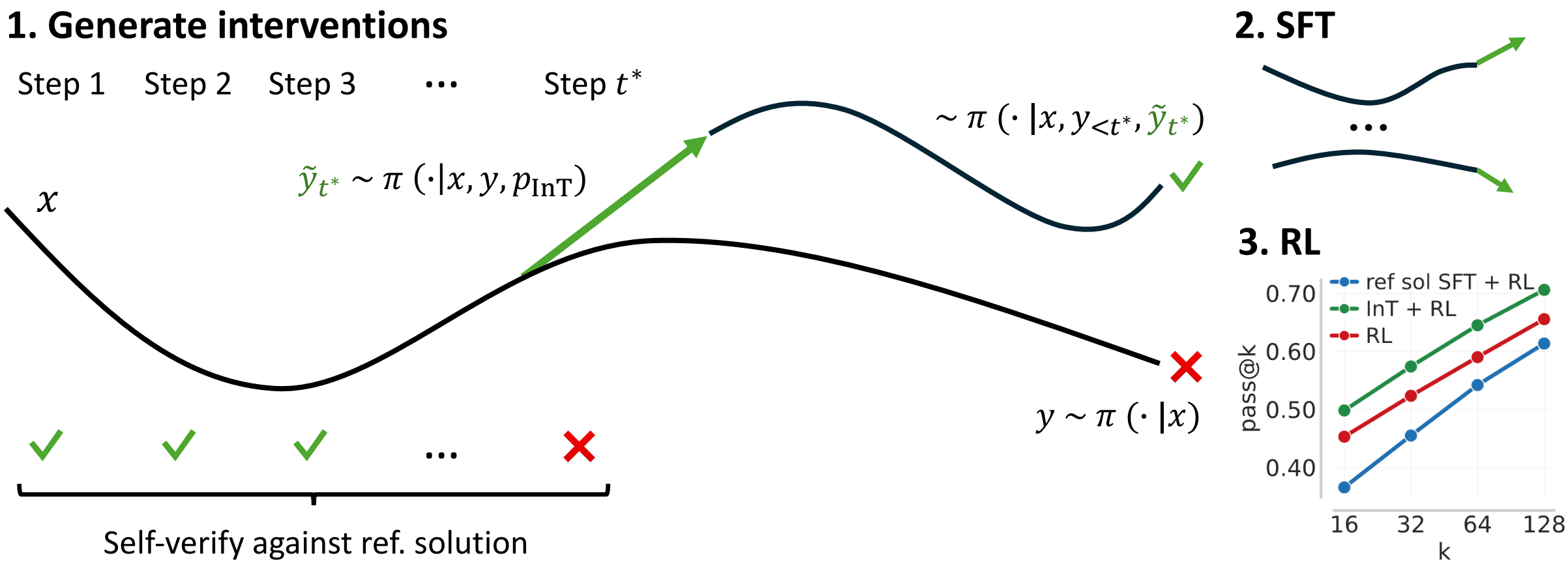
1. Generate interventions
Step 1
Step 2
Step 3
...
Step t^*
x
$\tilde{y}_{t^*} \sim \pi(\cdot|x, y, p_{\text{InT}})$
$\sim \pi(\cdot|x, y_{<t^*}, \tilde{y}_{t^*})$
$y \sim \pi(\cdot|x)$
...
Self-verify against ref. solution
2. SFT
...
3. RL
ref sol SFT + RL
InT + RL
RL
pass@k
0.70
0.60
0.50
0.40
16
32
64
128
k