

Motivation

- Introducing safety into anytime planner;
- Anytime search tree consists of solely the safe actions;
- Asymptotical guarantees.

Key idea

- Constrain propagated and posterior beliefs on the way down the tree;
- Clean values up the tree from impact of the dangerous actions and fix the visitation counts;
- Convergence in probability and with exponential rate proof under mild assumptions.

Problem Formulation

$$a_0^* \in \arg \max_{a_0 \in \mathcal{A}} Q^{\pi^*}(b_0, a_0; \rho_1) \quad \text{subject to}$$

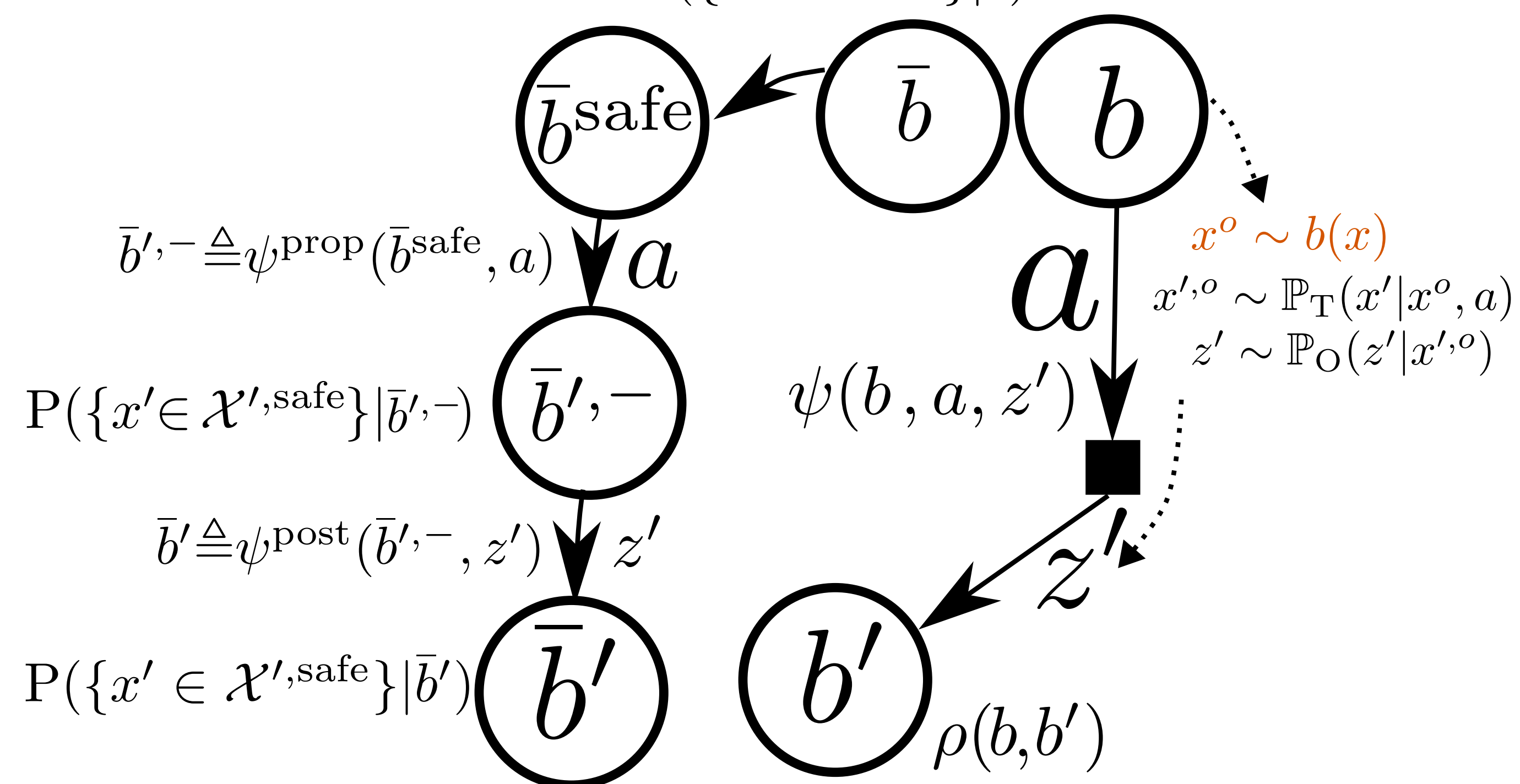
$$\underbrace{P(b_0; L \in \bigcap_{\ell=0}^L \Phi_\ell^\delta | b_0, a_0, \pi_{1:L-1}^*) = 1}_{\text{outer constraint}}$$

$$\Phi_\ell^\delta \triangleq \{b_\ell^-, b_\ell; b_\ell^- \in \mathcal{B}_\ell^-, b_\ell \in \mathcal{B}_\ell, \phi(b_\ell^-) \geq \delta, \phi(b_\ell) \geq \delta\}.$$

