Re-calibration of quantum devices by RL

An application to coherent-state receiver

arXiv:2404.10726

Matías Bilkis

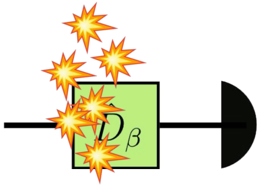
Computer Vision Center

IFLP CONICET UNLP

Joint work w/
Tomás Crota
Fernando Vilariño
Lorena Rebón
Mauricio Matera

This talk in a nutshell

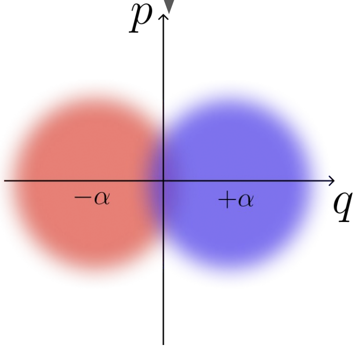
- What happens if the environment changes?
- Detect the change ?
  Adapt the policy ?
  What happens w/ Q-values?
  Which strategies can we adopt ?
- Binary quantum-state discrimination



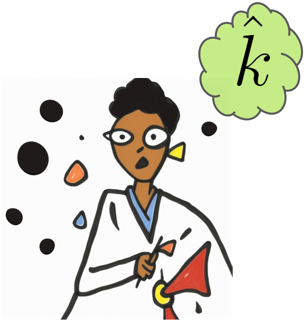
- Calibrate an optical receiver by RL

Coherent-state discrimination

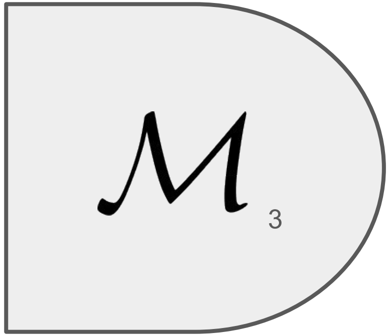
$$|\alpha_k\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{((-1)^k \alpha)^n}{\sqrt{n!}} |n\rangle$$



- Distinguish between $\{|\alpha\rangle, |-\alpha\rangle\}$
- POVM $\mathcal{M} = \{M_0, M_1\}$
- $p(n|\alpha_k) = \mathrm{Tr}\big(M_n |\alpha_k\rangle\langle\alpha_k|\big)$

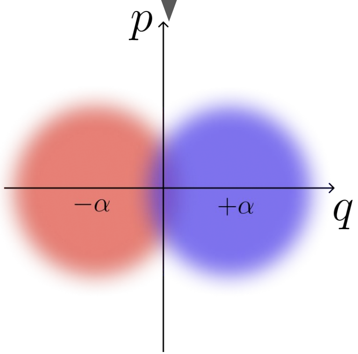


$n \sim p(n|\alpha_k)$



Success probability & Helstrom bound

$$|\alpha_k\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{((-1)^k \alpha)^n}{\sqrt{n!}} |n\rangle$$



- **Success probability;** $g(n)$ ***guessing rule***

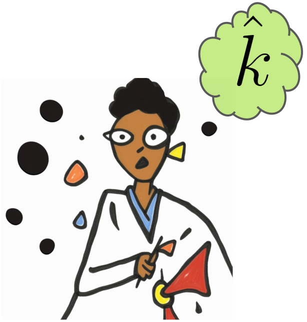
$$P_s(\mathcal{M}, g) = \sum_n p(n|\alpha_{\hat{k}}) \mathrm{pr}(\alpha_{\hat{k}})|_{\hat{k}=g(n)}$$

- **Helstrom bound**

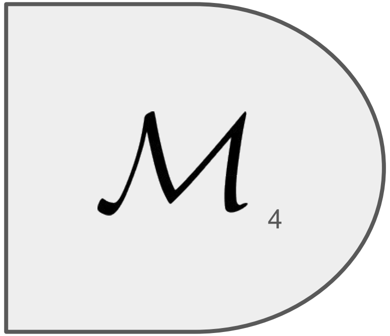
$$P_s(\mathcal{M}, g) \leq P_s(\mathcal{M}^*, g^*)$$
$$= \frac{1}{2}\Big(1 + ||p_0\rho_0 - p_1\rho_1||_1\Big)$$

