Towards Scalable and Safe Online Decision Making Under Uncertainty in Partially Observable Environments

Vadim Indelman

TECHNION Israel Institute of Technology

ANPL Autonomous Navigation and Perception Lab

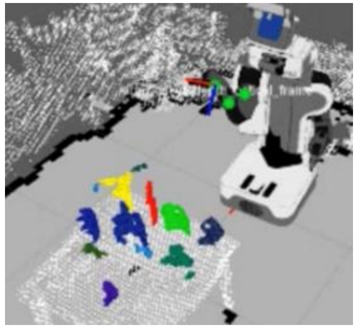
Stanford SISL Seminar

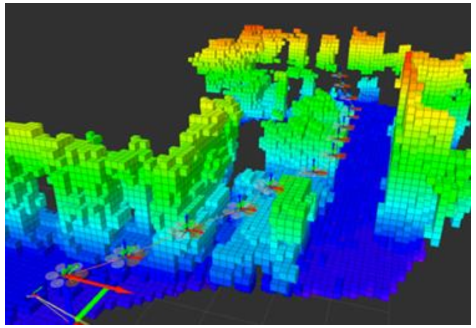
July 7, 2025

Advanced Autonomy

Involves autonomous navigation, active SLAM, informative gathering, active sensing, etc.






Advanced Autonomy

Perception and Inference

Where am I? What is the surrounding environment?

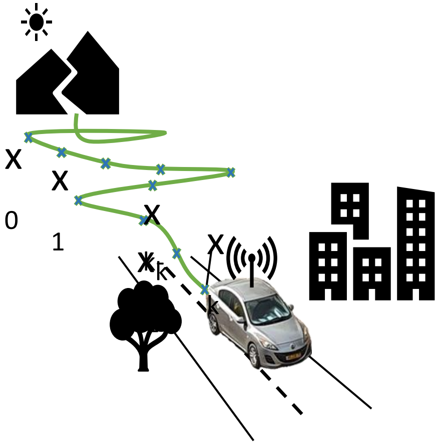
Key required capabilities

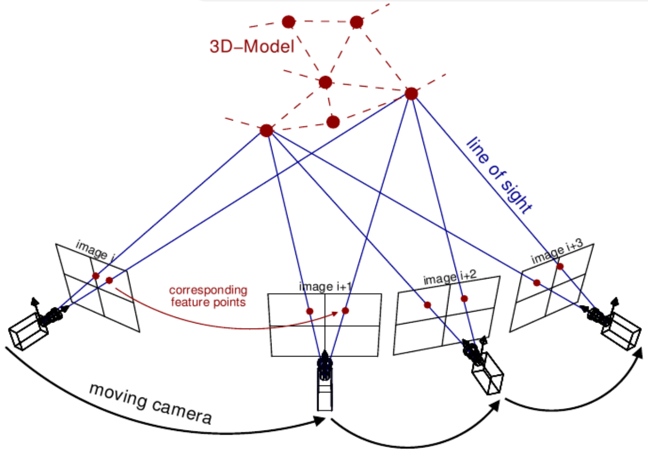
Decision-Making Under Uncertainty

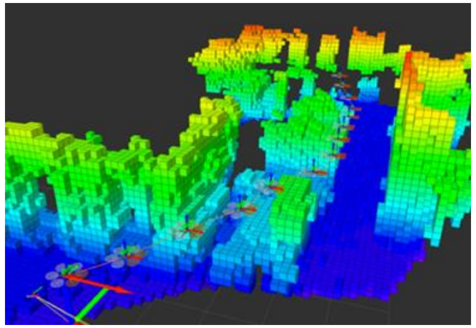
What should I be doing next?

Determine best action(s) to accomplish a task, account for different sources of uncertainty

Perception and Inference ⟷ **Decision-Making Under Uncertainty**







Partially Observable Markov Decision Process (POMDP)

- POMDP tuple:

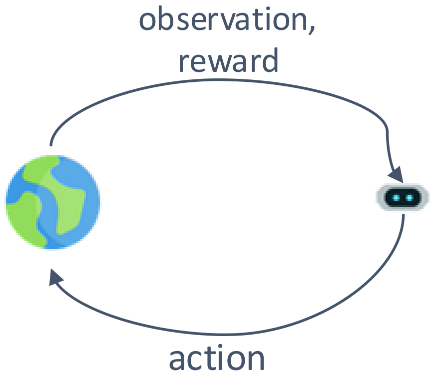
$$\langle \underbrace{\mathcal{X}, \mathcal{Z}, \mathcal{A}}, \underbrace{T, O}, \underbrace{\rho}, \underbrace{b_k} \rangle$$

state, observation, and action spaces

transition and observation models

Belief-dependent reward function

Belief at planning time instant k



- Value function

$$V^{\pi}(b_k) = \mathop{\mathbb{E}}_{z_{k+1:k+L}}\Big[\sum_{l=k}^{k+L} \underbrace{\rho}(b_l, \pi_l(b_l))\Big]$$

Belief-dependent reward function

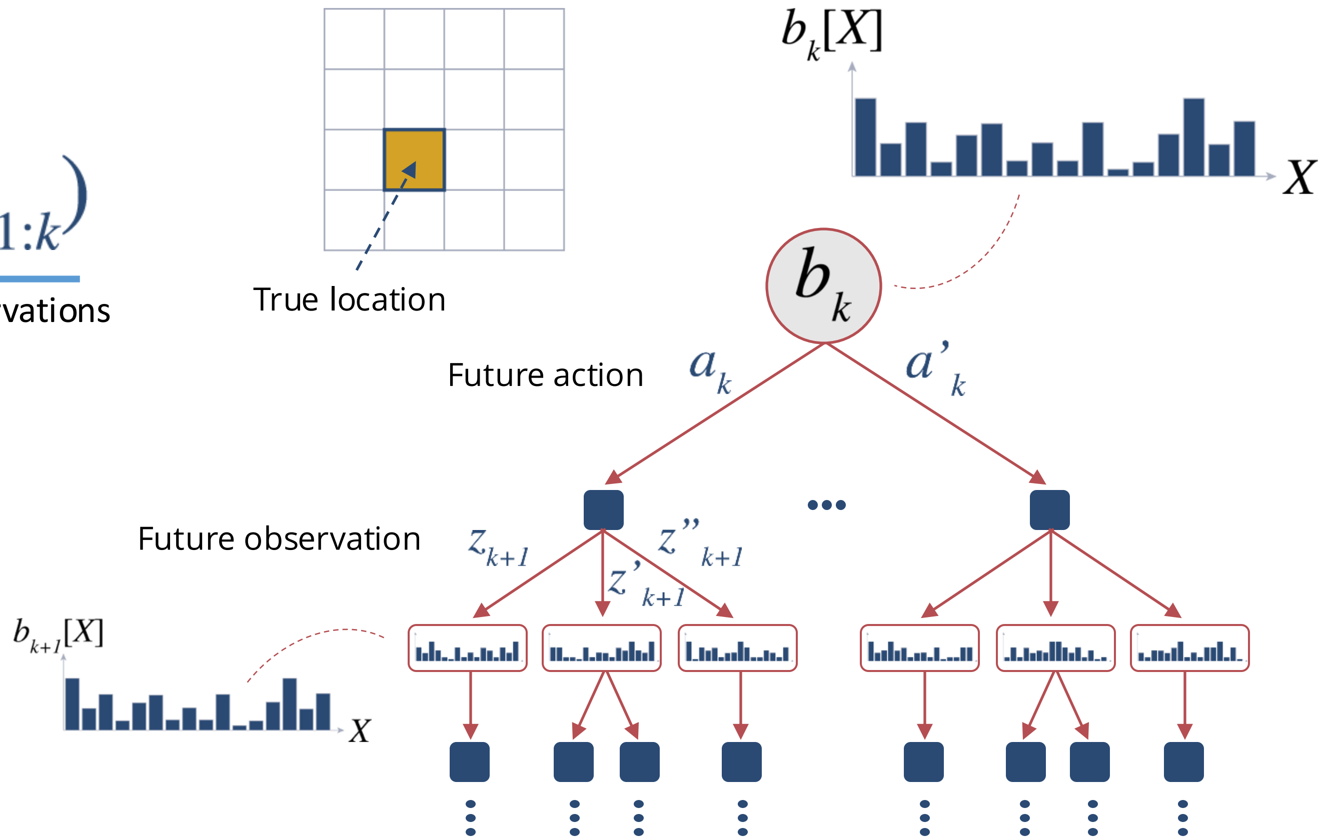
Challenge

Probabilistic Inference

Maintain a distribution over the state given data

$$b_k \triangleq b[X_k] = \mathbb{P}(\underbrace{X_k}_{\text{state}} \mid \underbrace{a_{0:k-1}}_{\text{actions}}, \underbrace{z_{1:k}}_{\text{observations}})$$

Example – grid world

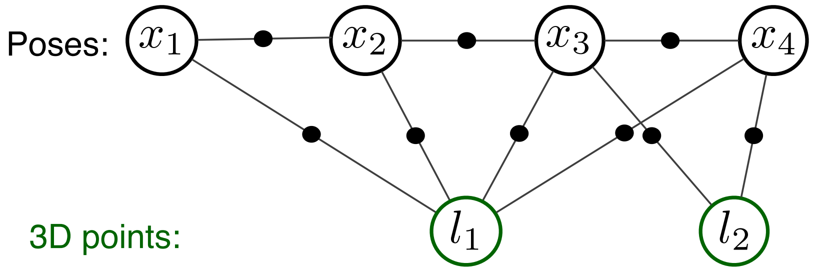


Decision-making under uncertainty

Involves reasoning about the entire observation and action spaces along planning horizon

Computationally intractable

More so, in high dimensional settings



ct autonomously online and efficiently
tasks in a safe and reliable fashion??

Agenda

Experience Reuse in POMDP Planning

POMDP Planning with Hybrid Beliefs

Simplification of POMDP with Formal Guarantees

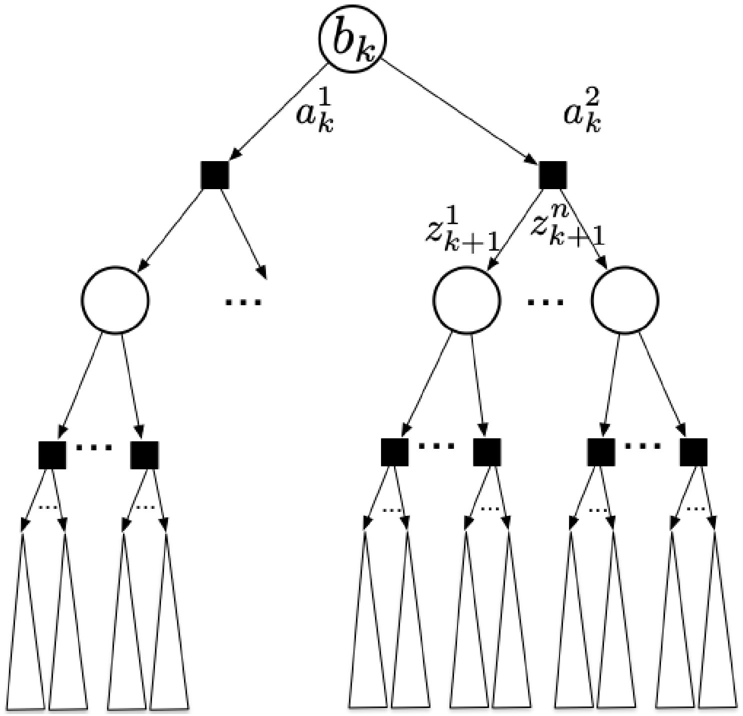
Multi-agent POMDP Planning with Inconsistent Beliefs

Agenda

- Experience Reuse in POMDP Planning
- POMDP Planning with Hybrid Beliefs
- Simplification of POMDP with Formal Guarantees
- Multi-agent POMDP Planning with Inconsistent Beliefs

Experience Reuse in POMDP Planning

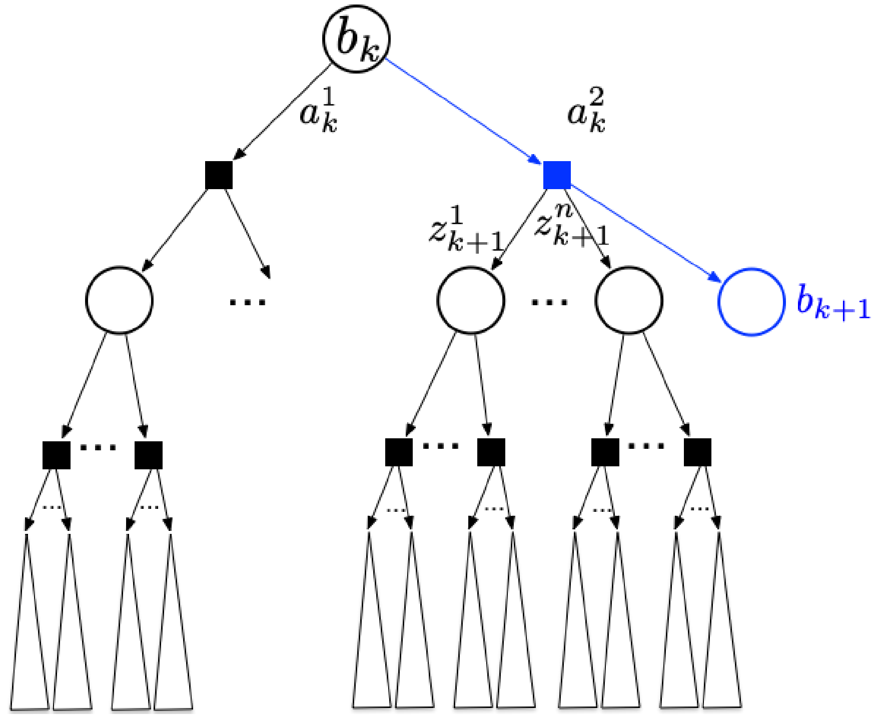
- Consider POMDPs with continuous state, action, and observation spaces



E. Farhi and V. Indelman, “iX-BSP: Incremental Belief Space Planning,” ICRA’19, arXiv’21.
M. Novitsky, M. Barenboim, and V. Indelman, “Previous Knowledge Utilization In Online Anytime Belief Space Planning,” arXiv’24.

Experience Reuse in POMDP Planning

- Consider POMDPs with continuous state, action, and observation spaces
- The probability of sampling the same belief/observation twice is zero

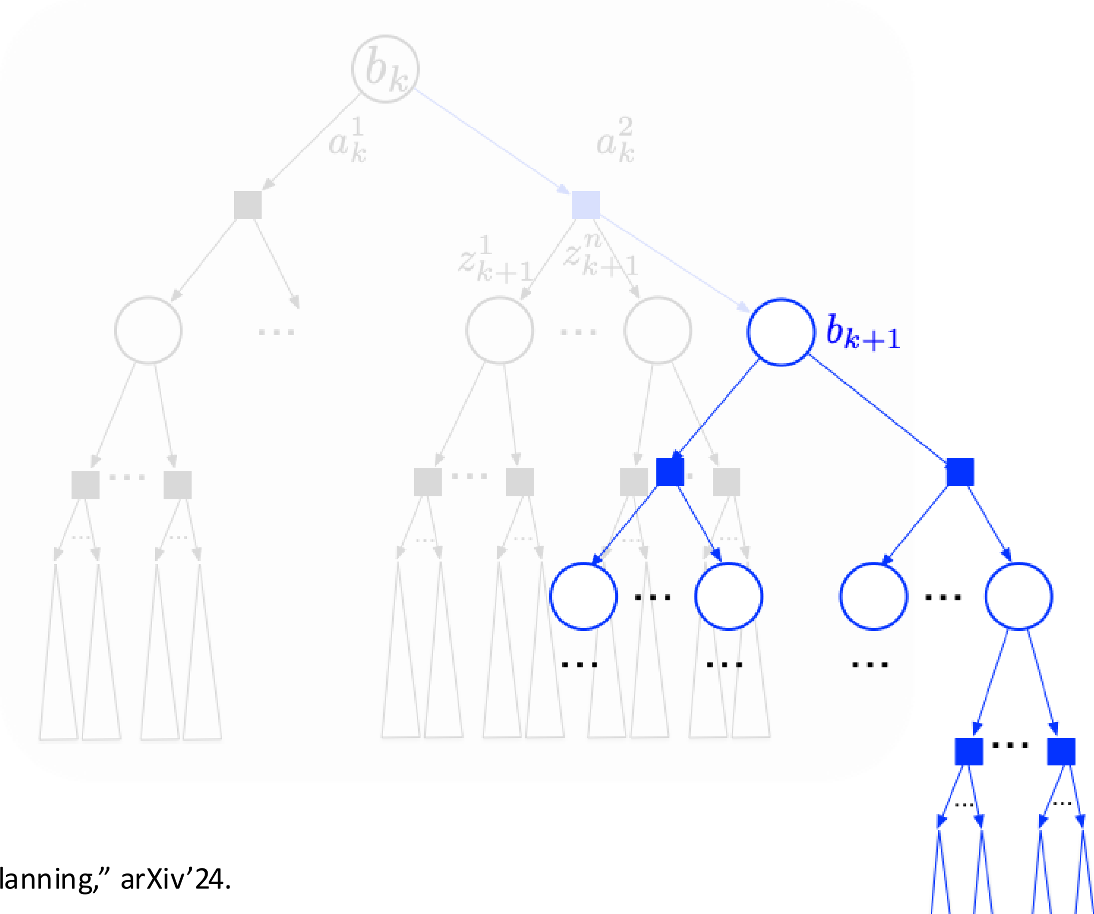


E. Farhi and V. Indelman, "iX-BSP: Incremental Belief Space Planning," ICRA'19, arXiv'21.
M. Novitsky, M. Barenboim, and V. Indelman, "Previous Knowledge Utilization In Online Anytime Belief Space Planning," arXiv'24.

