IMPROVING ROBUSTNESS OF QUANTUM FEEDBACK CONTROL WITH REINFORCEMENT LEARNING

Manuel Guatto, FZJ (PGI-8), email: m.guatto@fz-juelich.de

Obtaining reliable state preparation protocols is a key step towards practical implementation of many quantum technologies, and one of the main tasks in quantum control.
In this work, different reinforcement learning approaches are used to derive a feedback law for state preparation of a desired state in a target system.
In particular, we focus on the robustness of the obtained strategies with respect to different types and amount of noise.

Comparing the results indicates that the **learned controls are more robust to unmodeled perturbations** with respect to simple feedback strategy based on optimized population transfer, and that training on simulated nominal model retain the same advantages displayed by controllers trained on real data.

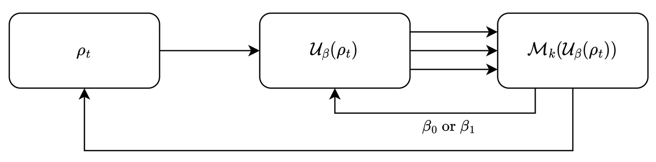
JÜLICH FORSCHUNGSZENTRUM

Guatto, arXiv:2401.17190

CONTROL AND LEARNING SCENARIOS

We train and validate our model on a qutrit system. We consider $\mathcal{N}_\alpha$ the noise superoperator, with the parameter α to weigh the amount of noise injected in the system and $\mathcal{M}^{\varepsilon}_{l}$ the measurement superoperator corresponding to the outcome l and the inaccuracy ε. We define β_t as the control law.

Basic Controller



Controller based on **deterministic feedback**

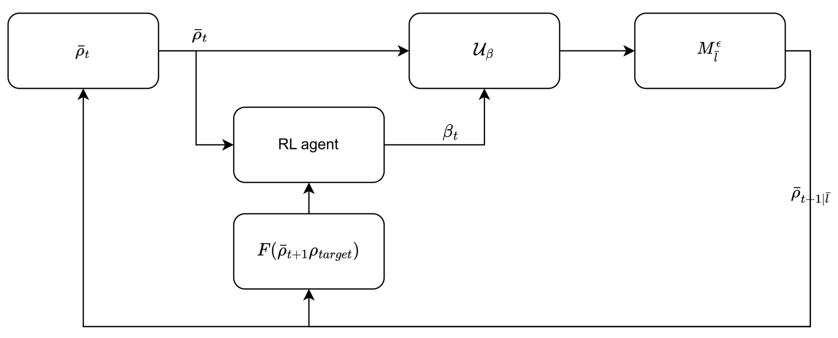
Controller input: l_t

Action:

$$\beta_0 = \arg\max_\beta tr(|2\rangle\langle 2|\,\mathcal{U}_\beta(|0\rangle\langle 0|)) = \arg\max_\beta tr(\langle 2|\,\mathcal{U}_\beta(|0\rangle\langle 0|)\,|2\rangle),$$

$$\beta_1 = \arg\max_\beta tr(|2\rangle\langle 2|\,\mathcal{U}_\beta(|1\rangle\langle 1|)) = \arg\max_\beta tr(\langle 2|\,\mathcal{U}_\beta(|1\rangle\langle 1|)\,|2\rangle).$$

Model Based Scenario (MBS)



RL agent trained with nominal dynamics **without noise**

State: $\begin{cases}\bar\rho(t+1) = \mathcal{M}^{\epsilon}_{\bar l(t)} \circ \mathcal{U}_{\beta_t}(\bar\rho(t)) \\ \bar\rho(0) = \rho_0\end{cases}$

