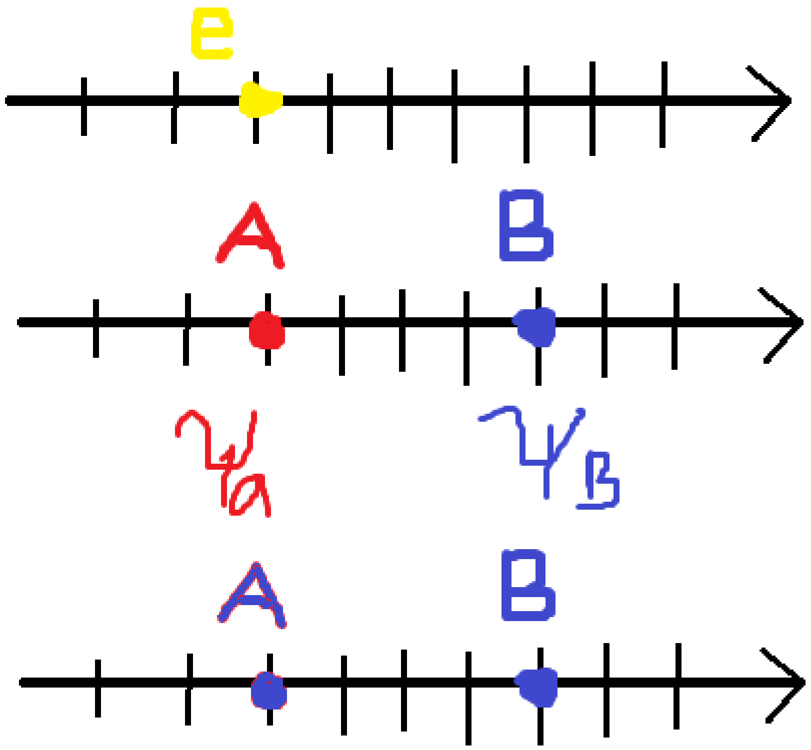
Creation and manipulation of quantized vortices in BEC using reinforcement learning

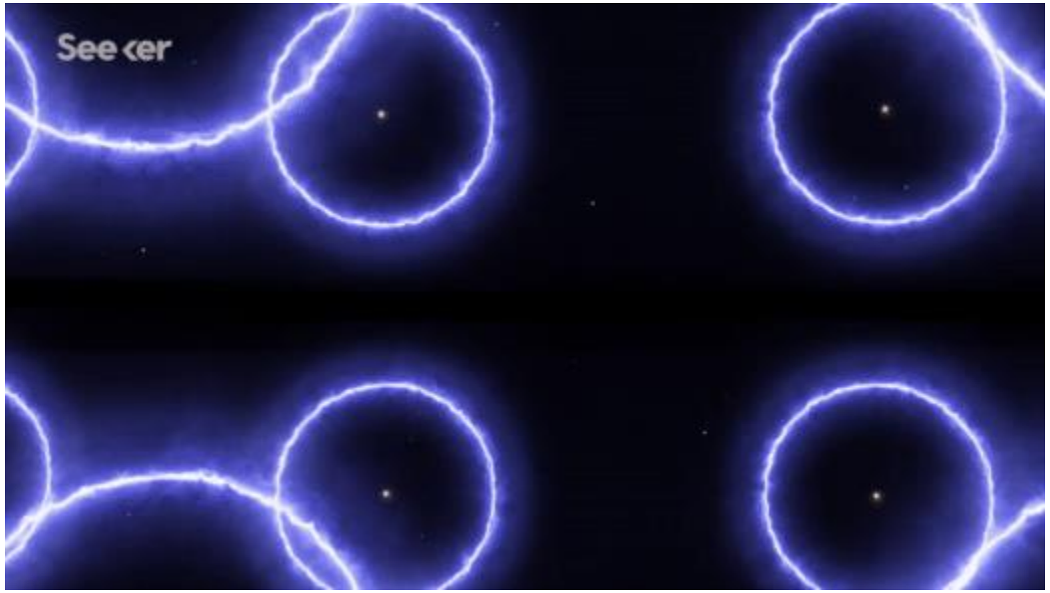
- Bosons and fermions
- Bose-Einstein Condensate
- Gross-Pitaevskii equation
- Environment and agent
- 2D system
- 3D system
- Conclusions and discussions

