This problem set will cover concepts from learning noisy quantum devices and learning shallow QNNs and SRE states. If you find that you are running low on time to finish all the problems, our recommendation is to try to aim for breadth rather than depth – e.g., it is better to complete a few parts, than to completely solve one of the three questions and skip the others. You could assume the result from the optional problem 2.C in answering other problems. Below we provide hints for the various problems in this assignment. While these may help you solve the problems more easily, you are not required to follow the hints as long as the proofs you provide are correct.

1 (45 + 10 Optional pts.) SPAM-Robust Learning of Hamiltonians

Introduction. Consider a noisy single-qubit quantum device that can perform the following operations:

- Noisy state preparation: $\rho_0 = \frac{1}{2}I + \sum_{P \in \{X,Y,Z\}} \alpha_P P \approx |0\rangle\langle 0|$ (you can assume the noise in the initial state is a very small constant).
- Noisy measurement (after Pauli twirling): $\mathcal{M}_0 = \{\mathcal{N}^\dagger(|0\rangle\langle 0|), \mathcal{N}^\dagger(|1\rangle\langle 1|)\}$ for a noise channel $\mathcal{N}(\rho) = (1-\lambda)\rho + \lambda \operatorname{tr}(\rho) I/2$ (you can assume the noise is a very small constant).
- Perfect single-qubit Clifford gates: $C \in \mathrm{Cl}_1$.
- Unknown Hamiltonian dynamics: $U(t) = e^{-it\sum_{Q\in\{X,Y,Z\}} \beta_Q Q}$.

The unknown parameters are $\alpha_X, \alpha_Y, \alpha_Z, \lambda, \beta_X, \beta_Y, \beta_Z$. Our goal is to understand what are learnable/unlearnable and how we can efficiently reconstruct the Hamiltonian parameters β_Q.

Part 1: Characterizing state preparation and measurement (SPAM). We begin by conducting experiments that do not involve the unknown Hamiltonian dynamics. The goal is to understand what is learnable among the unknown parameters $\alpha_X, \alpha_Y, \alpha_Z, \lambda$.

1.A. (5 pts.) Consider experiments of the form: (1) noisy state preparation, (2) many single-qubit Clifford gates, (3) noisy measurement (i.e., all experiments that do not involve the unknown Hamiltonian dynamics). Find a one-dimensional gauge transformation in the four-dimensional space defined by $\alpha_X, \alpha_Y, \alpha_Z, \lambda$ that leads to many different noisy devices that are impossible to distinguish.

1.B. (5 pts.) Prove that $\alpha_X, \alpha_Y, \alpha_Z$ can be learned to ε error from $\mathcal{O}(1/\varepsilon^2)$ experiments in a noisy quantum device with the promise that there is no noise on the measurement ($\lambda = 0$). Find the quantities that can be learned without such a strong promise and prove that they can be learned to ε error from $\mathcal{O}(1/\varepsilon^2)$ experiments.

Part 2: Transforming the device. We now consider all possible experiments that can be done by this noisy quantum device and develop two important transformation (the first one is a gauge transformation that removes the unlearnable from Part 1 and the second one is a transformation that expands the set of physical operations we can perform).

1.C. (5 pts.) Show that there is a gauge transformation from any device to a device with no measurement noise such that all experiments involving all operations (state prep, measurement, Clifford gates, Hamiltonian dynamics) remain unchanged. We can now operate in this device where the measurement is perfect and the noisy initial state is known up to ε error.

1.D. (10 pts.) We now develop the error analysis for Trotter formula. Given any Hamiltonian terms $h_1, \ldots, h_m$ with spectral norm bounded above by 1 over any number of qubits. Consider the following three unitaries:

- Ideal evolution: $V(t) = e^{-it\sum_{i=1}^m h_i}$.
- First-order Trotter: $V^{(1)}(t) = e^{-ith_m} \ldots e^{-ith_1}$.
- Second-order Trotter: $V^{(2)}(t) = e^{-i(t/2)h_1} \ldots e^{-i(t/2)h_m} e^{-i(t/2)h_m} \ldots e^{-i(t/2)h_1}$.

Prove that $\|V(t) - V^{(1)}(t)\|_\infty \leqslant \mathcal{O}(m^2t^2)$ and $\|V(t) - V^{(2)}(t)\|_\infty \leqslant \mathcal{O}(m^3t^3)$ as $t \to 0$ using the matrix Taylor expansion $e^X = I + X + \frac{X^2}{2} + \frac{X^3}{3!} + \ldots$. Use these norm bounds to show how we can simulate $V(\tau)$ up to ε error in spectral norm for any τ by composing $\mathcal{O}(\tau^{1.5}/\varepsilon^{0.5})$ number of e^{-ith_i} for varying i's and t's.

1.E. (5 pts.) Show how to construct the simplified unknown Hamiltonian dynamics: $U_Q(t) = e^{-it\beta_Q Q}$ for any $Q \in \{X, Y, Z\}$ using the Trotter formula. How many physical operations (in big-O notation) are needed to achieve ε-approximate simulation? (Hint: Use the reshaping technique from class.)

