Reinformcement Learning for Efficient Resource Allocation Between Bayesian Estimation and Operation in Quantum Computer Limited by Low-Frequency Noise

Jan A. Krzywda* and Evert van Nieuwenburg
⟨aQa^L⟩ Applied Quantum Algorithms, Lorentz Institute and Leiden Institute of Advanced Computer Science, Leiden University, P.O. Box 9506, 2300 RA Leiden, The Netherlands

I. INTRODUCTION

Intro [] Qubits are coupled to the environment. The coupling result is the fluctuations of qubit Hamiltonian parameters, which detoriate its control and result in the decoherence mechanisms. For the latter, the effect of the noise strongly depends on its dynamics, in particulair the timescale on which it changes. While fast fluctuations of the environment are relevant for qubit relaxation (set by T_1 time), and uncorrelated dephasing (T_2 time), additional contribution to dephasing comes from the slow fluctuations of the environment and set T_2^*, which is often the limiting factor for the qubit operation.

[**Mitigation techniques**] For the dephasing limited qubits, T_2-like dephasing due to uncorrelated noise, gives rise to phase flip channel, which in principle can be mittigated by Quantum Error Correction protocols []. More practically for near-term devices, similair goal can be achived using Quantum Error Mitigation techniques [], at the cost of measurments overhead. Utilizing the same methods for T_2^*-like dephasing due to slow fluctuations of the environment, is more challenging cite. The most common approach is to decorelate the noise using Pauli Twirling [], decouple the qubit from the environment using dynamical decoupling [], or use the dynamical error suppression techniques []. However, these methods are not always efficient, and can be costly in terms of resources.

[**Bayesian tracking**] Alternative approach to mittigate effects of slowely fluctuating paramters is to track their dynamics in real time and use feedback to counterpart their effects. For constant but unknown parameters, Hamiltonian Learning techniques can be employed []. What know?

[**RL for qubit control**] In this work we propose a Reinforcement Learning (RL) approach to mitigate the effect of the temporairly correlated noise on the qubit operation. The main idea is to use the slow fluctuations of the environment for learning, while the fast fluctuations are crucial for the qubit operation. The learning is done by the agent, that is responsible for the qubit control, and is based on the Bayesian estimation of the noise. The agent is trained to allocate the resources between the Bayesian estimation and the qubit operation, in order to maximize the qubit performance. The performance is measured by the qubit fidelity, which is the probability of the qubit to be in the desired state.

[Spin qubits] In spin qubits the relaxation time is orders of magnitude larger than dephasing time, due to small dipole moment of the two spin states []. For the dephasing the main contribution comes from the fluctuations of magnetic and electric fields, that are used to manipulate the qubit.

[**Outline**] The paper is organized as follows. In Sec. **??** we introduce the model of the qubit coupled to the environment, and the noise that affects the qubit operation. In Sec. **??** we introduce the Reinforcement Learning approach to mitigate the effect of the temporairly correlated noise on the qubit operation. In Sec. **??** we present the results of the RL approach, and compare it to the standard approach. In Sec. **??** we conclude the paper.

* j.a.krzywda@liacs.leidenuniv.nl