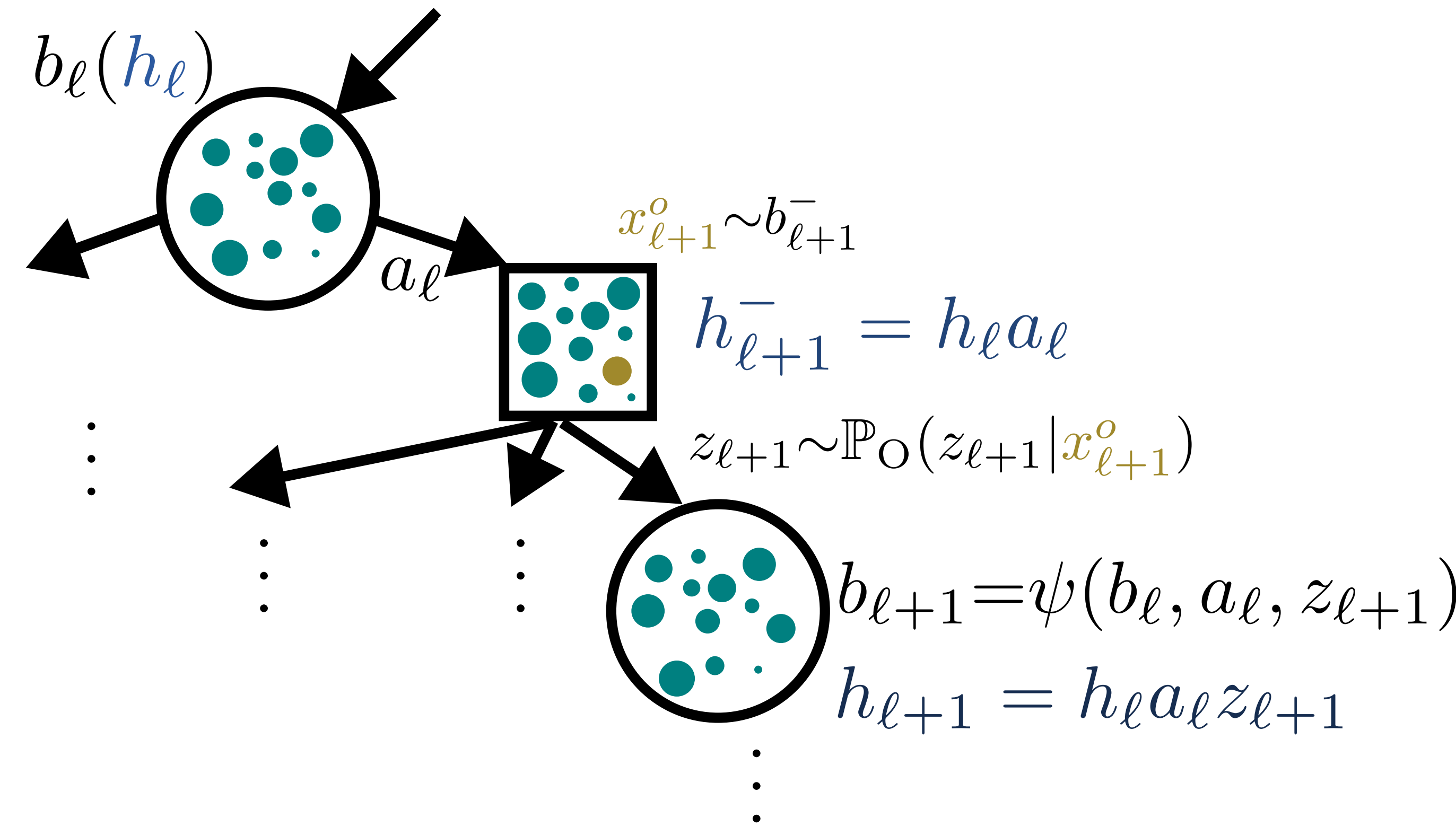
$$\phi(b_\ell) = P(\{x_\ell \in \mathcal{X}_\ell^{\text{safe}}\} | b_\ell) = \mathbb{E}_{x_\ell \sim b_\ell} [\mathbf{1}_{\{x_\ell \in \mathcal{X}_\ell^{\text{safe}}\}}] \quad \phi(b_\ell^-) = P(\{x_\ell \in \mathcal{X}_\ell^{\text{safe}}\} | b_\ell^-)$$