- **Optimal**: $|\mathrm{cat}\rangle \propto |\alpha\rangle + |-\alpha\rangle$

  Can be realized w/ lineal optics + on/off detectors:
- Feed(forward) protocol on consecutive Kennedy receivers

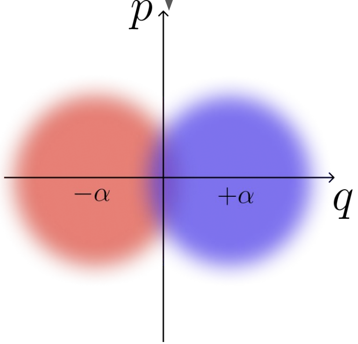


$$n \sim p(n|\alpha_k)$$

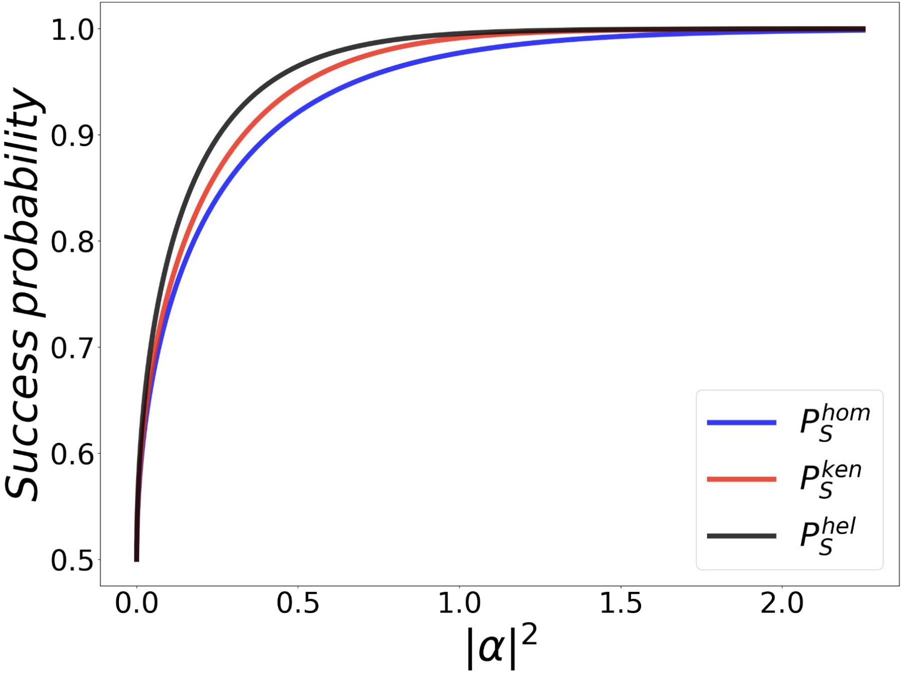


Kennedy receiver

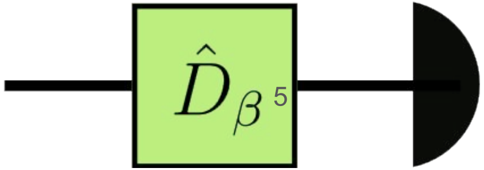
$$|\alpha_k\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{((-1)^k \alpha)^n}{\sqrt{n!}} |n\rangle$$



- Photon-detection measurement **(non-gaussian)** $\{|0\rangle\langle 0|, I - |0\rangle\langle 0|\}$
- $P_s^{ken} = \sum_{o=0,1} \max_k p_k p(n|\alpha^{(k)})$

