Efficient estimation of Lindbladians from trajectory data

Quanjun Lang, Bowen Li, Jianfeng Lu

Contents

1 Introduction

We study the problem of learning in open quantum systems. The goal is to learn the dynamics in the Lindblad quantum master equation (QME)

$$\frac{d}{dt}\rho = -i[H,\rho] + \sum_{k=1}^{N_J}(J_k\rho J_k^\dagger - \frac{1}{2}J_k^\dagger J_k\rho - \frac{1}{2}J_k^\dagger J_k\rho) \tag{1.1}$$ {eq_QME_main}

from observations of the trajectories. Here H is the system Hamiltonian and J_k's are jump operators for $k = 1,\dots N_J$. The density matrix $\rho \in \mathbb{C}^{N\times N}$ here characterizes the distribution of a quantum $|\psi(t)\rangle$ in the sense that $\mathbb{E}[|\psi(t)\rangle\langle\psi(t)|]$. Here $|\psi(t)\rangle \in \mathbb{C}^N$ itself satisfies the stochastic Schrödinger equation

$$d|\psi(t)\rangle = \left(-iH|\psi(t)\rangle - \frac{1}{2}\sum_{k=1}^{r} J_k^\dagger J_k|\psi(t)\rangle\right)dt + \sum_{k=1}^{r} J_k|\psi(t)\rangle\, dW_t^k \tag{1.2}$$

where W_t^k's are independent Brownian motions.

The above QME can be written as

$$\frac{d}{dt}\rho := \mathcal{L}\rho = \mathcal{L}_H\rho + \mathcal{L}_J\rho, \tag{1.3}$$ {eq_QME_Lindb

where $\mathcal{L}_H\rho = -i[H,\rho]$ and $\mathcal{L}_J\rho = \sum_{k=1}^{N_J}(J_k\rho J_k^\dagger - \frac{1}{2}J_k^\dagger J_k\rho - \frac{1}{2}J_k^\dagger J_k\rho)$. The super operator $\mathcal{L}$ is called the Lindbladian, and solutions of (1.3) can be expressed using semigroup generated by $\mathcal{L}$, namely

$$\rho(t) = e^{\mathcal{L}t}\rho(0). \tag{1.4}$$

Then the map $\mathcal{V}(t) = e^{\mathcal{L}t}$ is a complete positive trace-preserving (CPT) map for arbitrary time t. By the Choi-Kraus' Theorem, a mapping $\mathcal{V} : \mathcal{B}(\mathbb{R}^n) \to \mathcal{B}(\mathbb{R}^n)$ is completely positive and trace-preserving if and only if it can be expressed as

$$\mathcal{V}\rho = \sum_k A_k^\dagger \rho A_k \tag{1.5}$$

where $\sum_k A_k A_k^\dagger = I_{\mathcal{H}}$, and $\mathcal{B}(H)$ represents the space of bounded operators on $\mathbb{R}^n$.

The possible measurements of the density ρ are described by a Hermitian operator or *observable* O. The observation result is achieved by taking the expectation value of the outcome, which gives

$$\langle O\rangle := \mathrm{tr}(O\rho). \tag{1.6}$$

Using observables to infer the density ρ is called Quantum state tomography. As we will point out later, the full state information makes the estimation of H and J_k more efficient.

We aim to learn the Hamiltonian H and the jump operators J_k. We will first assume a discrete observation of density $\rho(t_l)^{(m)}$ from M independent copies, where $t_l = l\Delta t, l = 0, 1, \ldots, N_T$ is assumed to be an equidistant time grid, with $T := N_T\Delta t$. The problem is concluded as the following.

$$\textbf{Data: } \{\rho(t_l)^{(m)}\}_{m=1,l=1}^{M,N_T}; \quad \textbf{Goal: } \{H, J_1, \ldots, J_{N_J}\} \tag{1.7}$$

Related works

Quantum channel tomography

[**BL**: Literature review for Lindbladian learning/Prony method for system eigenvalues/Matrix completion]