Safe Beliefs

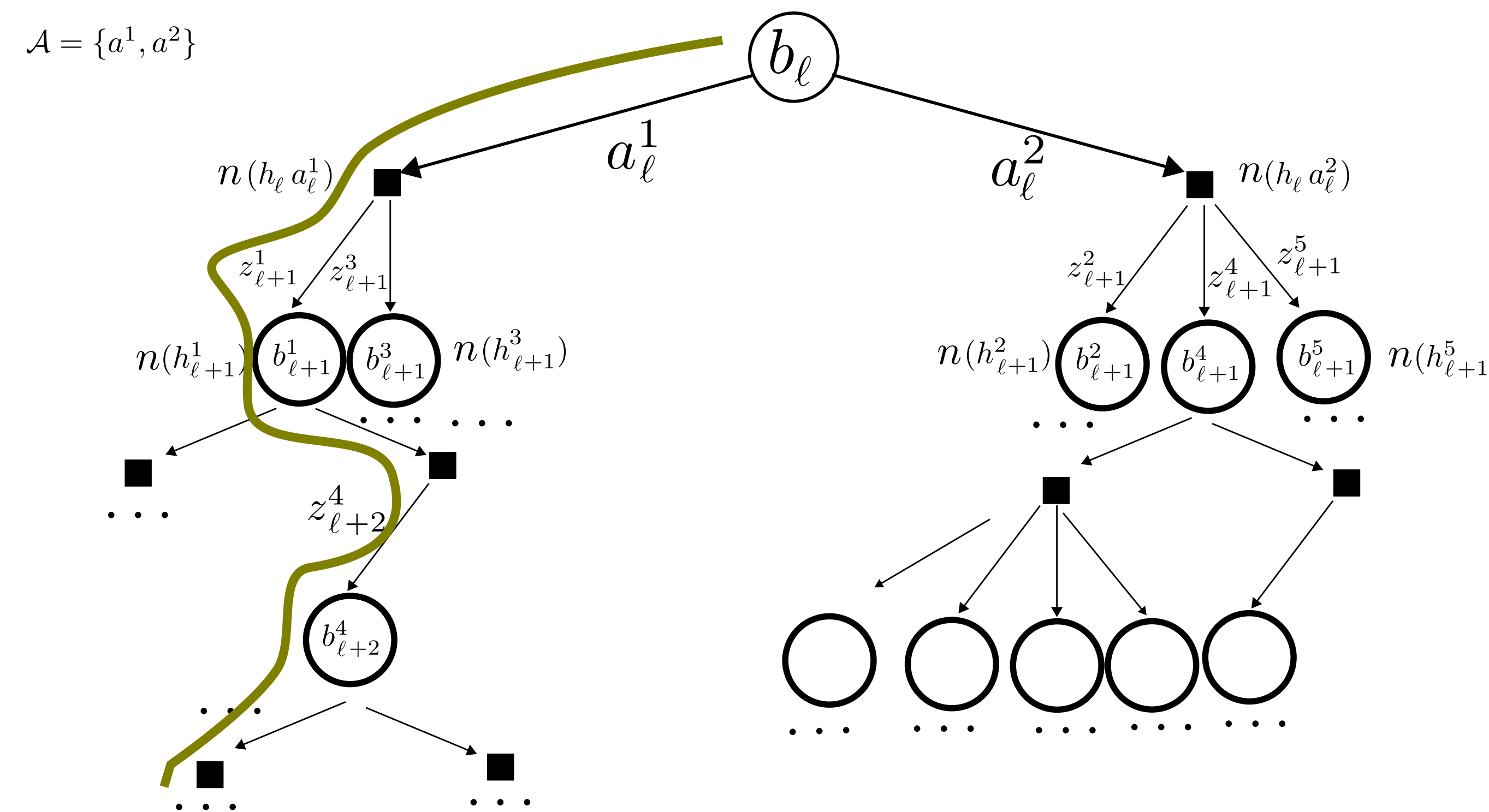
$$P(\{x \in \mathcal{X}^{\text{safe}}\} | \bar{b})$$



