Reinforcement Learning Using Quantum Optical Projective

Team 3

A. O. Parada Flores, D. J. Arbol Guerrero, F. Saravalle, M. Patraca Gonzalez, I. Tsoni

Index

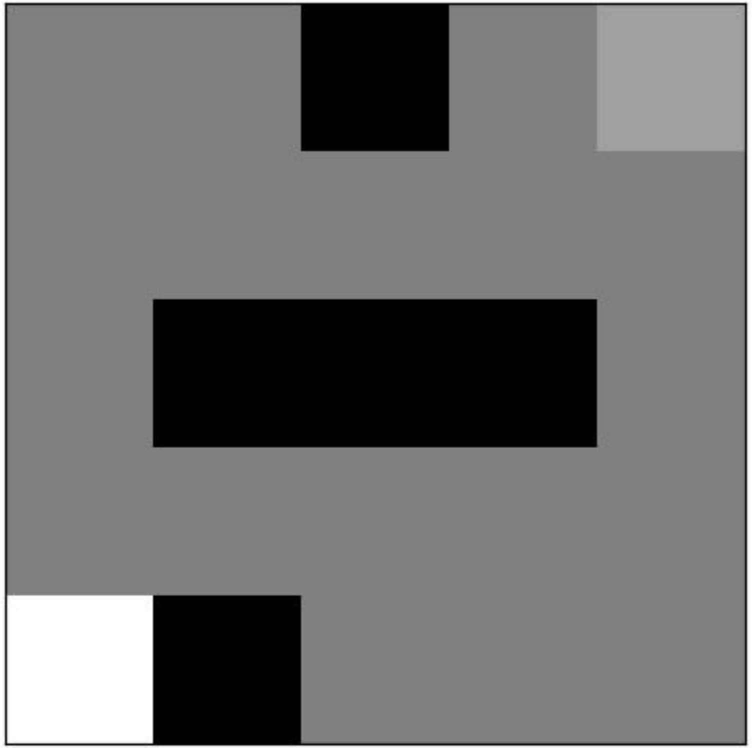
Problem statement

Legend

white: starting position
light grey: ending position
black: wall(not accessible)

Goal

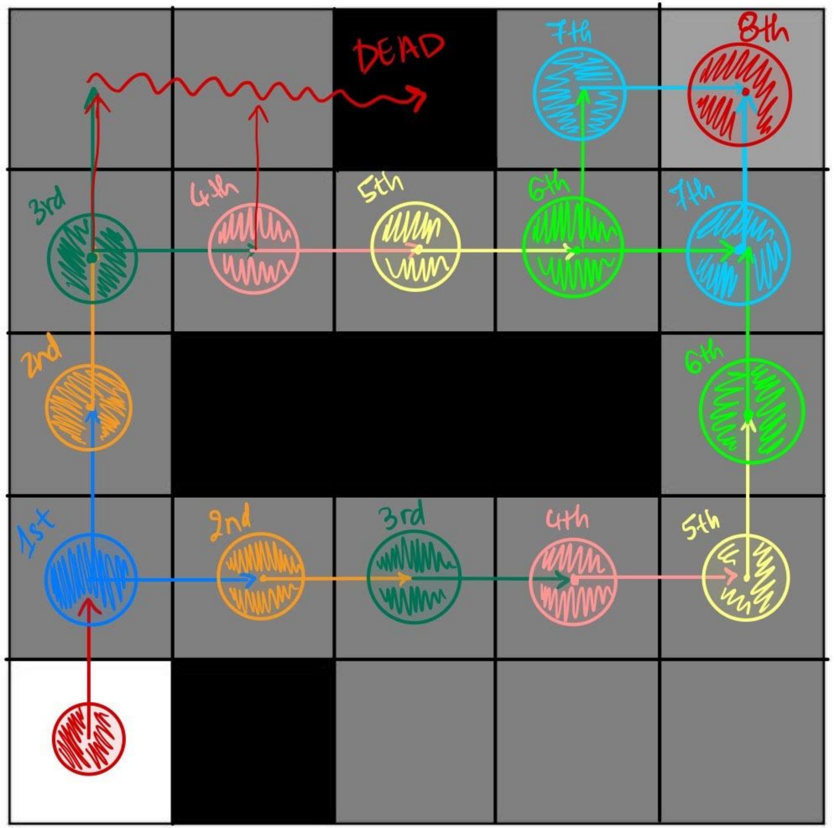
Finding the paths going from white box to the light grey one, without passing through the walls.



