Neutral atom quantum computing scheduling by Deep Reinforcement Learning

M. Peracci
Polimi

L. Moro
QBrain

M. Maronese
QBrain

E. Prati
Unimi, CNR

The Quantum Intelligence Lab - Milan

Università Statale Milano

Istituto di Fotonica e Nanotecnnologie

CNR-IFN Milano

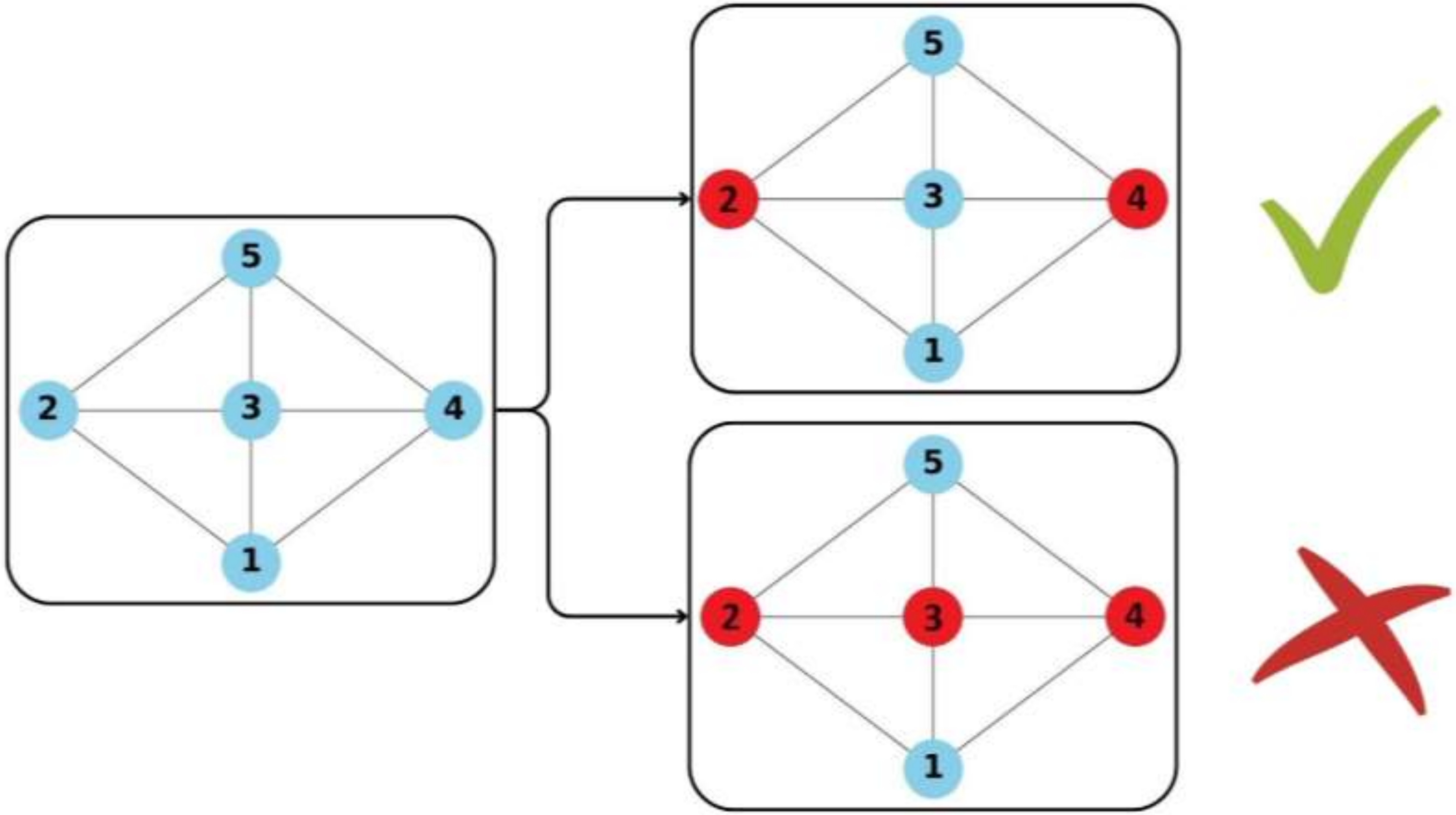
Politecnico Milano

Quantum Team

- Enrico Prati (Head, UNIMI)
- Paolo Zentilini (UNIMI)
- Luca Nigro (UNIMI)
- Francesco Monzani (UNIMI)
- Lorenzo Moro (POLIMI)
- Gabriele Agliardi (IBM POLIMI)
- Sebastiano Corli (POLIMI)
- Marco Maronese (UNIBO IIT)
- Lorenzo Rocutto (UNIBO)
- Andrea Zanetti (POLIMI)
- Davide Noè (Tohoku University)
- Matteo Di Giancamillo (POLIMI)
- Manuel Peracci (POLIMI)
- Rebecca Casati (UNIMI)
- Stefano Bruni (POLIMI)

The objective is to use the **Deep Reinforcement Learning** to decide the scheduling for the input pulse parameters for **Aquila neutral atoms quantum computer** in order to maximize the probability to find in the output a configuration that solves the **Maximum Independent Set** problem for an arbitrary graph.

