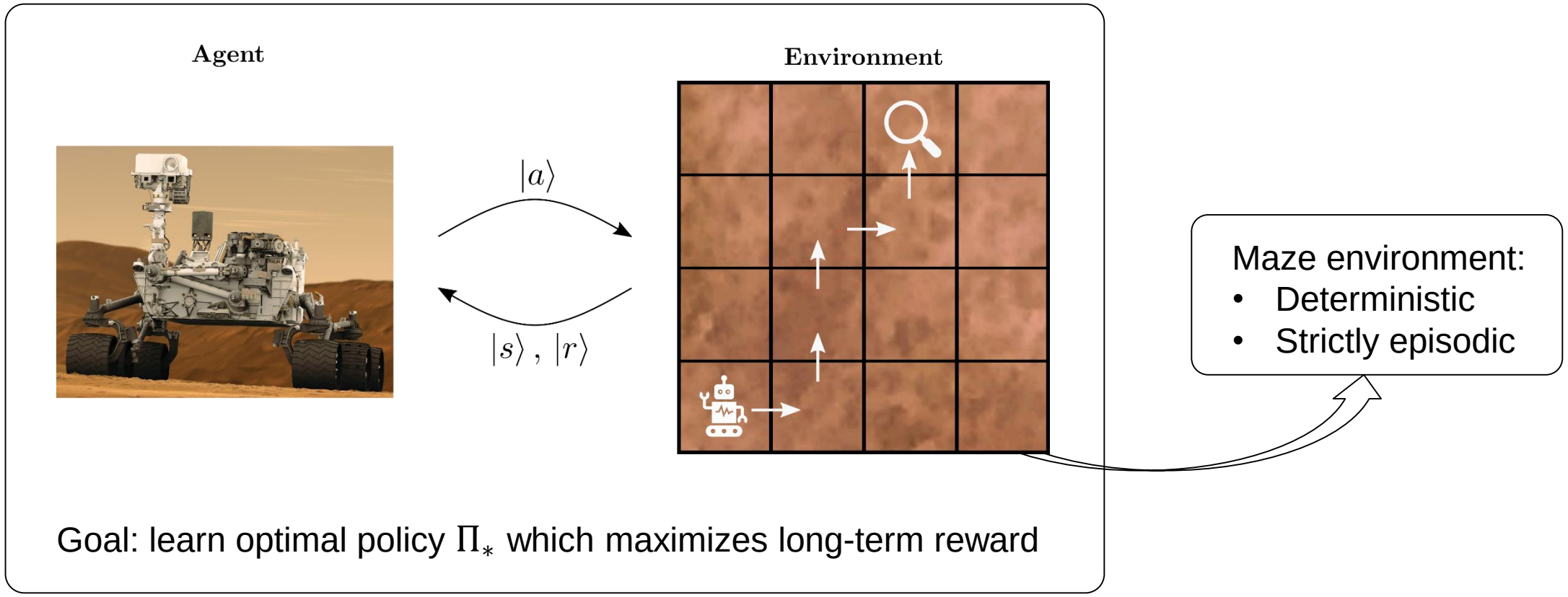
A Quantum Enhanced Learning Algorithm for Maze Problems

Oliver Sefrin and Sabine Wölk

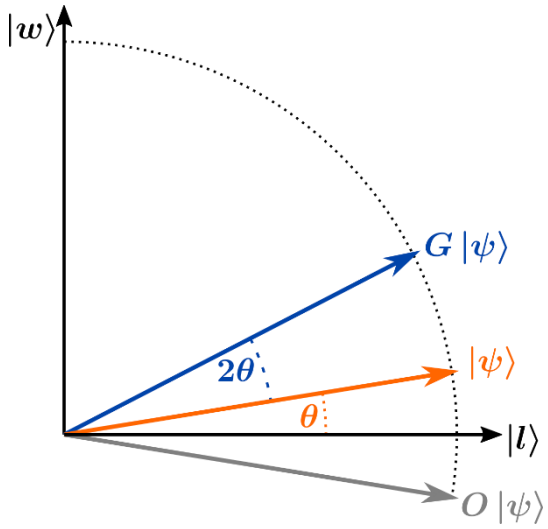
Institute of Quantum Technologies
German Aerospace Center (DLR), Ulm

Reinforcement Learning



Find rewarded actions „quicker“ using amplitude amplification!