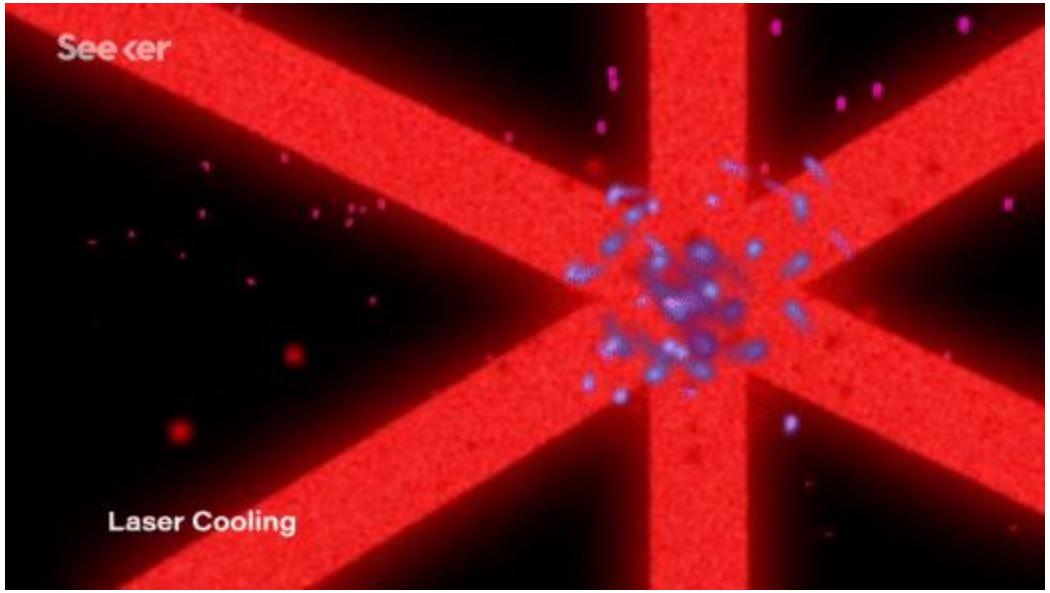
What are bosons?

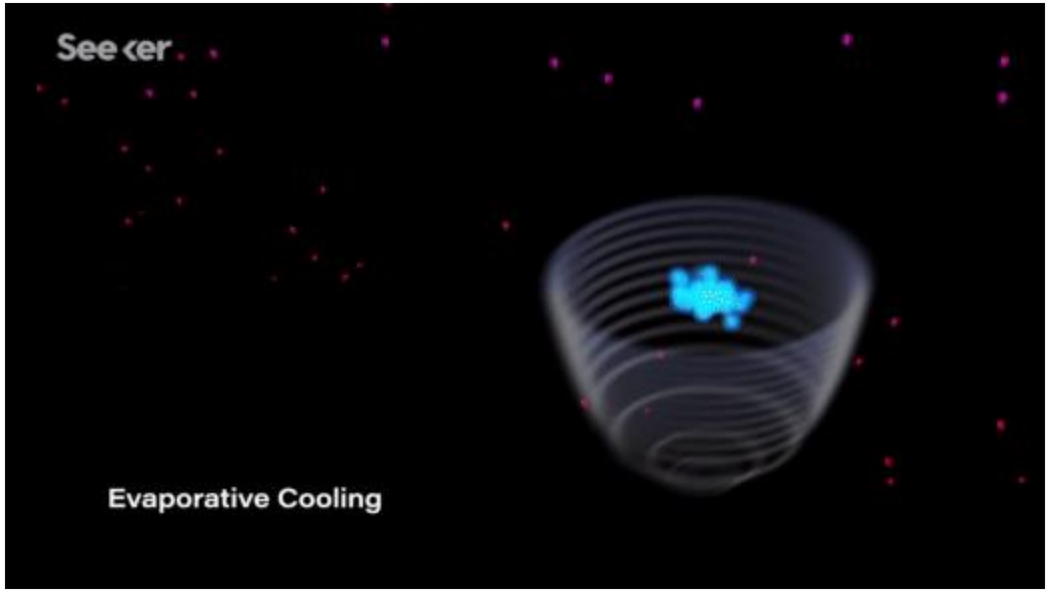
- Wave function $\psi(r1,r2)=+/-\psi(r2,r1)$
  A mathematical description of everything we know about a quantum system
- Quantum system