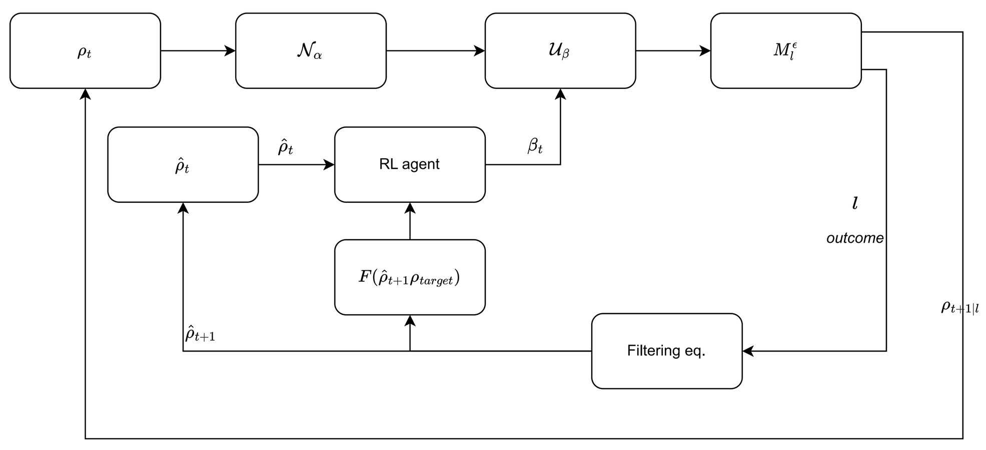
$$p(\bar l(t)) = tr(M^\dagger_{l_t} M_{l_t}(\bar\rho(t)))$$

Action: β_t

Reward:

$$F(\rho, \rho_{target}) = \mathrm{tr}\left[\sqrt{\sqrt{\rho}\,\rho_{target}\sqrt{\rho}}\right]^2$$

Data Based Scenario (DBS)



RL agent trained with **filtered dynamics**

State: $\begin{cases}\hat\rho(t+1) = \mathcal{M}^{\epsilon}_{l(t)} \circ \mathcal{U}_{\beta_t}(\hat\rho(t)) \\ \hat\rho(0) = \rho_0\end{cases}$

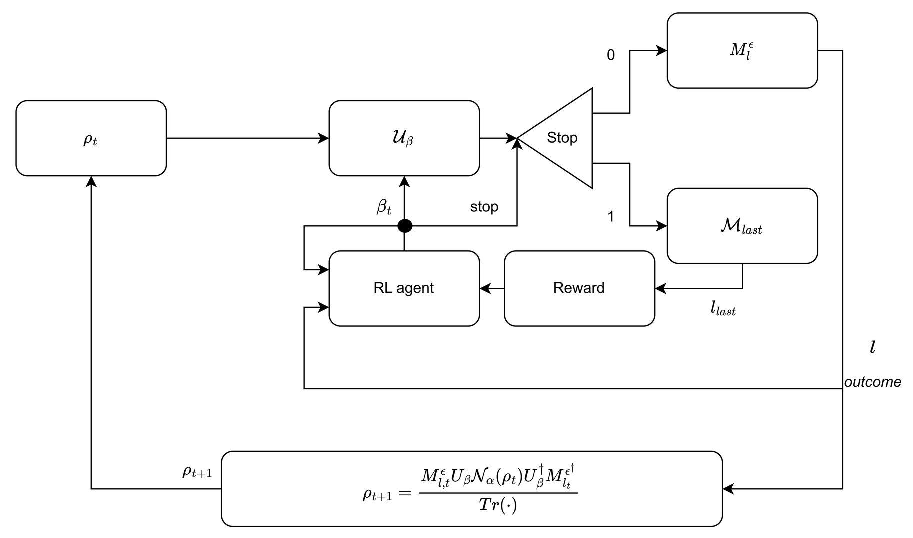
$$p(l(t)) = tr(M^\dagger_{l_t} M_{l_t} \mathcal{N}_\alpha(\rho(t)))$$

Action: β_t

Reward:

$$F(\rho, \rho_{target}) = \mathrm{tr}\left[\sqrt{\sqrt{\rho}\,\rho_{target}\sqrt{\rho}}\right]^2$$