Experience Reuse in POMDP Planning

- Consider POMDPs with continuous state, action, and observation spaces
- The probability of sampling the same belief/observation twice is zero

Online SOTA POMDP solvers typically perform calculations from **scratch at each planning session**

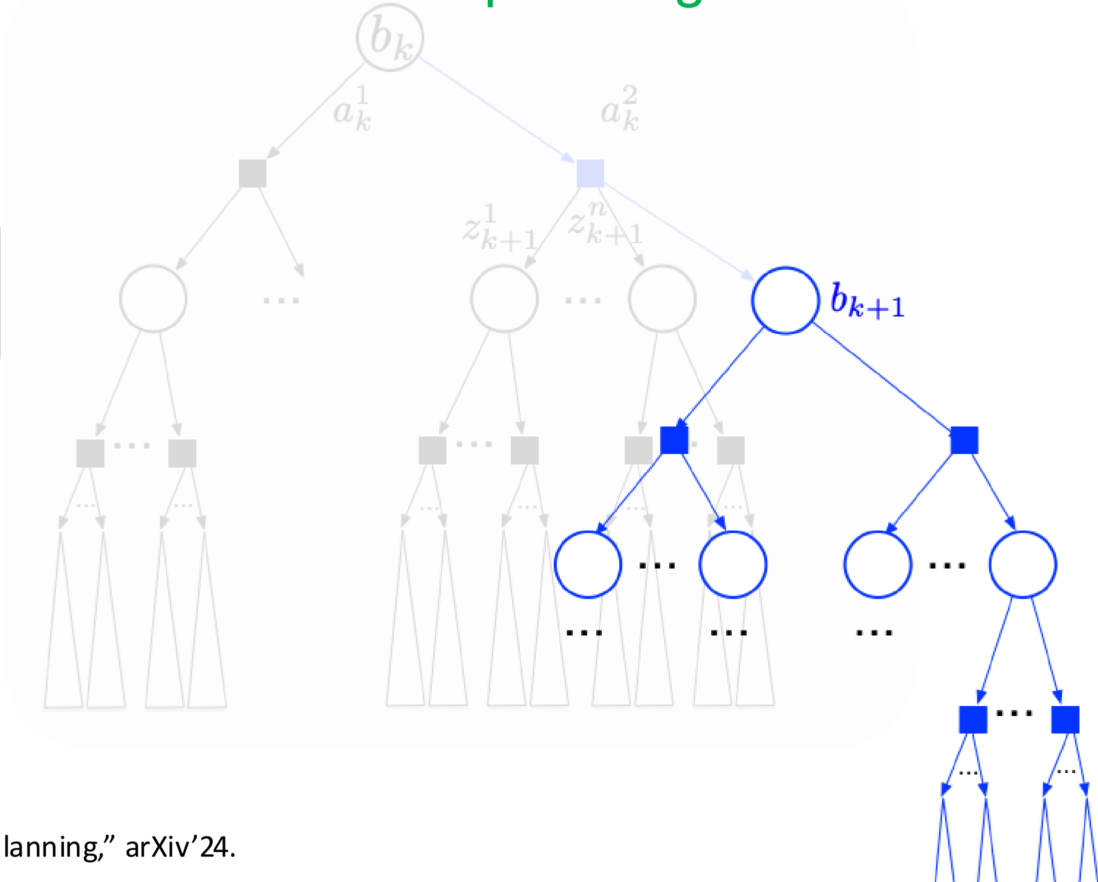


E. Farhi and V. Indelman, “iX-BSP: Incremental Belief Space Planning,” ICRA’19, arXiv’21.
M. Novitsky, M. Barenboim, and V. Indelman, “Previous Knowledge Utilization In Online Anytime Belief Space Planning,” arXiv’24.

Experience Reuse in POMDP Planning

- Consider POMDPs with continuous state, action, and observation spaces
- The probability of sampling the same belief/observation twice is zero
- Previously sampled beliefs can still provide useful info in the current planning session

Online SOTA POMDP solvers typically perform calculations from **scratch at each planning session**



E. Farhi and V. Indelman, “iX-BSP: Incremental Belief Space Planning,” ICRA’19, arXiv’21.
M. Novitsky, M. Barenboim, and V. Indelman, “Previous Knowledge Utilization In Online Anytime Belief Space Planning,” arXiv’24.

Experience Reuse in POMDP Planning

- Consider POMDPs with continuous state, action, and observation spaces
- The probability of sampling the same belief/observation twice is zero
- Previously sampled beliefs can still provide useful info in the current planning session

Key idea: Reuse previous trajectories/calculations to get an efficient estimation of

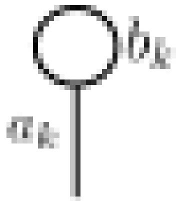
$$Q^{\pi}(b,a)=\mathbb{E}_{\pi}[\sum_{i=k}^{k+L-1}\gamma^{i-k}r(b_i,\pi_i(b_i),b_{i+1})\mid b_k=b,a_k=a]\triangleq\mathbb{E}_{\pi}[G\mid b_k=b,a_k=a]$$

- Instead of calculating each planning session from scratch (state of the art)

E. Farhi and V. Indelman, “iX-BSP: Incremental Belief Space Planning,” ICRA’19, arXiv’21.
M. Novitsky, M. Barenboim, and V. Indelman, “Previous Knowledge Utilization In Online Anytime Belief Space Planning,” arXiv’24.

Experience Reuse in POMDP Planning

- Consider a planning session at time instant k



$$Q^{\pi}(b_k, a_k)$$

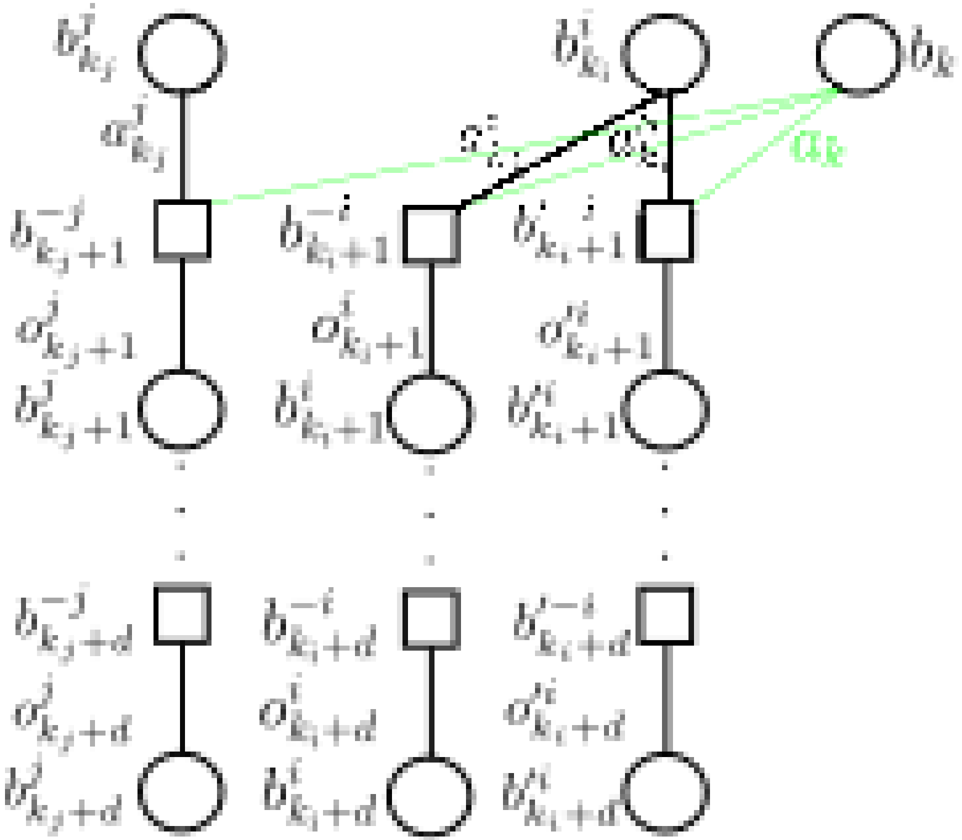
Current time

E. Farhi and V. Indelman, "iX-BSP: Incremental Belief Space Planning," ICRA'19, arXiv'21.
M. Novitsky, M. Barenboim, and V. Indelman, "Previous Knowledge Utilization In Online Anytime Belief Space Planning," arXiv'24.

Experience Reuse in POMDP Planning

- Consider a planning session at time instant k

$$Q^{\pi}(b_k, a_k)$$

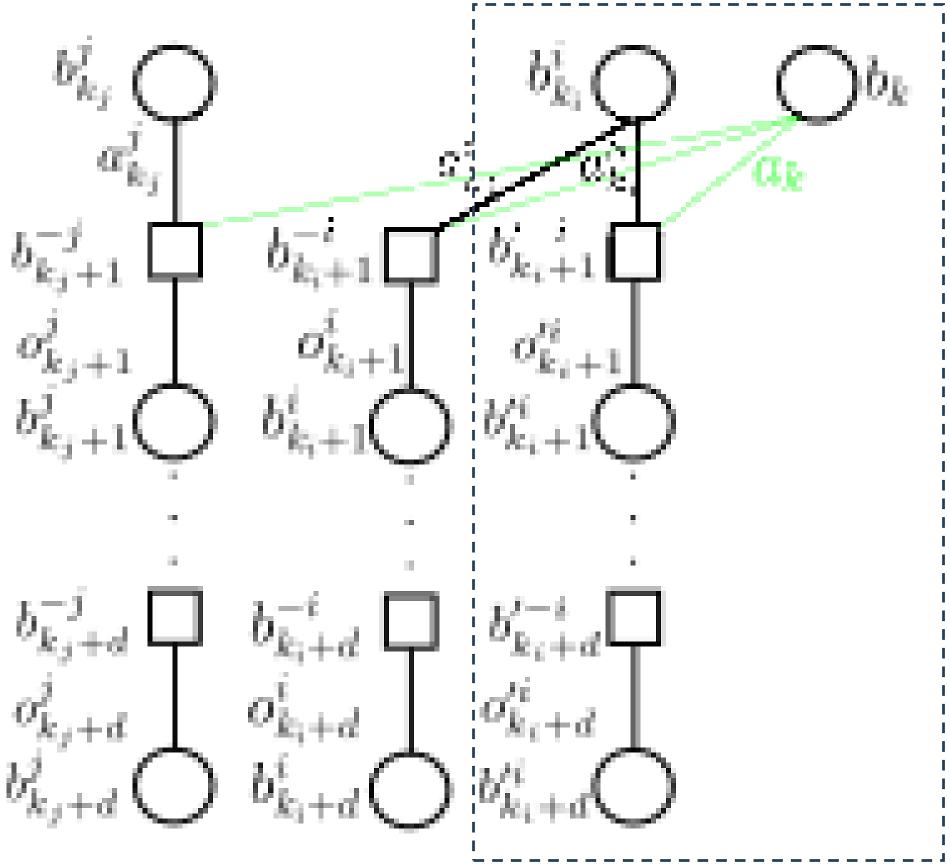


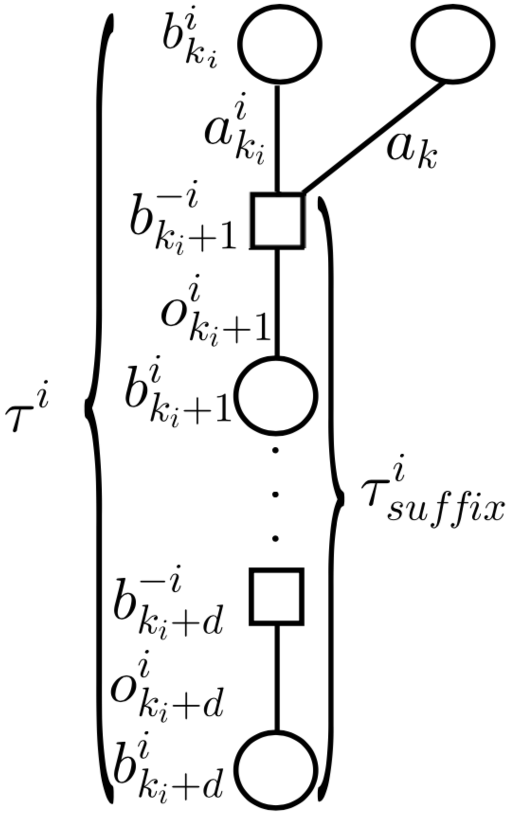
Previous data | Current time

E. Farhi and V. Indelman, "iX-BSP: Incremental Belief Space Planning," ICRA'19, arXiv'21.
M. Novitsky, M. Barenboim, and V. Indelman, "Previous Knowledge Utilization In Online Anytime Belief Space Planning," arXiv'24.