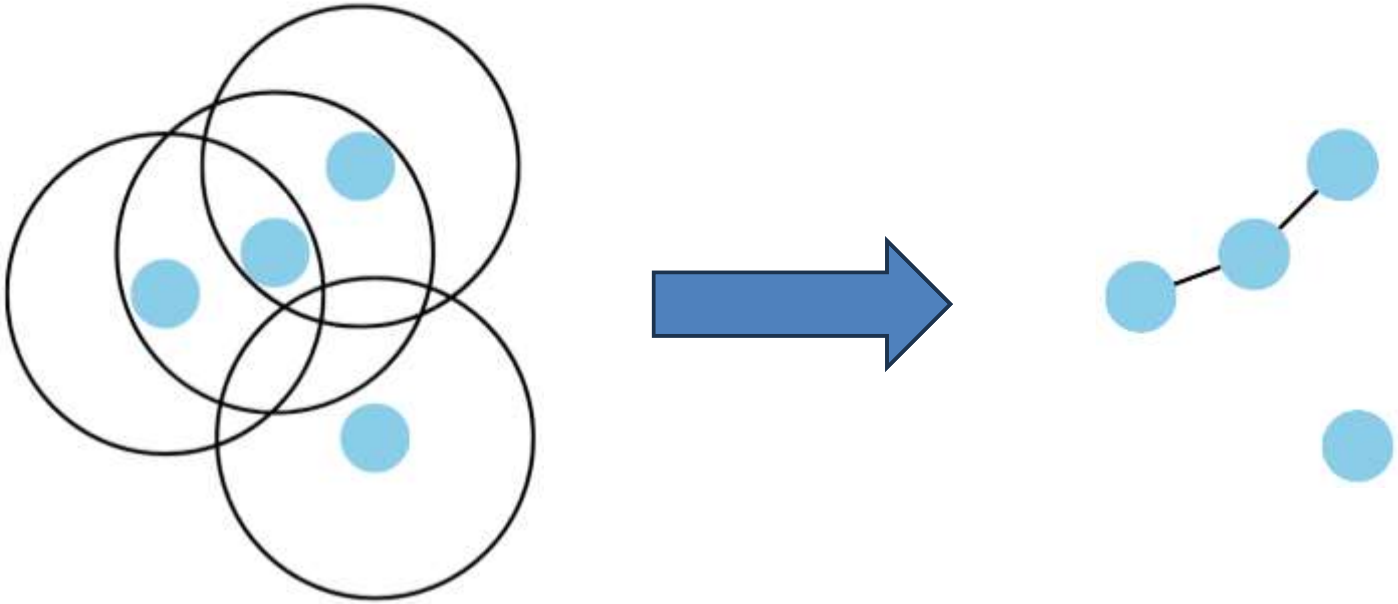
Maximum Independent Set problem



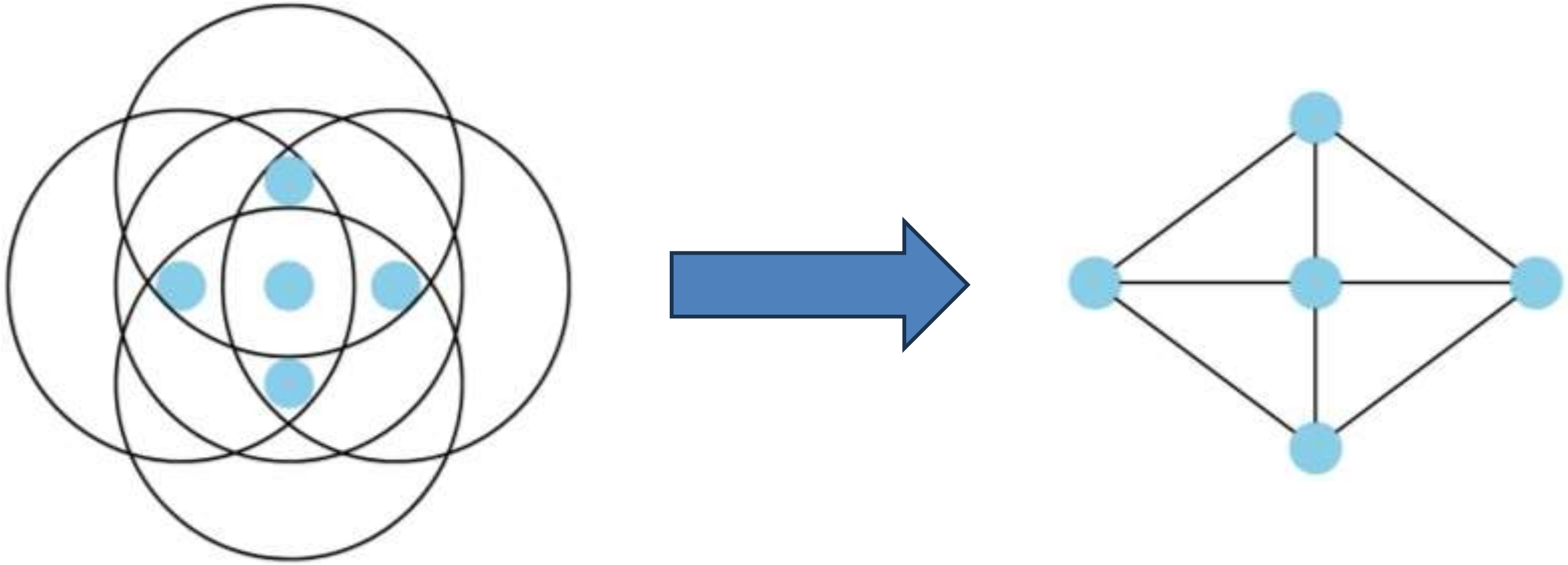
Maximum Independent Set (MIS) problem is the largest possible group of vertices within a graph such that no pair of selected vertices share an edge.

Diagonal-connected unit-disk grid graphs (DUGG)