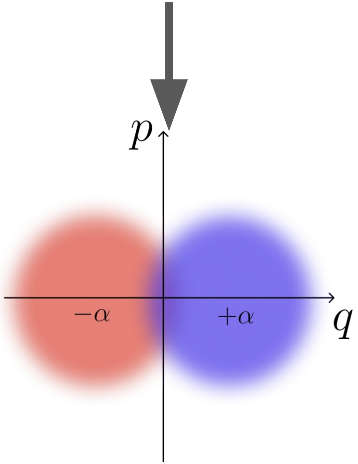


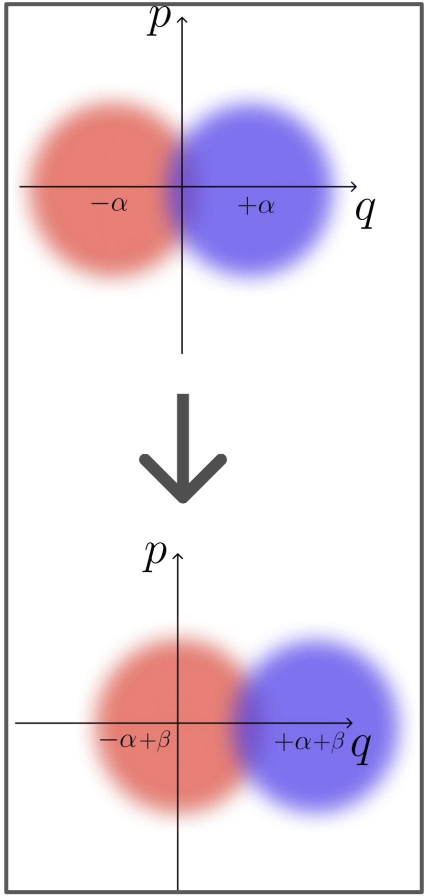
$n \sim p(n|\alpha_k)$

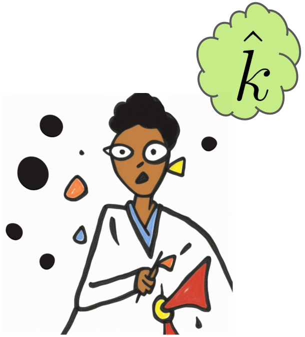


Kennedy receiver

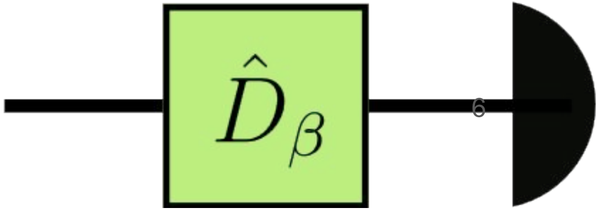
$$|\alpha_k\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{((-1)^k \alpha)^n}{\sqrt{n!}} |n\rangle$$