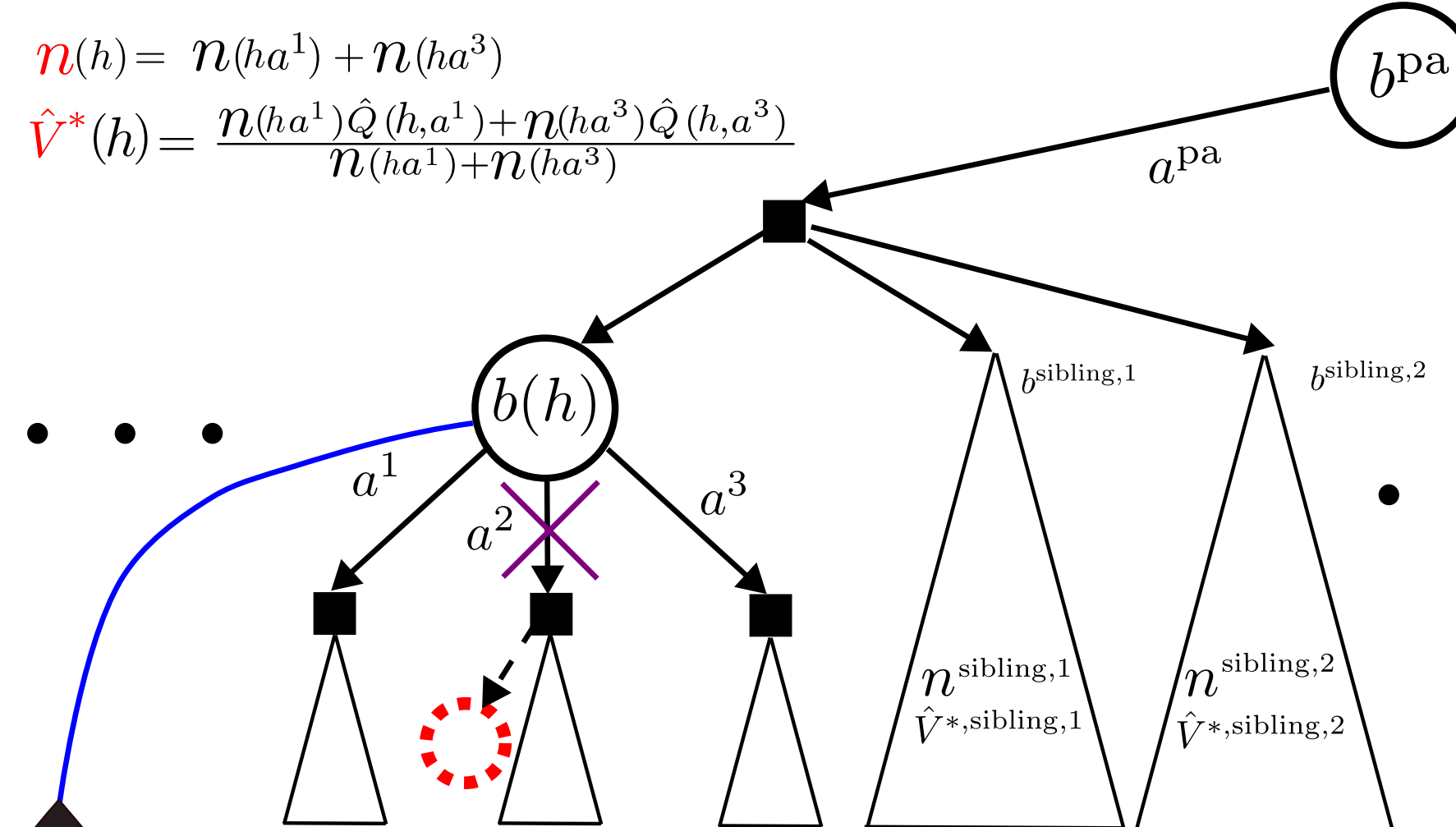
Sampling Beliefs



Asymmetric Policy Tree