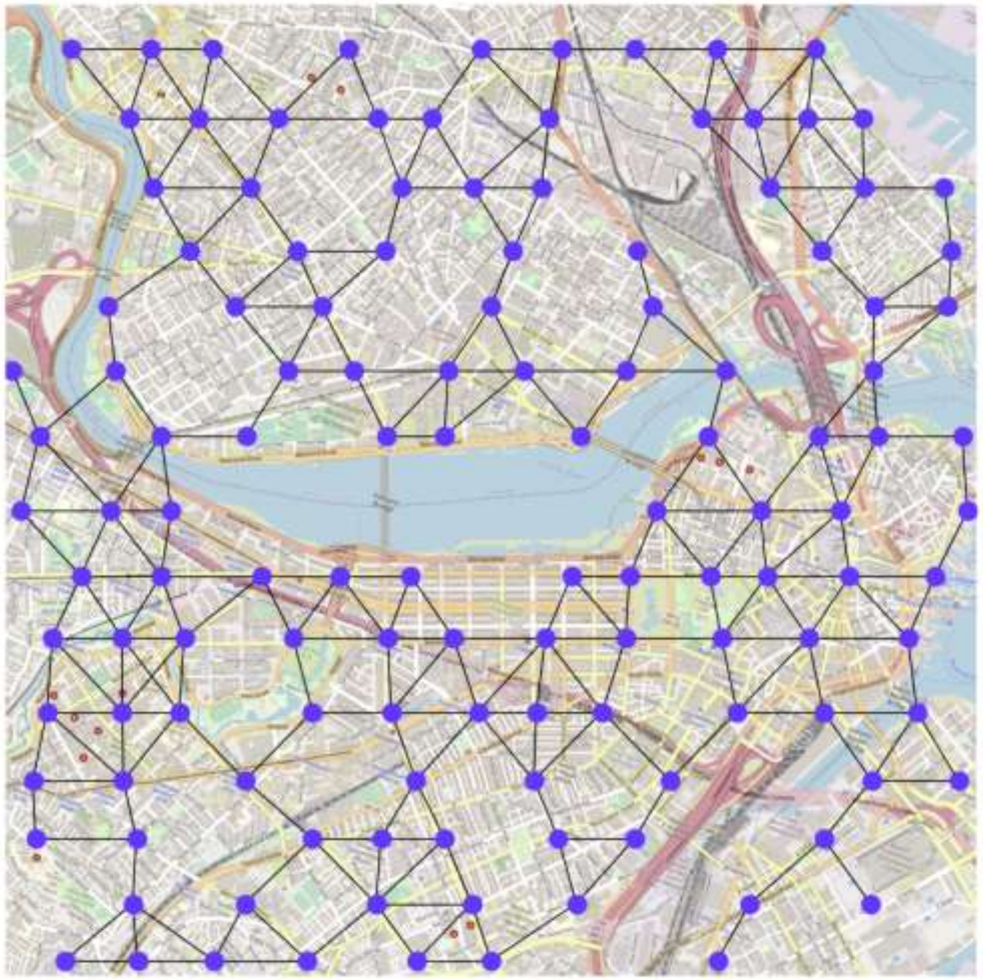
Unit disk graph



Diagonal-connected unit-disk grid graphs



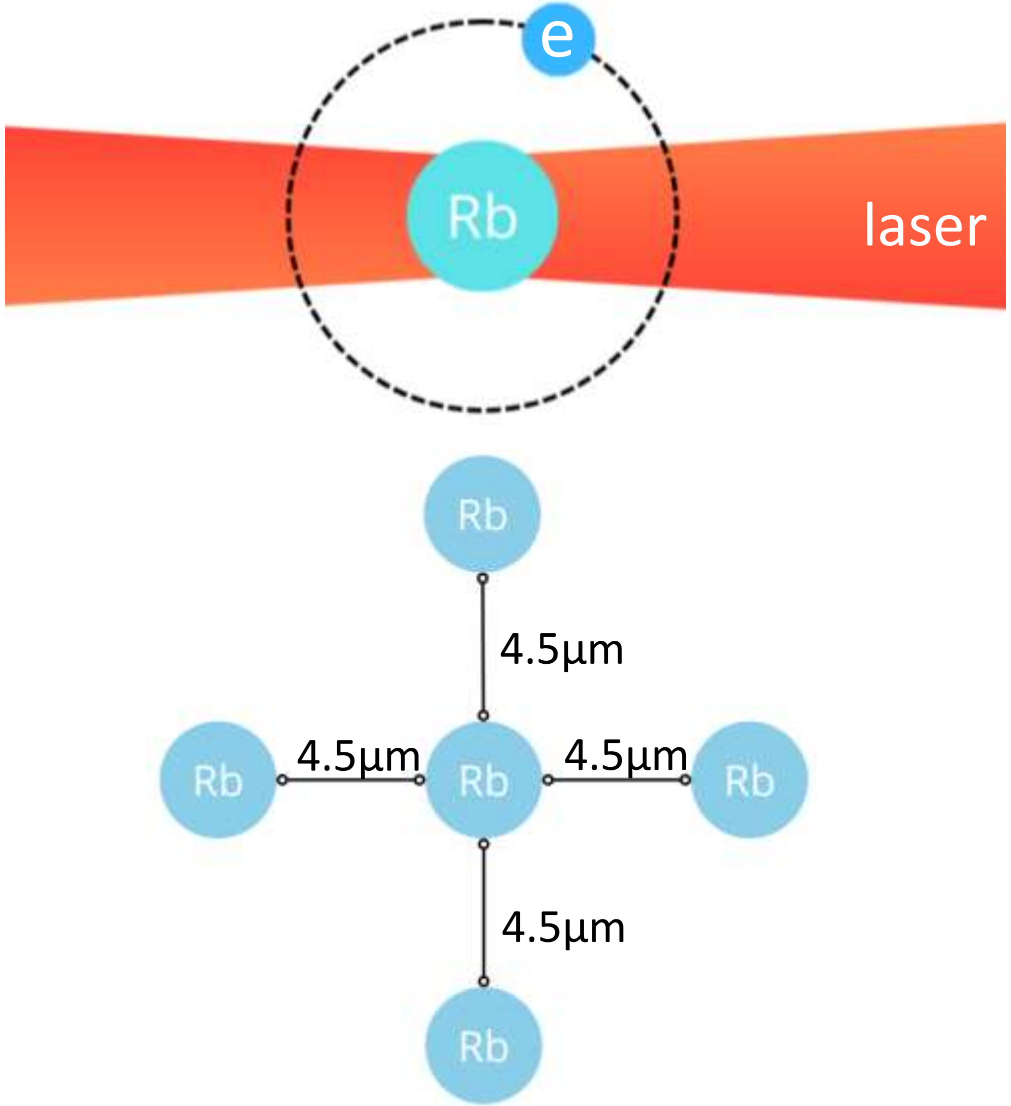
Real-world Maximum Independent Set problem



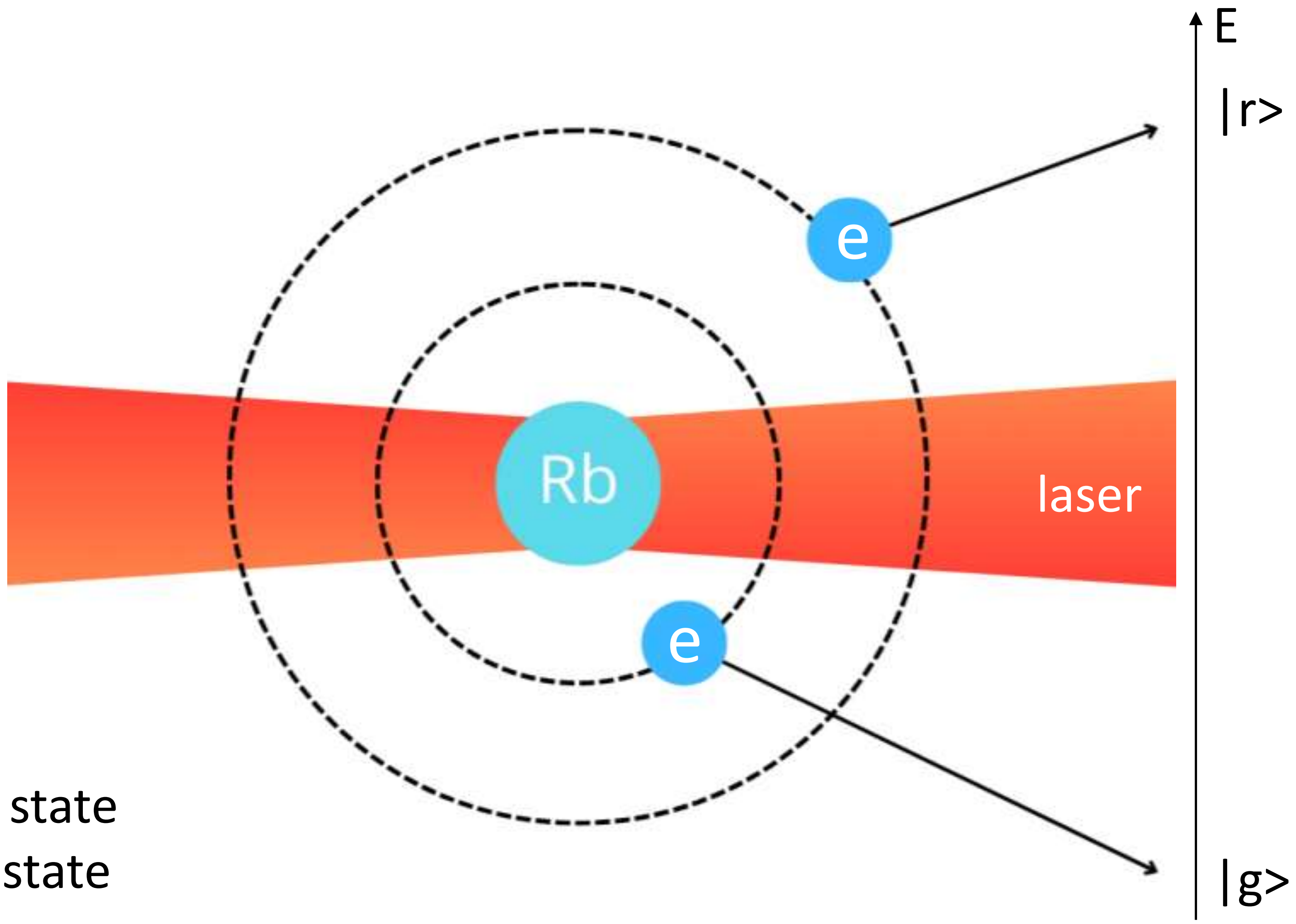
Antenna placement problem in Boston
From QuEraComputing Github