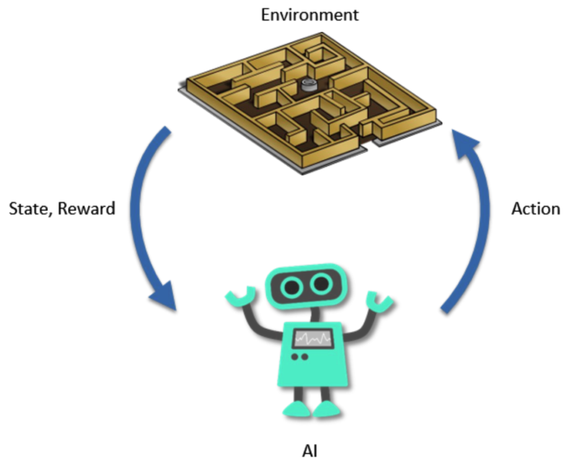


$$n \sim p(n|\alpha_k)$$

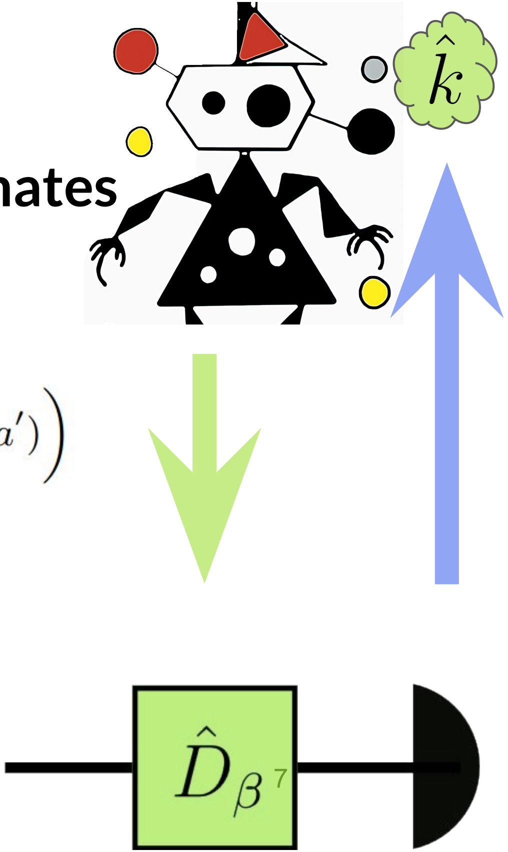


RL calibration

- **Estimate of how good a displacement is by trying it**
- **This is done by Q-learning, which updates these estimates**