Possible Path

<u>Simplification</u>
Possible moves:
right and up.

In the picture
all possible solutions
to the problem.



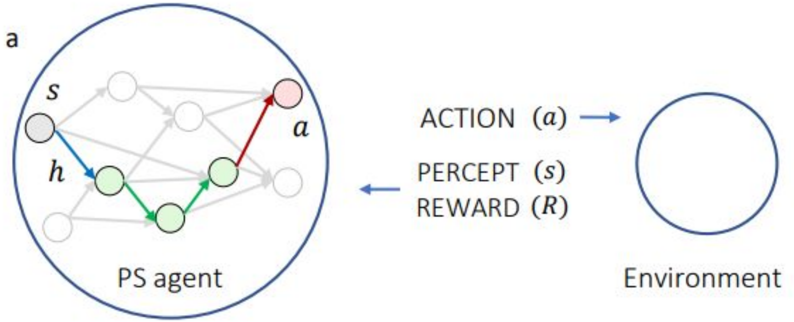
Classic Projective Simulation

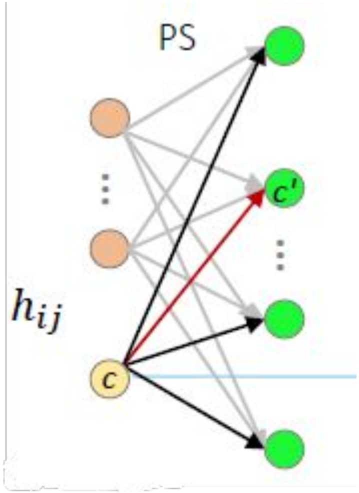
An “agent” navigates through an environment, doing a random walk, by projective his current state into a set of future states and decide which action to take.

$$p_{ij} = \frac{h_{ij}}{\sum_{j'} h_{ij'}}$$

Execute the action and evaluate with a reward system, which depends on the problem. Iterate to find the best paths.

$$h_{ij}^{(t+1)} = 1 + (1-\gamma)(h_{ij}^{(t)} - 1) + g_{ij}^{(t+1)} R$$



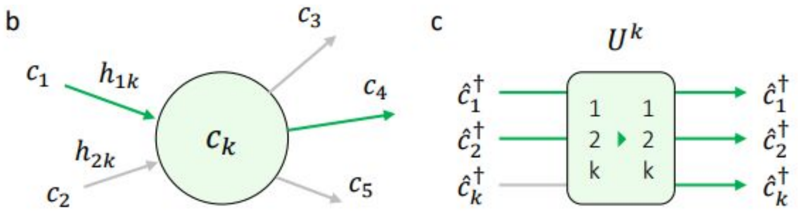


Quantum Projective Simulation

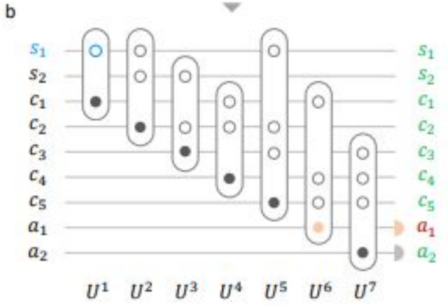
Decision-making is realized as a single-photon evolution in a mesh of tunable beam splitters with phase shifters.

The configuration is stored electronically in the phase shifters.

Each phase-shifter θ_{kl} at node (k,l) is set to implement the transition probabilities for the corresponding clip-to-clip connections.