BOSE-EINSTEIN CONDENSATE


Seeker


Seeker
Laser Cooling


Seeker
Evaporative Cooling

- Time dependent GP Equation $i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) + g|\Psi(\mathbf{r},t)|^2\right)\Psi(\mathbf{r},t)$

- Environment

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V_{\text{trap}}(\boldsymbol{r})\psi + V_G(\boldsymbol{r},t)\psi + g|\psi|^2\psi$$

- Agent
  CNN outputs 6 values, denoted as $Q(s_t,a)$, $a=0..5$, $0<=r<1$, $0.1<\epsilon<1$

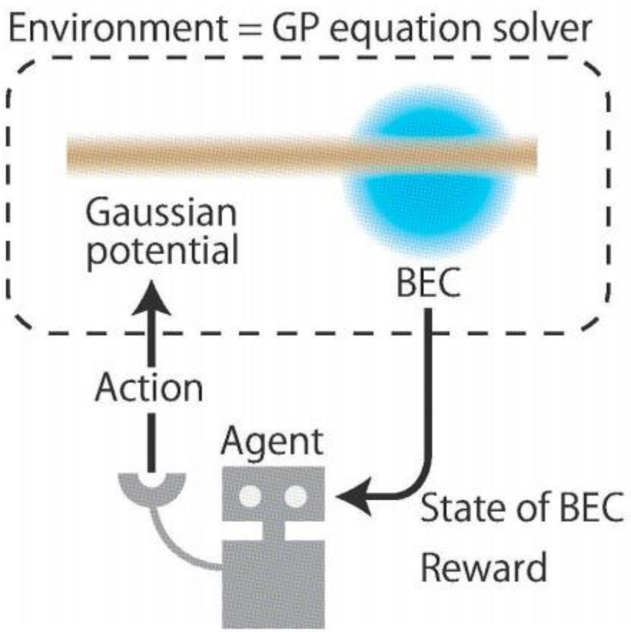
$$a_t = \begin{cases} \text{randomly chosen from 0 to 5} & (r < \epsilon) \\ \text{argmax}_a Q(s_t, a) & (r \geq \epsilon). \end{cases}$$

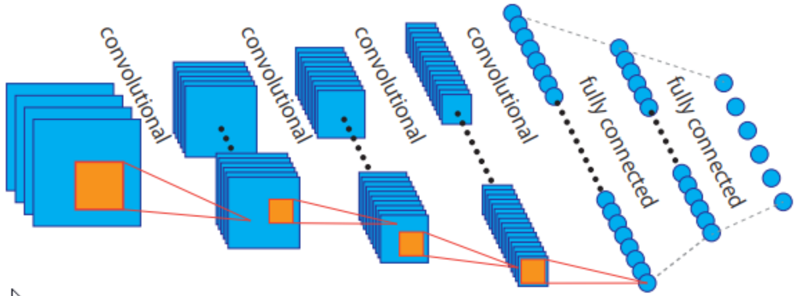
Bellman optimality equation

$$y_t = r_t + \gamma\hat{Q}(s_{t+\Delta t}, \text{argmax}_a Q(s_{t+\Delta t}, a)).$$

Algorithm 1

Initialize deep-Q network Q
Initialize target network as $\hat{Q} = Q$
for episode = 1, N_{episode} **do**
  Initialize environment and get initial observation s_0
  for $t = 0, T_{\text{end}}$ **do**
    Select action $a_t = \text{argmax}_a Q(s_t, a)$ with ϵ-greedy exploration
    Execute action a_t on environment and get reward r_t and next state $s_{t+\Delta t}$
    Store $(s_t, a_t, r_t, s_{t+\Delta t})$ in replay memory
    Sample minibatch of $(s_t, a_t, r_t, s_{t+\Delta t})$ from replay memory
    Set $y_t = r_t$ if $t = T_{\text{end}}$, otherwise $y_t = r_t + \gamma\hat{Q}(s_{t+\Delta t}, a')$, where $a' = \text{argmax}_a Q(s_{t+\Delta t}, a)$
    Train network using gradient of $L(y_t - Q(s_t, a_t))$
    Copy $\hat{Q} = Q$ every C steps
  end for
end for





