Quantum noise modeling through Reinforcement Learning

Simone Bordoni,[1,2,3] Andrea Papaluca,[4,2] Piergiorgio Buttarini,[1,2]
Alejandro Sopena,[5,2] Stefano Giagu,[1,3] and Stefano Carrazza[6,7,8,2]

[1]*Dep. of Physics, La Sapienza University of Rome, Piazzale Aldo Moro 2, Rome, 00185, Italy*
[2]*Quantum Research Centre, Technology Innovation Institute, Abu Dhabi, UAE.*
[3]*Sez. di Roma, INFN, Piazzale Aldo Moro 2, Rome, 00185, Italy*
[4]*School of Computing, The Australian National University, Canberra, ACT, Australia*
[5]*Instituto de Física Teórica, UAM-CSIC, Universidad Autónoma de Madrid, Cantoblanco, Madrid, Spain*
[6]*CERN, Theoretical Physics Department, CH-1211 Geneva 23, Switzerland.*
[7]*TIF Lab, Dipartimento di Fisica, Università degli Studi di Milano, Italy*
[8]*INFN, Sezione di Milano, I-20133 Milan, Italy.*

This work presents a novel machine learning approach to characterize the noise impacting a quantum chip and emulate it during simulations. By leveraging reinforcement learning, we train an agent to introduce noise channels that accurately mimic specific noise patterns. The proposed noise characterization method has been tested on simulations for small quantum circuits, where it consistently outperformed randomized benchmarking, a widely used noise characterization technique. Furthermore, we show a practical application of the algorithm using the well-known Grover's circuit.

Introduction. An important unresolved technological question concerns the practical applicability of Noisy Intermediate Scale Quantum (NISQ) [1] computers. The usability and reliability of these devices are hampered by errors that arise from gate infidelities, environmental interactions, thermal relaxation, measurement errors, and cross-talk [2, 3]. This work aims to develop a model capable of learning hardware-specific noise for use in circuit simulations. We use Reinforcement Learning (RL) to train an agent to add noise channels that replicate the noise of a specific quantum chip. This method minimizes heuristic assumptions about the noise model offering increased flexibility compared to conventional techniques. In this study, we present the results of the RL algorithm in simulated environments, serving as a proof of concept. Our efforts are currently directed towards gathering empirical data from quantum chips to validate the algorithm's efficacy under real operational conditions.

Learning a hardware-specific noise model is motivated by the limited availability of noise prediction techniques [4–6]. Moreover, the quantum computers currently accessible in the cloud are in high demand, leading to long waiting queues to test quantum algorithms. In this context, noise emulation of these devices emerges as an alternative to accelerate circuit testing.

Background. Reinforcement learning is a machine learning paradigm used to train an agent to make optimal decisions in a dynamic environment. It hinges on the fundamental concepts of policy and reward functions. The policy guides the agent's behavior, mapping the environment's states to actions. The reward function assigns a numeric value to state-action pairs, indicating their immediate desirability or associated cost. The training process involves finding an optimal policy that maximizes the expected long-term cumulative reward. During training, various episodes of agent-environment interaction are executed. At each episode's end, the reward is used to update the policy's weights, typically approximated by a Neural Network (NN). Recent years have seen the development of different optimization methods to enhance RL's convergence and stability during training. In our work, we achieved the best results using Proximal Policy Optimization (PPO) [7].

Quantum device noise presents a significant challenge in the NISQ era. In this work, we have considered both coherent and incoherent noise. Coherent noise preserves the state's purity and can be corrected once identified, in our work it is modeled using the single qubit rotation gates R_x and R_z. Incoherent noise is a non-invertible process that can be represented using the formalism of quantum channels. In this work, we utilize two significant incoherent noise channels. The depolarizing channel, which drives the state towards the maximally mixed state, is defined, for a single qubit, as:

$$\mathrm{Dep}(\rho) = (1-\lambda)\rho + \lambda\mathbb{I},$$

where λ is the depolarization parameter. The amplitude damping channel models the loss of energy from a qubit to the environment, and it is described by the map:

$$\mathrm{Damp}(\rho) = A_1\rho A_1^\dagger + A_2\rho A_2^\dagger,$$

where $A_1 = |0\rangle\langle 0| + \sqrt{1-\gamma}\,|1\rangle\langle 1|$, $A_2 = \sqrt{\gamma}\,|0\rangle\langle 1|$ and γ represents the decay probability from $|1\rangle$ to $|0\rangle$.