Notation and outline

We fix the notation used throughout this work. Let $\mathcal{H} \simeq \mathbb{C}^N$ be a finite-dimensional Hilbert space with $N = 2^n$ and $\mathcal{B}(\mathcal{H})$ be the space of bounded operators. We denote by $\mathcal{B}_{sa}(\mathcal{H})$ the subspace of self-adjoint operators and by $\mathcal{B}_{sa}^+(\mathcal{H})$ the cone of positive semidefinite operators. For simplicity, we write $A \geq 0$ (resp., $A > 0$) for a positive semidefinite (resp., definite) operator. We then denote by $\mathcal{D}(\mathcal{H}) := \{\rho \in \mathcal{B}_{sa}^+(\mathcal{H}) ;\ \mathrm{tr}\,\rho = 1\}$ the set of density operators (quantum states), and by $\mathcal{D}_+(\mathcal{H})$ the full-rank density operators. Moreover, let X^* be the adjoint operator of X and $\langle X, Y\rangle = \mathrm{tr}(X^*Y)$ be the Hilbert-Schmidt inner product on $\mathcal{B}(\mathcal{H})$. The adjoint of a superoperator $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ with respect to the inner product $\langle\cdot,\cdot\rangle$ is denoted by $\Phi^\dagger$.

2 Preliminary

2.1 Quantum Markovian dynamics

2.2 Observable and measurements

2.3 Pauli basis and Kossakowski matrix

We can construct the orthonormal basis of the space of traceless matrices which consists of three types, $\{F_i\} = (\{S^{(j,k)}\}, \{J^{(j,k)}\}, \{D^l\})$. Let e_j denote the j-th standard basis vector in $\mathbb{C}^N$ and $E^{(j,k)} := e_j e_k^\top$ denote the matrix in $\mathbb{C}^{N\times N}$ with all entries being zeros, except for the (j,k)-th being one. For $j \neq k$, there are $N(N-1)$ basis matrices, given by

$$S^{(j,k)} = \frac{1}{\sqrt{2}}\left(E^{(j,k)} + E^{(k,j)}\right), 1 \leq j < k \leq N \tag{2.1}$$

$$J^{(j,k)} = \frac{-i}{\sqrt{2}}\left(E^{(j,k)} - E^{(k,j)}\right), 1 \leq j < k \leq N \tag{2.2}$$

For the case of $j = k$, we construct $N - 1$ matrices explicitly with the traceless constraint.

$$D^{(l)} = \frac{1}{\sqrt{l(l+1)}} \left(\sum_{k=1}^{l} E^{(k,k)} - lE^{(l+1,l+1)} \right), l = 1, \ldots, N-1 \tag{2.3}$$

Given the orthonormal basis $\{F_i\}$, we have the following theorem.

Theorem 2.1. *A linear operator $\mathcal{L} : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$ is the generator of a completely positive semigroup of $\mathbb{C}^{N\times N}$ if it can be expressed in the following form*

$$\mathcal{L}\rho = -i[H, \rho] + \frac{1}{2} \sum_{j,k=1}^{N^2-1} c_{kl} \left([F_k, \rho F_l^*] + [F_k \rho, F_l^*] \right) \tag{2.4}$$ {eq_GKS}

where $H = H^*, \mathrm{tr}(H) = 0, \mathrm{tr}(F_k) = 0, \mathrm{tr}(F_j F_k^*) = \delta_{jk}, j, k = 1, 2, \ldots, N^2 - 1$, and $C = (c_{jk})$ is a complex positive matrix.

The above form is often called the GKS (Gorini-Kossakowski-Sudarshan) form of the generator $\mathcal{L}$ and C the Kossakowski matrix. Since the matrix C is positive, it can be diagonalized with a unitary matrix U, therefore

$$UCU^\dagger = \mathrm{diag}(\gamma_1, \ldots, \gamma_{N^2-1}) \tag{2.5}$$

where the eigenvalues γ_i are non-negative. The jump operators can be defined through

$$F_j = \sum_{k=1}^{N^2-1} u_{kj} A_k. \tag{2.6}$$

Note that A_k orthonormal, i.e. $\mathrm{tr}(A_k A_l) = \delta_{kl}$. We have the following diagonal form

$$L\rho = -i[H, \rho] + \sum_{k=1}^{N^2-1} \gamma_k \left(A_k \rho A_k - \frac{1}{2} A_k^\dagger A_k \rho - \frac{1}{2} A_k^\dagger A_k \right). \tag{2.7}$$

Let $J_k = \sqrt{\gamma_k} A_k$, we recover the form of Lindblad equation as in (1.1). Note that the generator is invariant under the following transformations

- Unitary transformations of the set of jump operators,
$$\sqrt{\gamma_i} A_i \to \sqrt{\gamma_i'} A_i' = \sum_{j=1} v_{ij} \sqrt{\gamma_j} A_j$$
where v_{ij} is a unitary matrix.
- Inhomogeneous transformations
$$A_i \to A_i' = A_i + a_i \tag{2.8}$$
$$H \to H' = H + \frac{1}{2i} \sum_j \gamma_j \left(\overline{a_j} A_j - a_j A_j^\dagger \right) + b, \tag{2.9}$$
where a_i are complex numbers and b is real.

