\text{shadow}} \le 4^{k(O)}||O||^2_\infty\,, \tag{21}$$

where $k(O)$ is the weight, or locality, of O; this is naturally suited to O that are Pauli strings of low-local bounded weight, and only requires single-qubit gates. In this case, we can examine the estimator for the expectation value of a Pauli string $O = \prod_{j=1}^{n} P_j$:

$$\hat{o} = \mathrm{Tr}(O\hat{\rho}) \tag{22}$$

$$= \mathrm{Tr}\left(\bigotimes_{j=1}^{n} \left(3P_j U_j^\dagger |\hat{b}_j\rangle\langle\hat{b}_j| U_j - P_j\right)\right) \tag{23}$$

$$= \prod_{j=1}^{n} \mathrm{Tr}\Big(3P_j U_j^\dagger |\hat{b}_j\rangle\langle\hat{b}_j| U_j - P_j\Big) \tag{24}$$

If $P_j = I$, then the j-th contribution to the product is

$$3\,\mathrm{Tr}\Big(U_j^\dagger |\hat{b}_j\rangle\langle\hat{b}_j| U_j\Big) - 2 = 1 \tag{25}$$

If $P_j \in \{X, Y, Z\}$, then the j-th contribution to the product is

$$3\langle\hat{b}_j| U_j P_j U_j^\dagger |\hat{b}_j\rangle = 3\delta_{U_j, P_j} \tag{26}$$

The RHS means that the LHS evalutes to 3 iff the j-th single-qubit Clifford of U diagonalizes the j-th single-qubit Pauli of O. Since the Clifford group normalizes the Pauli group, we know that $U_j P_j U_j^\dagger \in \mathcal{P}$, so either P_j is diagonalized to become Z, or otherwise it must go to X or Y. We know that $\langle\hat{b}_j|Z|\hat{b}_j\rangle = \pm 1$, or otherwise 0 if we replace Z with X or Y. Thus, the *entire* estimate $\hat{o}$ is 0 unless the unitary U used to build a given snapshot $\hat{\rho}$ "aligns with" the observable O, in terms of their single-qubit bases. In this way, it is easy to tabulate the contributions from all of the shadows to a particular $\hat{o}$ without ever explicitly writing down operators, avoiding costly classical space overhead. For observables that are Pauli strings as just discussed, the dependence of the error scaling on the locality can be improved to $3^{k(O)}$ [2]. Now the exponential scaling in this case is apparent, because we randomly generate

3.1.3 Error scaling

Here we provide a bit more analysis regarding the error scaling of the shadow tomography procedure. Let $\hat{o}_i(N, K)$ denote the estimate of $\mathrm{Tr}(O_i\rho)$ where K sets of N shadows each are used in the median-of-mean (MoM) estimation. If we just directly average without MoM (so $K = 1$), then

$$\hat{o}_i(N, 1) = \frac{1}{N}\sum_{j=1}^{N} \mathrm{Tr}(O_i\hat{\rho}_j) \tag{27}$$

Chebyshev's inequality [3] states that for N i.i.d random variables $X_1, \ldots, X_N$ with $\mathbb{E}(X_i) = \mu_i$ and $Var(X_i) = \sigma_i^2$, then for any $\epsilon > 0$:

$$\mathrm{prob}\left(\left|\frac{\sum_{i=1}^{N} X_i}{N} - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{N\epsilon^2} \tag{28}$$

If we define this probability as δ, then we have

$$N \leq \frac{\sigma^2}{\delta \epsilon^2} \tag{29}$$

which quantifies the number of samples needed to obtain, with probability δ, a sample mean which differs from the true mean by no more than ϵ (for the worst case, we would just set this to an equality). To improve the scaling in δ, the authors propose using median-of-means estimation (MoM):

$$\hat{o}_i(N, K) = \text{median}\big\{\hat{o}_i^{(1)}(N, 1), \ldots, \hat{o}_i^{(K)}(N, 1)\big\} \tag{30}$$

where $\hat{o}_i^{(j)}(N, 1)$ denotes the sample mean of the j-th batch of N estimates. For the median of these values to differ from the true mean by more than ϵ, more than half of the values must themselves differ by at least as much - since these events are independent, the probability of this occurring should be exponentially suppressed in the number of MoM batches, K. Specifically, the MoM estimator obeys [4]