Quantum Amplitude Amplification



Environment unitary U_{Env}:

$$U_{\mathrm{Env}} |\vec{a}\rangle_A |0\rangle_S |0\rangle_R = |\vec{a}\rangle_A |\vec{s}(\vec{a})\rangle_S |r(\vec{a})\rangle_R$$

Phase kick-back oracle[1] O_{Env}:

$$O_{\mathrm{Env}} |\vec{a}\rangle_A |0\rangle_S |-\rangle_R = (-1)^{r(\vec{a})} |\vec{a}\rangle_A |0\rangle_S |-\rangle_R$$

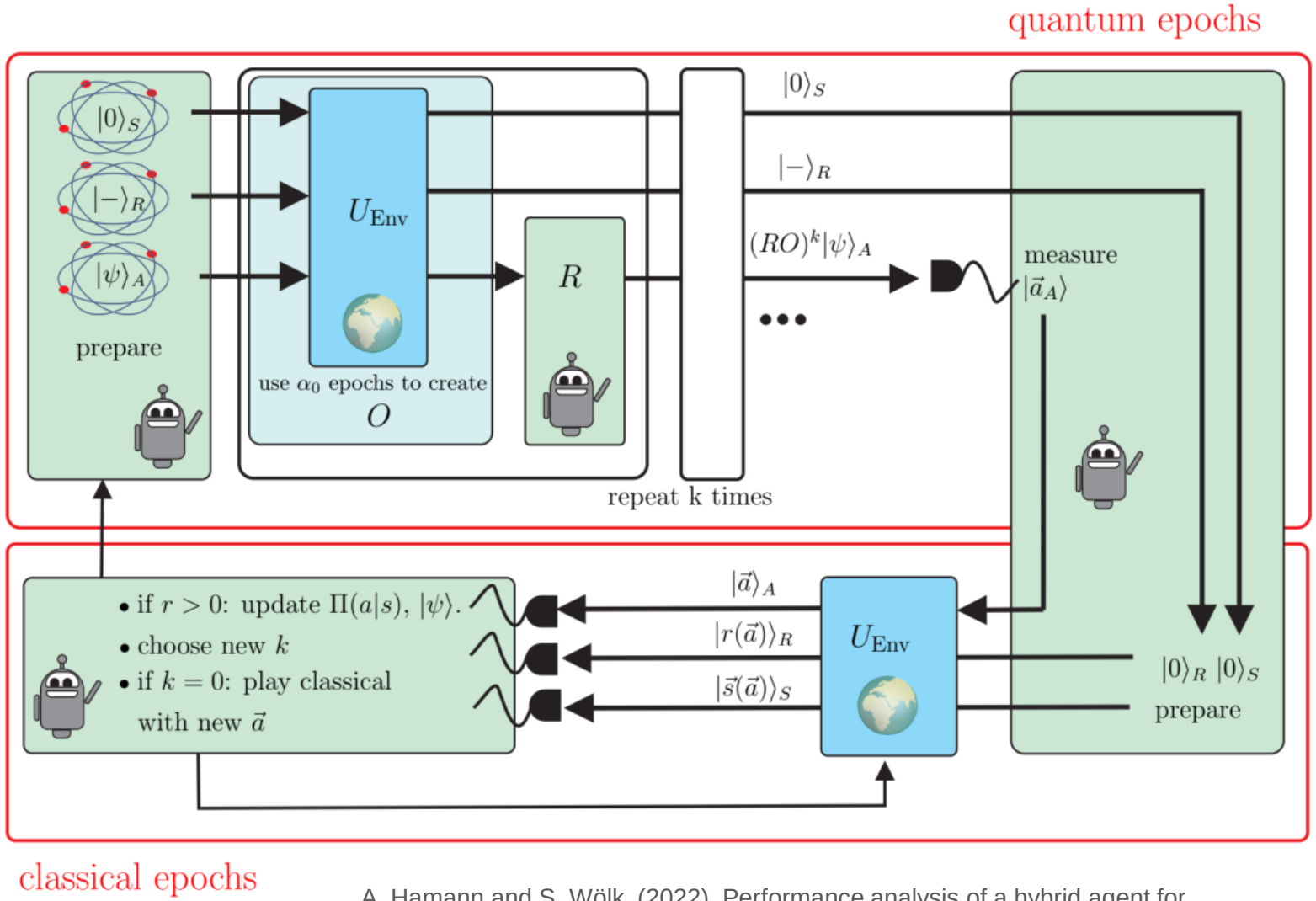
Grover operator:

$$G = (I - 2|\psi\rangle_A\langle\psi|)O_{\mathrm{Env}} \qquad \text{with} \qquad |\psi\rangle_A = \sum_{\{\vec{a}\}} \sqrt{\Pi(\vec{a})}\, |\vec{a}\rangle_A$$

$$G(p_{\mathrm{init}}, k) = \sin^2\left[(2k+1)\arcsin(\sqrt{p_{\mathrm{init}}})\right]$$

[1] V. Dunjko, J. M. Taylor, and H. J. Briegel. (2016). Quantum-enhanced machine learning. *Physical review letters*, 117(13):130501.

Quantum Enhanced Learning Algorithm



Quadratic speed-up in learning time:

$$\langle T\rangle_q \leq 9/4\sqrt{\langle J\rangle\langle T\rangle_{cl}}$$

Quantum advantage up to threshold probability

Experimental Proof of Concept:

V. Saggio et al. (2021). Experimental quantum speed-up in reinforcement learning agents. *Nature*, 591(7849):229–233.

A. Hamann and S. Wölk. (2022). Performance analysis of a hybrid agent for quantum-accessible reinforcement learning. *New Journal of Physics*, 24(3):033044.