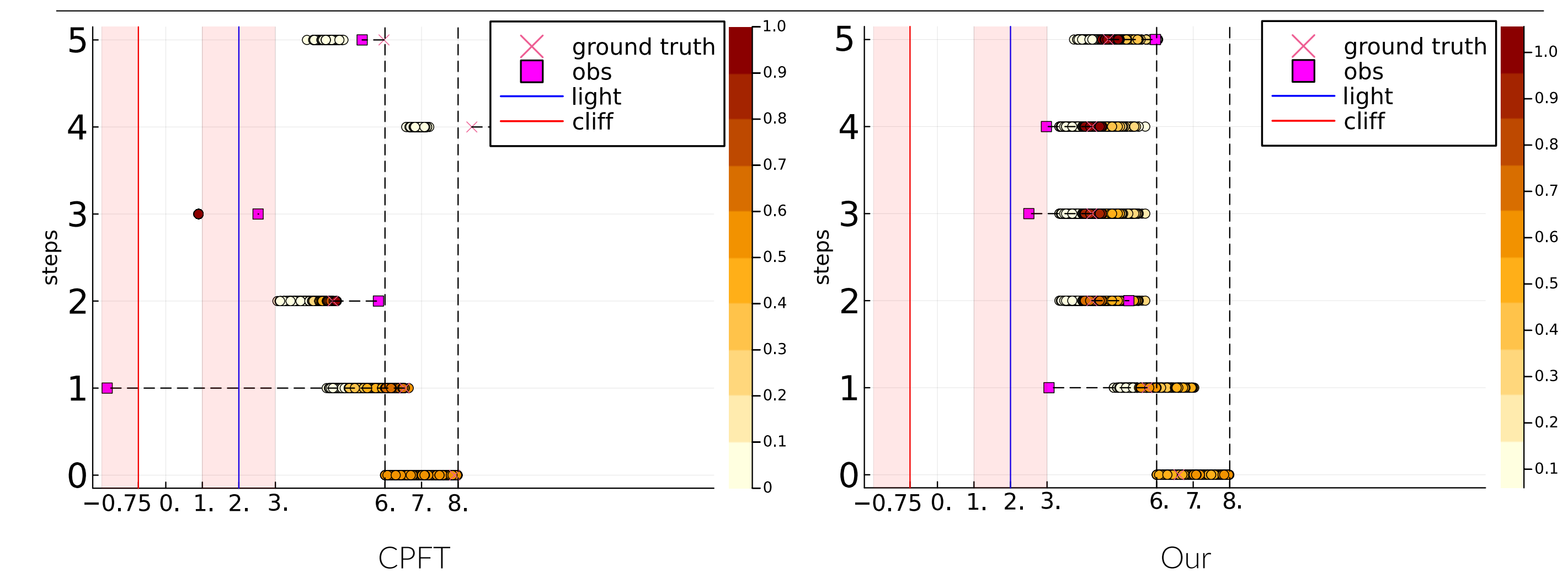


Cleaning Logic

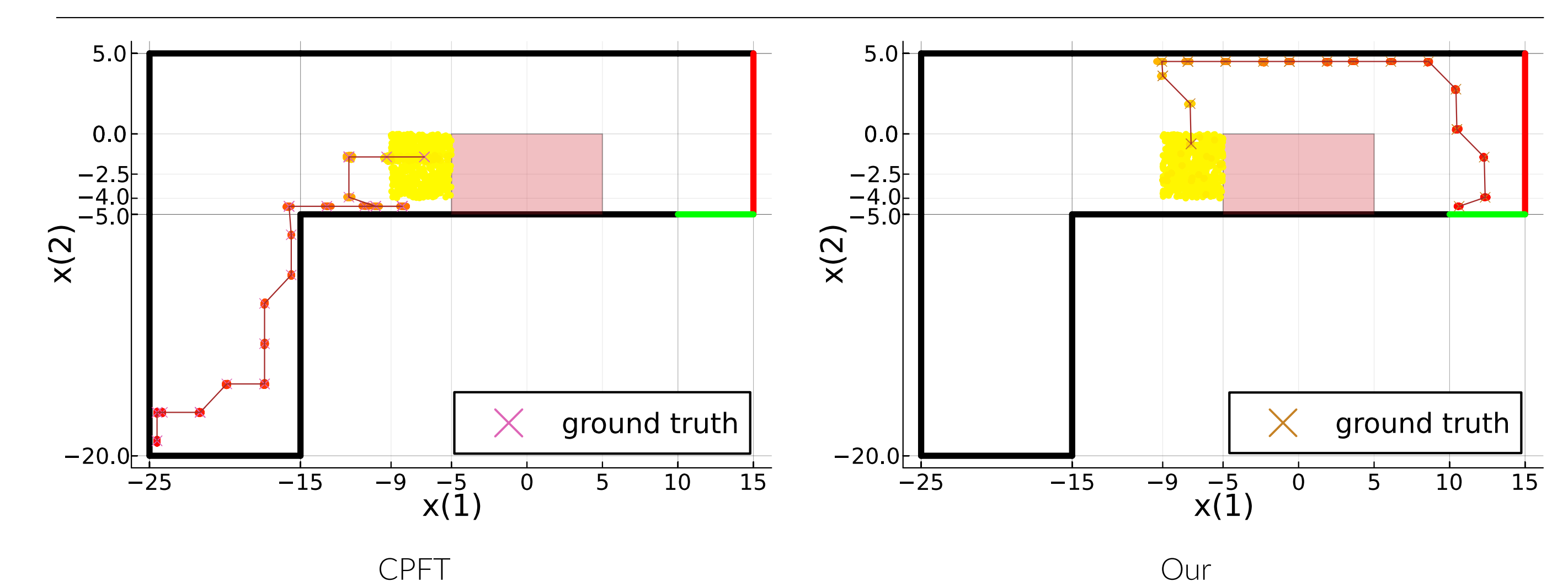