2D layer shape	$64^2 \times 4$	$15^2 \times 32$	$7^2 \times 64$	$3^2 \times 128$	256	128	6
filter size	8^2	3^2	3^2	3^2			
stride	4	2	2	2			
3D layer shape	$32^3 \times 5$	$15^3 \times 32$	$7^3 \times 64$	$3^3 \times 128$	256	128	6
filter size	4^3	3^3	3^3	3^3			
stride	2	2	2	2			

Results

- 2D systems

$$i\hbar \frac{\partial \psi_\perp}{\partial t} = -\frac{\hbar^2}{2m}\nabla_\perp^2\psi_\perp + V_{\mathrm{trap}}(x,y)\psi_\perp + V_G(x,y,t)\psi_\perp + g_\perp|\psi_\perp|^2\psi_\perp.$$

$$\begin{aligned} a &= 0 & \xi &\to \xi + 0.15, \\ a &= 1 & \xi &\to \xi - 0.15, \\ a &= 2 & \eta &\to \eta + 0.15, \\ a &= 3 & \eta &\to \eta - 0.15, \\ a &= 4 & A_\perp &\to A_\perp + 2, \\ a &= 5 & A_\perp &\to A_\perp - 2, \end{aligned}$$

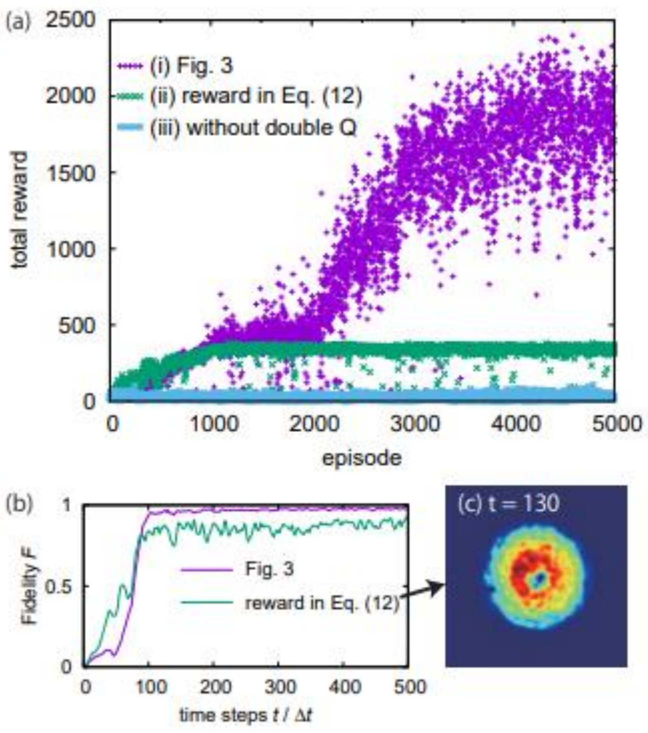
  Initial parameters $\xi(0)=0$, $\eta(0)=2$, $A_\perp(0)=20$.

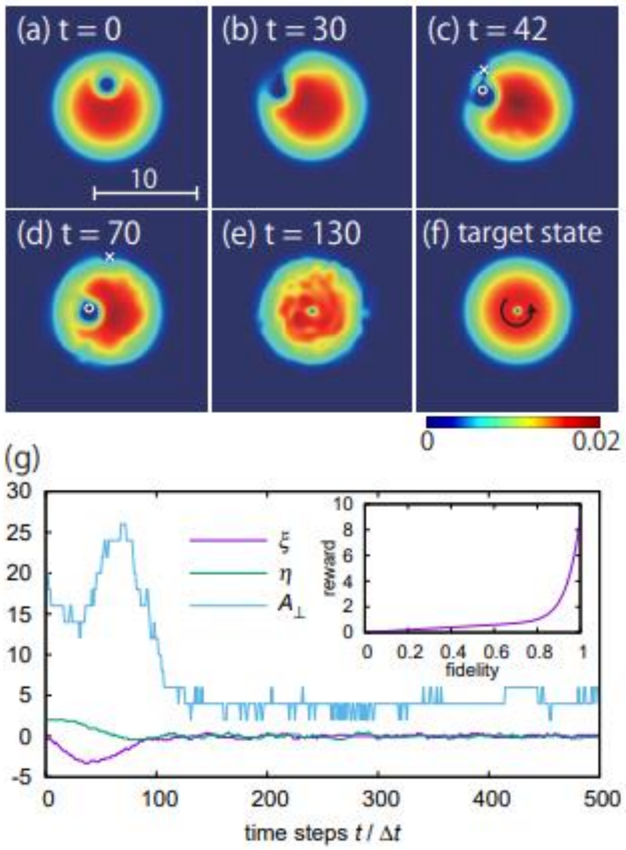
- Fidelity $F(t) = \left| \int \psi^*_{\mathrm{target}}(x,y)\psi_\perp(x,y,t)dxdy \right|^2$