The Search Length Dilemma

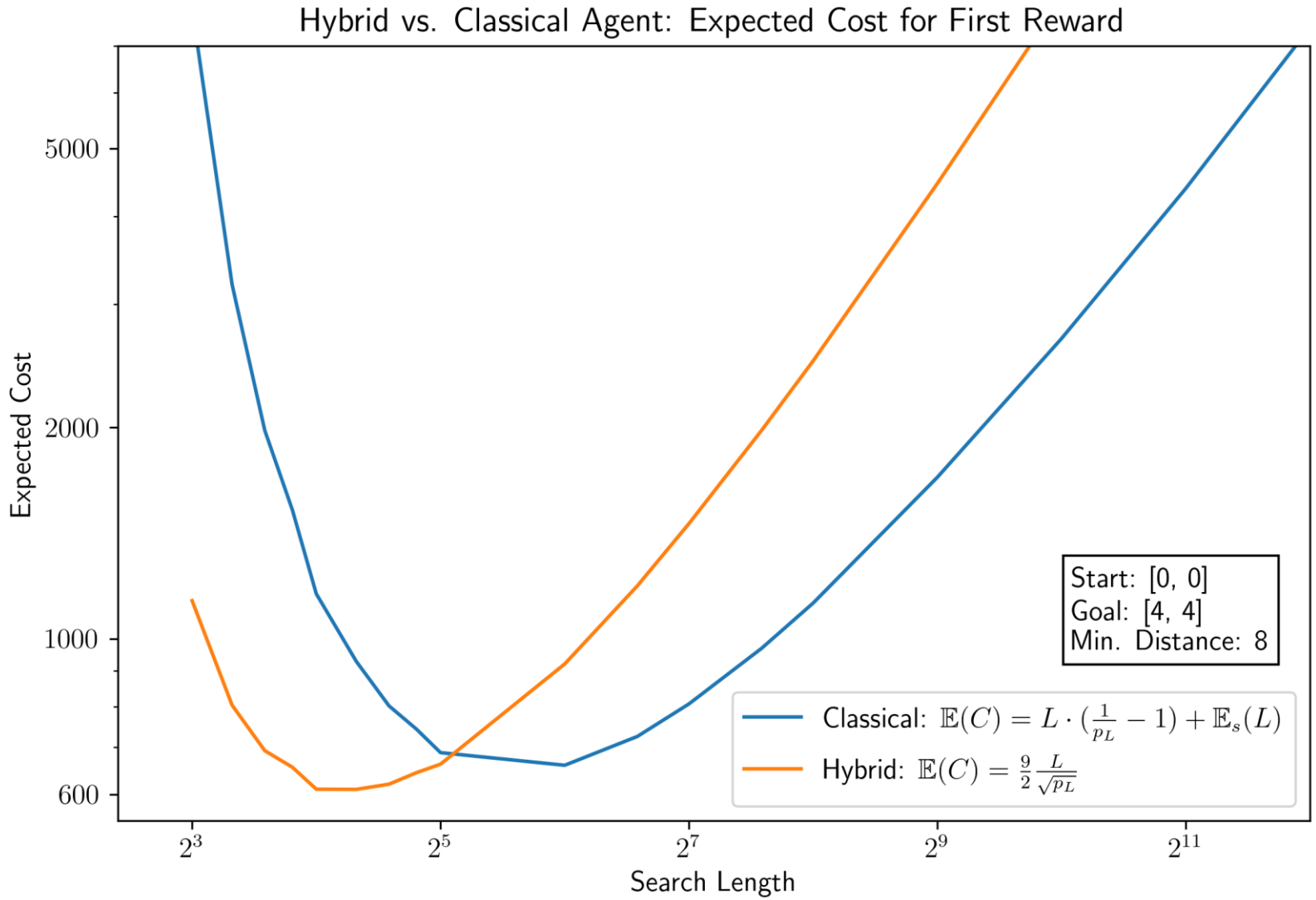
Restriction: fixed *search length L*

Cost $= N_{\text{queries}} * L = N_{\text{steps}}$

Minimal *search length* $\not\Rightarrow$ minimal *cost*

Minimal distance possibly unknown

Search length selection strategy required!



Thank You!

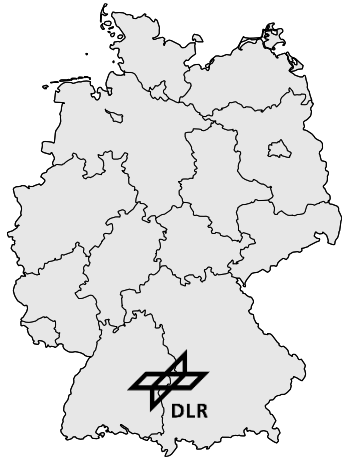
Oliver Sefrin (PhD Student)
oliver.sefrin@dlr.de

PD Dr. Sabine Wölk
sabine.woelk@dlr.de

Key Points:

Quantum enhanced RL algorithm:
- View finding rewarded actions as quantum search problem
- Find rewards quadratically faster to speed up RL

Next step:
- Search length selection strategy for unknown goal distances



PS: we have open positions! Contact us for more details.