Experience Reuse in POMDP Planning

- **Key idea**: multiple importance sampling (MIS) estimator



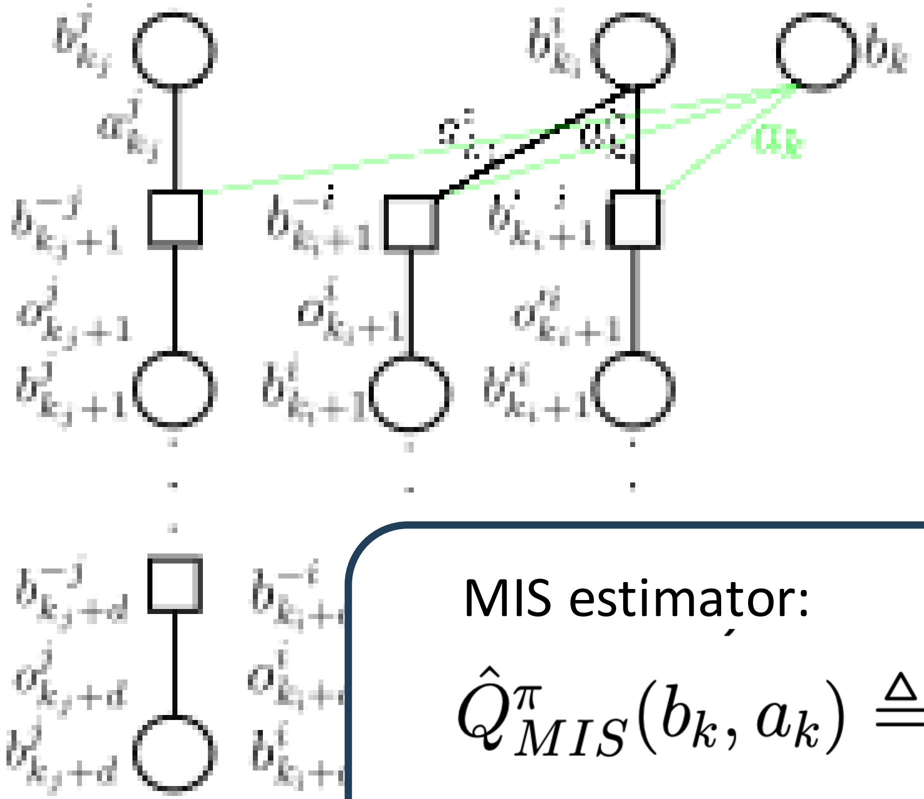


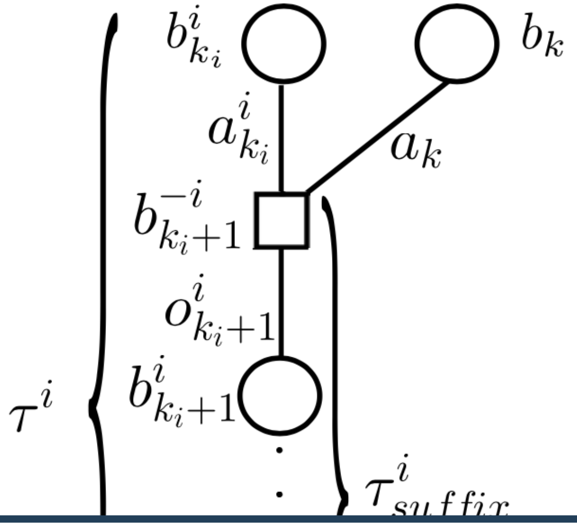
$$w_i \triangleq \frac{\mathbb{P}(\tau^i_{suffix} | b_k, a_k, \pi)}{\mathbb{P}(\tau^i_{suffix} | b^i_{k_i}, a^i_{k_i}, \pi)}$$

E. Farhi and V. Indelman, "iX-BSP: Incremental Belief Space Planning," ICRA'19, arXiv'21.
M. Novitsky, M. Barenboim, and V. Indelman, "Previous Knowledge Utilization In Online Anytime Belief Space Planning," arXiv'24.

Experience Reuse in POMDP Planning

- **Key idea**: multiple importance sampling (MIS) estimator





$$w_i \triangleq \frac{\mathbb{P}(\tau^i_{suffix}|b_k, a_k, \pi)}{\mathbb{P}(\tau^i_{suffix}|b^i_{k_i}, a^i_{k_i}, \pi)}$$

MIS estimator:

$$\hat{Q}^{\pi}_{MIS}(b_k, a_k) \triangleq \sum_{m=1}^{M} \sum_{l=1}^{n_m} \frac{\mathbb{P}(\tau^{l,m}_{suffix}|b_k, a_k, \pi)\overbrace{\tilde{G}^{l,m}}^{\text{return}}}{\sum_{j=1}^{M} n_j \cdot \mathbb{P}(\tau^{l,m}_{suffix}|b^j_{k_j}, a^j_{k_j}, \pi)}.$$

E. Farhi and V. Indelman, "iX-BSP: Incremental Belief Space Planning," ICRA'19, arXiv'21.
M. Novitsky, M. Barenboim, and V. Indelman, "Previous Knowledge Utilization In Online Anytime Belief Space Planning," arXiv'24.

Experience-Based Value Function Estimation

MIS estimator:

$$\hat{Q}^{\pi}_{MIS}(b_k, a_k) \triangleq \sum_{m=1}^{M} \sum_{l=1}^{n_m} \frac{\mathbb{P}(\tau^{l,m}_{suffix} | b_k, a_k, \pi)\tilde{G}^{l,m}}{\sum_{j=1}^{M} n_j \cdot \mathbb{P}(\tau^{l,m}_{suffix} | b^j_{k_j}, a^j_{k_j}, \pi)}.$$

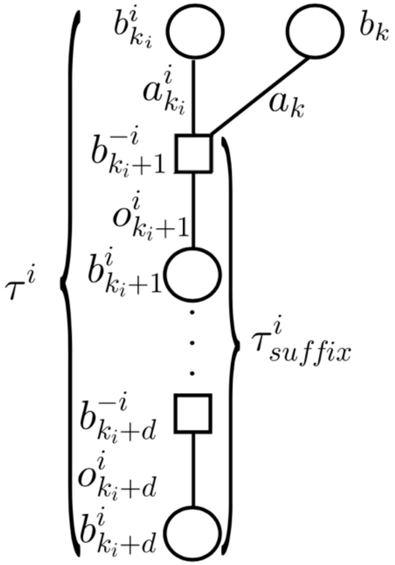
Theorem 1

$$\frac{\mathbb{P}(\tau^i_{suffix} | b_k, a_k, \pi)}{\mathbb{P}(\tau^i_{suffix} | b^i_{k_i}, a^i_{k_i}, \pi)} = \frac{\mathbb{P}(b^{-i}_{k_i+1} | b_k, a_k)}{\mathbb{P}(b^{-i}_{k_i+1} | b^i_{k_i}, a^i_{k_i})}$$

Proof.

$$\frac{\mathbb{P}(\tau^i_{suffix} | b_k, a_k, \pi)}{\mathbb{P}(\tau^i_{suffix} | b^i_{k_i}, a^i_{k_i}, \pi)} = \frac{\mathbb{P}(b^{-i}_{k_i+1}, o^i_{k_i+1}, \ldots, b^i_{k_i+L} | b_k, a_k, \pi)}{\mathbb{P}(b^{-i}_{k_i+1}, o^i_{k_i+1}, \ldots, b^i_{k_i+L} | b^i_{k_i}, a^i_{k_i}, \pi)} =$$

$$\frac{\mathbb{P}(b^{-i}_{k_i+1} | b_k, a_k)}{\mathbb{P}(b^{-i}_{k_i+1} | b^i_{k_i}, a^i_{k_i})} \cdot \cancel{\frac{\mathbb{P}(o^i_{k_i+1}, \ldots, b^i_{k_i+L} | b^{-i}_{k_i+1}, \pi)}{\mathbb{P}(o^i_{k_i+1}, \ldots, b^i_{k_i+L} | b^{-i}_{k_i+1}, \pi)}} = \frac{\mathbb{P}(b^{-i}_{k_i+1} | b_k, a_k)}{\mathbb{P}(b^{-i}_{k_i+1} | b^i_{k_i}, a^i_{k_i})} \quad \square$$

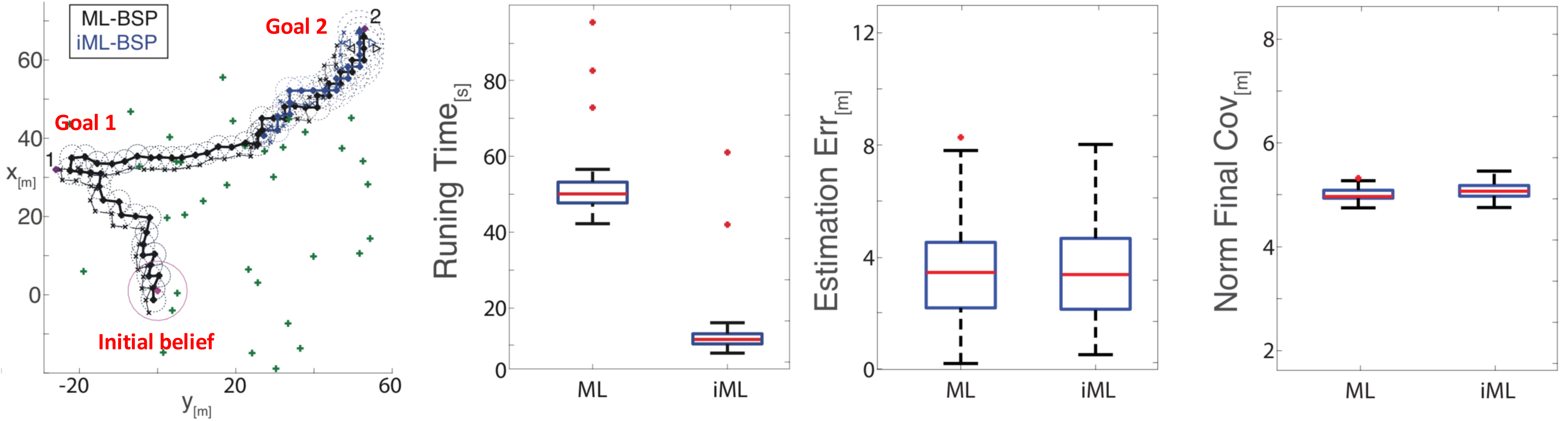


M. Novitsky, M. Barenboim, and V. Indelman, “Previous Knowledge Utilization In Online Anytime Belief Space Planning,” arXiv’24.

Incremental Belief Space Planning

Basic simulation – autonomous navigation in unknown environments:

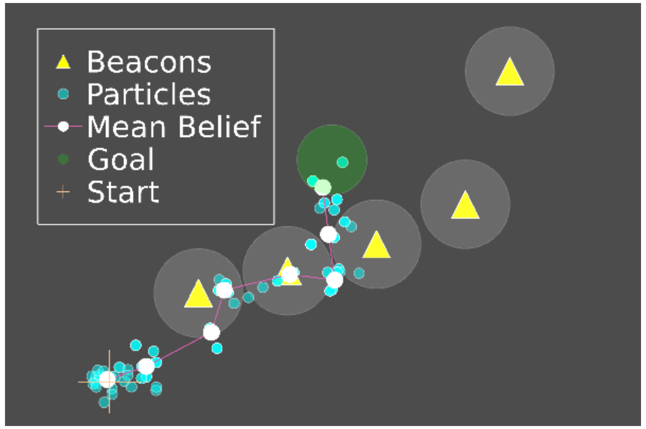
ML-BSP: BSP with ML observations
(one sample per look ahead step)

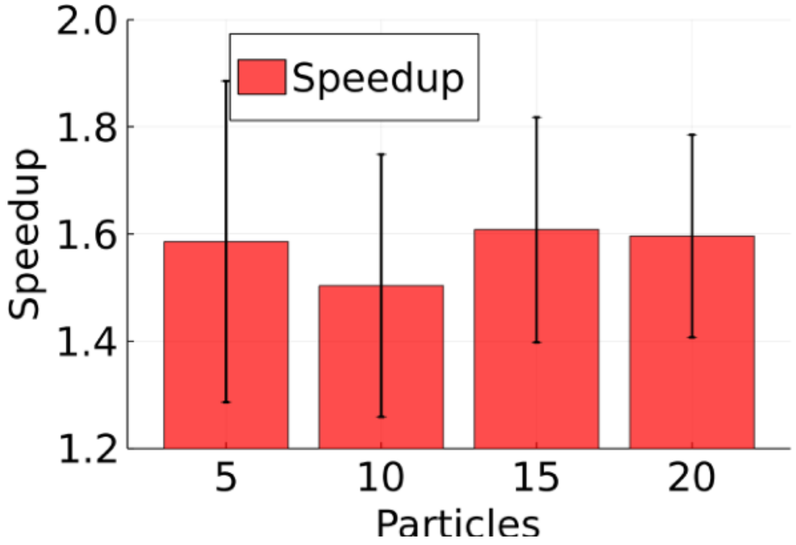


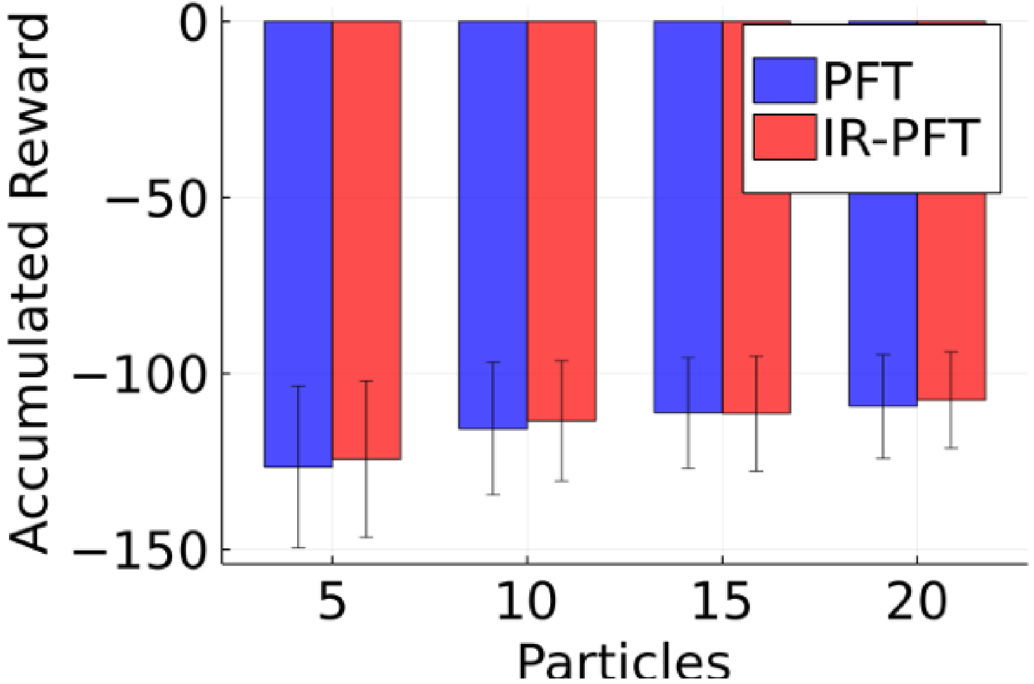
E. Farhi and V. Indelman, "iX-BSP: Incremental Belief Space Planning," ICRA'19, arXiv'21.
M. Novitsky, M. Barenboim, and V. Indelman, "Previous Knowledge Utilization In Online Anytime Belief Space Planning," arXiv'24.

Incremental Reuse Particle Filter Tree (IR-PFT)

- Extend PFT-DPW[1] , incorporating trajectories from previous planning sessions for fast estimation of $Q(b_k, a_k)$



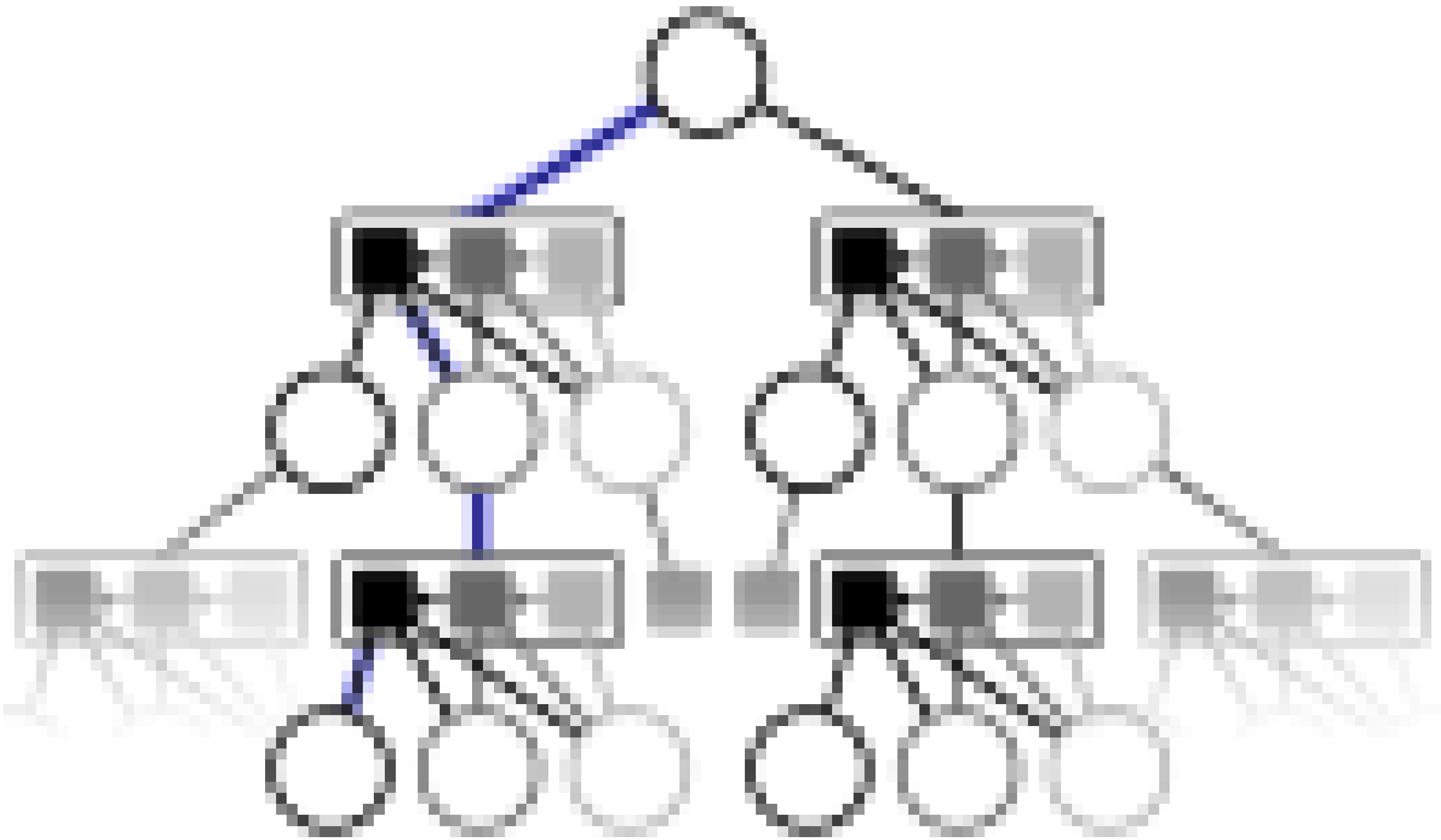




[1]Z. Sunberg and M. Kochenderfer. "Online algorithms for POMDPs with continuous state, action, and observation spaces." ICAPS, 2018.

