

# 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  $|\psi^*\rangle$  from an initial state  $|\psi_i\rangle$  in a finite time  $T$  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  $\Delta t = \frac{T}{N}$  and some dephasing noise  $\eta_t$  drawn from a zero-mean Gaussian distribution is present. Together this yields the Hamiltonian in Eqn. 1.

$$H_t = -h_t^x \sigma_x - (h_t^z + \eta_t) \sigma_z \quad (1)$$

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

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

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 minimum possible magnetic field strength within which the control can be tuned. The set of controls can be simplified to

$$\mathbb{U} := \{h = (h^x, h^z) \in \{-1, 0, 1\} \times \{-1, 0, 1\}\}.$$

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

$$F(\psi_0, \psi_1) = |\langle \psi_0 | \psi_1 \rangle|^2 \quad (3)$$

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

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

An algorithm will be designed solve the Markov decision problem:

$$\max_{\gamma \in \Gamma} E^\gamma \left[ \sum_{t=0}^T c(|\psi_t\rangle, h_t) \right].$$

## 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  $H_0, \dots, H_m$  are  $n \times n$  Hermitian matrices.

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

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

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

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

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

For the specific Hamiltonian introduced in Eqn. 1,  $A$  is zero and

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

Since  $B_1$  and  $B_2$  are skew-Hermitian with zero trace and linearly independent, they form a basis for the Lie-algebra of  $2 \times 2$  skew-Hermitian matrices with zero trace. As a result, the overall system in Eqn. 2 is controllable.