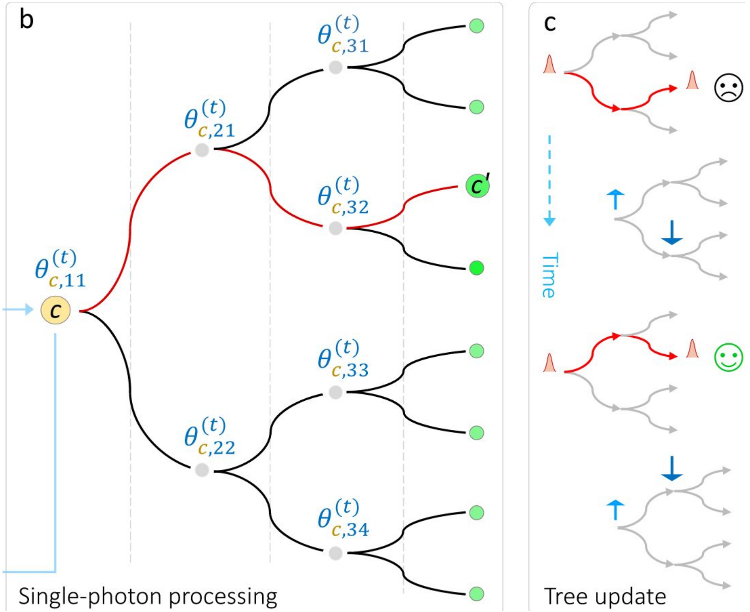


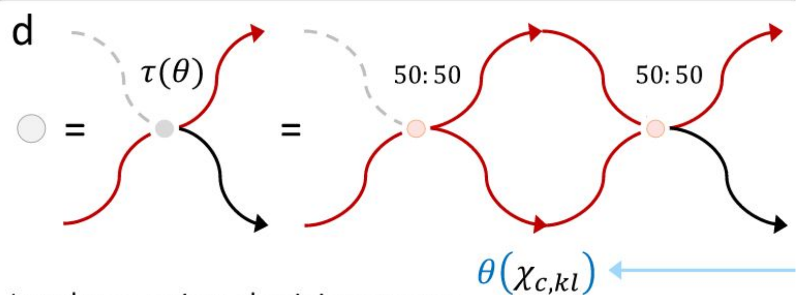
$$\hat{c}_j^\dagger(\tau_n) = \sum_{i=1}^{|C|} U_{ji}^{ECM}\, \hat{c}_i^\dagger(\tau_0); \qquad U^{ECM}(\vec{\theta}) = \prod_{k=1}^{n} U^k(\vec{\theta})$$



Circuit Building

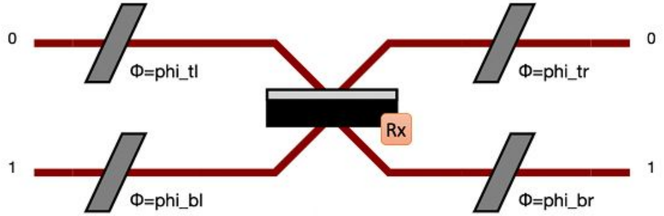
Arbitrary probabilistic transitions can be implemented on a photonic platform with a cascade of beamsplitters, whose transmissivities reproduce the distribution probability.





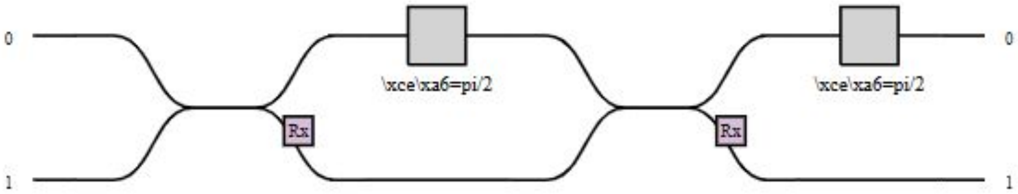
Circuit Building

Principal components: Beam Splitter (BS), Phase Shifters (PS).



$$\begin{pmatrix} e^{i(\phi_{tl}+\phi_{tr})}\cos(\frac{\theta}{2}) & ie^{i(\phi_{tr}+\phi_{bl})}\sin(\frac{\theta}{2}) \\ ie^{i(\phi_{tl}+\phi_{br})}\sin(\frac{\theta}{2}) & e^{i(\phi_{br}+\phi_{bl})}\cos(\frac{\theta}{2}) \end{pmatrix}$$

Mach-Zehnder Interferometer



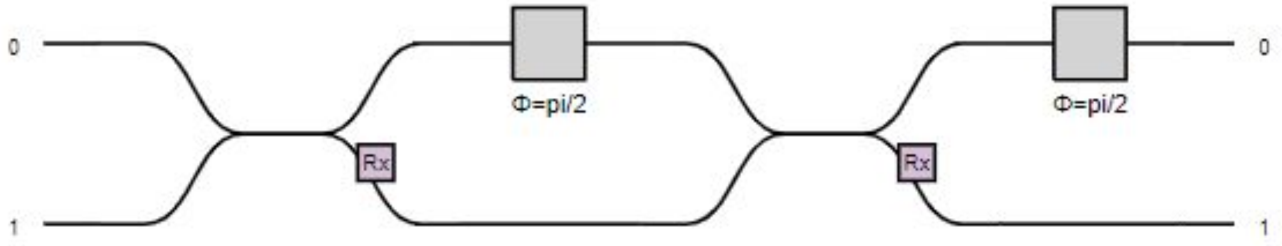
Implementation

- Grid: 2D array nxn.
- Reward: 0 for each step, 10 for reaching the end and -20 for crossing a wall.
- Creation of the optical circuit.
- Evaluation of the loss function to determine the best path.