By the above transformations, it is always possible to choose traceless Hamiltonian and traceless orthogonal jump operators. For the Hamiltonian H, since it is a Hermitian matrix

$$h_i := \mathrm{tr}(HF_i) \tag{2.10}$$

is a real number for all basis matrices F_i. By the transformation condition 1, we cannot uniquely determine jump operators. Together with the low rank assumption that $N_J \ll N^2 - 1$, we aim to find the following unique representation of the Lindbladian,

$$\{h_i\}_{i=1}^{N^2-1}, \qquad h_i \in \mathbb{R} \tag{2.11}$$
$$C \in \mathbb{C}^{(N^2-1)\times(N^2-1)}, \qquad rank(C) = N_J \tag{2.12}$$

Note that C is assumed to be Hermitian, therefore we can find matrix $U \in \mathbb{C}^{(N^2-1)\times N_J}$ such that $C = UU^\dagger$.

In this case, the Lindbladian can be written as

$$\mathcal{L}\rho = \sum_{j=1}^{N^2-1} h_j g_j(\rho) + \sum_{k,l=1}^{N^2-1} c_{k,l} G_{k,l}(\rho) \tag{2.13}$$

where $g_j(\rho) = -i[F_j, \rho]$ and $G_{k,l}(\rho) = \frac{1}{2}[F_k, \rho F_l^\dagger] + [F_k \rho, F_l^\dagger]$.

3 A framework of learning Lindbladian

The given data is assumed to be observed on a sparse time grid. To estimate the derivative of the trajectory, the first step is to use the Prony method for each entry of the density ρ. We then use the alternating least squares to reconstruct the Hamiltonian and jump operators by exploiting its low-rank structure.

3.1 Loss function

Assume that the initial density $\rho(0)$ follows a distribution μ_0, and then we can define the loss function as

$$\mathcal{E}(\widehat{\mathcal{L}}) = \mathbb{E}_{\mu_0}\left[\int_0^T \left\|\dot{\rho}_t - \widehat{\mathcal{L}}\rho_t\right\|_F^2 dt\right] \tag{3.1}$$

where $\|A\|_F$ represents the Frobenius norm of A and $\widehat{\mathcal{L}}$ represents a candidate Lindbladian. Given M independent density trajectory observed on a uniform discrete mesh $\{t_l\}_{l=1}^L$, $t_l = l\Delta t$, namely the following data

$$\textbf{Data:}\{\rho_{t_l}^m\}_{m,l}^{M,L} \tag{3.2}$$

We also assume a Gaussian random noise for each observation point. Given the observation data, it is natural to define the loss that approximates the above integral since it represents the negative log-likelihood. Namely, we have

$$\mathcal{E}(\widehat{\mathcal{L}}) = \frac{1}{ML}\sum_{m=1}^{M}\sum_{l=1}^{L}\left\|\dot{\rho}_{t_l}^m - \widehat{\mathcal{L}}\rho_{t_l}^m\right\|_F^2 \tag{3.3}$$

3.2 One stage algorithm

Note that the candidate Lindbladian is determined by $\widehat{\mathcal{L}}$ as $(\widehat{h}, \widehat{U})$, we have the loss functions

$$\mathcal{E}(\widehat{h}, \widehat{C}) = \frac{1}{ML}\sum_{m=1}^{M}\sum_{l=1}^{L}\left\|\dot{\rho}_{t_l}^m - \left(\sum_j \widehat{h}_j g_j(\rho_{t_l}^m) + \sum_{k,l} c_{kl} G_{kl}(\rho_{t_l}^m)\right)\right\|_F^2 \tag{3.4}$$