Neutral Atoms Quantum Computers

The Aquila magneto-optical trap in QuEra's facilities

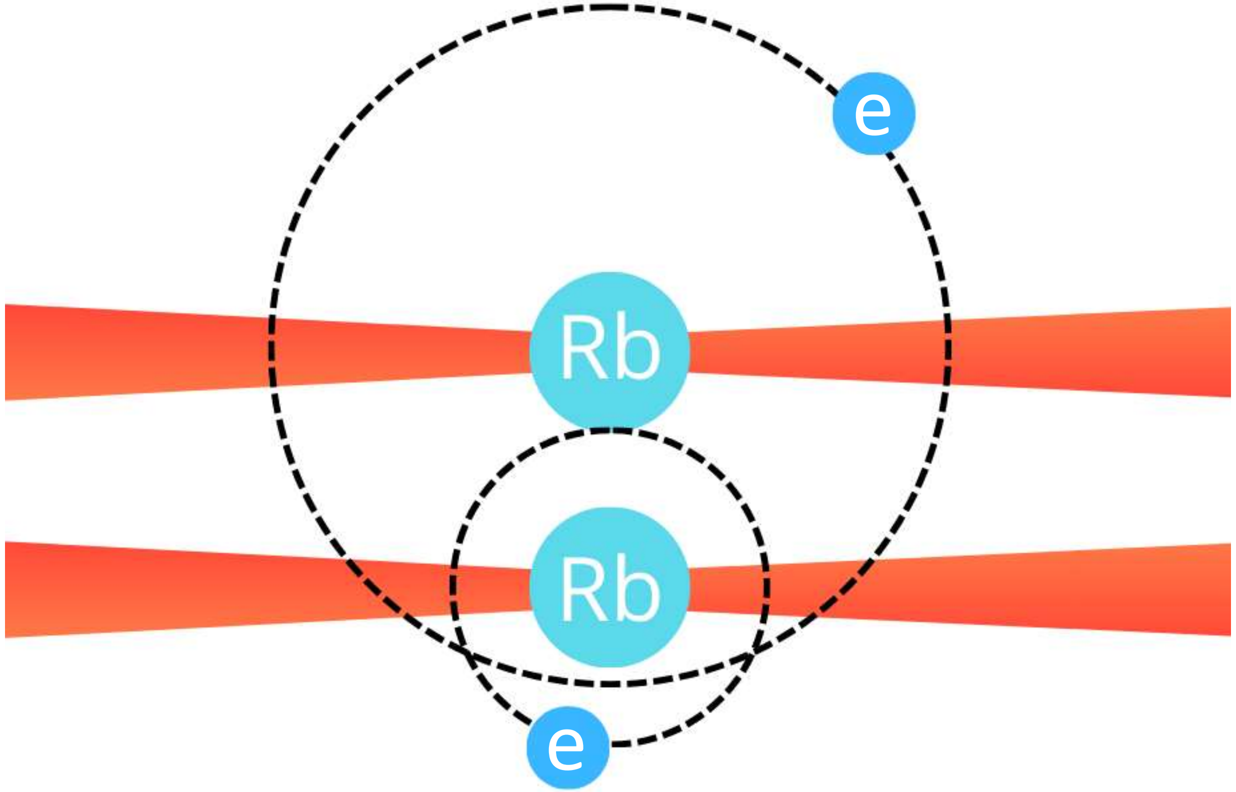


Ground and Rydberg states



|r> = Rydberg state
|g> = Ground state

Rydberg Blockade



$\{|gg\rangle, |gr\rangle, |rg\rangle\}$

Two atoms within a certain distance cannot be both in the Rydberg state

↓

Similar to the constraints on the MIS problem

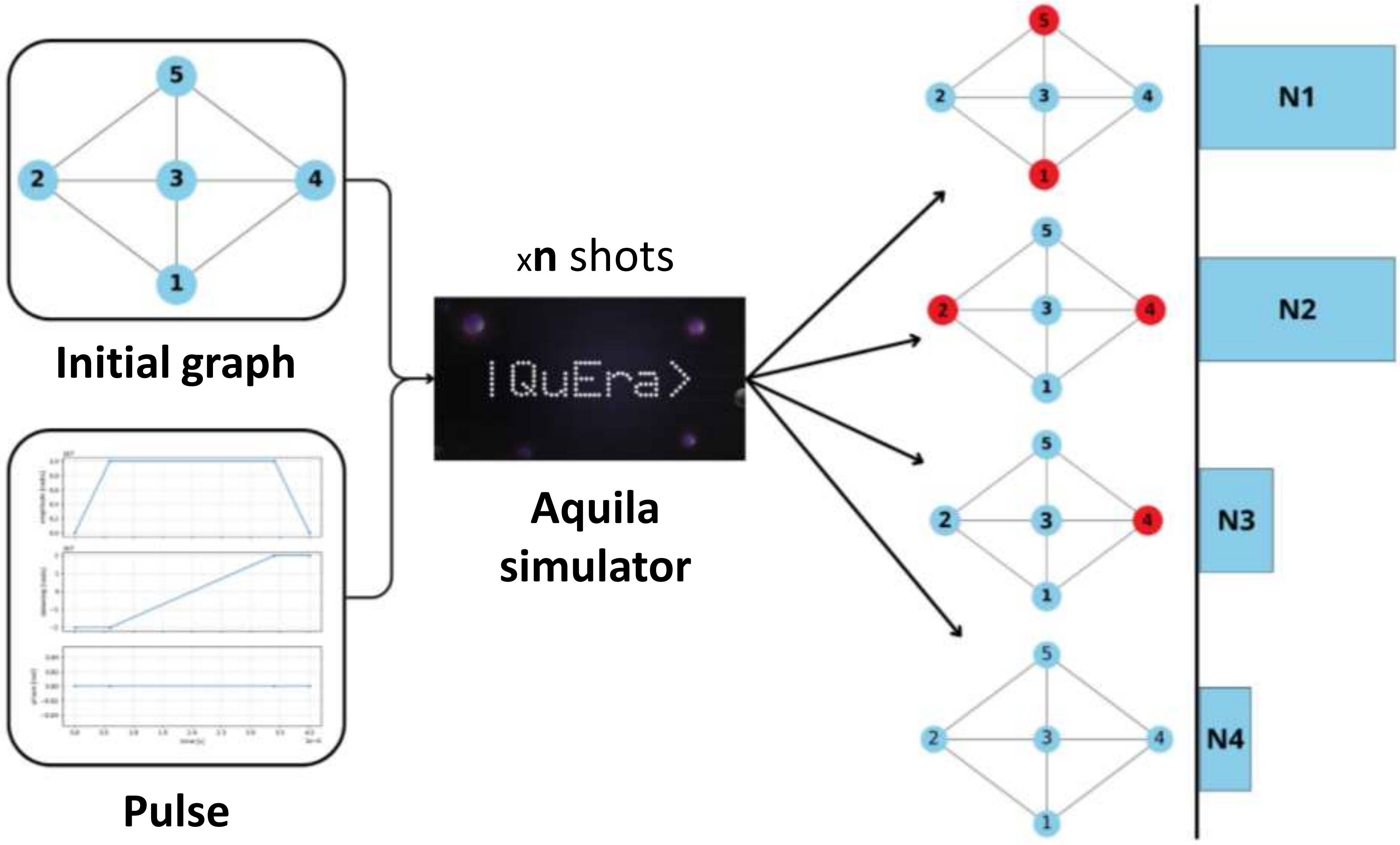
Analog Quantum Mode

$$H(t) = \frac{\Omega(t)}{2}\sum_i e^{i\phi(t)}|g_i\rangle\langle r_i| + e^{-i\phi(t)}|r_i\rangle\langle g_i| - \Delta(t)\sum_i \hat{n}_i + \sum_{i<j} \underbrace{\frac{C_6}{|\vec{x}_i - \vec{x}_j|^6}}_{\text{Van Der Waals interactions}} \hat{n}_i\hat{n}_j.$$

$\Omega(t)$	Rabi Frequency
$\Delta(t)$	Detuning
$\phi(t)$	Phase
$\vec{x}_i$	The position of the atoms
$\hbar$	Equal to 1
$\hat{n}_i = \lvert r_i\rangle\langle r_i\rvert$	Counts Rydberg excitations

Source: *Jonathan Wurtz et al. Aquila: QuEra's 256-qubit neutral-atom quantum computer, 2023*