$$\text{prob}(|\hat{o}_i(N, K) - \text{Tr}(O_i\rho)| > \epsilon) \leq e^{-2K(1/2 - \sigma_i^2/N\epsilon^2)^2} \leq e^{-K/2} \leq 2e^{-K/2} \tag{31}$$

where the second inequality only holds if $N \leq \sigma_i^2/\epsilon^2$. Since each estimate $\hat{o}_i(N, K)$ is independent, if we want *all* M estimates to be within ϵ with probability $1 - \delta$, then

$$\delta = \text{prob(at least one exceeds } \epsilon) \leq \sum_{i=1}^{M} \text{prob}(|\hat{o}_i(N, K) - Tr(O_i\rho)| > \epsilon) \tag{32}$$

$$\leq 2Me^{-K/2} \tag{33}$$

via a "union-bound". So in the worst case $K = 2\log 2M/\delta$. This, combined with $N = \sigma_i^2/\epsilon^2$, gives the bounds of equations S13 and S14 in [2]. The variances σ_i^2 are the shadow norms of the operators O_i. [2]

3.1.4 Derandomization

TODO

4 Shadow State Tomography for Open Systems

We will consider the following open system models of N qubits, in interaction with a bosonic bath.

The total Hamiltonian of the system is given by $H = H_S + H_{SB} + H_B$, where

$$H_{SB} = g \sum_{i=1}^{N} A_i \otimes B_i \,. \tag{34}$$

[2] I'm not sure where the factor of 34 comes from, or why they introduce an extra factor of 2.

In the interaction picture, $B_i(t) = e^{itH_B} B_i e^{-itH_B}$, the bath operator correlation functions $\langle B_i(t) B_j(t') \rangle \equiv \mathrm{Tr}(\rho_B B_i(t) B_j(t'))$ define the $\gamma_{ij}(\omega)$ s in the dynamics given by the following Lindbladian

$$\mathcal{L}[\bullet] = -i[H_S + H_{LS}, \bullet] + \sum_{i,j,\omega} \gamma_{ij}(\omega) \left(A_j(\omega) \bullet A_i^\dagger(\omega) - \frac{1}{2} \{ A_i^\dagger(\omega) A_j(\omega), \bullet \} \right) . \tag{35}$$

H_{LS} is a Lamb shift Hamiltonian [5], and $e^{itH_S} A_i e^{-itH_S} = \sum_\omega e^{-it\omega} A_i(\omega)$, define the operators $A_i(\omega)$.

4.1 A simple model for dephasing

By setting $A_i = Z^{(i)}$, and $H_S = \sum_{i=1}^N \omega_i Z^{(i)}$, we see that the only relevant ω in Eq. 35, is zero. In this case, $A_i(\omega) = A_i(\omega)\delta_{\omega\,0} = Z^{(i)}\delta_{\omega\,0}$. Furthermore, the Lamb shift Hamiltonian will be given by

$$H_{LS} = \sum_{i,j} S_{ij} Z^{(i)} Z^{(j)} . \tag{36}$$

where

$$S_{ij} = \int_{-\infty}^{\infty} \gamma_{ij}(\omega') \mathcal{P} \left(-\frac{1}{\omega'} \right) . \tag{37}$$

Furthermore, for simplicity we assume that $\gamma_{ij}(0) = \tilde{\gamma}_i \delta_{ij}$. This assumption holds in Harmonic oscillator baths and other similar models in which different bath operators B_i and B_j are uncorrelated. Therefore, the Lamb shift Hamiltonian becomes proportional to the identity operator, and hence negligible in the subsystem dynamics. The Lindbladian reduces to

$$\mathcal{L}[\bullet] = -i[H_S, \bullet] + \sum_i \tilde{\gamma}_i \left(Z^{(i)} \bullet Z^{(i)} - \bullet \right) . \tag{38}$$

Thus, to specify the model further, we must be given a set of rates $\{\tilde{\gamma}_i\}$.

4.2 Expectation values of local operators

In order to ascertain the initial expectation values of local operators, we can perform shadow tomography on the system. Shadow state tomography will allow us to efficiently compute the outcome of all local operators when the dynamics have not significantly altered the initial state of the system. More specifically, in the Heisenberg picture, one can see that the locality of an operator A might increase due to the dynamics generated by the system Hamiltonian. The first term in the Lindbladian, given in Eq. 38 would not be locality preserving if the support of the term $[H_S, O(t)]$ is different than that of $O(t)$. In this work, we investigate the importance of preserving locality when using shadow state tomography using random single body Pauli unitaries.