QOMDP Scenario



RL agent trained with measurement outcomes

State: l_t

Action: β_t + *stop*

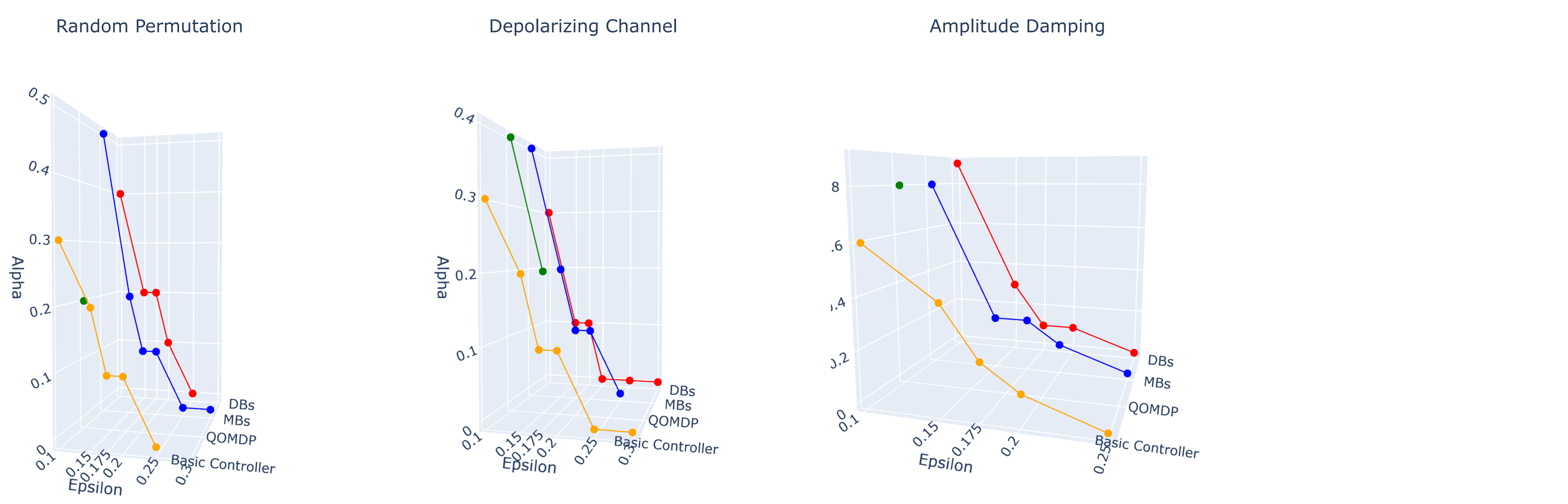
Reward:

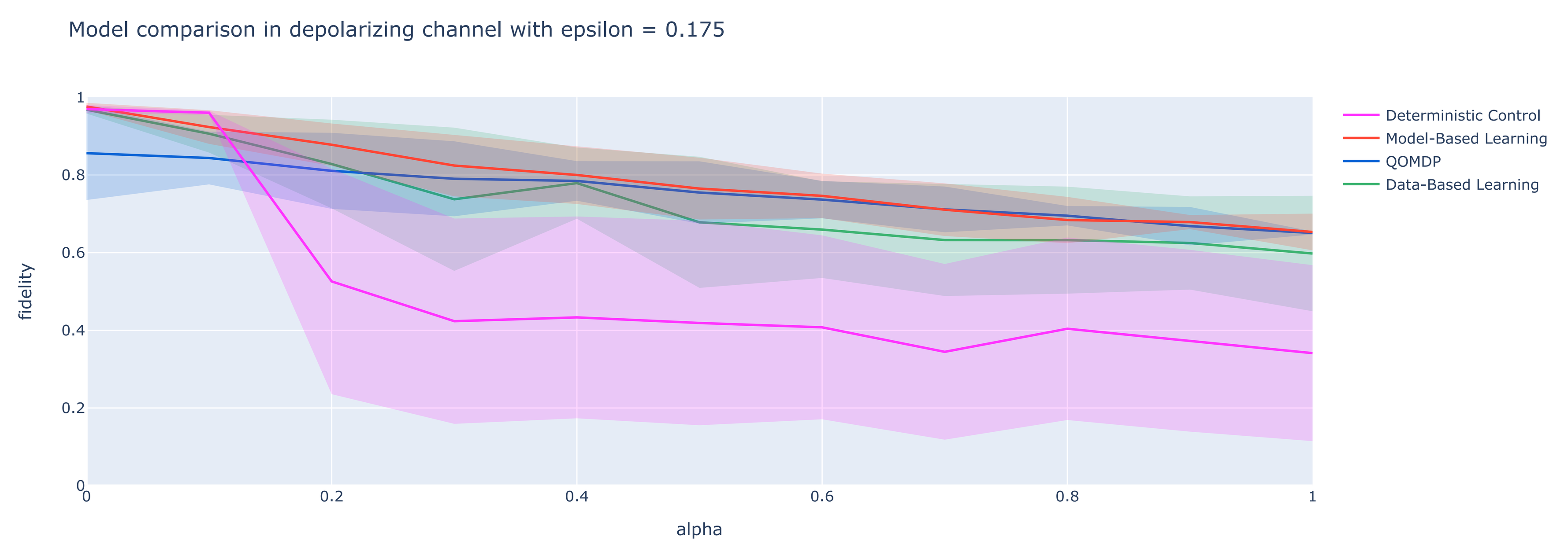
$$R(stop, l_{last}, done) = \begin{cases} 0, & \text{if } stop = 0\; done = \text{False} \\ -1, & \text{if } stop = 0\; done = \text{True} \\ r_{fun}(l_{last}), & \text{if } stop = 1 \end{cases}$$