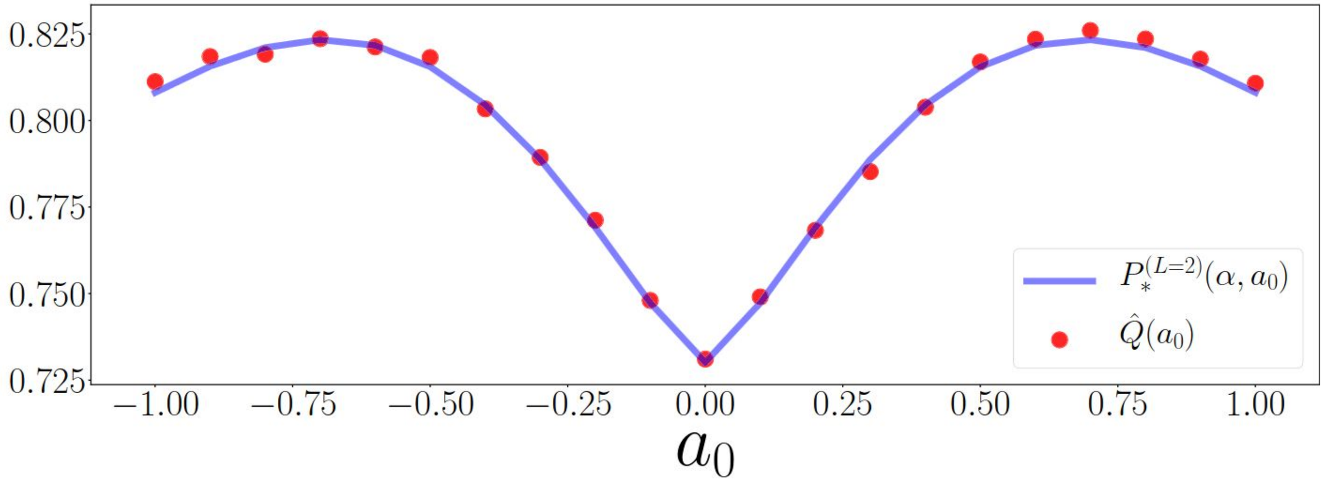


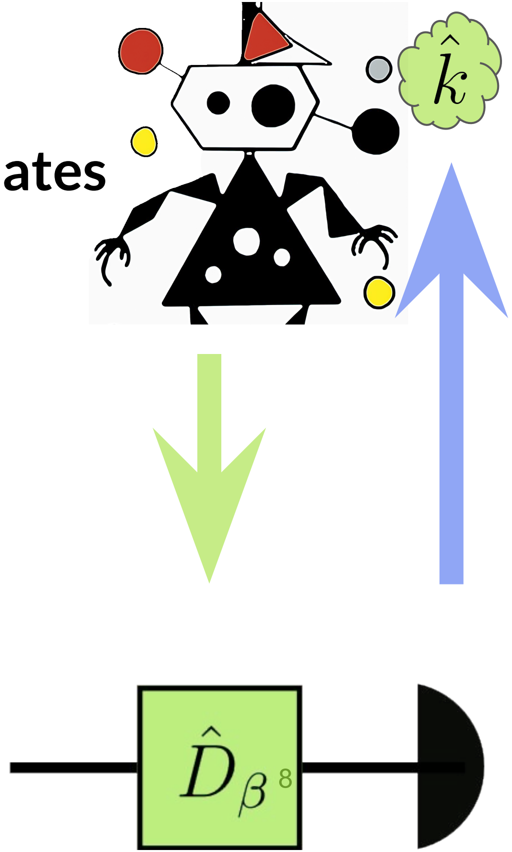
$$\hat{Q}(s_\ell, a_\ell) \leftarrow (1-\tilde{\alpha})\hat{Q}(s_\ell, a_\ell) + \tilde{\alpha}\left(r_{\ell+1} + \gamma \max_{a' \in \mathcal{A}(s_{\ell+1})} \hat{Q}(s_{\ell+1}, a')\right)$$



RL calibration

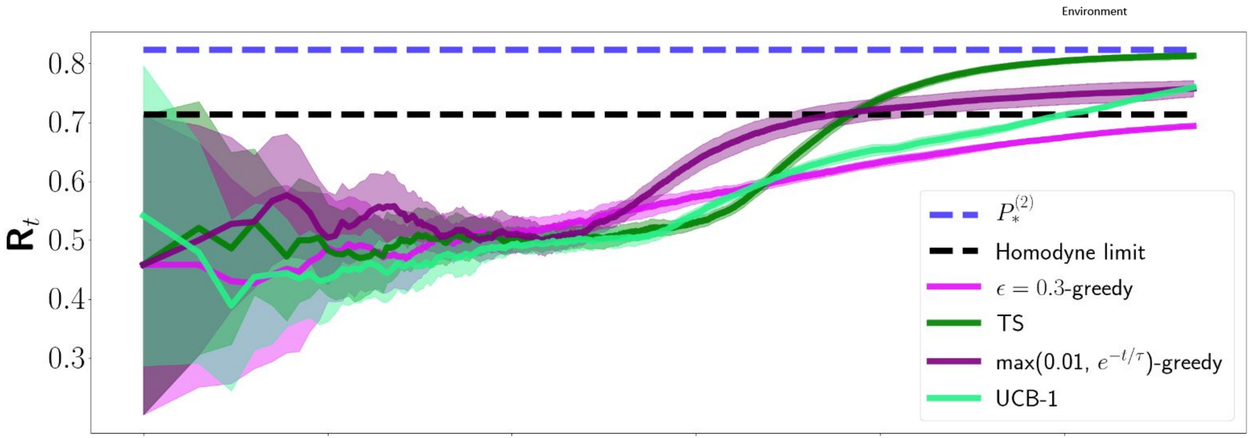
- Estimate of how good a displacement is by trying it
- This is done by Q-learning, which updates these estimates