5 Analysis

One of the main goals of this paper is to emphasize that shadow state tomography is applicable on open quantum systems, given by master equations, describing the evolution of the system. As we have seen, shadow tomography can be used to efficiently estimate observables of a given state ρ, but now we would like to use it to track an observable over time. Generally, there are two ways we can proceed, based on the equivalence between the Schrödinger and Heisenberg pictures

$$\langle O(t)\rangle = \mathrm{Tr}(O\rho(t)) = \mathrm{Tr}(O(t)\rho_0)\,. \tag{39}$$

We can replace ρ with its shadow estimate to obtain

$$\langle O(t)\rangle = \mathbb{E}[\mathrm{Tr}(O\hat{\rho}(t))] = \mathbb{E}[\mathrm{Tr}(O(t)\hat{\rho}_0)]\,. \tag{40}$$

Defining a total duration T and a resolution Δt, in the middle term we take a snapshot at each timestep $t_i = i\Delta t$ where $0 \leq i \leq L$ and $L = T/\Delta t$ and build a shadow estimate of O over time. A set of L shadows $\hat{\rho}_i = \hat{\rho}(t_i)$ can be used to provide one snapshot of the entire trajectory of $\langle O(t)\rangle$ with a temporal resolution Δt. On the RHS, working in the Heisenberg picture, we could also fix an initial state $\rho_0 = \rho(t_0) = \rho(0)$. Given an *assumed* model for the open system dynamics, the shadow estimate of ρ_0 alone is sufficient to approximately predict the dynamics of $O(t)$. Thus, one can trade-off overhead in quantum measurements (i.e. needing to generate shadows at each time) and classical post-processing (i.e. needing to classically evolve only the initial shadow).

Working in the Heisenberg picture, the evolution of a Hermitian operator O, through the following master equation

$$\dot{O}(t) = \mathcal{L}^\dagger[O(t)] = i[H(t), O(t)] + \sum_\alpha \gamma_\alpha \left(L_\alpha^\dagger O(t) L_\alpha - \frac{1}{2}\{L_\alpha^\dagger L_\alpha, O(t)\}\right)\,. \tag{41}$$

In order to show that the complexity of shadow state tomography will not be violated by choosing to measure expectation values of operators at various times, we must require that the shadow norm of the operator $O(t)$ remains bounded by a time-independent constant B, for all times $t > 0$. We show that for two different choices of unitaries for shadow state tomography, the corresponding bounds are indeed time-independent, given certain constraints.

- **Random Clifford Unitaries:** The shadow norm using random Clifford unitaries satisfy the bound [6]

$$\left\| O(t) - \frac{\mathrm{Tr}(O(t))}{2^n}\mathbb{I}\right\|_{\text{shadow, } \mathcal{U}_C} \leq 3\,\mathrm{Tr}\big(O(t)^2\big)\,.$$

- **Random one-body Paulis**: The shadow-norm for Pauli unitaries are bounded by [6]

$$\|O(t)\|_{\text{shadow},\mathcal{U}_P} \leq 4^{k(t)}\|O(t)\|_\infty\,,$$

where $k(t)$ is the support of the operator at time t. Clearly, in order for the shadow norm of $O(t)$ to be bounded by a constant B, we must require that $k'(t) \leq 0$. The models we are considering satisfy this property, as $\textit{support}([L_\alpha, O(0)]) = \textit{support}([H, O(0)]) = \textit{support}(O(0))$. Since, the Lindblad master equation is trace-preserving, the operator norm of the operator will also be bounded above by $\mathrm{Tr}(O^2)$.

5.1 Boundedness of $\mathrm{Tr}\big(O^2(t)\big)$

Here, we will show that open system dynamics, given by Lindblad master equation, preserves the inequality $\mathrm{Tr}(O^2(t)) \leq \mathrm{Tr}(O^2(0))$. Note that the Lindblad master equation in Heisenberg picture, denoted by $\mathcal{L}^\dagger[\bullet]$, is a unital map, and as a consequence, any unital CPTP can be written as a mapping using Kraus operators, which preserves the "purity" of an operator O [5]. For the models we are considering, we will specifically show that $\mathrm{Tr}(O^2)$ is indeed non-increasing by proving $\partial_t \mathrm{Tr}(O^2(t)) \leq 0$ for all t