  Aim to increase fidelity, reward r_t is monotonically increasing $r_t = F(t) + 8[F(t)]^{16}$

- Total reward $R_{\mathrm{total}} = \sum_{n=0}^{500} r_{n\Delta t},$

- Target state: Single counterclockwise vortex at the center





Results

- 3D systems

$$V_G(\mathbf{r},t) = A(t)\exp\left\{\frac{y^2}{d_y^2(t)} + \frac{[z-\zeta(t)]^2}{d_z^2}\right\}$$

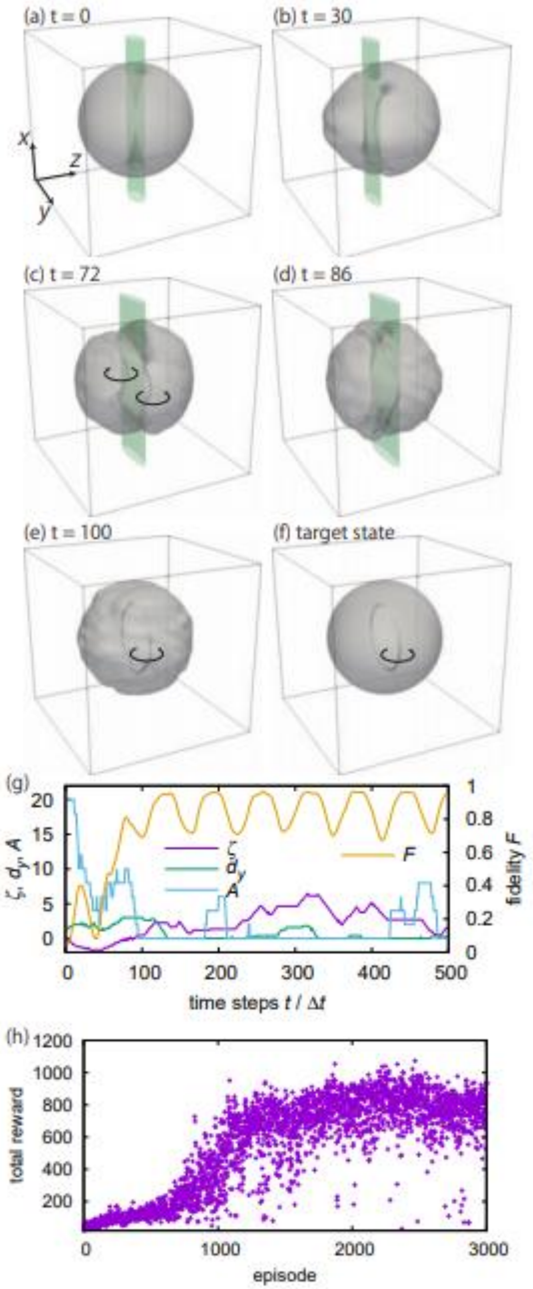
Initial parameters $\zeta(0)=0$, $d_y(0)=1$, $d_z=0.5$, $A(0)=20$.

$$\begin{aligned}
a &= 0 & \zeta &\to \zeta + 0.15,\\
a &= 1 & \zeta &\to \zeta - 0.15,\\
a &= 2 & d_y &\to d_y + 0.2,\\
a &= 3 & d_y &\to d_y - 0.2,\\
a &= 4 & A &\to A + 2,\\
a &= 5 & A &\to A - 2,
\end{aligned}$$

- Fidelity $F(t) = \left|\int \psi^*_{\text{target}}(\mathbf{r})\psi(\mathbf{r},t)d\mathbf{r}\right|^2$

  Reward has same form as before.

- Target state: Stationary state with a vortex ring.



Conclusions and discussions