Note that the loss is quadratic in h, but we have not utilize the low rank property of $\widehat{C}$. By taking $\widehat{C} = \widehat{U}\widehat{U}^\dagger$, the loss becomes quartic in $\widehat{U}$. From the later algorithm, we relax the condition of $\widehat{C} = \widehat{U}\widehat{U}^\dagger$ to be $\widehat{C} = \widehat{U}\widehat{V}^\dagger$ and add an extra constraint $\|\widehat{U} - \widehat{V}\|_F^2$. Therefore the loss becomes

$$\mathcal{E}(\widehat{h}, \widehat{U}, \widehat{V}) = \frac{1}{ML}\sum_{m=1}^{M}\sum_{l=1}^{L}\left\|\dot{\rho}_{t_l}^m - \left(\sum_j \widehat{h}_j g_j(\rho_{t_l}^m) + \sum_{k,j,l} u_{k,j} v_{j,l}^\dagger G_{kl}(\rho_{t_l}^m)\right)\right\|_F^2 + \lambda\|\widehat{U} - \widehat{V}\|_F^2 \tag{3.5}$$

3.2.1 Alternating Least Squares

We use the alternating least squares (ALS) to minimize the loss function $\mathcal{E}$. For the triplet $(\widehat{h}, \widehat{U}, \widehat{V})$, the loss functions is quadratic for each one if the other two terms are fixed. It is direct to construct linear regression. When N is large, we can apply gradient descent for each step of the least square.

3.3 Two stage algorithm

The previous method refers to estimate the parameters of the Lindbladian operator over the basis matrices $\{F_i\}$, which guaranteed the uniqueness. But the computation of the projection of ρ on to the basis can be expensive, therefore we introduce the two stage algorithm, which refers to exploiting the low rank structure of the Lindbladian operator first, and then perform the decomposition over the basis $\{F_i\}$.

3.3.1 Vectorization

Using the fact that

$$\mathbf{vec}(AXB) = (B^\top \otimes A)\mathbf{vec}(X) \tag{3.6}$$

where **vec** is the vectorization of a matrix in $\mathbb{C}^{N\times N}$ to a vector in $\mathbb{C}^{N^2}$ in the column-first order, we can write the vectorized version of the equation (1.1).

$$\mathbf{vec}(\rho)'(t) = \left\{-i(I\otimes H - H^\top \otimes I) + \sum_{k=1}^{N_J}\left[\overline{J_k}\otimes J_k - \frac{1}{2}I\otimes(J_k^\dagger J_k) - \frac{1}{2}\overline{(J_k^\dagger J_k)}\otimes I\right]\right\}\mathbf{vec}(\rho)(t) := \mathbf{L}\,\mathbf{vec}(\rho)(t). \tag{3.7}$$

where the super operator is denoted as $\mathbf{L}$. Note that $\mathbf{L}$ has only $N_K = 2 + 3N_J$ Kronecker products in the summation, where the subscript K represents the number of Kronecker terms. We can write $\mathbf{L}$ as

$$\mathbf{L} = \sum_{k=1}^{N_K} \mathbf{V}_k \otimes \mathbf{U}_k \tag{3.8}$$

where $\mathbf{V}_k$ and $\mathbf{U}_k$ are matrices in $\mathbb{C}^{N\times N}$.

And the corresponding operator $\mathcal{L}$ is written as

$$\mathcal{L}\rho = \sum_{k=1}^{\widetilde{r}} \mathbf{U}_k \rho \mathbf{V}_k^\top \tag{3.9}$$

3.3.2 Reshaping

We apply the reshaping operator $\mathcal{R}$, defined as in [VLP93]. If matrix $A \in \mathbb{C}^{N^2\times N^2}$, define the rearrangement of A (relative to the blocking parameters N) by

$$A = \begin{bmatrix} A_{11} \dots A_{1n} \\ \vdots \ddots \vdots \\ A_{n1} \dots A_{nn} \end{bmatrix}, \quad \mathcal{R}(A) = \begin{bmatrix} \widetilde{A}_1 \\ \vdots \\ \widetilde{A}_n \end{bmatrix}, \quad \widetilde{A}_j = \begin{bmatrix} \mathbf{vec}(A_{1j})^\top \\ \vdots \\ \mathbf{vec}(A_{nj})^\top \end{bmatrix} \tag{3.10}$$

This rearrangement operator $\mathcal{R}$ gives

$$\mathcal{R}(A \otimes B) = \mathbf{vec}(A)\mathbf{vec}^\top(B), \tag{3.11}$$

Therefore,

$$\mathbf{E} := \mathcal{R}(\mathbf{L}) = \sum_{k=1}^{\widetilde{r}} \mathbf{vec}(\mathbf{V}_k)\mathbf{vec}^\top(\mathbf{U}_k). \tag{3.12}$$

That is to say, the matrix $\mathbf{E}$ has low rank N_K compared to N^2.