$$\begin{aligned}
\partial_t \mathrm{Tr}\big(O^2(t)\big) &= 2\,\mathrm{Tr}\Big(\dot{O}(t)O(t)\Big) = 2\,\mathrm{Tr}\big(\mathcal{L}^\dagger[O(t)]O(t)\big) \\
&= 2\,\mathrm{Tr}\left(\left(i[H(t), O(t)] + \sum_\alpha \gamma_\alpha (L_\alpha^\dagger O(t) L_\alpha - \frac{1}{2}\{L_\alpha^\dagger L_\alpha, O(t)\})\right) O(t)\right) \\
&= 2\sum_\alpha \gamma_\alpha\,\mathrm{Tr}\left(\left(L_\alpha^\dagger O(t) L_\alpha - \frac{1}{2}\{L_\alpha^\dagger L_\alpha, O(t)\}\right) O(t)\right) \\
&= 2\sum_\alpha \gamma_\alpha\,\mathrm{Tr}\big(\big(L_\alpha^\dagger O(t) L_\alpha O(t) - L_\alpha^\dagger L_\alpha O^2(t)\big)\big) \\
&= 2\sum_\alpha \gamma_\alpha\,\mathrm{Tr}\left(L_\alpha^\dagger [O(t), L_\alpha] O(t)\right). \qquad (42)
\end{aligned}$$

Even though one can prove that for a more general set of Lindblad operators L_α, $\partial\,\mathrm{Tr}(O^2(t)) \leq 0$, it is quite straightforward to show that when $L_\alpha^\dagger = L_\alpha$(as it is for our simple dephasing model) or more generally when $[L_\alpha, L_\alpha^\dagger] = 0$, $\partial_t\,\mathrm{Tr}(O^2(t)) \leq 0$. To see this, note that we can write

$$\mathrm{Tr}\big(L_\alpha^\dagger [O, L_\alpha] O\big) = -\frac{1}{2}\left(\mathrm{Tr}\big([L_\alpha, O]^\dagger [L_\alpha, O]\big) + \mathrm{Tr}\big([L_\alpha^\dagger, L_\alpha] O^2\big)\right). \qquad (43)$$

Thus, when $[L_\alpha, L_\alpha^\dagger] = 0$,

$$\mathrm{Tr}\big(L_\alpha^\dagger [O, L_\alpha] O\big) = -\frac{1}{2}\left(\mathrm{Tr}\big([L_\alpha, O]^\dagger [L_\alpha, O]\big)\right) \leq 0. \qquad (44)$$

So, it follows that $\partial_t\,\mathrm{Tr}(O^2(t)) \leq 0$, since $\gamma_\alpha > 0$ for all α.

5.2 Locality of operator $O(t)$

The second part of the Lindbladian $\mathcal{L}^\dagger[\bullet]$, i.e. the dynamics generated by the Lindblad operators L_α is locality preserving, as both Lindblad operators for a specific α act on the same qubit. However, for different models, such as a model in which γ_{ij} are non-zero for

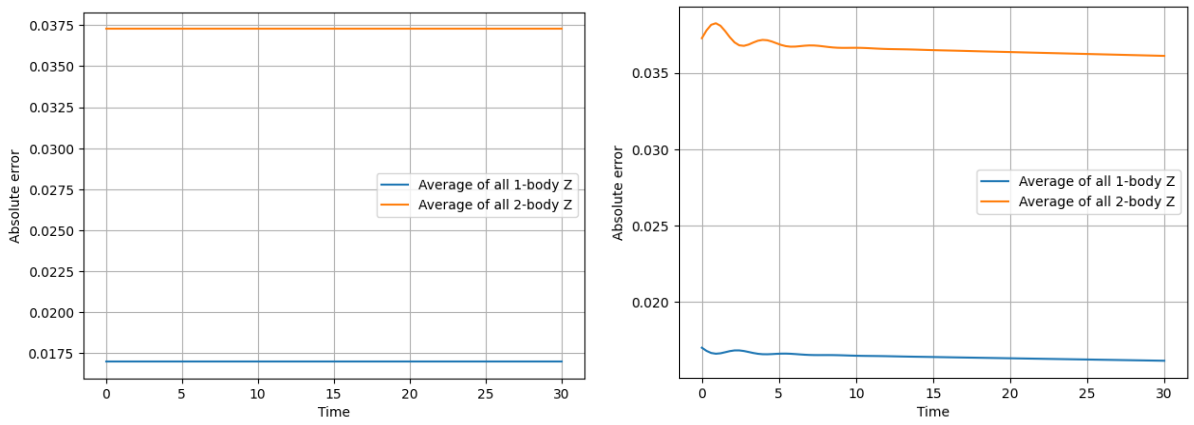
$i \neq j$, this would no longer be true. This is because a single body operator such as X_i, would then accrue two-body contributions to its evolution right away and its locality would immediately increase as open system dynamics initiate. For simplicity, we have relegated the responsibility of locality-altering dynamics to the system Hamiltonian. In the simple case of a bare Hamiltonian $H_S = \sum \omega_i Z^{(i)}$, the locality of operators would be preserved by the Lindbladian. However, an interacting subsystem of the form $H'_S = H_S + \sum_{\langle i,j \rangle} X_i X_j$ would alter the locality of many local operators such as Z or Y.