Aquila simulations and MIS probability

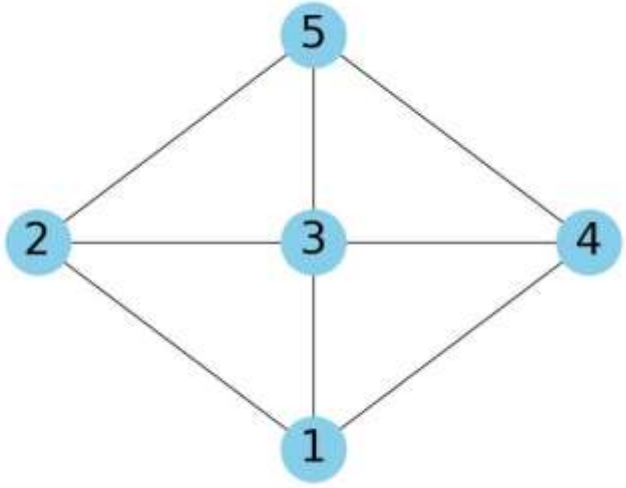


MIS probability

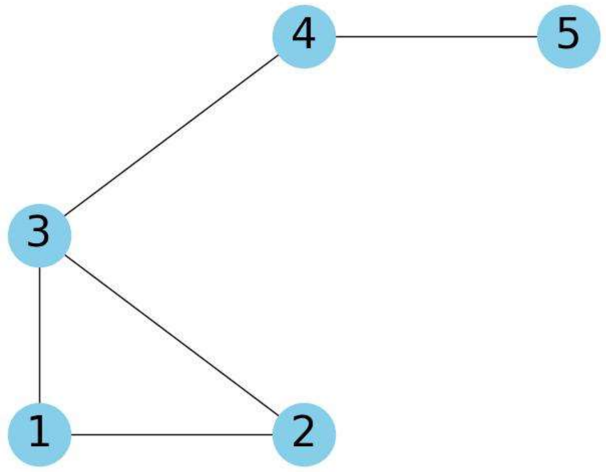
$$p = \frac{N1+N2}{n}$$

n = N1 + N2 + N3 + N4

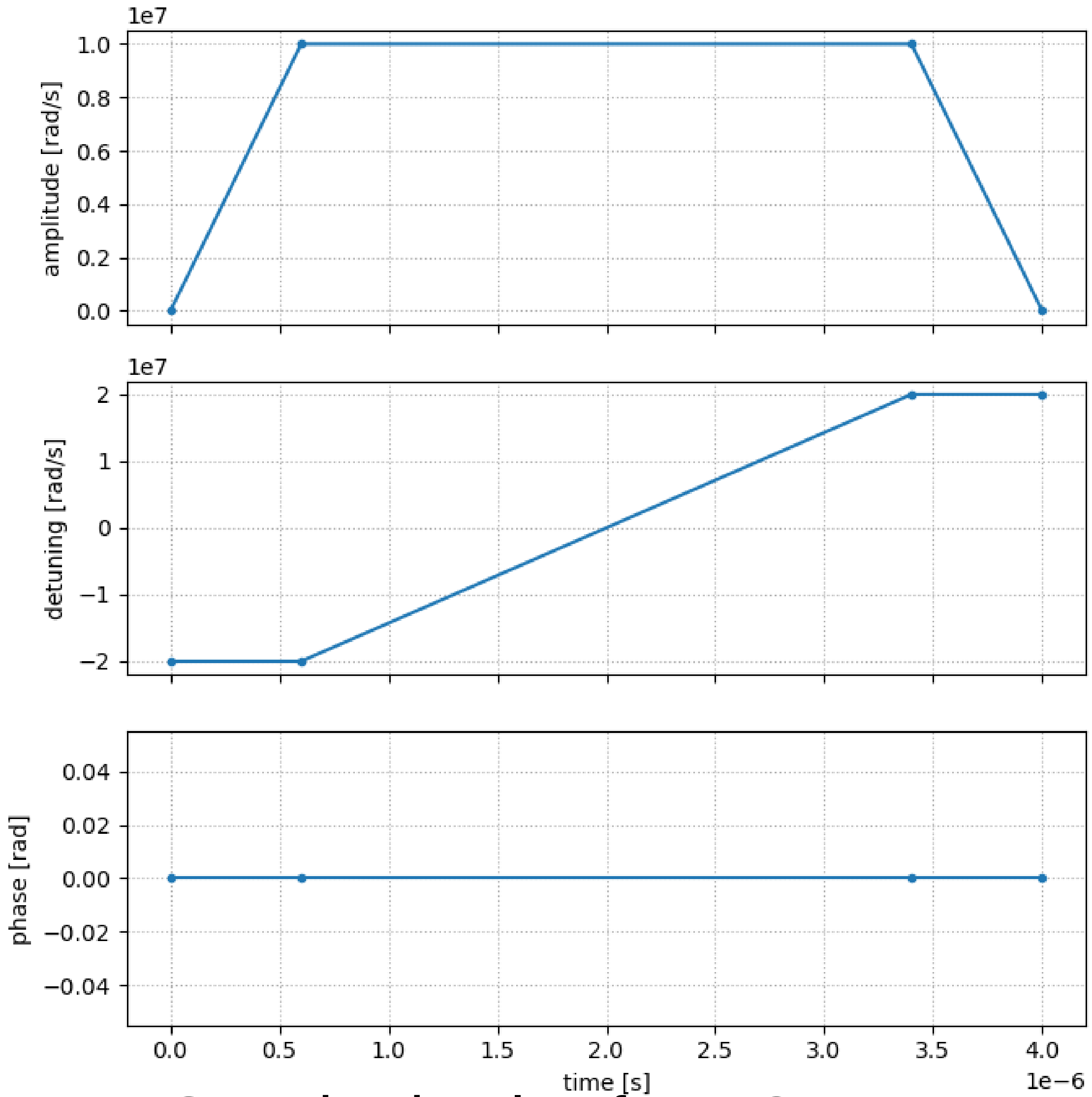
Standard Pulse from Quera



MIS probability