3.3.3 Low rank

In order to use the rearranged low-rank property, we vectorize the density matrix ρ and decompose the Frobenius norm as

$$\mathcal{E}(\widehat{\mathcal{L}}) = \frac{1}{ML} \sum_{m=1}^{M} \sum_{l=1}^{L} \sum_{i,j=1}^{N} \left| (\dot{\rho}_{t_l}^m)_{i,j} - (\widehat{\mathcal{L}}\rho_{t_l}^m)_{i,j} \right|^2. \tag{3.13}$$

Notice that

$$(\mathbf{U}_k \rho \mathbf{V}_k^\top)_{i,j} = \mathbf{vec}(\mathbf{U}_k)^\top P_{ij} \mathbf{vec}(\mathbf{V}_k) = \langle P_{ij}, \mathbf{vec}(\mathbf{U}_k)\mathbf{vec}(\mathbf{V}_k)^\top \rangle_F \tag{3.14}$$

where $P_{ij} = e_{ij} \otimes \rho$, and e_{ij} is the $n \times n$ matrix that are all zeros except for the (i, j) entry equals to 1, and $\langle \cdot, \cdot \rangle_F$ represents the Frobenius inner product of $\mathbb{R}^{n^2\times n^2}$. Hence

$$(\widehat{\mathcal{L}}\rho_{t_l}^m)_{i,j} = \left(\sum_{k=1}^{\widetilde{r}} \mathbf{U}_k \rho_{t_l}^m \mathbf{V}_k \right)_{i,j} = \sum_{k=1}^{\widetilde{r}} \langle P_{t_l,ij}^m, \mathbf{vec}(\mathbf{U}_k)\mathbf{vec}(\mathbf{V}_k)^\top \rangle_F = \langle P_{t_l,ij}^m, \widehat{\mathbf{E}} \rangle_F \tag{3.15}$$

and $P_{t_l,ij}^m = e_{ij} \otimes \rho_{t_l}^m$. Therefore the loss function can be written with respect to the reshaped unknown matrix $\widehat{\mathbf{E}}$

$$\mathcal{E}(\widehat{\mathcal{L}}) = \mathcal{E}(\widehat{\mathbf{E}}) = \frac{1}{ML} \sum_{m=1}^{M} \sum_{l=1}^{L} \sum_{i,j=1}^{n} \left| (\rho_{t_l}^m)_{ij} - \langle P_{t_l,ij}^m, \widehat{\mathbf{E}} \rangle_F \right|^2. \tag{3.16}$$

3.3.4 Connection to Matrix sensing

In the language of matrix sensing, the matrix $P_{t_l,ij}^m$ can be considered as the measurement matrices, with the corresponding output $(\rho_{t_l}^m)_{ij}$. Now we define the operator $\mathcal{A} : \mathbb{R}^{n^2\times n^2} \to \mathbb{R}^{MLn^2}$ as

$$\mathcal{A}(\widehat{\mathbf{E}}) = \mathbf{vec}\left[(\langle P_{t_l,ij}^m, \widehat{\mathbf{E}} \rangle_F)_{m,l,i,j} \right] \tag{3.17}$$

and the vector $\mathbf{y} \in \mathbb{R}^{MLn^2}$ as $\mathbf{vec}\left[(\rho_{t_l}^m)_{ij}\right]$, we have the loss function written as

$$\mathcal{E}(\widehat{E}) = \left\| \mathcal{A}(\widehat{\mathbf{E}}) - \mathbf{y} \right\|^2 \tag{3.18}$$

up to a constant.