6 Numerics and Results

We use QuTiP [7] to define a $N = 5$ qubit model where $H_S = \sum_i \omega_i Z^{(i)} + H_{\text{int}}$, with $\omega_i = \omega = -1/2$. All qubits undergo Z-dephasing with strength $\tilde{\gamma}_i = \gamma = 0.1$. For a fixed, random, initial pure state, we consider two models: a non-interacting setting where $H_{\text{int}} = 0$, and an pairwise interacting model where $H_{\text{int}} = \sum_i g_i X_i X_{i+1}$ and where the qubits have a ring topology. The coupling strengths g_i are made uniform, i.e. $g_i = g = 0.01$, so the system is translationally invariant except for the random initial state. We leverage the code associated with [2] to simulate the generation and measurement of shadows, and for now we are only using sample means to build estimates of observables instead of the median-of-means method which appears in the complexity bounds given in Sec. 3; this means that effectively $K = 1$ for our simulations, because we did not see any numerical significant differences between the two approaches [3].

For dephasing in non-interacting model above, we are guaranteed that the bound of the shadow norm of the operator remains independent of the system size and so the number of copies required to get a good estimate of the measurement outcome is logarithmic in the system dimension. However, for the interacting model, the locality of local observables changes. In this case, the Clifford unitary sampling (which we did not implement) should maintain the same operator bound for different times; however, sampling from an ensemble of Pauli operators will have a time dependent upper bound, as the operator support of the observable O could grow as a function of time. This does not immediately guarantee that one could estimate the outcomes of local observables for $\rho_S(t)$ efficiently from copies of $\rho_S(0)$, but we will show that in practice, the error scaling is apparently better than the worst-case scaling, at least for the observables we have investigated so far.

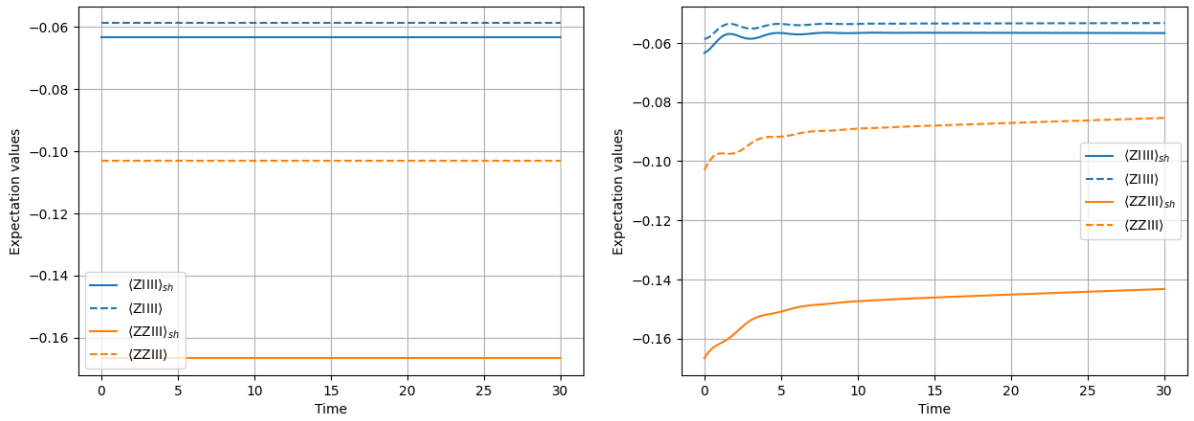
[3] Furthermore, one would still obtain an exponential bound in the sampling error via Hoeffding's inequality even in the usual sample mean case.