0.996 +- e-16

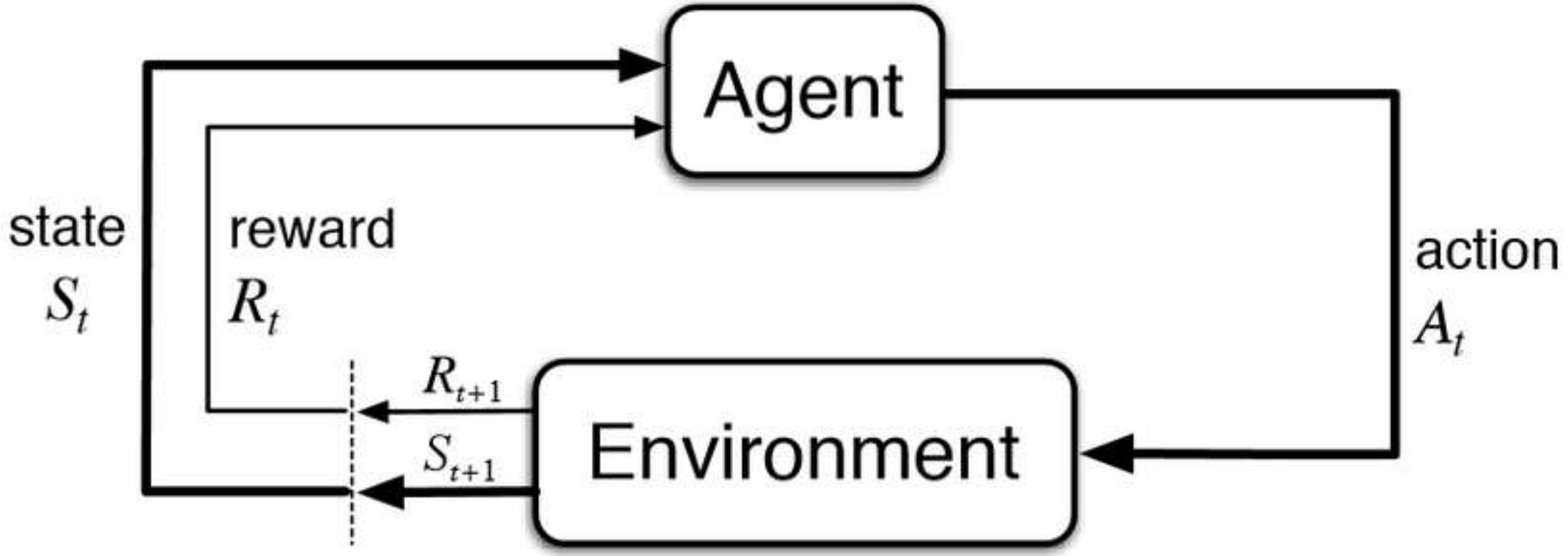


MIS probability
0.004 +- e-20



Standard pulse from Quera notebooks

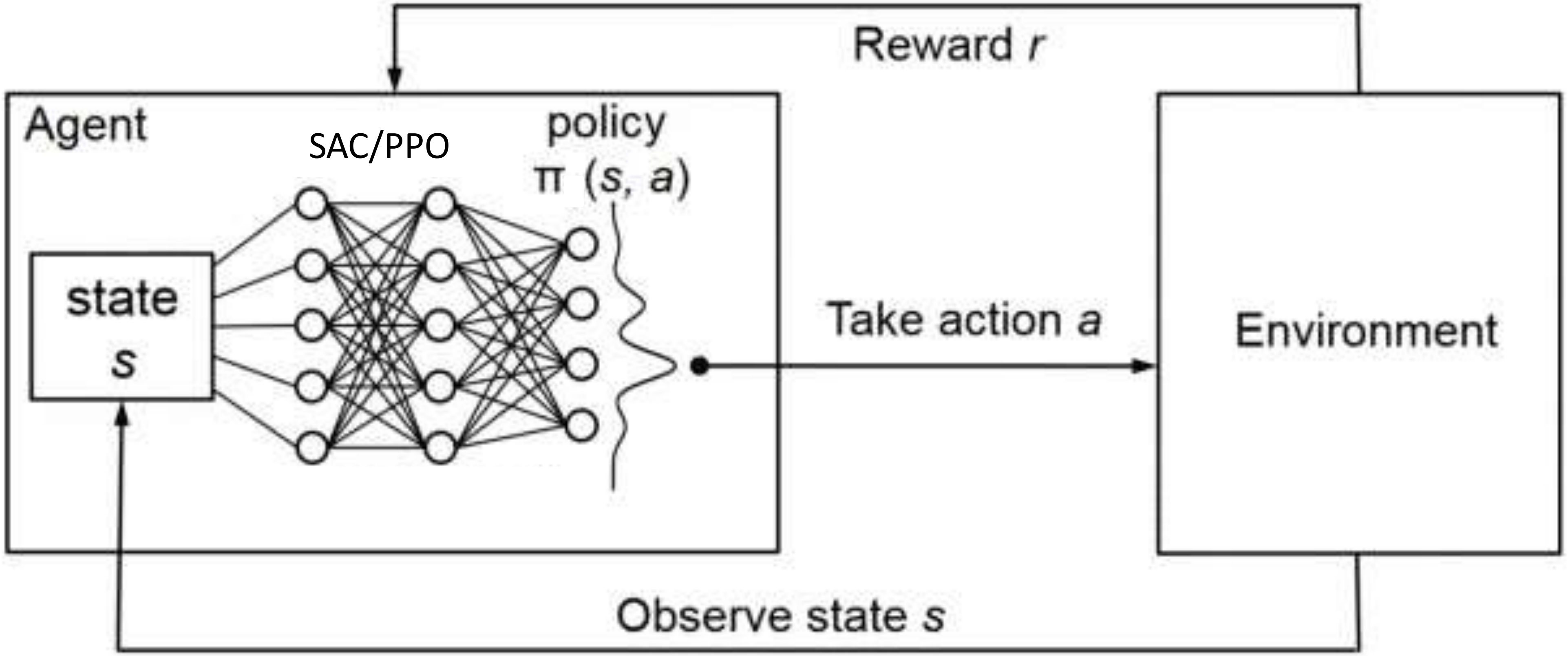
Reinforcement Learning



$\max_{\pi} E_{\pi}[\sum_{n=0}^{\infty} \gamma^{n} R_{t+k+1} | St = s]$ for each s ∈ S

