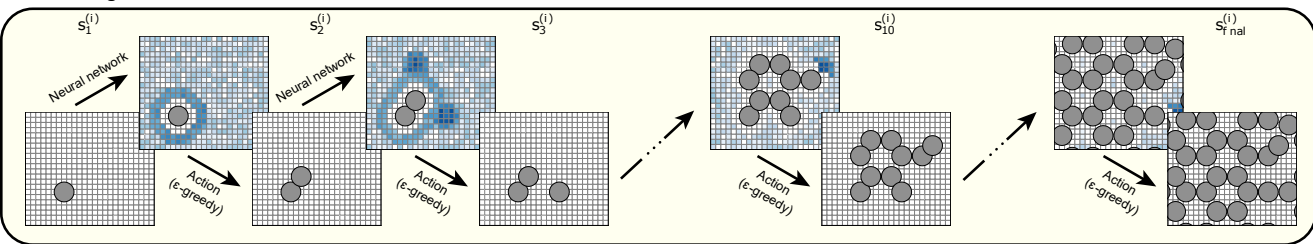
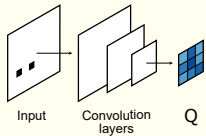


## Building Phase

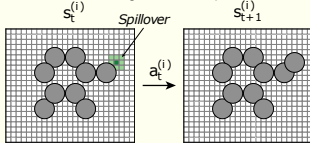


## Training Phase

Neural network weight update

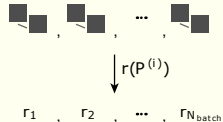


Training data example

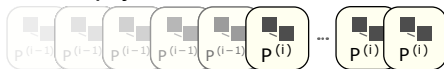


$$Q_{\text{target}}(s_t^{(i)}, a_t^{(i)}) = r(P^{(i)})$$

Calculating training batch rewards



Replay buffer



## Evaluation Phase

DFT calculation