E. Farhi and V. Indelman, “iX-BSP: Incremental Belief Space Planning,” ICRA’19, arXiv’21.
M. Novitsky, M. Barenboim, and V. Indelman, “Previous Knowledge Utilization In Online Anytime Belief Space Planning,” arXiv’24.

Action-Gradient Monte Carlo Tree Search for Non-Parametric Continuous (PO)MDPs



I. Lev-Yehudi, M. Novitsky, M. Barenboim, R. Benchetrit, and V. Indelman, "Action-Gradient Monte Carlo Tree Search for Non-Parametric Continuous (PO)MDPs", arXiv'25.

Agenda

Experience Reuse in POMDP Planning

POMDP Planning with Hybrid Beliefs

Semantic Risk Awareness

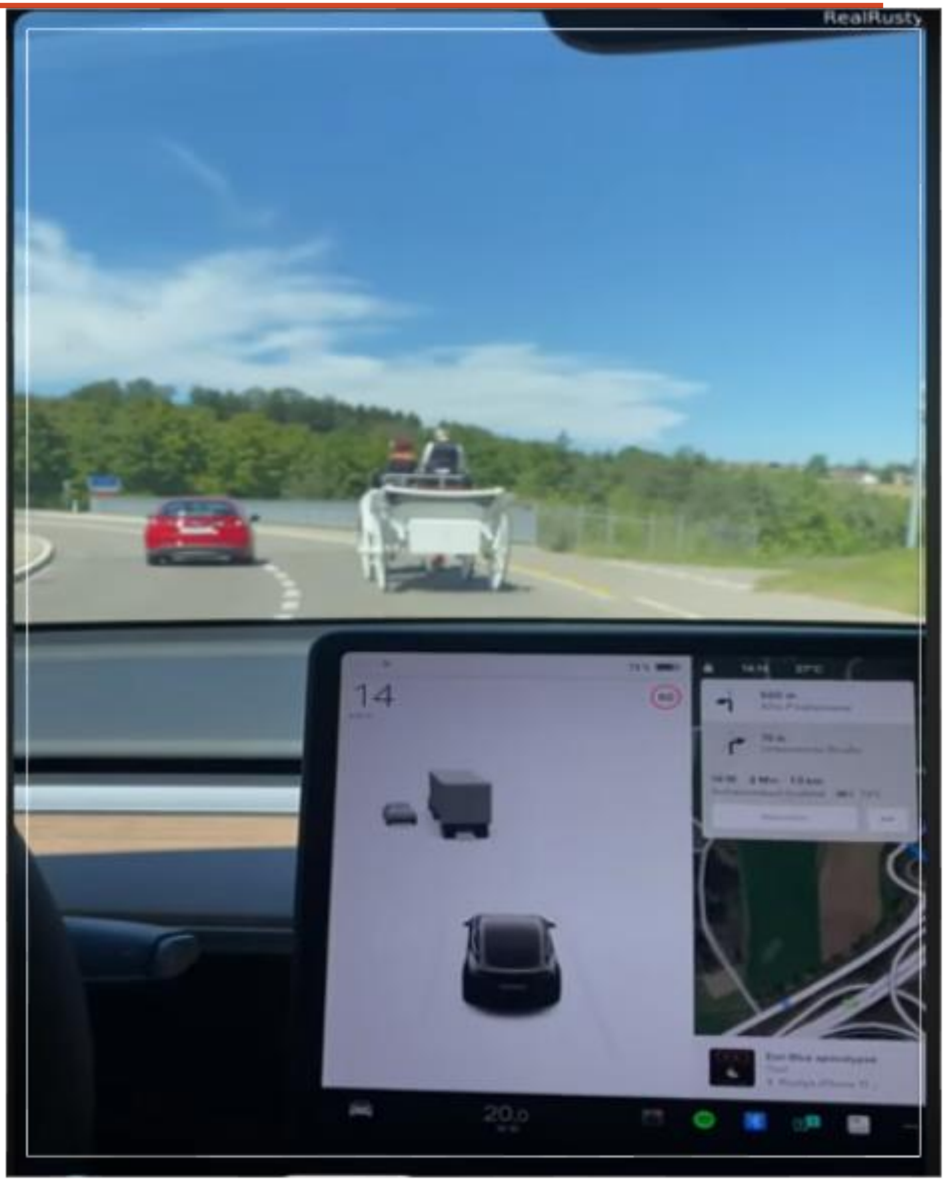
Ambiguous Environments

Simplification of POMDP with Formal Guarantees

Multi-agent POMDP Planning with Inconsistent Beliefs

Semantic Perception & SLAM

- Usually, semantics and geometry are considered **separately**
- Cannot use coupled observation models or priors
- Can lead to absurd & **unsafe** performance

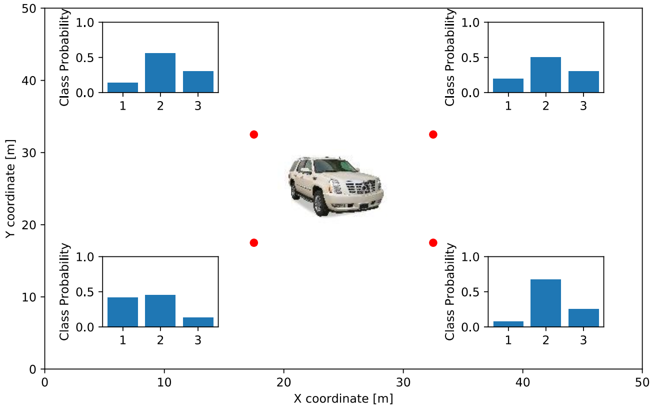


TECHNION
Israel Institute
of Technology

ANPL
Autonomous Navigation
and Perception Lab

Coupled Models

- View-dependent semantic observation model:



$$\mathbb{P}(z^s \mid c, \mathcal{X}^{rel})$$

Object class

Semantic observation (from a classifier)

Agent's viewpoint relative to object

- Class and poses can be coupled via learned prior probabilities
- Reward/constraint can depend on both classes and poses

Y. Feldman and V. Indelman, "Bayesian Viewpoint-Dependent Robust Classification under Model and Localization Uncertainty," ICRA'18.
V. Tchuiev, Y. Feldman, and V. Indelman, "Data Association Aware Semantic Mapping and Localization via a Viewpoint Dependent Classifier Model," IROS'19.
V. Tchuiev and V. Indelman, "Epistemic Uncertainty Aware Semantic Localization and Mapping for Inference and Belief Space Planning," Artificial Intelligence, 2023.
T. Lemberg and V. Indelman, "Online Hybrid-Belief POMDP with Coupled Semantic-Geometric Models and Semantic Safety Awareness", arXiv'25.

Hybrid Belief

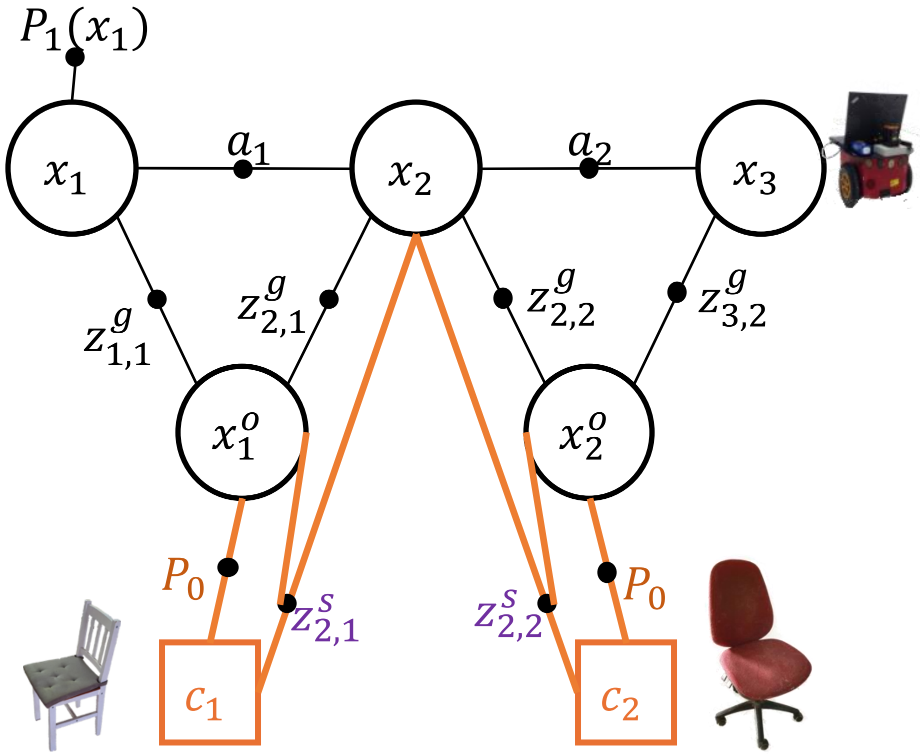
- **Hybrid Belief** at time instant k:

$$b[X_k, C] = \mathbb{P}(X_k, C \mid \mathcal{H}_k)$$

Robot's and objects' poses

Objects' classes

History (actions, geometric & semantic observations)



- Classes and agent poses are <u>dependent</u>
- Classes of different objects are <u>dependent</u>

- As opposed to:
    - Per-frame classification
    - Modeling semantic observations as viewpoint **independent**

Y. Feldman and V. Indelman, “Bayesian Viewpoint-Dependent Robust Classification under Model and Localization Uncertainty,” ICRA’18.
V. Tchuiev, Y. Feldman, and V. Indelman, “Data Association Aware Semantic Mapping and Localization via a Viewpoint Dependent Classifier Model,” IROS’19.
V. Tchuiev and V. Indelman, “Epistemic Uncertainty Aware Semantic Localization and Mapping for Inference and Belief Space Planning,” Artificial Intelligence, 2023.
T. Lemberg and V. Indelman, “Online Hybrid-Belief POMDP with Coupled Semantic-Geometric Models and Semantic Safety Awareness”, arXiv’25.

POMDP Planning with Hybrid Semantic-Geometric Beliefs

- Value function

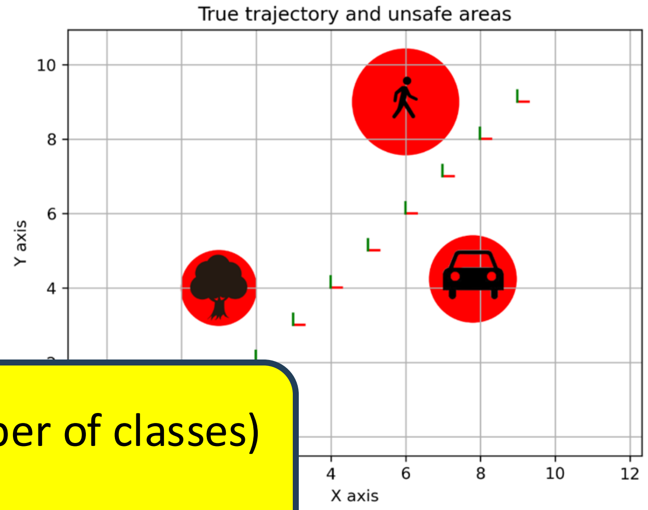
$$V^{\pi}(b_k) = \mathbb{E}_{z_{k+1:k+L}}\left[\sum_{l=k}^{k+L-1} \rho(b_l, \pi_l(b_l), b_{l+1})\right]$$

- **Semantic Risk Awareness**

$$\mathbb{P}_{safe} \triangleq \mathbb{P}(\{\wedge_{t=k+1}^{L} x_t \notin \mathcal{X}_{unsafe}(C, X^o)\} \mid \boxed{b_k[x_k, C, X^o]}, \pi)$$

Objects' classes

Objects' poses



The number of classification hypotheses is M^N (N: number of objects, M: number of classes)
How to sample w/o pruning hypotheses? How to estimate $\mathbb{P}_{safe}$?

T. Lemberg and V. Indelman, "Online Hybrid-Belief POMDP with Coupled Semantic-Geometric Models and Semantic Safety Awareness," arXiv'25.

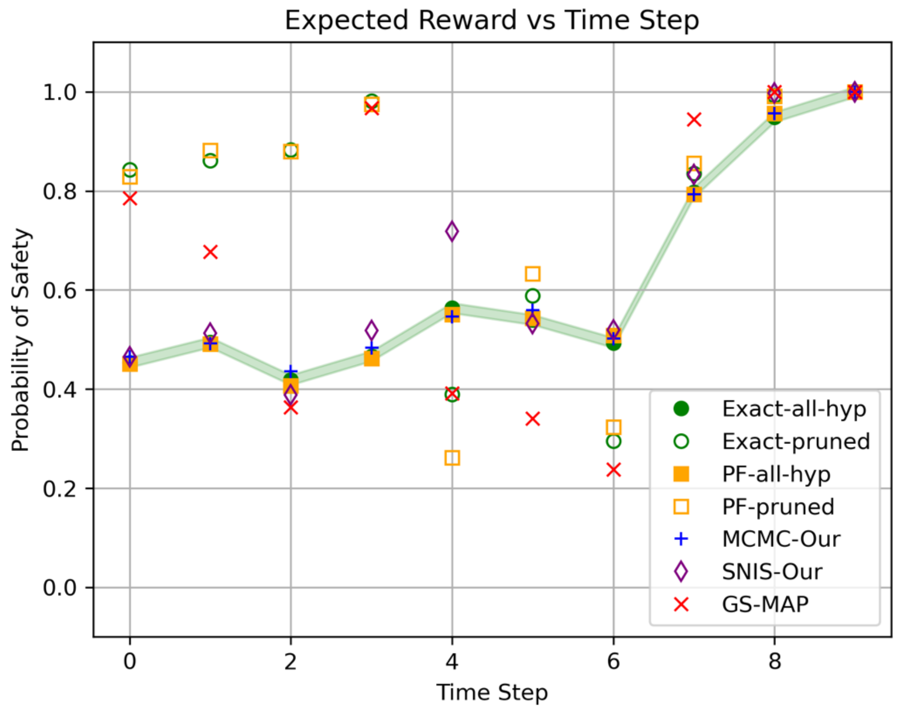
POMDP Planning with Hybrid Semantic-Geometric Beliefs

Experiments - Estimation of $\mathbb{P}_{safe}$ with different methods

- **Exact-all-hyp** – belief computed exactly
- **Exact-pruned** – pruned version
- **PF-all-hyp** – Particle filter
- **PF-pruned** – pruned version

Our methods

- **MCMC-Our** – MCMC samples
- **SNIS-Our** – self-normalized importance sampling
- **GS-MAP** – separate semantic and geometric



T. Lemberg and V. Indelman, "Online Hybrid-Belief POMDP with Coupled Semantic-Geometric Models and Semantic Safety Awareness," arXiv'25.