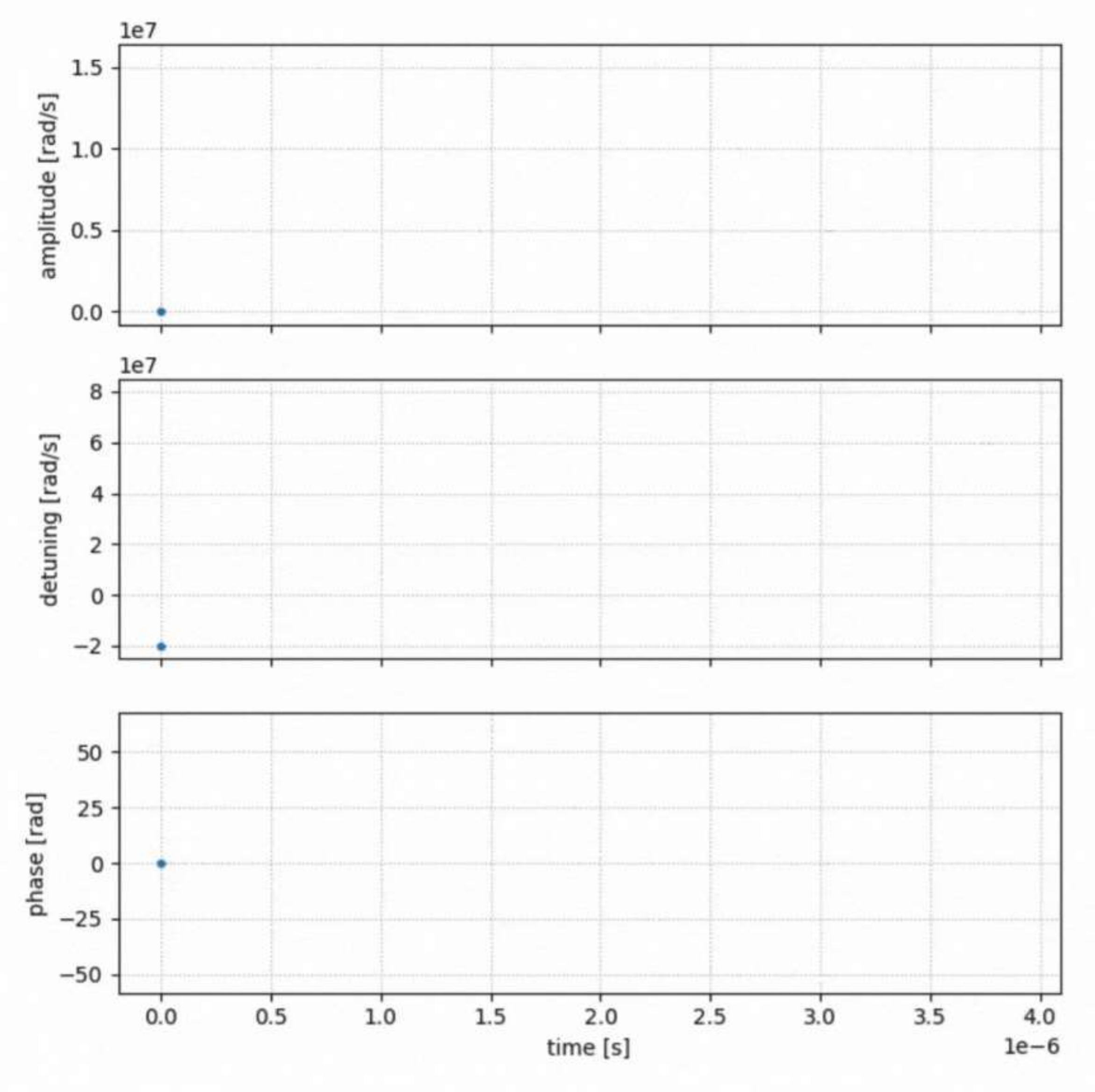
π(a|s) is the policy $\quad \gamma \in [0,1]$ discount factor