Part 3: Learning the Hamiltonian. We are now ready to learn the unknown Hamiltonian.

1.F. (5 PTS.) Show how to combine the perfect single-qubit Clifford gates and the noisy initial state ρ_0 in the gauge-transformed device to create the state $\rho_P = \frac{I+\alpha_P P}{2}$ for any $P \in \{X, Y, Z\}$.

1.G. (10 PTS.) Using the new physical operations we have created from the previous subproblems, prove that we can learn β_Q to ε error for all $Q \in \{X, Y, Z\}$ using a total of $\mathcal{O}(1/\varepsilon^2)$ experiments.

1.H. (10 OPTIONAL PTS.) Prove that we can learn β_Q to ε error for all $Q \in \{X, Y, Z\}$ using a total of $\tilde{\mathcal{O}}(1/\varepsilon)$ experiments.

2 (55 + 10 OPTIONAL PTS.) LEARNING LONG-RANGE ENTANGLED STATES

Introduction. In the course, we see how to learn short-range entangled (SRE) states. The goal of this problem is to devise an algorithm for learning and producing n-qubit quantum states on a 2D grid that possess long-range entanglement (LRE). The general algorithmic outline is the following: (1) learn the local reduced density matrices (RDMs) of the unknown state over constant-size patches, (2) use these RDMs to produce some useful objects (local inversions in the class; recovery maps for this problem), (3) show that these objects are sufficient to reconstruct the unknown state.

Part 1: Classical Markov Chains. Let us recall classical Markov chains with the following warm-up problems.

2.A. (5 PTS.) Consider a classical probability distribution p over n bits. Let $X_1, \ldots, X_n$ be the random variables associated with the n bits. The distribution p is a classical Markov chain if it satisfies the following:

$$p(X_i = x_i | X_1 = x_1, \ldots, X_{i-1} = x_{i-1}) = p(X_i = x_i | X_{i-1} = x_{i-1})$$

for any $x_1, \ldots, x_{i-1}$ such that $p(X_1 = x_1, \ldots, X_{i-1} = x_{i-1}) > 0$. We are given samples from the distribution p. Show how to efficiently learn the 2-bit marginal distributions $p(X_{i-1} = x_{i-1}, X_i = x_i)$. Show how given complete knowledge of the 2-bit marginal distributions enable us to reconstruct p in the sense that we can sample from p.

2.B. (10 PTS.) We can generalize the notion of Markov chains to higher dimensions. Suppose the classical probability distribution p is now a classical Markov chain (also known as Markov random field) over n bits laid out on a 2D grid satisfying

$$p(X_C = x_C | X_A = x_A, X_B = x_B) = p(X_C = x_C | X_B = x_B),$$

for any arbitrarily large subset of bits A and any geometrically-local constant-size patches B, C, s.t., the minimum distance between A and C is larger than a constant, and any x_A, x_B, s.t., $p(X_A = x_A, X_B = x_B) > 0$. We are given samples from the distribution p. Show how to efficiently learn the marginal distributions $p(X_B = x_B, X_C = x_C)$ and how given complete knowledge of these marginal distributions we can reconstruct p. You can use a pictorial proof for the reconstruction step. You should think about how to satisfy the constraint that A and C must be separated.

From these warm-up problems, you can see that marginals can be used to produce these useful objects that serve as a recovery map that can sequentially grow the probability distribution to reconstruct the unknown distribution p that we wish to learn. We will now take this understanding to the context of quantum states.

Part 2: Quantum Markov Chains. Consider an unknown state ρ_0 that is a quantum Markov chain (QMC). A QMC state ρ is a quantum generalization of a classical Markov chain. Many well-known representative states (those with zero correlation length) in a quantum phase of matter are QMC states. For example, the $|0...0\rangle$ state representing trivial phase is a QMC state; the toric code ground state on a 2D plane encoding $k = 0$ logical qubits representing the toric code phase is also a QMC state.

Throughout this problem, we will focus on n qubits laid out on a 2D grid. The same algorithm generalizes to higher dimension, but we will focus solely on 2D lattices for simplicity. Similar to classical Markov chain that enables one to sample bits in patch C using information only in patch B, a QMC state guarantees that for any arbitrarily large subset of qubits A and any geometrically-local constant-size patches B, C, s.t., the minimum distance between A and C is larger than a constant, there exists a channel $\Phi_{B \to BC}$ (known as the recovery map) taking qubits in patch B to patch BC such that the following holds:

$$\rho_{ABC} = (\mathbb{1}_A \otimes \Phi_{B \to BC})(\rho_{AB}),$$

where $\rho_{AB} = \mathrm{tr}_C(\rho_{ABC})$ and ρ_{ABC} are the reduced density matrices (RDMs) of the QMC state ρ_0.