For simplified noise modeling, we can employ a technique known as Randomized Benchmarking (RB) [8, 9]. RB allows to efficiently estimate the average error magnitude across a set of quantum gates, with resource requirements scaling polynomially with the number of qubits. Employing RB as a noise predictor involves extracting the average gate error and introducing a depolarizing channel with a depolarizing error equal to this parameter after each gate. This noise model serves as a basic benchmark for other, more sophisticated, noise characterization techniques.

Methodology. The RL algorithm requires training, testing, and evaluation datasets, which consist of ensembles of random quantum circuits and their corresponding final states as Density Matrices (DMs). These DMs serve as ground truth labels during the training phase of the algorithm. All circuits for training are composed of Clifford gates extracted from the set of native gates $\{R_x, R_z, CZ\}$ implemented in the quantum devices of the Technology Innovation Institute of Abu Dhabi [10]. Clifford gates are chosen due to their lower simulation cost and their large use in randomized benchmarking. For the performance evaluation, we tested the algorithm on non-Clifford circuits to show that it maintains its ability to generalize. We have tested the RL algorithm with different noise models and using different number of qubits. Specifically, in this work we will report the results obtained on three qubits circuits using the noise model described in the following. A depolarizing channel with parameter $\lambda = 0.02$ is applied after each R_z and CZ gate, and an amplitude damping channel with decay parameter $\gamma = 0.03$ is applied after each R_x and CZ gate. A coherent $R_x(\theta')$ error with angle $\theta' = 0.04 \cdot \theta$ is introduced after each $R_x(\theta)$ gate. Similarly, a coherent $R_z(\theta')$ error is added after each $R_z(\theta)$ gate, with $\theta' = 0.03 \cdot \theta$. This noise model is not intended to be realistic but to test the proposed algorithm on a gate dependent noise model.

To train the RL agent, we need to represent a quantum circuit as an array that can be readily processed by the policy neural network. In the following we refer to this array as the Quantum Circuit Representation (QCR). The QCR has the shape of [*qubits*, *depth*, *encoding*]. The first dimension corresponds to the circuit's qubits, while the second dimension represents the circuit's moments. The *encoding* dimension encodes the information regarding gates and noise channels acting on a specific qubit at a specific circuit moment. To enable the agent to adapt to circuits of varying depths, we introduced the concept of a kernel size similar to the kernels used in convolutional neural networks. The kernel size (k) establishes a "window" that restricts the number of circuit moments the agent can observe at any given time. For instance, with $k = 3$, the agent only observes the current moment and the immediately preceding and following ones. The window's center starts from the first moment and slides one position at each step until the circuit's end is reached. This approach is based on the heuristic assumption that a gate's noise is most influenced by its temporally proximate gates.

The training, conducted using the PPO algorithm, divides the policy NN into two components: the actor NN, responsible for action selection, and the critic NN, tasked with reward prediction. The configuration of the policy NN includes an initial feature extractor composed of a convolutional layer. The feature extractor's output is then passed to the actor and critic NNs, each composed of two dense layers. The total number of trainable parameters is of the order of 10^4.