Source: Andrew G. Barto, Reinforcement Learning: An Introduction, 2018



PPO - Proximal Policy Optimization
SAC - Soft Actor-Critic

Reinforcement Learning Steps



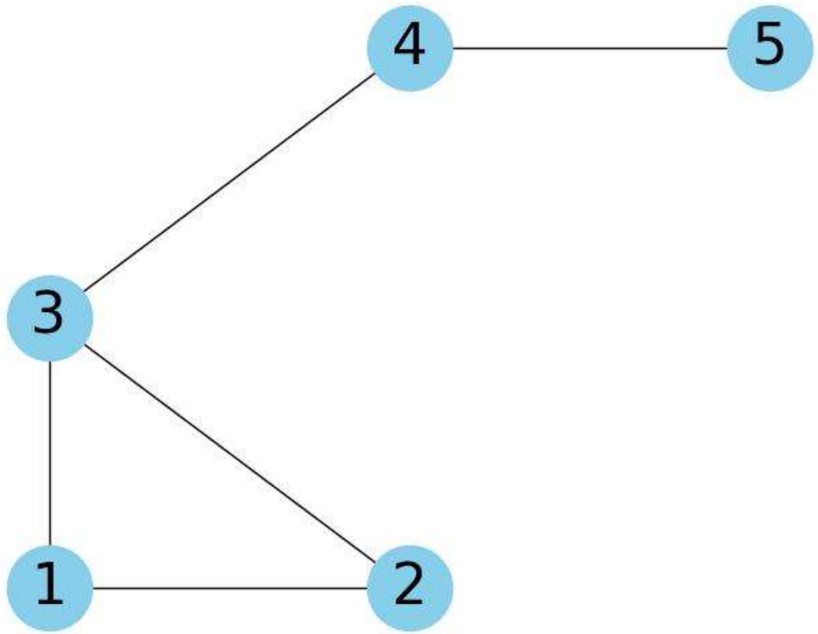
Reinforcement Learning Reward and State

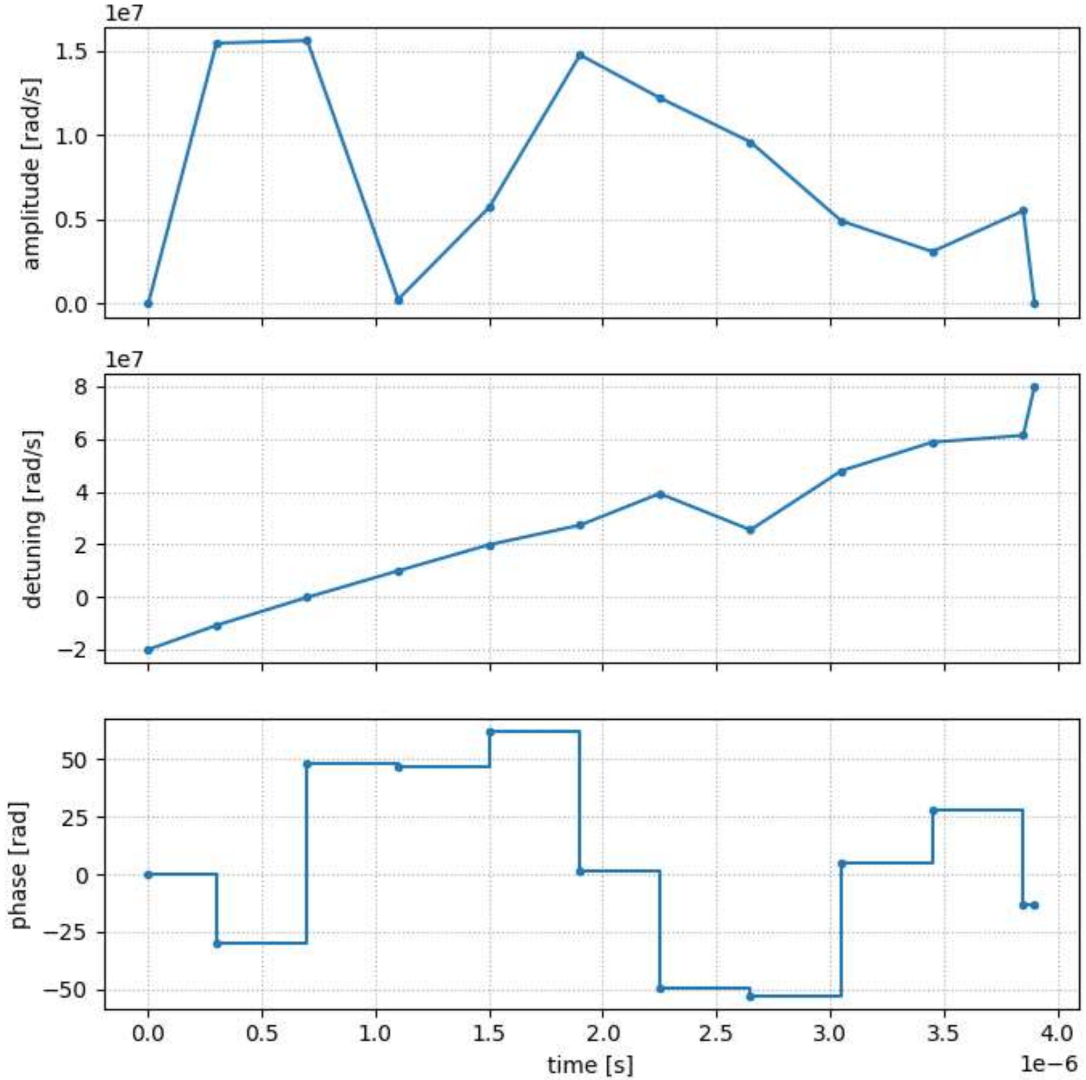
$$r = s(w_1 r_e + w_2 r_p + rx)$$

- s ϵ [0,1]
- w_1 and w_2 are tunable hyperparameters such that w1 +w2 = 1
- r_xϵ [-1,0] is a penalization reward if the initial graph is different from the current one
- r_e ϵ [-1,0] is the energy reward
- r_p is the probability reward

$$H(t) = \frac{\Omega(t)}{2} \sum_i e^{i\phi(t)} |g_i\rangle\langle r_i| + e^{-i\phi(t)} |r_i\rangle\langle g_i| - \Delta(t) \sum_i \hat{n}_i + \sum_{i<j} \frac{C_6}{|\vec{x}_i - \vec{x}_j|^6} \hat{n}_i \hat{n}_j.$$

Deep RL pulse may drastically increase MIS probability



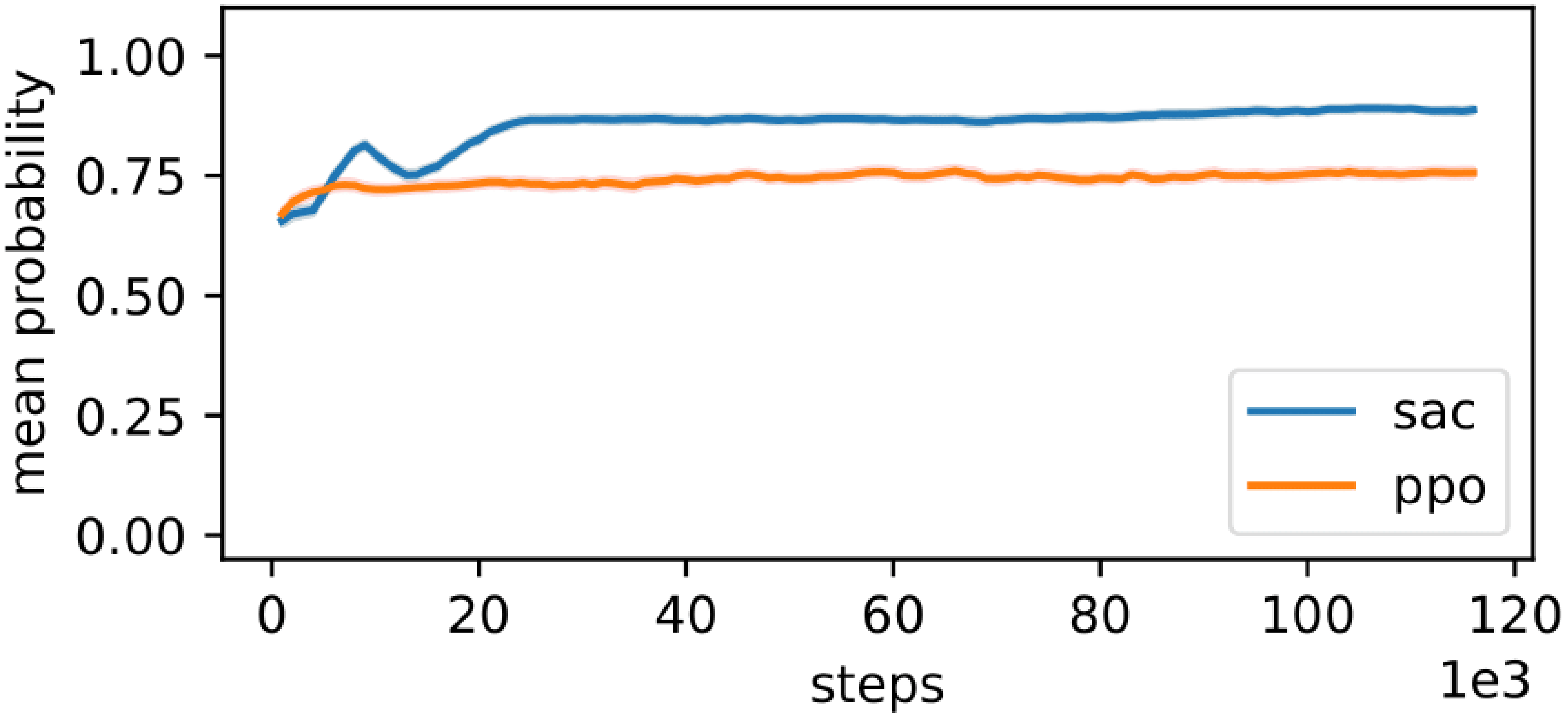


MIS probability with standard pulse
0.004 +- e-20

MIS probability With Deep RL pulse
0.89 +- 0.01

Pulse obtained with Deep RL

SAC makes learning faster wrt PPO



Conclusions

- Non decreasing detuning makes learning faster (Detuning < 0 favourites ground states, detuning > 0 favourites Rydberg states)
- Adding memory to the state makes learning faster
- Statistics: MIS probability of RL decreases with number of node similar to standard pulse