POMDP Planning with Hybrid Semantic-Geometric Beliefs

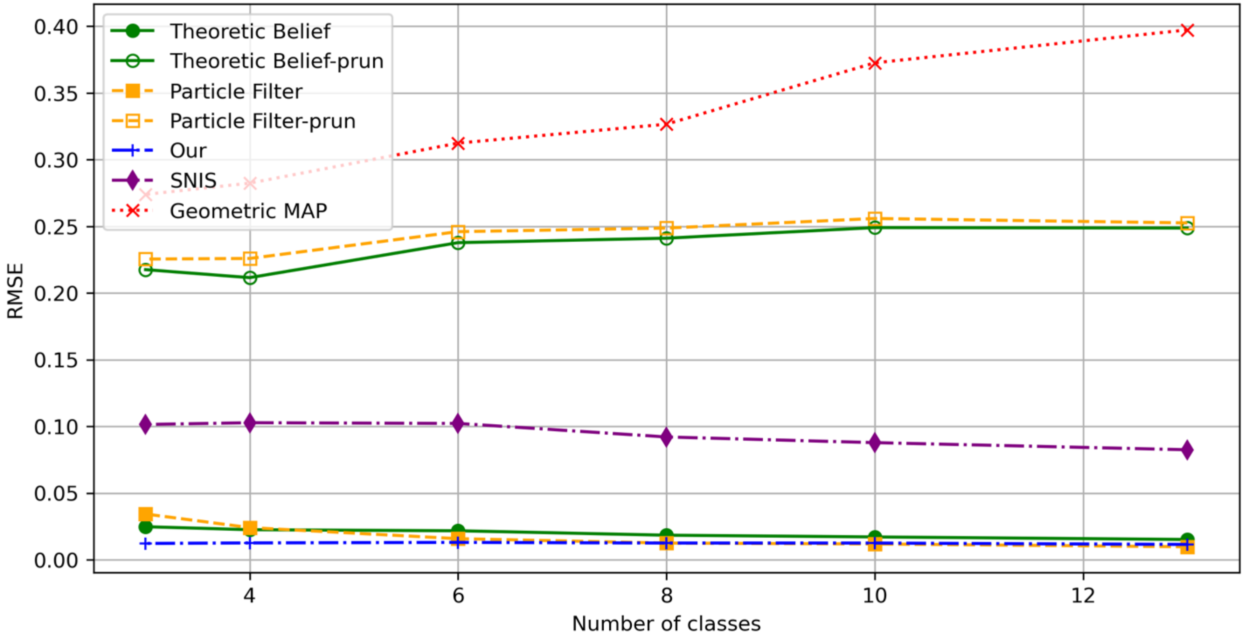
Experiments - Estimation of $\mathbb{P}_{safe}$ with different methods

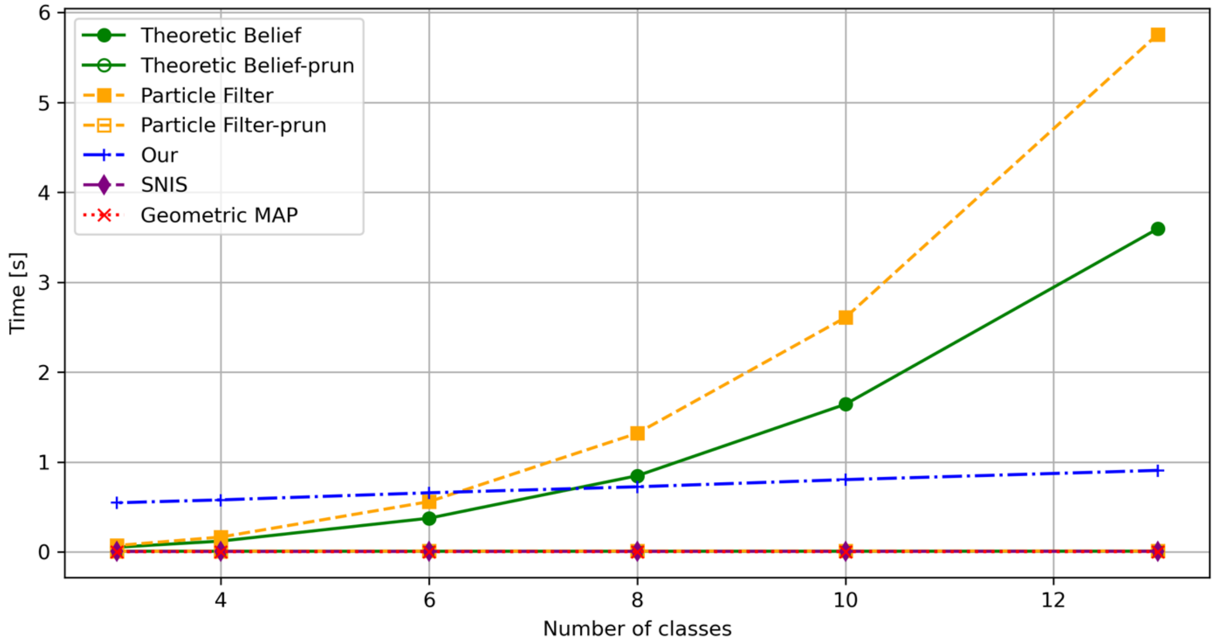
- **Exact-all-hyp** – belief computed exactly
- **Exact-pruned** – pruned version
- **PF-all-hyp** – Particle filter
- **PF-pruned** – pruned version

Our methods

- **MCMC-Our** – MCMC samples
- **SNIS-Our** – self-normalized importance sampling
- **GS-MAP** – separate semantic and geometric

Sensitivity to number of classes





T. Lemberg and V. Indelman, "Online Hybrid-Belief POMDP with Coupled Semantic-Geometric Models and Semantic Safety Awareness," arXiv'25.

POMDP Planning with Hybrid Semantic-Geometric Beliefs

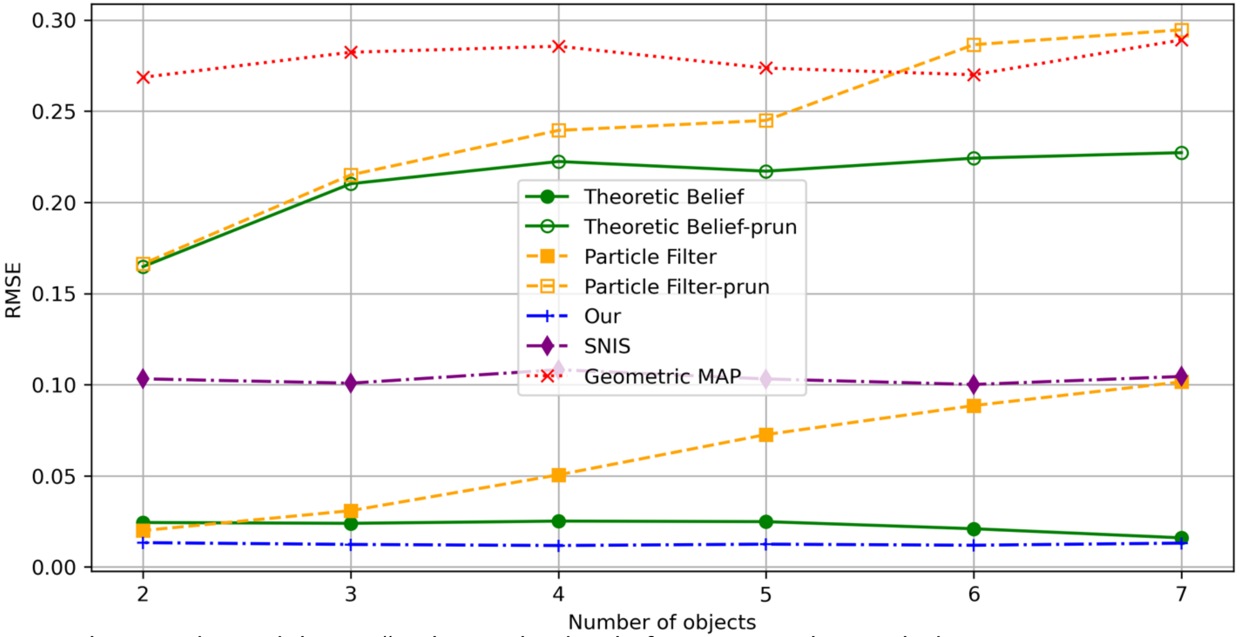
Experiments - Estimation of $\mathbb{P}_{safe}$ with different methods

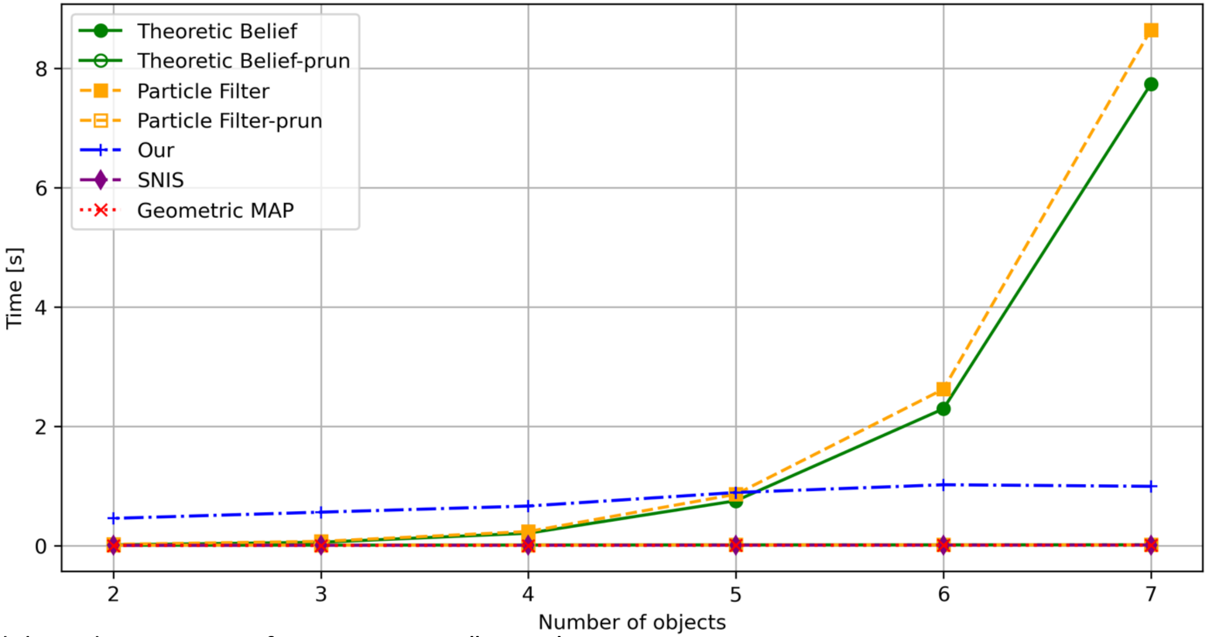
- **Exact-all-hyp** – belief computed exactly
- **Exact-pruned** – pruned version
- **PF-all-hyp** – Particle filter
- **PF-pruned** – pruned version

Our methods
- **MCMC-Our** – MCMC samples
- **SNIS-Our** – self-normalized importance sampling

- **GS-MAP** – separate semantic and geometric

Sensitivity to number of objects





T. Lemberg and V. Indelman, "Online Hybrid-Belief POMDP with Coupled Semantic-Geometric Models and Semantic Safety Awareness," arXiv'25.

Agenda

Experience Reuse in POMDP Planning

POMDP Planning with Hybrid Beliefs

Simplification of POMDP with Formal Guarantees

Multi-agent POMDP Planning with Inconsistent Beliefs

Semantic Risk Awareness

Ambiguous Environments

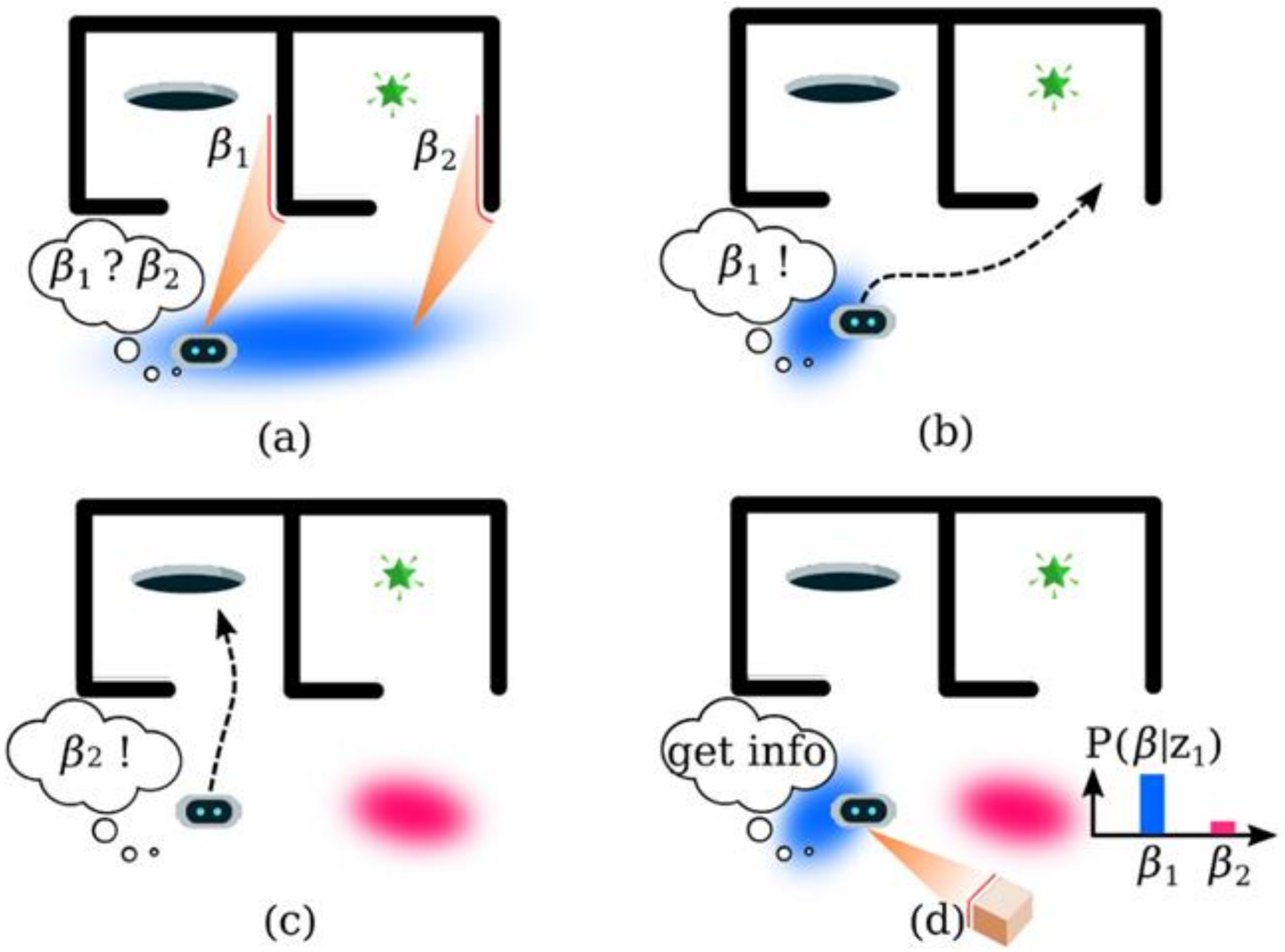
Ambiguous Scenarios

- Have to reason about data association hypotheses within inference and planning

An observation:
(e.g. LIDAR)

How should the agent act?