The formal definition of a quantum Markov chain (QMC) is the following: a quantum state ρ_{ABC} is said to be a QMC if and only if the conditional mutual information $I(A:C|B) = 0$ on ρ_{ABC}. A powerful fact is that for any state ρ_{ABC} with $I(A:C|B) = 0$, there always exists a decomposition of the qubit system B into a direct sum (Cartesian product) of many tensor product subspaces $b_j^L \otimes b_j^R$ such that:

$$\rho_{ABC} = \bigoplus_j p_j \left(\rho_{Ab_j^L} \otimes \rho_{b_j^R C} \right),$$

where p_j is a probability distribution (with $p_j > 0$) and $\bigoplus_j$ is the direct sum (in the matrix representation, they form a block diagonal matrix). You can understand this decomposition as a probabilistic mixture of different orthogonal subspaces within B, such that these subspaces are in a tensor product form $b_j^L \otimes b_j^R$, where b_j^L is entangled only with A while b_j^R is entangled only with C. We begin by proving some fundamental properties of the recovery map $\Phi_{B\to BC}$, then we show how to find them locally, and how to reconstruct the full state ρ_0 using these recovery maps.

2.C. (10 OPTIONAL PTS.) **Existence of $\Phi_{B\to BC}$.** Prove that there exists a purification $|\psi_{ABCE}\rangle$ of a QMC state ρ_{ABC} for a purifying register E and exists a CPTP map $\Phi_{B\to BC}$, such that

$$|\psi_{ABCE}\rangle\langle\psi_{ABCE}| = (\mathbb{1}_{AE} \otimes \Phi_{B\to BC})(\mathrm{Tr}_C(|\psi_{ABCE}\rangle\langle\psi_{ABCE}|)). \tag{1}$$

This immediately implies the existence of a recovery map from B to BC for ρ_{ABC}.

2.D. (10 PTS.) **Basic properties of $\Phi_{B\to BC}$.** Prove that for any purification $|\psi_{BCE}\rangle$ of ρ_{BC}, the recovery map $\Phi_{B\to BC}$ satisfying Eq. (1) also satisfies

$$|\psi_{BCE}\rangle\langle\psi_{BCE}| = (\mathbb{1}_{E} \otimes \Phi_{B\to BC})(\mathrm{Tr}_C(|\psi_{BCE}\rangle\langle\psi_{BCE}|)). \tag{2}$$

Prove that conversely the recovery map $\Phi_{B\to BC}$ satisfying Eq. (2) for some purification also satisfies Eq. (1).

2.E. (5 PTS.) **Finding the recovery map $\Phi_{B\to BC}$ for ρ_{ABC}.** Given that A can be arbitrarily large, we cannot perform quantum state tomography on ρ_{ABC}. We can only perform quantum state tomography on ρ_{BC} to learn this constant-size patch. Show that one can find a recovery map $\Phi_{B\to BC}$ for ρ_{ABC} from only the knowledge of ρ_{BC}. That is, show that $\rho_{ABC} = (\mathbb{1}_A \otimes \Phi_{B\to BC})(\rho_{AB})$.

2.F. (5 PTS.) **Reconstructing the unknown state ρ_0.** Show pictorially how to reconstruct the state ρ_0 on the 2D plane using the recovery maps $\Phi_{B\to BC}$ for the constant-size patches B, C.

Part 3: LRE States. Now, suppose the unknown LRE state of interest is $\rho = U\rho_0 U^\dagger$, where U is an unknown shallow, depth-d local quantum circuit of a QMC state ρ_0 (such as the unique toric code ground state on a 2D plane with uniform boundary). The shallow quantum circuit breaks the QMC property. Instead, we consider a slightly weaker notion of recovery map. For the state ρ, for an arbitrarily large subsystem A and (sufficiently large) constant-size patches B, C, such that A, C is separated by a large enough constant distance, a new extending map $\Gamma : \mathbb{D}_B \to \mathbb{D}_{B'C}$ exists, where B' is a slightly shrunken subsystem of B (so $B' \subsetneq B$), to correctly "grow" the state, i.e.,

$$\rho_{AB'C} = (\mathbb{1}_A \otimes \Gamma_{B\to B'C})(\rho_{AB}).$$

Note that B' only need to be shrunken near the edges of the boundary between the two neighboring (sufficiently large) constant-size patches B and C. You only need to consider A, B, C that arises from the reconstruction procedure in 2.H (which are similar to 2.B and 2.E but slightly different due to the weaker extending map).

2.G. (10 PTS.) **Proving the existence of Γ.** Show how the existence of Φ for ρ_0 (from Part 2) and the shallowness of U together imply the existence of Γ. (Hint: The proof is constructive. You can build Γ from $\Phi, U, U^\dagger$. Think about how one could reveal a subsystem via local inversion, use the recovery map to grow the state, perform inverse of the local inversion, and trace out the qubits in the neighborhood of B, C)

2.H. (5 PTS.) **Finding the map Γ.** Show that one can find the new extending map Γ from $\rho_{\tilde{B}\tilde{C}}$, where $\tilde{B}$ and $\tilde{C}$ are slightly thickened version of B and C. (Hint: use similar ideas as the recovery map for QMC states)

2.I. (10 PTS.) **Reconstructing the unknown state ρ.** Show how to create the state ρ on the 2D plane using the map Γ. (Hint: The algorithm is the same sequential sweep as classical Markov chain and QMC state. The only difference is that the extending map is slightly more complex because the extension eats the edges. Similar to all previous subproblems, you only need to write a pictorial proof.)