3.4 Prony fitting of the trajectory

[**BL**: We use Prony method to estimate the spectrum of the system and the derivative of the trajectory. We then use the alternating least squares to reconstruct the Lindbladian by exploiting its low-rank structure.] To increase the accuracy of the estimation of the derivatives, we use the Prony method to fit the trajectories $\langle A\rangle$. Since the solution can be written as $\rho(t) = e^{t\mathbf{L}}\rho(0)$, and $\mathbf{L}$ contains the jump operators, it converges to a steady state exponentially. In the case the $\mathbf{L}$ is diagonalizable, $\mathbf{L} = U\Sigma U^{-1}$, where $\Sigma = \mathrm{diag}\{\lambda_1, \dots, \lambda_{n^2}\}$(possibly repeated). Then

$$\rho(t) = Ue^{t\Sigma}U^{-1}\rho(0) \tag{3.19}$$

4 temp

4.0.1 Trajectory based settings

Assume that we have M independent density trajectory observed on a uniform discrete mesh $\{t_l\}_{l=1}^L$, $t_l = l\Delta t$, namely the following data

$$\textbf{Data:}\{\rho_{t_l}^m\}_{m,l}^{M,L} \tag{4.1}$$

We also assume a Gaussian random noise for each observation point. Our goal is to estimate the Hamiltonian H and the jump operators C_k. By the correspondence between $\{H, C_k\}$ and $\{\mathbf{V}_k, \mathbf{U}_k\}$, it is equivalent to estimate $\mathbf{L}$, and therefore, $\mathcal{R}(\mathbf{L})$ from the data.

We can assume that the initial density $\rho(0)$ follows a distribution μ_0, and then we can define the loss function as

$$\mathcal{E}(\widehat{\mathcal{L}}) = \mathbb{E}_{\mu_0}\left[\int_0^T \left\|\dot{\rho}_t - \widehat{\mathcal{L}}\rho_t\right\|_F^2 dt\right] \tag{4.2}$$

where $\|A\|_F$ represents the Frobenius norm of A. Given the observation data, it is natural to define the loss that approximates the above integral since it represents the negative log-likelihood of the unknown matrix $\mathbf{L}$. Namely, we have

$$\mathcal{E}(\widehat{\mathcal{L}}) = \frac{1}{ML}\sum_{m=1}^{M}\sum_{l=1}^{L}\left\|\dot{\rho}_{t_l}^m - \widehat{\mathcal{L}}\rho_{t_l}^m\right\|_F^2 \tag{4.3}$$

In order to use the rearranged low-rank property, we vectorize the density matrix ρ and decompose the Frobenius norm as

$$\mathcal{E}(\widehat{\mathcal{L}}) = \frac{1}{ML}\sum_{m=1}^{M}\sum_{l=1}^{L}\sum_{i,j=1}^{n}\left|(\dot{\rho}_{t_l}^m)_{i,j} - (\widehat{\mathcal{L}}\rho_{t_l}^m)_{i,j}\right|^2. \tag{4.4}$$

Notice that

$$(\mathbf{U}_k\rho\mathbf{V}_k^\top)_{i,j} = \mathbf{vec}(\mathbf{U}_k)^\top P_{ij}\mathbf{vec}(\mathbf{V}_k) = \langle P_{ij}, \mathbf{vec}(\mathbf{U}_k)\mathbf{vec}(\mathbf{V}_k)^\top\rangle_F \tag{4.5}$$

where $P_{ij} = e_{ij}\otimes\rho$, and e_{ij} is the $n\times n$ matrix that are all zeros except for the (i,j) entry equals to 1, and $\langle\cdot,\cdot\rangle_F$ represents the Frobenius inner product of $\mathbb{R}^{n^2\times n^2}$. Hence

$$(\widehat{\mathcal{L}}\rho_{t_l}^m)_{i,j} = \left(\sum_{k=1}^{\widetilde{r}}\mathbf{U}_k\rho_{t_l}^m\mathbf{V}_k\right)_{i,j} = \sum_{k=1}^{\widetilde{r}}\langle P_{t_l,ij}^m, \mathbf{vec}(\mathbf{U}_k)\mathbf{vec}(\mathbf{V}_k)^\top\rangle_F = \langle P_{t_l,ij}^m, \widehat{\mathbf{E}}\rangle_F \tag{4.6}$$

and $P_{t_l,ij}^m = e_{ij}\otimes\rho_{t_l}^m$. Therefore the loss function can be written with respect to the reshaped unknown matrix $\widehat{\mathbf{E}}$

$$\mathcal{E}(\widehat{\mathcal{L}}) = \mathcal{E}(\widehat{\mathbf{E}}) = \frac{1}{ML}\sum_{m=1}^{M}\sum_{l=1}^{L}\sum_{i,j=1}^{n}\left|(\rho_{t_l}^m)_{ij} - \langle P_{t_l,ij}^m, \widehat{\mathbf{E}}\rangle_F\right|^2. \tag{4.7}$$