Unsafe!

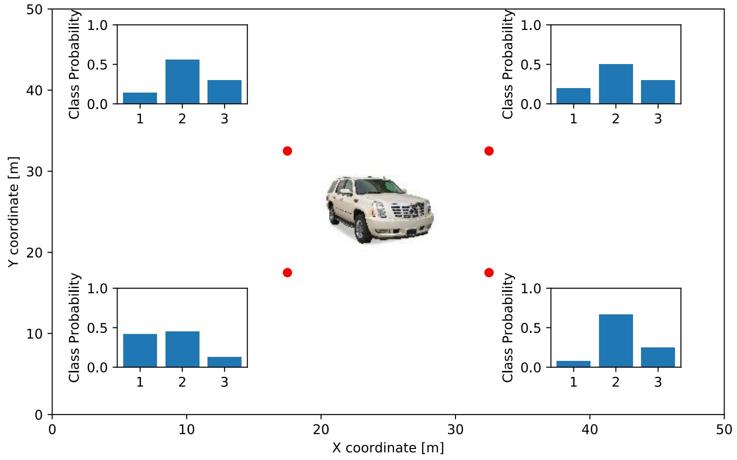


(a) (b) (c) (d)

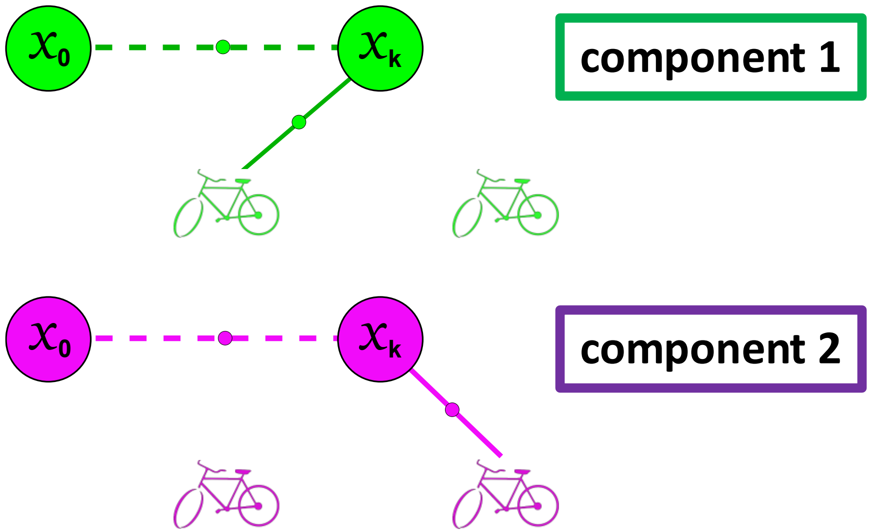
S. Pathak, A. Thomas, and V. Indelman, "A Unified Framework for Data Association Aware Belief Space Planning and Perception", IJRR'18.

Autonomous Semantic Perception & Ambiguous Environments

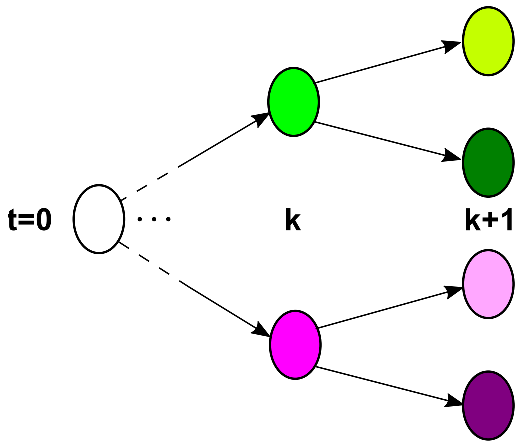
Viewpoint dependent semantic models



Data association hypotheses

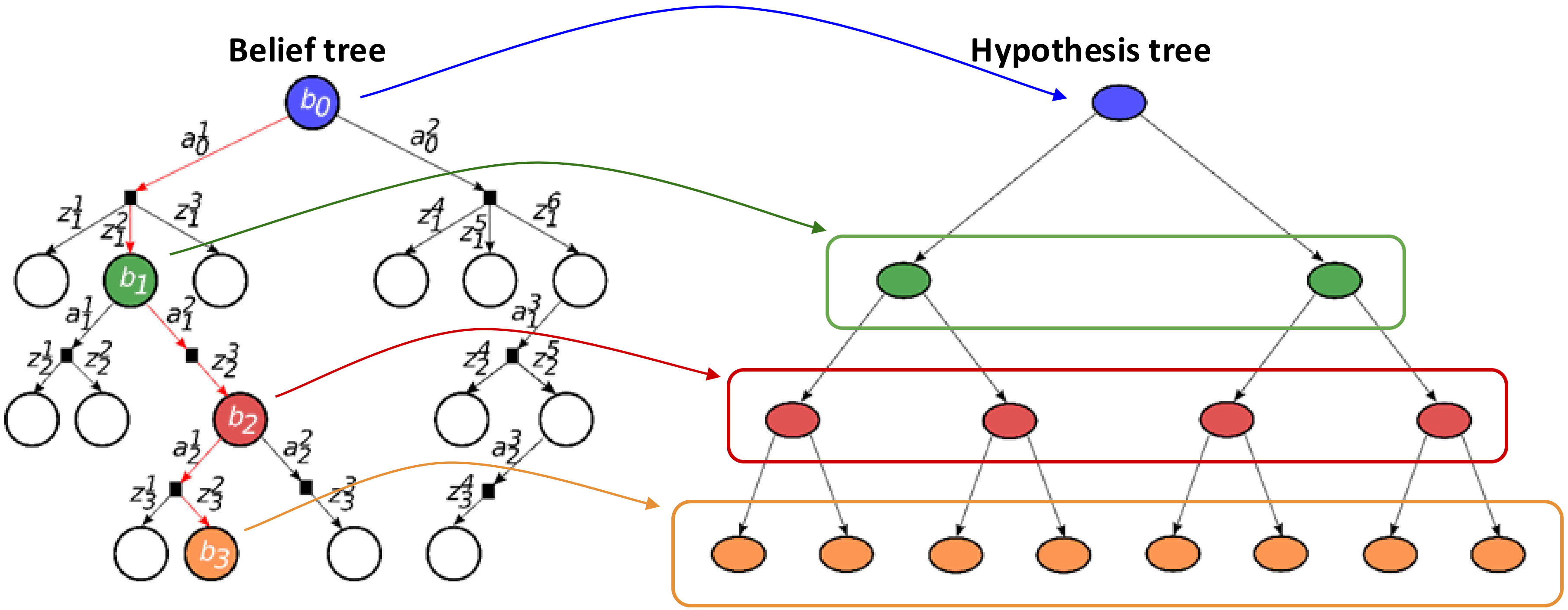


- Hybrid beliefs (over continuous and discrete RVs)
- The number of hypotheses can grow exponentially
- Impact on **safe** decision making?



Continuous-Discrete State Spaces - the Challenge

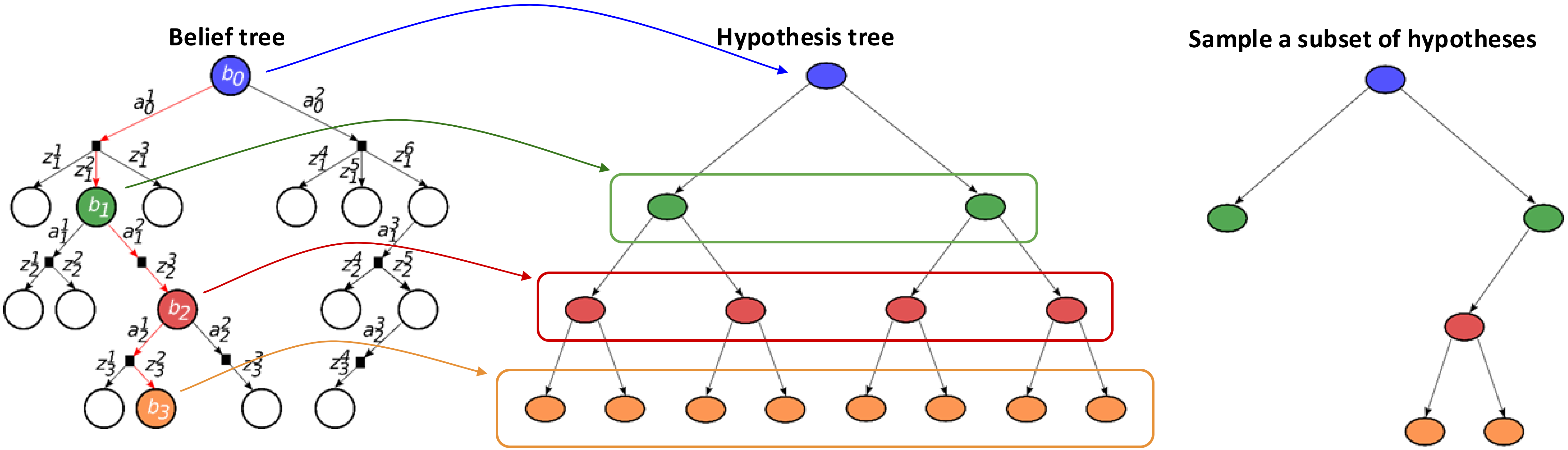
- The number of hypotheses may grow **exponentially** with the planning horizon!



M. Barenboim, M. Shienman, and V. Indelman, “Monte Carlo Planning in Hybrid Belief POMDPs,” IEEE RA-L’23.

Continuous-Discrete State Spaces - the Challenge

- The number of hypotheses may grow **exponentially** with the planning horizon!



Impact on **safe** decision making?

M. Barenboim, M. Shienman, and V. Indelman, “Monte Carlo Planning in Hybrid Belief POMDPs,” IEEE RA-L’23.

Simplification of POMDP with Hybrid Beliefs

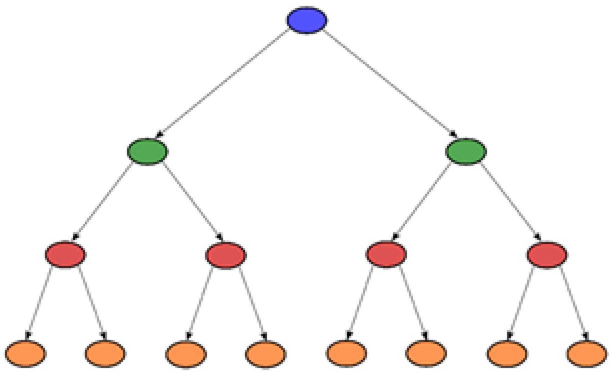
- Deterministic bound to relate the full set of hypotheses to a subset thereof,

Corollary

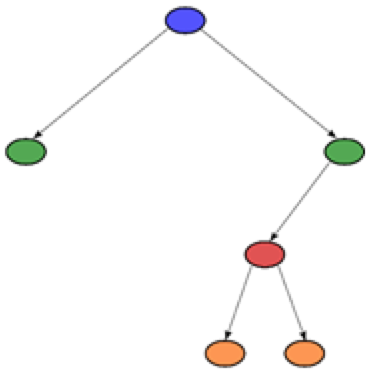
For any policy π, and selection of hypotheses set $\{\beta^i_{0:\mathcal{T}}\}_{i=0}^{|\mathcal{B}|}$ the following holds,

$$\left|V^\pi(b_0) - \bar{V}^\pi(\bar{b}_0)\right| \le \mathcal{R}_{max}\left[\mathcal{T}\delta_0^\beta + \sum_{k=1}^{\mathcal{T}}\sum_{\tau=1}^{k}\bar{\mathbb{E}}_{z_{1:\tau}}\left[\delta_\tau^\beta\right]\right].$$

Full tree



Any subset



Importantly, the bound relies on the available hypotheses

Can bound the theoretical value with access only to the simplified tree

Bounds can be evaluated online

M. Barenboim, I. Lev-Yehudi, and V. Indelman, “Data Association Aware POMDP Planning with Hypothesis Pruning Performance Guarantees,” IEEE RA-L’23.

Agenda

Experience Reuse in POMDP Planning

POMDP Planning with Hybrid Beliefs

Simplification of POMDP with Formal Guarantees

Multi-agent POMDP Planning with Inconsistent Beliefs

Simplification of Decision-Making Problems

Concept:

- Identify and solve a **simplified (computationally) easier** decision-making problem
- Provide (adaptive) performance guarantees

Specific simplifications include:

- Sparsification of Gaussian beliefs (high dim. state)
- Topological metric for Gaussian beliefs (high dim. state)
- Utilize a subset of samples (nonparametric beliefs)
- Utilize a subset of hypotheses (hybrid beliefs)
- Simplified models and spaces
- Simplification of Risk-Averse POMDP Planning
- Simplification in a multi-agent setting

[Indelman RAL'16; Elimelech & Indelman IJRR'22; Sztyglic & Indelman IROS'22, Zhitnikov & Indelman AIJ'22, TRO'24; Shienman & Indelman ICRA'22; Kitanov & Indelman IJRR'24; Zhitnikov et al. IJRR'24, Barenboim & Indelman NIPS'23; Lev-Yehudi, Barenboim & Indelman AAAI'24, Kong & Indelman ISRR 2024]

Simplification of Decision-Making Problems

Concept:

- Identify and solve a **simplified (computationally) easier** decision-making problem
- Provide (adaptive) performance guarantees

Specific simplifications include:

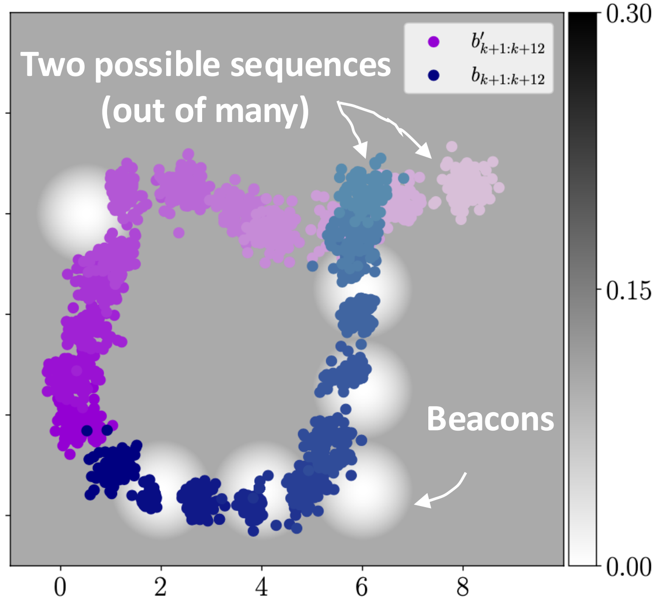
- Sparsification of Gaussian beliefs (high dim. state)
- Topological metric for Gaussian beliefs (high dim. state)
- Utilize a subset of samples (nonparametric beliefs)
- Utilize a subset of hypotheses (hybrid beliefs)
- Simplified models and spaces
- Simplification of Risk-Averse POMDP Planning
- Simplification in a multi-agent setting

[Indelman RAL'16; Elimelech & Indelman IJRR'22; Sztyglic & Indelman IROS'22, Zhitnikov & Indelman AIJ'22, TRO'24; Shienman & Indelman ICRA'22; Kitanov & Indelman IJRR'24; Zhitnikov et al. IJRR'24, Barenboim & Indelman NIPS'23; Lev-Yehudi, Barenboim & Indelman AAAI'24, Kong & Indelman ISRR 2024]

Simplification of POMDPs with Nonparametric Beliefs

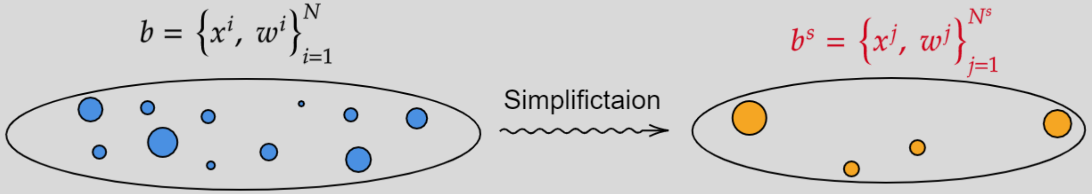
- Value function

$$V_k^{\pi}(b_k) \equiv J_k(b_k, \pi) = \mathbb{E}\{\sum_{l=0}^{L-1} r(b_{k+l}, \pi_{k+l}(b_{k+l})) + r(b_{k+L})\}$$



