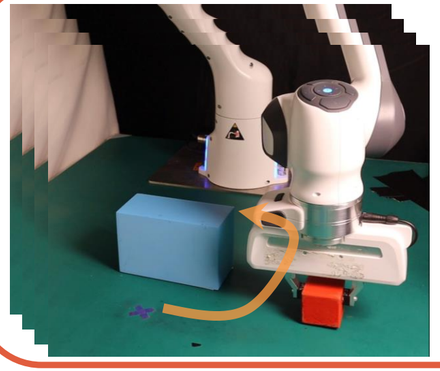
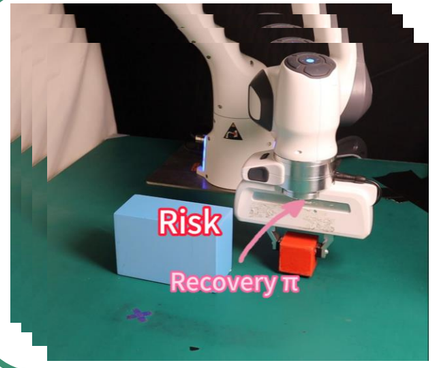
Offline Phase



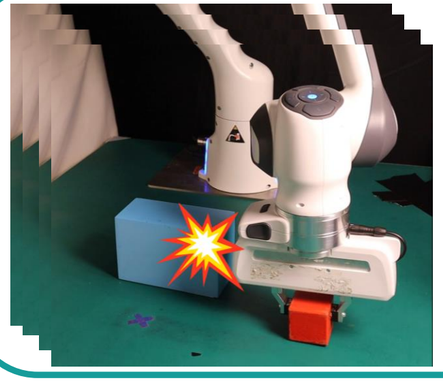
Task Demos & Policy

π_T



Recovery Demos & Policy

π_R



Failure Demos & World Model

WM

Online Phase

Safe Evaluation

C_H^π

$< \varepsilon_{safe}$

$> \varepsilon_{safe}$

a_{task}

π_T

s

Env

a_{task}

a_{recovery}

π_R

$\mathcal{D}$

a_{task}

Safe Exploration

π_T

Task policy update

$\mathcal{D}$