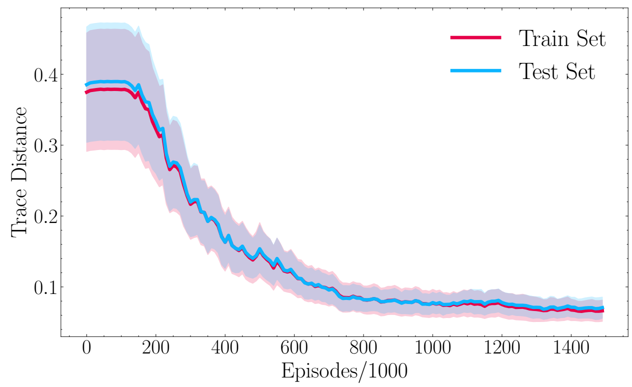

FIG. 1. Average trace distance between DMs generated by the model and the ground truth DMs during training for 1.5×10^6 episodes. Error bars report the standard deviation.

Each episode of the training process begins with the agent receiving a randomly selected quantum circuit from the training set. For each circuit moment, the agent observes the QCR and takes an action: any combination of the selected set of noise channels (depolarizing, amplitude damping and coherent errors R_x and R_z), together with their corresponding noise parameters, is inserted in that circuit moment. The agent receives the reward at the end of each circuit. The reward is a function of the Trace Distance (TD) between the ground truth DM (ρ_{true}) and the DM of the noisy circuit generated by the agent (ρ_{agent}):

$$\text{Reward}(\rho_{\text{agent}}, \rho_{\text{true}}) = \frac{1}{\alpha \text{TD}(\rho_{\text{agent}}, \rho_{\text{true}})^2 + \epsilon},$$

where α is an hyperparameter and ϵ is a small parameter introduced to prevent numerical instabilities. The trace distance is a common metric used in quantum information to measure the distinguishability between two quantum states. After numerous episodes, the agent is expected to learn the optimal placement of noise channels in a noise-free circuit to reconstruct the final density matrix of the real noisy circuit. Once trained, the proposed algorithm should be capable of generalizing to previously unseen circuits, thereby enabling realistic noisy simulations.

Results. To train the RL agent, we generated a dataset of 500 random circuits allocating 400 circuits for training and reserving the remaining 100 for testing. Figure 1 reports the average trace distance over the training and test sets obtained by the agent during training for 1.5×10^6 episodes. The RL agent effectively learns to simulate the noise, exhibiting no signs of overfitting.

To assess the model's generalization capability, we evaluate it on random circuits of varying depths. Figure 2 reports the performance of the RL agent compared with the randomized benchmarking method evaluated using the average trace distance. We have also included two limit cases: the circuit where no noise has been added

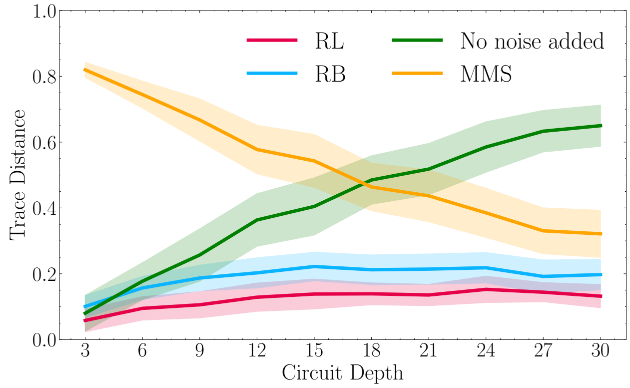

FIG. 2. Performance comparison of the RL agent with respect to the RB noise model. The limit cases of circuits without noise and the maximally mixed state (MMS) are also reported. Error bars report the standard deviation.