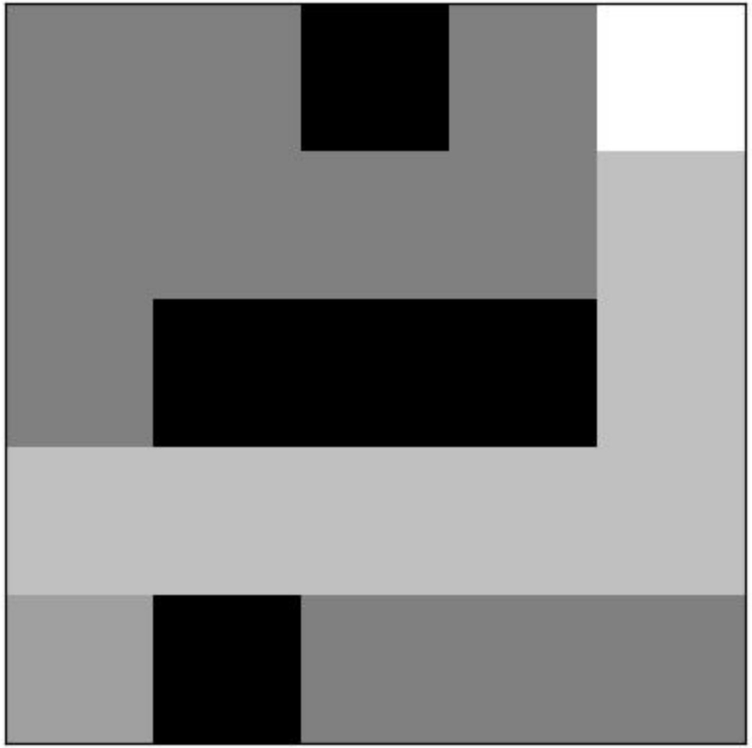
Same implementation with noise, emission probability 0.3