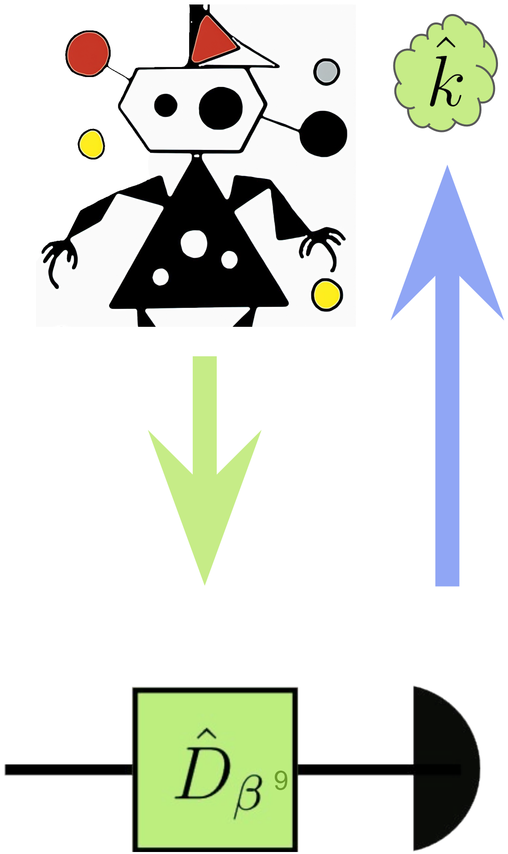




RL calibration

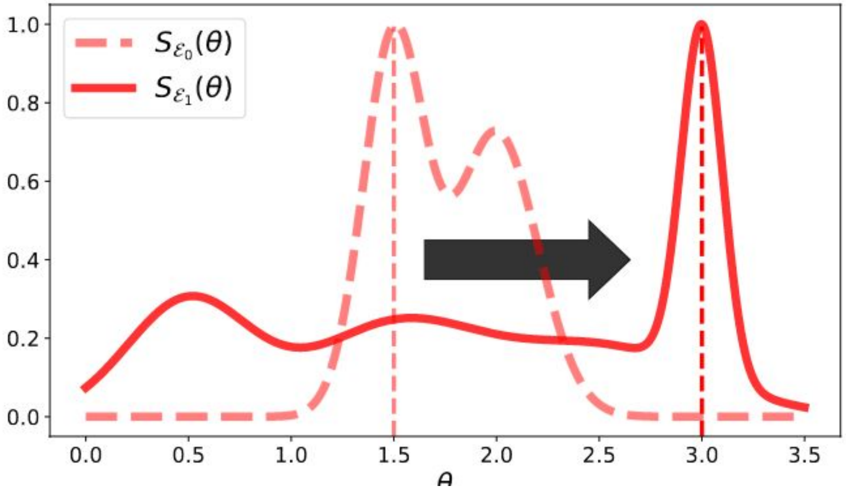
- Agent guesses, reward 1/0 if ok/wrong
- Noise robustness ([PRR, 033295 (2020)])





But if the environment changes?

$$\mathcal{E}_0 \rightarrow \mathcal{E}_1$$

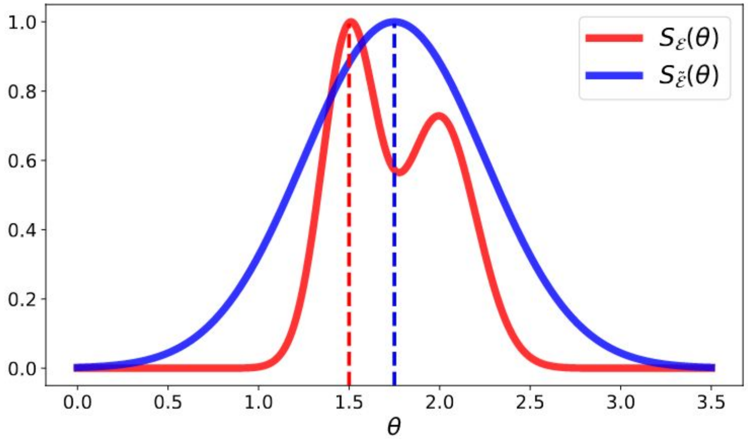


De-calibration witness

- How to realize a change occurs?
- Monitor the reward $\rightarrow$ inaccessible when “deploying”
- Example: monitor empirical outcome probabilities

Q-values

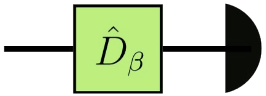
- Would them adapt? → Standard Q-learning yes & no (granted exploration)
- Can we do some sort of approximation?
- [Other approaches → learn map] $\mathcal{E}_0 \rightarrow \mathcal{E}_1$
- Here: effective re-initialization of the Q-values + *fine-tune* w/ RL