Results

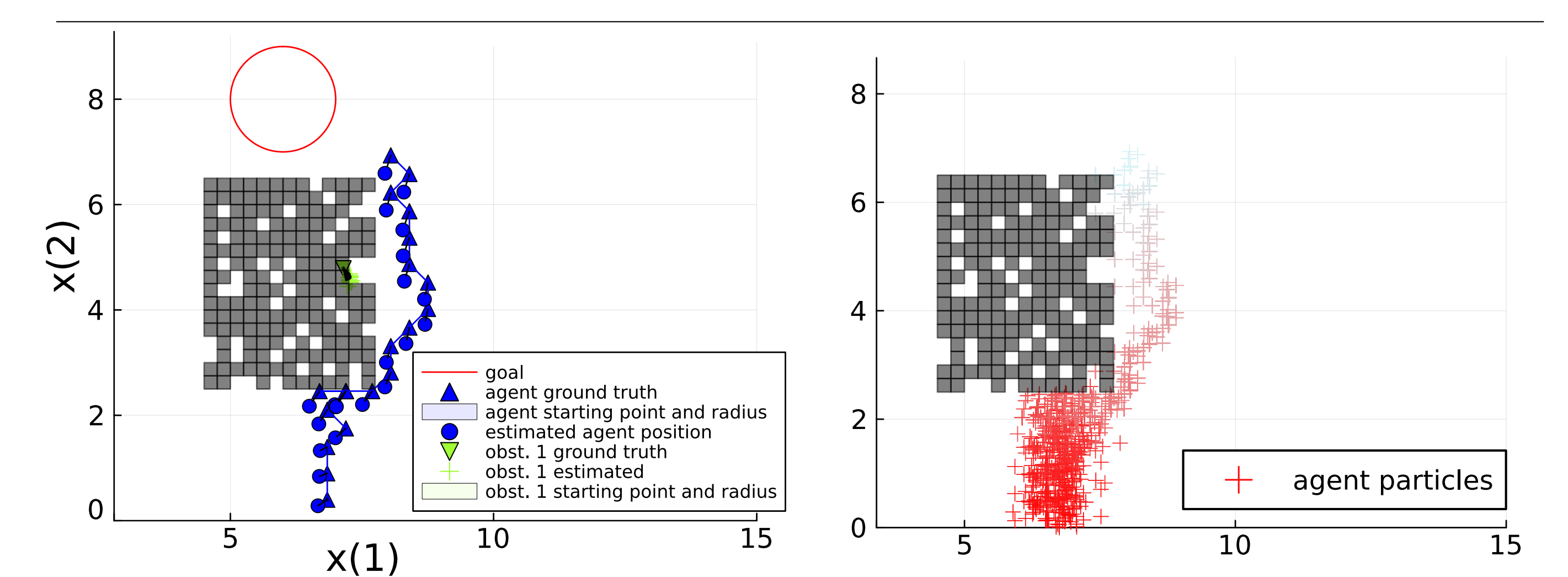
Light Dark



Roomba



SLAM



Push-box 2D