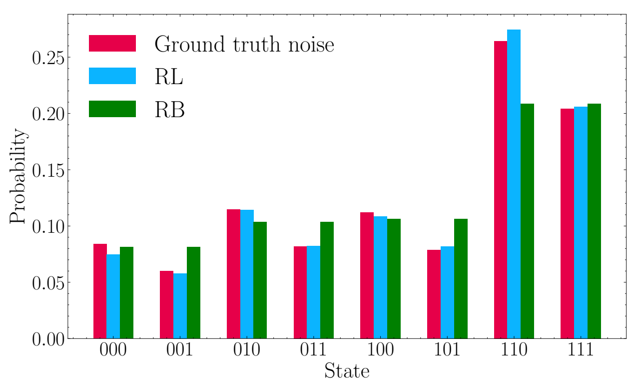

FIG. 3. Outcome probabilities for a noisy three qubit Grover's circuit obtained with the RL algorithm and the RB noise model. The ground truth noisy outcomes are reported for comparison.

and the maximally mixed state. The RL agent demonstrates its adaptability to circuits of different depths, consistently outperforming RB. This result suggests that, while RB categorizes all noise sources as depolarizing, our algorithm can discern the specific characteristics of the noise. The improvement is especially pronounced on shorter circuits. As the circuit depth increases, the noise approximates to a global depolarizing channel, reducing the relative advantage of the RL agent.

We have studied the proposed algorithm using different noise models and a number of qubits spanning from one to three. The results obtained are similar to the ones presented in this section. The RL agent is particularly efficient in the presence of many coherent errors.

We finally report the application of the RL agent for a famous use case: the circuit to implement Grover's algorithm [11]. In particular, we have considered a circuit to find the target state $|11\rangle$ using an ancillary qubit, for a total of three qubits. The circuit, transpiled into native gates, is composed of 40 gates, of which 7 are CZ gates, for a total of 25 circuit moments. Figure 3 reports the outcome probabilities of the noisy Grover's circuit obtained with the RL algorithm and the RB model. It is possible to observe that the RL algorithm has a better capability to reconstruct a realistic outcome, in particular for the peak probability of the state $|110\rangle$. The fidelity between the ground truth DM and the one obtained with the RL model is 0.975, while the fidelity obtained with the RB model is 0.949. The result obtained on the Grover's circuit constitutes an interesting generalization test of the RL agent as this circuit has a particular structure with respect to the random circuits used for training.

Conclusions. In this work we have demonstrated that is is possible to use reinforcement learning for replicating specific noise models. The presented RL model has showed outstanding generalization properties, consistently outperforming randomized benchmarking in a simulated environment. While we are working on testing the algorithm on real quantum hardware, we are considering potential future applications not only reproducing a specific noise pattern. Using the knowledge of the noise for its mitigation could be an interesting approach.

The current model's limitation is its scalability to circuits with many qubits. Scaling the model would necessitate a significant increase in the number of actions, complicating and lengthening the training process. Additionally, on quantum hardware, obtaining the ground truth density matrices for circuits with many qubits via quantum state tomography requires exponentially more measurements. We are considering potential solutions to these challenges. One approach could involve training the model with probability distributions derived from measurements, rather than density matrices. To address the first issue, we could partition large circuits into smaller ones, facilitating parallel training of multiple smaller models. While these ideas require further validation, this work demonstrates that machine learning's application to learn noise patterns within small quantum circuits is a promising proof of concept that could lead to future advancements.

Acknowledgment. This work is partially supported by ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union – NextGenerationEU.

[1] Preskill, J. *10.22331/q-2018-08-06-79*
[2] Kjaergaard, M. et al. *10.1146/annurev-conmatphys-031119-050605*
[3] Zhao, P. et al. *10.1103/PRXQuantum.3.020301*
[4] Zlokapa, A. et al. *10.48550/arXiv.2005.10811*
[5] Georgopoulos, K. et al. *10.1103/PhysRevA.104.062432*
[6] Harper, R. et al. *10.1038/s41567-020-0992-8*
[7] Schulman, J et al. *10.48550/arXiv.1707.06347*
[8] Emerson, J. et al. *10.1088/1464-4266/7/10/021*
[9] Heinrich, M. et al. *10.48550/arXiv.2212.06181*
[10] https://github.com/qiboteam/qibolab_platforms_qrc
[11] Grover, L. *10.48550/arXiv.quant-ph/9605043*