Up-Right optimization

Optical circuit



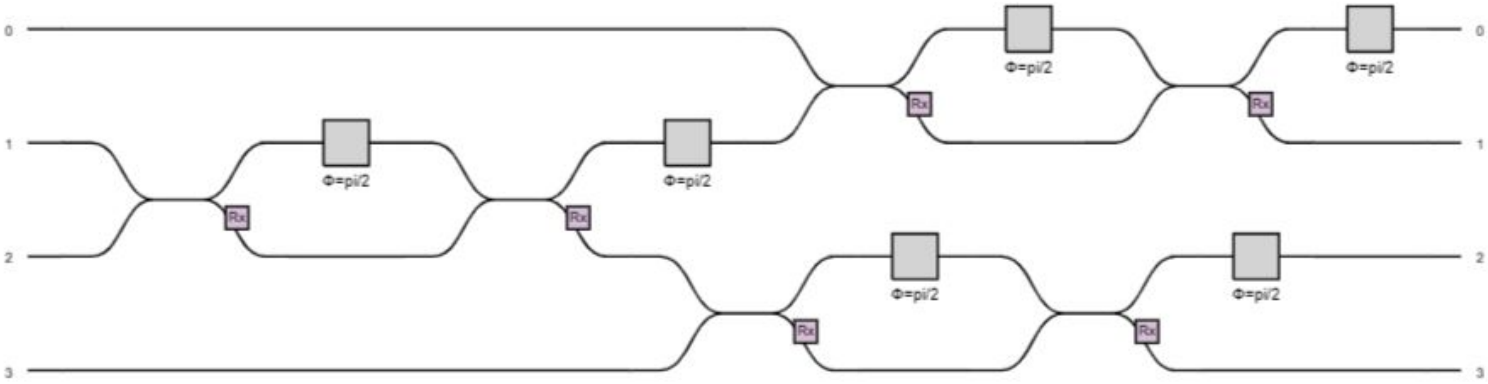
Up-Right after optimization



With or without noise we nearly get the same result

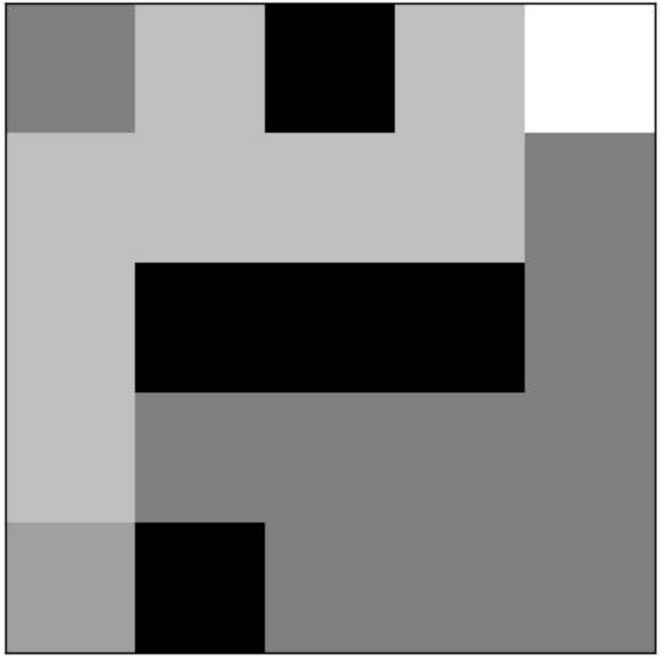
Generalization

Optical circuit

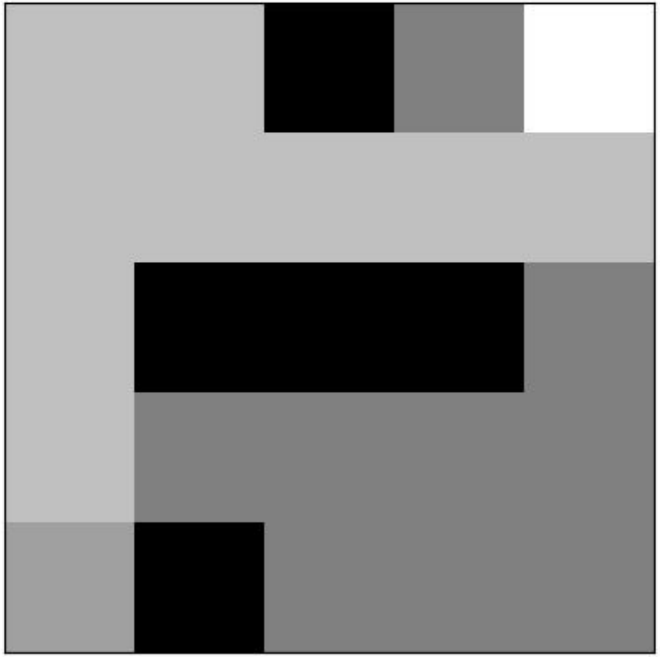


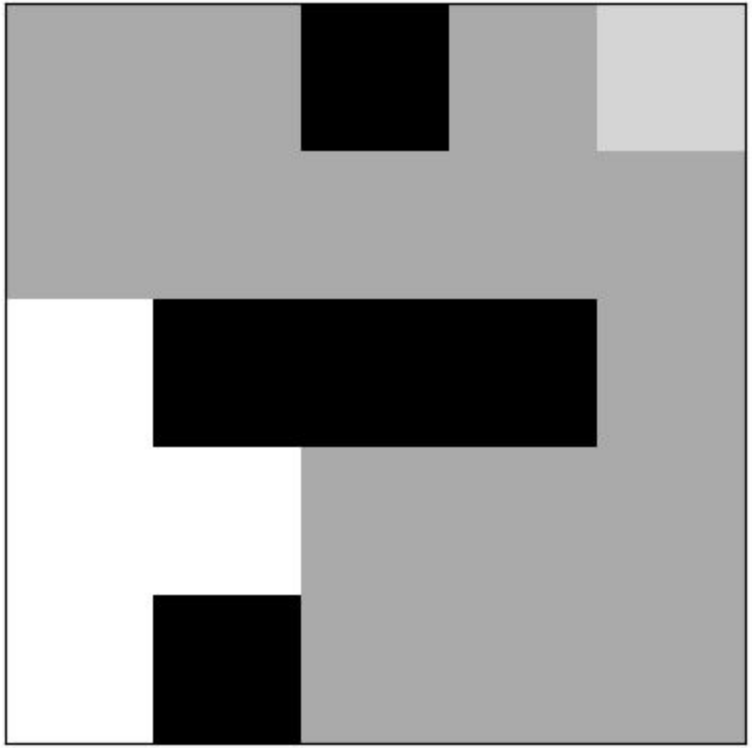
All possible directions - after optimization

With noise



Without noise





(0, 0) {0: 0.03619361979906159, 1: 0.962371789997033, 2: 0.0012302937247732286, 3: 0.0002042964791322087}

Thank you!