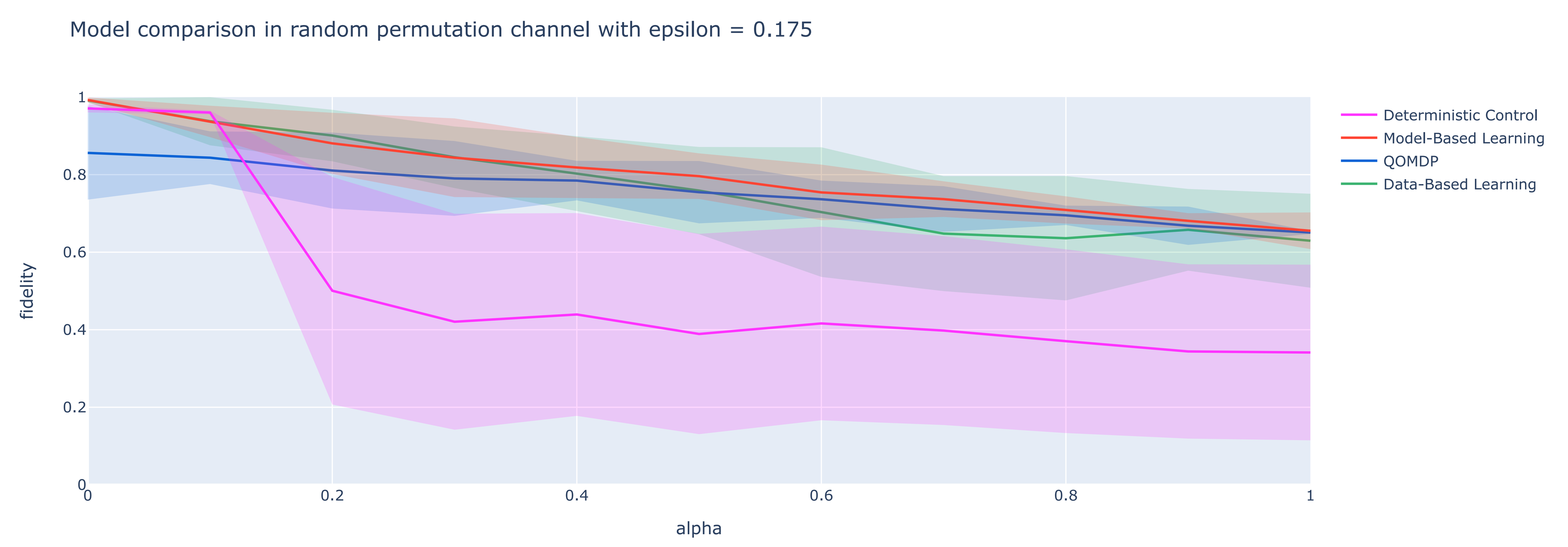
$$r_{fun}(l_{last}) = \begin{cases} +1, & \text{if } l_{last} = l_{target} \\ -1, & \text{otherwise} \end{cases},$$

RESULTS

In this figure we represent the highest value of noise intensity α that can be handled by each controller while maintaining a fidelity of at least 0.9 in order to prepare state ρtarget = |2X2|. Missing points in the curve indicate that the target fidelity cannot be guaranteed.







Examples of necessary number of timesteps to achieve fidelity ≥ 0.9