(a) Average over all 1 and 2-body Z observables, non interacting case

(b) Average over all 1 and 2-body Z observables, interacting case

Figure 1: Plots of absolute errors of 1 and 2 body operators. Absolute error is defined as the absolute difference between the expectation value of the exact $\rho_S(0)$ and the shadow state.



(a) Selected 1 and 2-body Z observables over time, non interacting case

(b) Selected 1 and 2-body Z observables over time, interacting case

Figure 2: The label $\langle \ldots \rangle_{\text{sh}}$ refers to the expectation value of the operator using shadow state.

7 Discussion and Future Directions

Even though we expected that the estimation errors of measurement outcomes of local operators would worsen as the locality of the operators grow as a function of time, the plots in the previous section illustrate that the expectation value of operators can be approximated efficiently without loss of accuracy for seemingly arbitrary long times.

This could have multiple explanations. If one writes an operator $O(t) = \sum_{l=1}^{N} \sum_i \alpha_i^{(l)}(t) A_i^l$, where A_i^l are l-body operators, then it might be that the coefficients for larger l values become non-zero, but with a very negligible weight compared to those of smaller ones. So, even though the operator's locality increases, most of its support is still on the original locality.

Further, the fact that the shadow norm is bounded above by an exponential function of the locality of an operator, does not suggest that the shadow norm will saturate this bound.

One possible way to further investigate the role of locality and errors obtained from shadow tomography would be to perform a similar analysis for larger interaction strength. It might be that the time-scale for interactions to imprint their effects on raising the locality of the operator was larger than the total time we have simulated. So, further investigation is required to shed light on this matter.

Another approach is to decrease the number of shadows used in estimating the observables, in order to increase the sampling error enough to hit the upper bound; this would require adapting the code to use the derandomization scheme of [8] to allow for estimates built from a very small number of shadows, which are tailored to the desired observables. The reason this is difficult with the current code is because with only a few random shots (i.e. random basis measurements), it is unlikely one will perform any measurements that are actually suitable for estimating the desired observables, since the shadows are independent of the observables in the present case.

References

[1] Scott Aaronson. *Shadow Tomography of Quantum States*. arXiv:1711.01053 [quant-ph]. Nov. 2018. URL: http://arxiv.org/abs/1711.01053 (visited on 04/26/2023).

[2] Hsin-Yuan Huang, Richard Kueng, and John Preskill. "Predicting many properties of a quantum system from very few measurements [SUPPLEMENT]". en. In: *Nature Physics* 16.10 (Oct. 2020), pp. 1050–1057. ISSN: 1745-2473, 1745-2481. DOI: 10.1038/s41567-020-0932-7. URL: http://www.nature.com/articles/s41567-020-0932-7 (visited on 04/26/2023).

[3] Michel Goemans. "Chernoff bounds, and some applications". en. In: ().

[4] Matthieu Lerasle. *Lecture Notes: Selected topics on robust statistical learning theory*. en. arXiv:1908.10761 [cs, math, stat]. Aug. 2019. URL: http://arxiv.org/abs/1908.10761 (visited on 05/10/2023).

[5] Daniel Lidar. *Open Quantum Systems*. May 2023. (Visited on 05/01/2023).

[6] Jonathan Kunjummen et al. "Shadow process tomography of quantum channels". In: *Physical Review A* 107.3 (Mar. 2023). arXiv:2110.03629 [quant-ph], p. 033319. ISSN: 2469-9926, 2469-9934. DOI: 10.1103/PhysRevA.107.033319. URL: http://arxiv.org/abs/2110.03629 (visited on 04/26/2023).

[7] J.R. Johansson, P. D. Nation, and F. Nori. "QuTiP 2: A Python framework for the dynamics of open quantum systems". In: *Comp. Phys. Comm.* 184.1234 (2013). DOI: 10.1016/j.cpc.2012.11.019.

[8] Hsin-Yuan Huang, Richard Kueng, and John Preskill. "Efficient Estimation of Pauli Observables by Derandomization". en. In: *Physical Review Letters* 127.3 (July 2021), p. 030503. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.127.030503. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.030503 (visited on 05/05/2023).