4.0.2 Connection to Matrix sensing

In the language of matrix sensing, the matrix $P_{t_l,ij}^m$ can be considered as the measurement matrices, with the corresponding output $(\rho_{t_l}^m)_{ij}$. Now we define the operator $\mathcal{A}: \mathbb{R}^{n^2\times n^2}\to\mathbb{R}^{MLn^2}$ as

$$\mathcal{A}(\widehat{\mathbf{E}}) = \mathbf{vec}\left[(\langle P_{t_l,ij}^m, \widehat{\mathbf{E}}\rangle_F)_{m,l,i,j}\right] \tag{4.8}$$

and the vector $\mathbf{y}\in\mathbb{R}^{MLn^2}$ as $\mathbf{vec}\left[(\rho_{t_l}^m)_{ij}\right]$, we have the loss function written as

$$\mathcal{E}(\widehat{E}) = \left\|\mathcal{A}(\widehat{\mathbf{E}}) - \mathbf{y}\right\|^2 \tag{4.9}$$

up to a constant.

4.0.3 Using observables Not finished yet

Assume that we have M independent density trajectory, a set of observables $\{A_q\}_{q=1}^Q$ so that the density trajectory is observed on a uniform discrete mesh $\{t_l\}_{l=1}^L$, $t_l = l\Delta t$, namely the following data

$$\textbf{Data:}\{\langle A_q\rangle^m(t_l)\}_{m,q,l}^{M,Q,L} \tag{4.10}$$

We also assume a Gaussian random noise for each observation point. Our goal is to estimate the Hamiltonian H and the jump operators C_k. By the correspondence between $\{H, C_k\}$ and $\{\mathbf{V}_k, \mathbf{U}_k\}$, it is equivalent to estimate $\mathbf{L}$, and therefore, $\mathcal{R}(\mathbf{L})$ from the data.

It is natural to define the loss function as

$$\mathcal{E}(\widehat{\mathbf{L}}) = \frac{1}{Q}\sum_{q=1}^{Q}\int_0^T \left\|\langle A_q\rangle'(t) - 1\right\| \tag{4.11}$$

5 Theoretical analysis

5.1 Spectrum and derivative estimation

[**BL**: We analyze the spectrum/derivative estimation error in terms of measurement number/sample complexity and total Lindbladian simulation time]

5.2 Lindbladian estimation

[**BL**: We analyze the scaling of the Lindbladian reconstruction error in the derivative estimation error]

5.2.1 Problem settings

Suppose we observe data from

$$y = \mathcal{A}(M) + z \tag{5.1}$$

where M is an unknown $n_1 \times n_2$ matrix, $\mathcal{A}: \mathbb{R}^{n_1\times n_2} \to \mathbb{R}^m$ is a linear mapping, and z is an m-dimensional noise term. The goal is to recover a good approximation of M while requiring as few measurements as possible. Each measurement is interpreted as

$$[\mathcal{A}(M)]_i = \langle A_i, M\rangle_F = y_i \tag{5.2}$$

where A_i is a $n_1 \times n_2$ matrix.

We first introduce the isometry constants of the linear map $\mathcal{A}$.

Definition 5.1. *For each integer $r = 1, 2, \ldots, n$, the isometry constant δ_r of $\mathcal{A}$ is the smallest quantity such that*

$$(1-\delta_r)\|X\|_F^2 \le \|\mathcal{A}(X)\|^2 \le (1+\delta_r)\|X\|_F^2 \tag{5.3}$$

holds for all matrices X of rank at most r.

We say that $\mathcal{A}$ satisfies the RIP at rank r if δ_r is bounded by a sufficiently small constant between 0 and 1.

5.2.2 Gaussian Measurement

An essential example shows that the Gaussian measurement has RIP.

Definition 5.2. *$\mathcal{A}$ is a Gaussian measurement operator if each measurement matrix A_i, $1 \le i \le m$, contains i.i.d. $\mathcal{N}(0, 1/m)$ entries, and A_i are independent from each other.*

Theorem 5.1. *[CP, Theorem 2.3] Fix $0 \le \delta < q$ and let $\mathcal{A}$ be a random measurement ensemble obeying the following condition: for any given $X \in \mathbb{R}^{n_2\times n_2}$ and any fixed $0 < t < 1$,*

$$\mathbb{P}\left(\left|\|A(X)\|^2 - \|X\|_F^2\right| > t\|X\|_F^2\right) \le C\exp(-cm) \tag{5.4}$$

for fixed constants $C, c > 0$ (which may depend on t). Then if $m \ge Dnr$, $\mathcal{A}$ satisfies the RIP with isometry constant $\delta_r \le \delta$ with probability exceeding $1 - Ce^{-dm}$ for fixed constants $D, d > 0$.