Simplification:

- Utilize a **subset** of samples for planning
- Information-theoretic reward (entropy)
- Analytical (**cheaper**) bounds over the reward

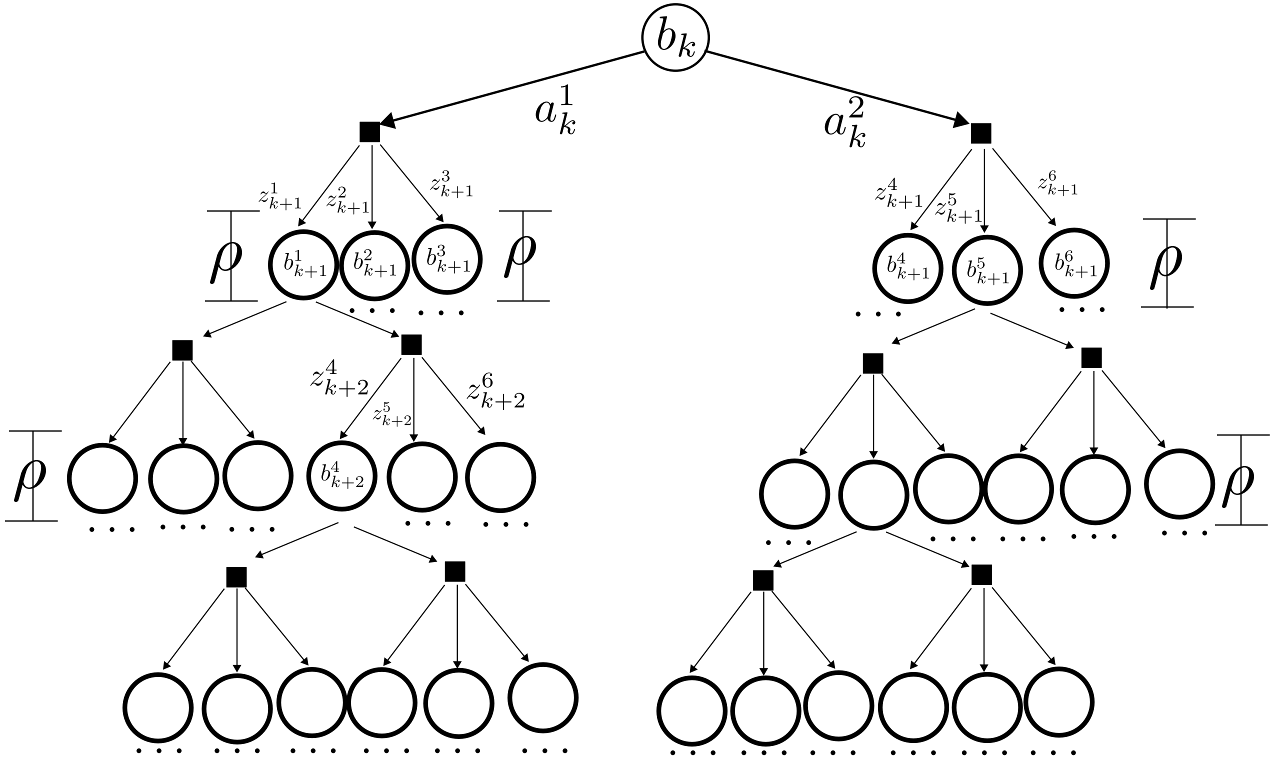


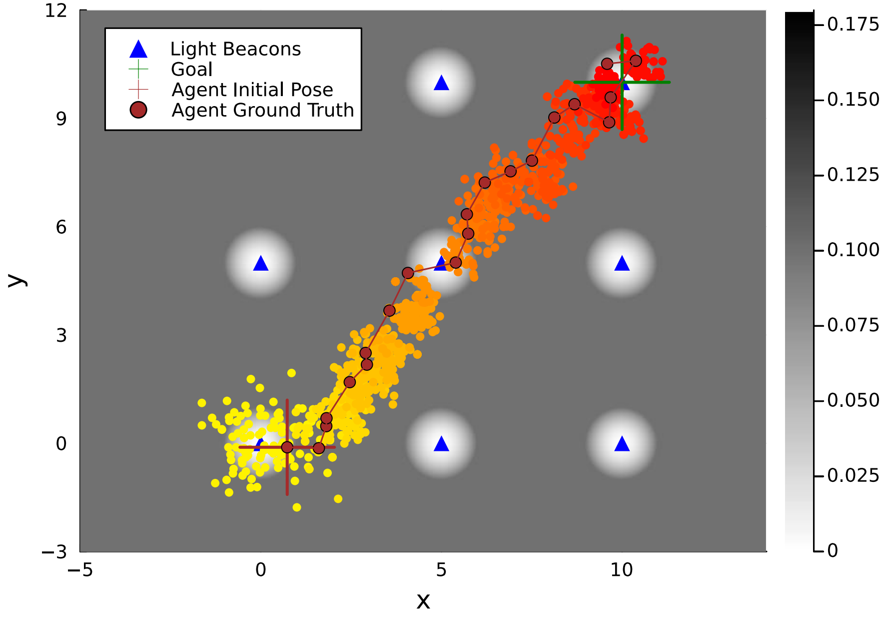
$$lb(b, b^s, a) \leq r(b, a) \leq ub(b, b^s, a)$$

O. Sztyglic and V. Indelman, “Speeding up POMDP Planning via Simplification”, IROS’22.
A. Zhitnikov, O. Sztyglic, and V. Indelman, “No Compromise in Solution Quality: Speeding Up Belief-dependent Continuous POMDPs via Adaptive Multilevel Simplification”, IJRR’24.

Simplification of POMDPs with Nonparametric Beliefs

- **Adaptive multi-level** simplification in a Sparse Sampling setting:



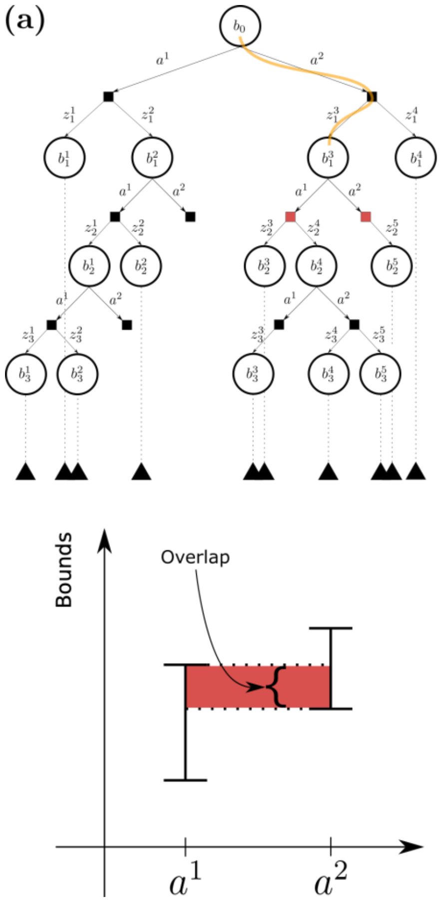


Typical speedup of 20% - 50%,
Same performance!

O. Sztyglic and V. Indelman, “Speeding up POMDP Planning via Simplification”, IROS’22.
A. Zhitnikov, O. Sztyglic, and V. Indelman, “No Compromise in Solution Quality: Speeding Up Belief-dependent Continuous POMDPs via Adaptive Multilevel Simplification”, IJRR’24.

Simplification of POMDPs with Nonparametric Beliefs

- **Adaptive multi-level** simplification in an MCTS setting:



O. Sztyglic and V. Indelman, “Speeding up POMDP Planning via Simplification”, IROS’22.
A. Zhitnikov, O. Sztyglic, and V. Indelman, “No Compromise in Solution Quality: Speeding Up Belief-dependent Continuous POMDPs via Adaptive Multilevel Simplification”, IJRR’24.

Simplification of Decision-Making Problems

Concept:

- Identify and solve a **simplified (computationally) easier** decision-making problem
- Provide (adaptive) performance guarantees

Specific simplifications include:

- Sparsification of Gaussian beliefs (high dim. state)
- Topological metric for Gaussian beliefs (high dim. state)
- Utilize a subset of samples (nonparametric beliefs)
- Utilize a subset of hypotheses (hybrid beliefs)
- Simplified models and spaces
- Simplification of Risk-Averse POMDP Planning
- Simplification in a multi-agent setting

[Indelman RAL'16; Elimelech & Indelman IJRR'22; Sztyglic & Indelman IROS'22, Zhitnikov & Indelman AIJ'22, TRO'24; Shienman & Indelman ICRA'22; Kitanov & Indelman IJRR'24; Zhitnikov et al. IJRR'24, Barenboim & Indelman NIPS'23; Lev-Yehudi, Barenboim & Indelman AAAI'24, Kong & Indelman ISRR 2024]

POMDPs with Deterministic Guarantees

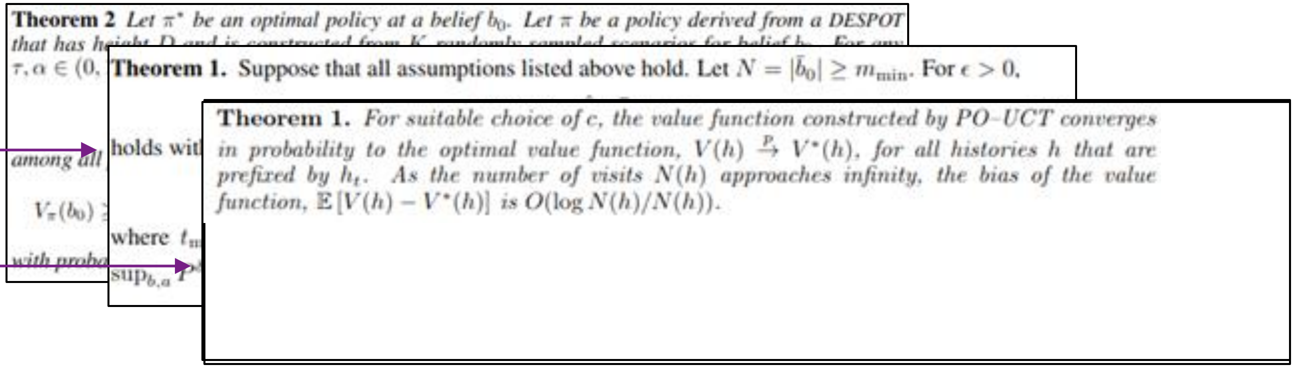
SOTA sampling based approaches come with probabilistic theoretical guarantees

DESPOT

AdaOPS

POMCP

…


Theorem 2 *Let π^* be an optimal policy at a belief b_0. Let π be a policy derived from a DESPOT*

Theorem 1. Suppose that all assumptions listed above hold. Let $N = |\bar{b}_0| \geq m_{\min}$. For $\epsilon > 0$,

Theorem 1. *For suitable choice of c, the value function constructed by PO-UCT converges in probability to the optimal value function, $V(h) \xrightarrow{p} V^*(h)$, for all histories h that are prefixed by h_t. As the number of visits $N(h)$ approaches infinity, the bias of the value function, $\mathbb{E}[V(h) - V^*(h)]$ is $O(\log N(h)/N(h))$.*

Can we get deterministic guarantees?

We show that deterministic guarantees are indeed possible!

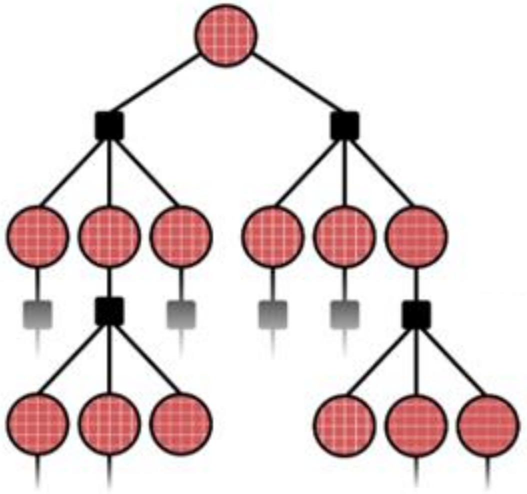
M. Barenboim and V. Indelman, “Online POMDP Planning with Anytime Deterministic Guarantees,” NeurIPS’23.

Online POMDP Planning with Anytime Deterministic Guarantees

Concept:

Instead of solving the original POMDP, consider a simplified version of that POMDP.

$$\mathcal{M} \rightsquigarrow \bar{\mathcal{M}}$$



Derive a mathematical relationship between the solution of the simplified, and the theoretical POMDP.

$$V(b) \quad ? \quad \bar{V}(b)$$

M. Barenboim and V. Indelman, “Online POMDP Planning with Anytime Deterministic Guarantees,” NeurIPS’23.

Online POMDP Planning with Anytime Deterministic Guarantees

- Given a POMDP: $\mathcal{M} = \langle \mathcal{X}, \mathcal{Z}, \mathcal{A}, b_0, \mathcal{P}_T, \mathcal{P}_Z, \rho, \gamma \rangle$
- Define a **simplified** POMDP,

$$\bar{\mathcal{M}} = \langle \underbrace{\bar{\mathcal{X}}, \bar{\mathcal{Z}}}, \mathcal{A}, \underbrace{\bar{b}_0, \bar{\mathcal{P}}_T, \bar{\mathcal{P}}_Z}, \rho, \gamma \rangle$$

$$\bar{\mathcal{X}}(H_t) \subset \mathcal{X}$$
$$\bar{\mathcal{Z}}(H_t) \subset \mathcal{Z}$$

$$\bar{b}_0(x) \triangleq \begin{cases} b_0(x) & , x \in \bar{\mathcal{X}}_0 \\ 0 & , otherwise \end{cases}$$

$$\bar{\mathbb{P}}(x_{t+1} \mid x_t, a_t) \triangleq \begin{cases} \mathbb{P}(x_{t+1} \mid x_t, a_t) & , x_{t+1} \in \bar{\mathcal{X}}(H^-_{t+1}) \\ 0 & , otherwise \end{cases}$$

$$\bar{\mathbb{P}}(z_t \mid x_t) \triangleq \begin{cases} \mathbb{P}(z_t \mid x_t) & , z_t \in \bar{\mathcal{Z}}(H_t) \\ 0 & , otherwise \end{cases}$$

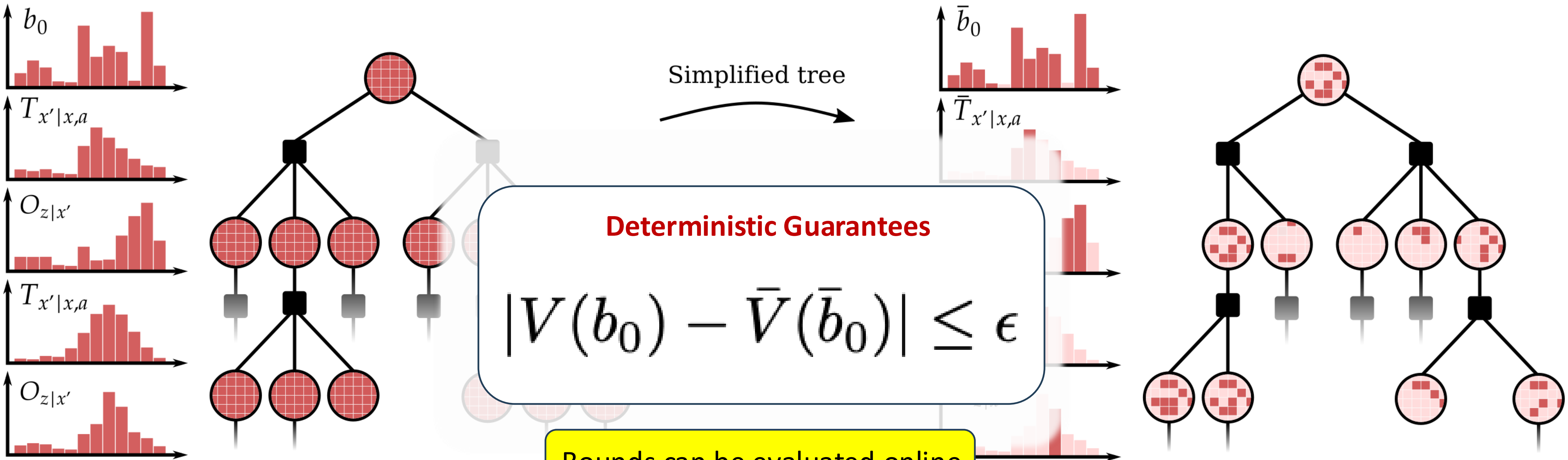
- Simplified value function

$$\bar{V}^\pi(\bar{b}_t) \triangleq r(\bar{b}_t, \pi_t) + \bar{\mathbb{E}}_{z_{t+1:\mathcal{T}}}\left[\bar{V}^\pi(\bar{b}_{t+1})\right]$$