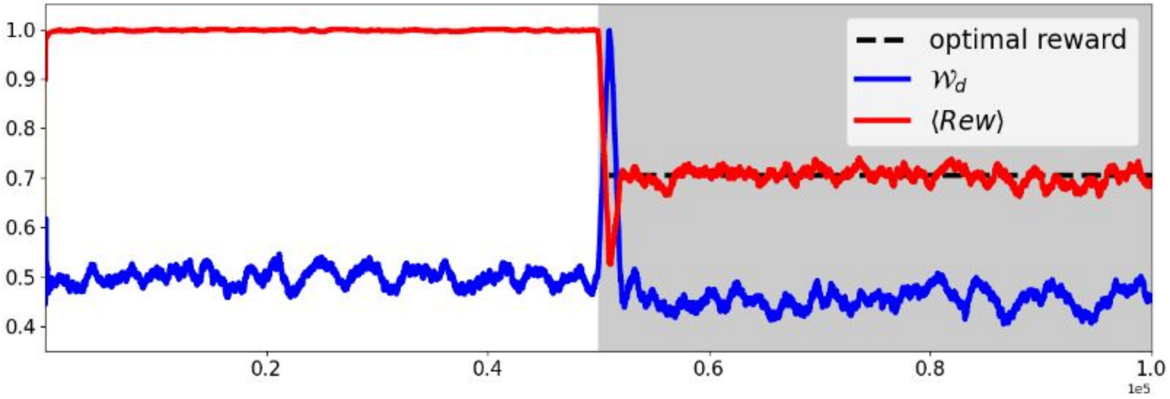
Case-study

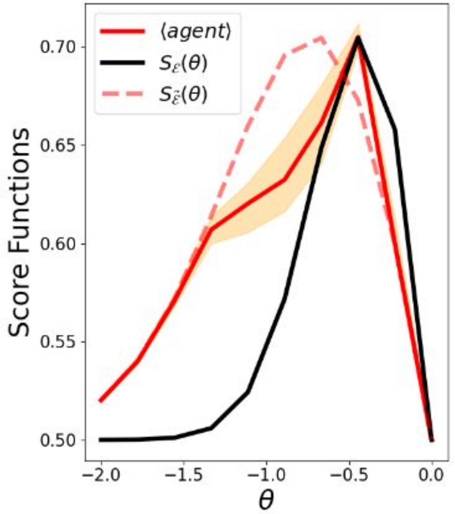
- **The displacements are faulty** $\beta \to \lambda\beta$

  and intensity |α|^2 changes



- **Monitor empirical outcome probabilities (decalibration witness)** $\hat{p}(n=1)$
- **Effective model: success probabilities (of un-faulty device!)**





Conclusion & questions

- **We study how RL adapts to a sudden change in environment, and develop simple strategies to complement Q-learning**

- **We tackle a very simple case, the Kennedy receiver**

- **Alternative scenarios & problems to try this machinery?**
  ***e.g.* quantum control**

I'm happy to talk :)

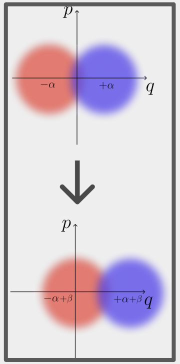
mbilkis@cvc.uab.cat

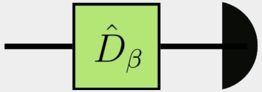
Matías Bilkis
Computer Vision Center

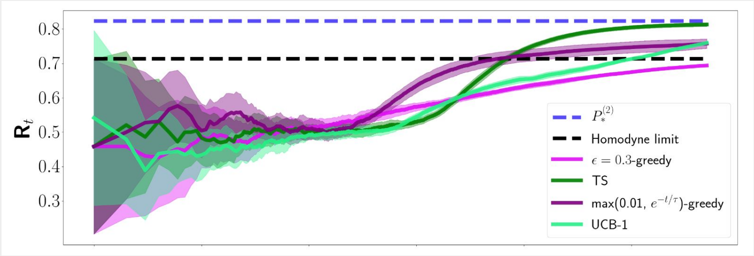
arXiv:2404.10726

Re-calibration of quantum devices by RL

An application to coherent-state receiver







But if the environment changes?

$$\mathcal{E}_0 \rightarrow \mathcal{E}_1$$