1 Problem Description

1.1 Model

The previous section has provided the background necessary to present the stochastic control problem we aim to solve using reinforcement learning. Consider the problem of preparing a quantum state $|\psi_f\rangle$ from an initial state $|\psi_i\rangle$ in a finite time τ in n steps. We study the Hamiltonian introduced in Eqn. **??** for spin in a magnetic field. At each time t, a constant magnetic field $h_t = (h_t^x, h_t^z)$ is applied for the time step $\Delta t = \frac{\tau}{n}$ and some dephasing noise η_t drawn from a zero-mean Gaussian distribution is present. Together this yields the Hamiltonion in Eqn. 1.

$$H_t = \eta_t \sigma_z - h_t^x \sigma_x - h_t^z \sigma_z \tag{1}$$

Overall, the state evolves according to the model in Eqn. 2.

$$|\psi_{t+1}\rangle = e^{-iH_t \Delta t} |\psi_t\rangle \tag{2}$$

The set of states is given by

$$\mathbb{X} := \left\{ |\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |1\rangle : 0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi \right\}$$

which correspond to points on the Bloch sphere. In practice, there is a maximum and minumum possible magnetic field strength within which the control can be tuned so the set of controls can be simplified to

$$\mathbb{U} := \{h = (h^x, h^z) : h_i \in \{h_{min}, h_{min} + \Delta, h_{min} + 2\Delta, \ldots, h_{max}\},\ i = x, z\}$$

where $\Delta = (h_{max} - h_{min})/M$, so that each component of the control field can take $M + 1$ values including h_{min} and h_{max} for some $M \in \mathbb{Z}_{>0}$.

The fidelity between the two states $|\psi_0\rangle$ and $|\psi_1\rangle$ is given by in Eqn. 3.

$$F(\psi_0, \psi_1) = |\langle\psi_0|\psi_1\rangle|^2 \tag{3}$$

Consider the cost function in Eqn. 4, which rewards fidelity between the current state $|\psi_t\rangle$ and the target state $|\psi_f\rangle$.

$$c(|\psi_t\rangle, h_t) = \begin{cases} 100F(\psi_t, \psi_f)^3 & 0 \leq F(\psi_t, \psi_f) < 0.99 \\ 5000 & 0.99 \leq F(\psi_t, \psi_f) \leq 1 \end{cases} \tag{4}$$

Altogether, the model in Eqn. 2, the state space $\mathbb{X}$, the action space $\mathbb{U}$, and the cost function in Eqn. 4 constitue a Markov control problem.

1.2 Controllability

Before proceeding with the algorithm design, we first verify the controllability of the system. In many situations, the Schrödinger equation introduced in Section **??** takes the form of the system in Eqn. 5, where $H_0, \ldots, H_m$ are $n \times n$ Hermitian matrices.

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = \left(H_0 + \sum_{i=1}^{m} H_i u_i(t) \right) |\psi\rangle \tag{5}$$

Such a system is (pure state) controllable if for every pair of initial and final states, $|\psi_i\rangle$ and $|\psi_f\rangle$, there exist control functions $u_1(t), \ldots, u_m(t)$ and a time $\tau > 0$ such that the solution of Eqn. 5 at time τ, with initial condition $|\psi_i\rangle$, is $|\psi(\tau)\rangle = |\psi_f\rangle$ **d2007introduction**.

Eqn. 5 can be transformed into the equivalent system given in Eqn. 6.

$$\frac{\partial |\psi\rangle}{\partial t} = A |\psi\rangle + \sum_{i=1}^{m} B_i |\psi\rangle u_i(t) \tag{6}$$

Under the assumption that $u_1, \ldots, u_m$ are piecewsie continuous functions, there is an equivalent characterization of controllability. Namely, the system is controllable if the matrices $B_1, \ldots, B_m$ generate the Lie-algebra of $n \times n$ skew-Hermitian matrices with zero trace **d2007introduction**.

For the specific Hamiltonian introduced in Eqn. 1,

$$A = 0 \text{ and } B_1 = i\sigma_x = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \text{ and } B_2 = i\sigma_z = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Since B_1 and B_2 are skew-Hermitian with zero trace and linearly independent, they form a basis for the Lie-algebra of 2×2 skew-Hermitian matrices with zero trace (for details see **aless2010topo**). As a result, the overall system in Eqn. 2 is controllable.