If $\mathcal{A}$ is a Gaussian random measurement ensemble, $\|\mathcal{A}(X)\|^2$ is distributed as $m^{-1}\|X\|_F^2$ times a chi-squared random variable with m degrees of freedom and from standard concentration inequalities,

$$\mathbb{P}\left(\left|\|A(X)\|^2 - \|X\|_F^2\right| > t\|X\|_F^2\right) \le 2\exp\left(-\frac{m}{2}(t^2/2 - t^3/3)\right) \tag{5.5}$$

Same result holds if $\mathcal{A}$ is a random projection, or A_i contains sub-Gaussian entries.

5.2.3 Pauli Measurement

Let $M \in \mathbb{C}^{n\times n}$ be an unknown matrix of rank at most r. Let $W_1, \ldots, W_{n^2}$ be an orthonormal basis for $\mathbb{C}^{n\times n}$, with respect to the Frobenius inner product. We choose m basis elements, $S_1, \ldots, S_m$, i.i.d. uniformly at random from $\{W_1, \ldots, W_{n^2}\}$. We then observe the coefficients $\langle S_i, M\rangle$.

Definition 5.3. *We say the basis $\{W_1, \ldots, W_{n^2}\}$ is **incoherent** if the W_i all have small operator norm,*

$$\|W_i\| \leq K/\sqrt{n}. \tag{5.6}$$

where K is a constant.

In the special case that W_i is given by Pauli matrices $P_1 \otimes \cdots \otimes P_n/\sqrt{d}$, they are incoherent with $\|W_i\| \leq K/\sqrt{d}$ with $K = 1$.

The linear operator $\mathcal{A}$ is constructed by taking the i-th measurement as S_i.

Definition 5.4. *For a matrix X, its Schatten p-norm is defined as*

$$\|X\|_p = \left(\sum_i \sigma_i(X)^p\right)^{1/p}, \tag{5.7}$$

where $\sigma_i(X)$ is the i-th singular value of X. In particular, $\|X\|_* = \|X\|_1$ is the trace norm or nuclear norm.

Theorem 5.2. *[Liu11, Theorem 2.1] Fix some constant $0 \leq \delta < 1$. Let $\{W_1, \ldots, W_{n^2}\}$ be an orthonormal basis for $\mathbb{C}^{n\times n}$ that is incoherent. Let $m = CK^2 \cdot rd\log^6 d$, for some constant C that depends only on $\delta, C = O(1/\delta^2)$. Let $\mathcal{A}$ be defined as above. Then with high probability over the selection of $\{S_1, \ldots, S_m\}$, $\mathcal{A}$ satisfies the RIP over the set of all $X \in \mathbb{C}^{d\times d}$ such that $\|X\|_* \leq \sqrt{r}\|X\|_F$. Furthermore, the failure probability is exponentially small in $\delta^2 C$.*

6 Numerical experiments

Given a discrete time mesh $t_n = n\Delta t$, generate data $\rho(t_n)$ using the Python package mesolve. Try to learn the quantum channels of $e^{t_n \mathbf{L}}$ using the ALS code, and then compare the eigenvalues of each result.

References

[CP] Emmanuel J Candes and Yaniv Plan. ``Tight Oracle Bounds for Low-rank Matrix Recovery from a Minimal Number of Random Measurements''. In: ().

[Liu11] Yi-Kai Liu. *Universal Low-Rank Matrix Recovery from Pauli Measurements*. 2, 2011. arXiv: 1103.2816 [quant-ph, stat]. URL: http://arxiv.org/abs/1103.2816 (visited on 02/29/2024). preprint.

[VLP93] C. F. Van Loan and N. Pitsianis. ``Approximation with Kronecker Products''. In: *Linear Algebra for Large Scale and Real-Time Applications*. Ed. by Marc S. Moonen, Gene H. Golub, and Bart L. R. De Moor. Dordrecht: Springer Netherlands, 1993, pp. 293–314. (Visited on 03/20/2024).