M. Barenboim and V. Indelman, "Online POMDP Planning with Anytime Deterministic Guarantees," NeurIPS'23.

Online POMDP Planning with Anytime Deterministic Guarantees

- Deterministic guarantees (assuming discrete spaces)



Deterministic Guarantees

$$|V(b_0) - \bar{V}(\bar{b}_0)| \leq \epsilon$$

Bounds can be evaluated online

M. Barenboim and V. Indelman, "Online POMDP Planning with Anytime Deterministic Guarantees," NeurIPS'23.

Online POMDP Planning with Anytime Deterministic Guarantees

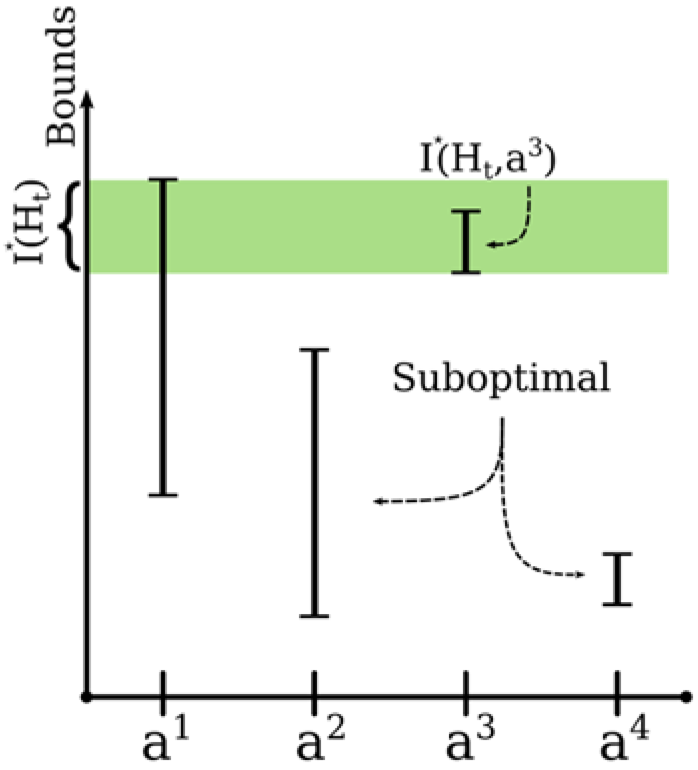
Importantly, the bounds can be calculated during planning.

How can we use them?

- **Pruning of sub-optimal branches**
  - → Made possible by the deterministic guarantees
- **Stopping criteria for the planning phase**
  - → Made possible by the deterministic guarantees
- **Finding the optimal solution in finite time**
  - → Without recovering the theoretical tree

Deterministic Guarantees

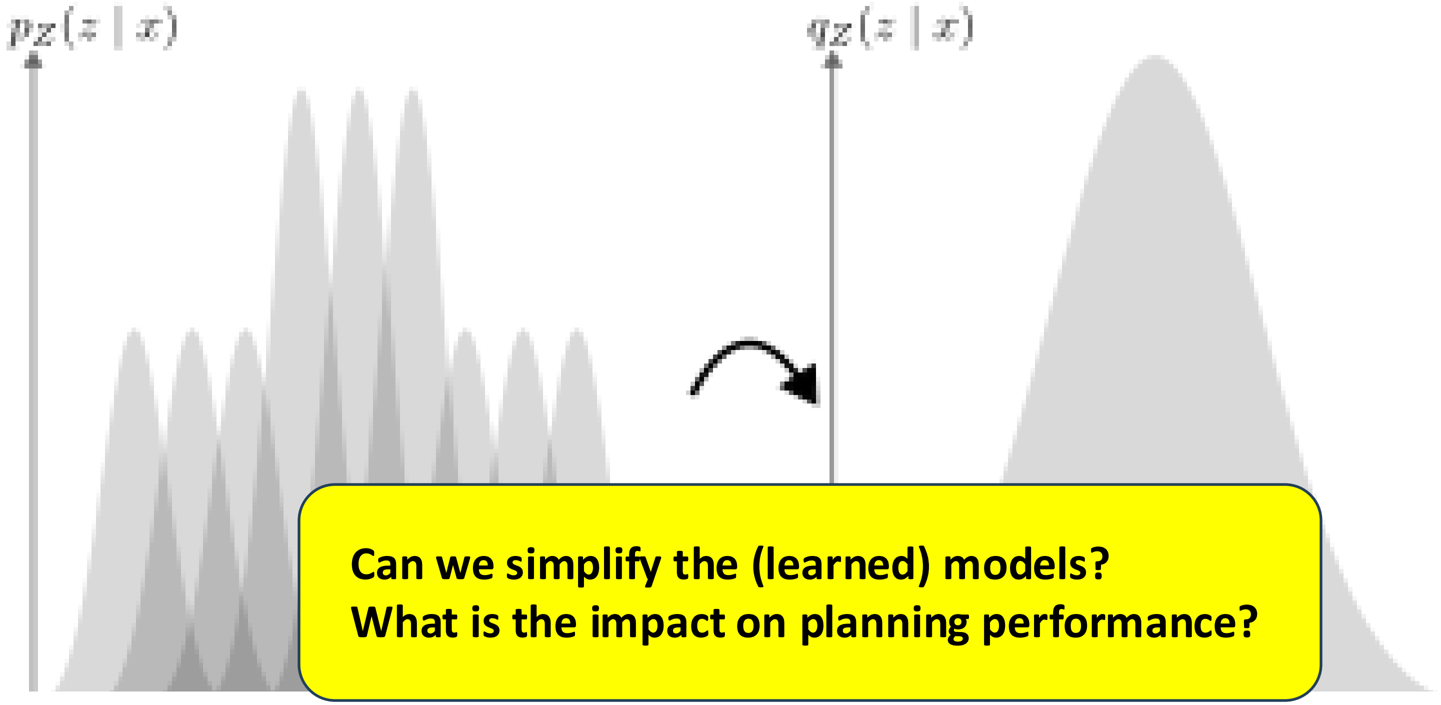
$$|V(b_0) - \bar{V}(\bar{b}_0)| \leq \epsilon$$



M. Barenboim and V. Indelman, "Online POMDP Planning with Anytime Deterministic Guarantees," NeurIPS'23.

Simplifying Complex Observation Models with Probabilistic Guarantees

- We replace the (learned) observation model p_Z with a cheaper model q_Z
  - Simpler GMM, Shallower Neural Network, etc.
  - Example:



Can we simplify the (learned) models?
What is the impact on planning performance?

Simplified models

$p_\theta(z \mid x)$

Original, **expensive**

$q_\phi(z \mid x)$

Simplified, **cheap**

I. Lev-Yehudi, M. Barenboim, and V. Indelman, "Simplifying Complex Observation Models in Continuous POMDP Planning with Probabilistic Guarantees and Practice," AAAI'24.

Simplifying Complex Observation Models with Probabilistic Guarantees

- We replace the (learned) observation model p_Z with a cheaper model q_Z
- Simplified action-value function: $Q^{q_Z}_{\mathbf{P}}$

Corollary 3

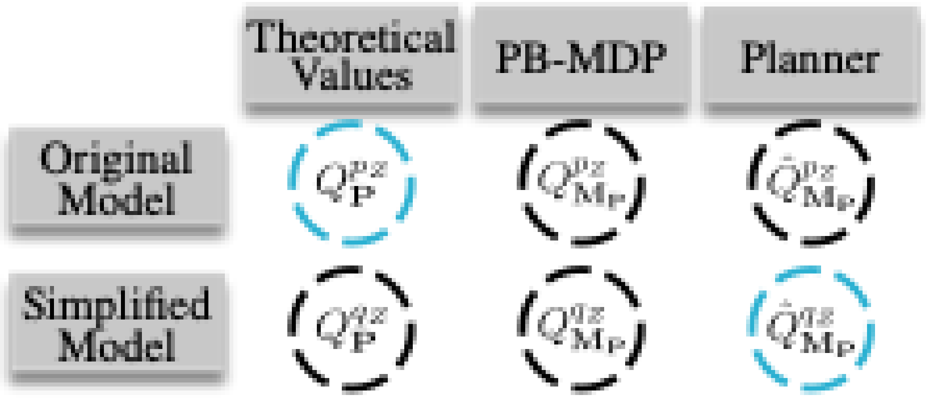
For arbitrary $\varepsilon, \delta > 0$ there exists a number of particles for which

$$\left|\underline{Q^{p_Z}_{\mathbf{P}}(b_t, a)} - \underline{\hat{Q}^{q_Z}_{\mathbf{M_P}}(\bar{b}_t, a)}\right| \leq \hat{\Phi}_{\mathbf{M_P}}(\bar{b}_t, a) + \varepsilon$$

with probability of at least $1 - \delta$ for any guaranteed planner

Theoretical Q function of the POMDP, with **original** models

Estimator of the Q function of a **particle-belief** POMDP, with **simplified** models

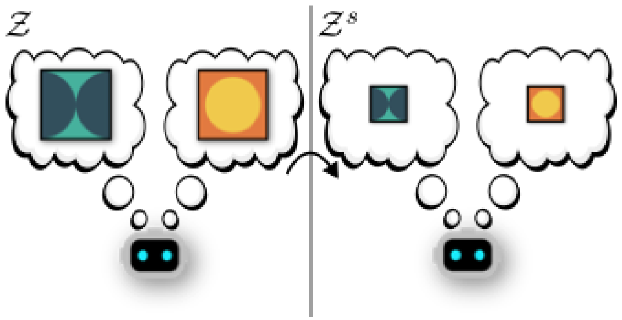


- Importance sampling
- Separate calculations to offline/online

I. Lev-Yehudi, M. Barenboim, and V. Indelman, "Simplifying Complex Observation Models in Continuous POMDP Planning with Probabilistic Guarantees and Practice," AAAI'24.

Simplified POMDP Planning with an Alternative Observation Space

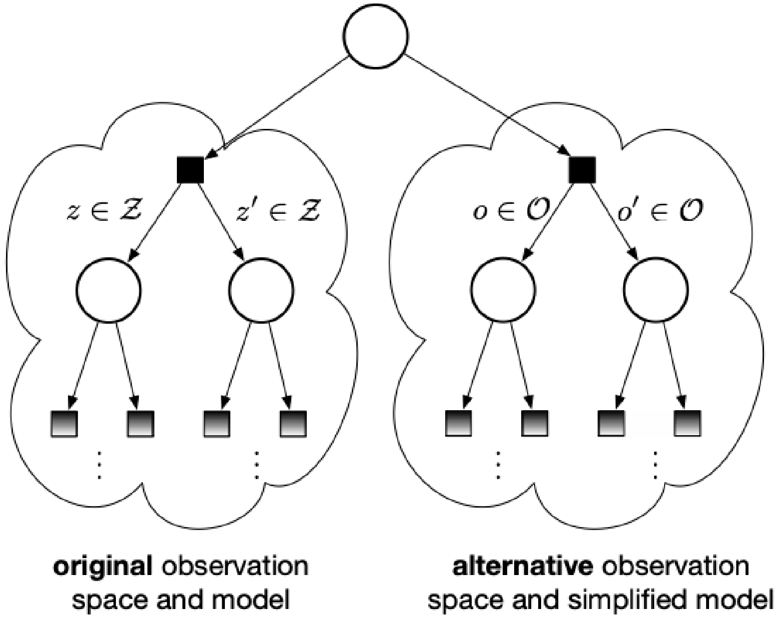
- Switch to an alternative observation space and model



Model Definition

POMDP tuple: $\langle \mathcal{X}, \mathcal{A}, \mathcal{Z}, \mathbb{P}_T, \mathbb{P}_Z, b_k, r \rangle \rightarrow \langle \mathcal{X}, \mathcal{A}, \mathcal{O}, \mathbb{P}_T, \mathbb{P}_O, b_k, r \rangle$

- Only at certain levels and branches of the tree

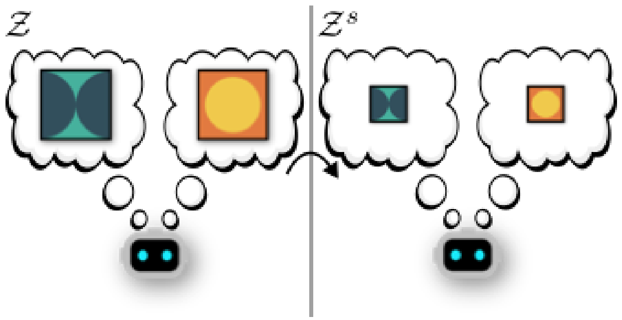


D. Kong and V. Indelman, “Simplified Belief Space Planning with an Alternative Observation Space and Formal Performance Guarantees,” ISRR’24.

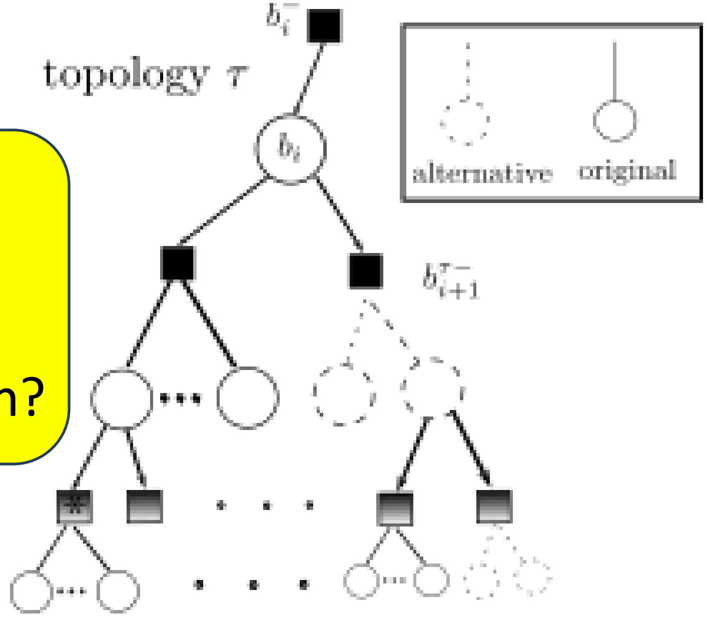
Simplified POMDP Planning with an Alternative Observation Space

- Switch to an alternative observation space and model

Model Definition
POMDP tuple: $\langle \mathcal{X}, \mathcal{A}, \mathcal{Z}, \mathbb{P}_T, \mathbb{P}_Z, b_k, r \rangle \rightarrow \langle \mathcal{X}, \mathcal{A}, \mathcal{O}, \mathbb{P}_T, \mathbb{P}_O, b_k, r \rangle$



- Only at certain levels and branches of the tree



Main questions addressed:

- How to decide online where to simplify in belief tree?
- How to provide formal performance guarantees?
- How to adaptively transition between the different levels of simplification?

D. Kong and V. Indelman, "Simplified Belief Space Planning with an Alternative Observation Space and Formal Performance Guarantees," ISRR'24.

Simplification of Decision-Making Problems

Concept:

- Identify and solve a **simplified (computationally) easier** decision-making problem
- Provide performance guarantees

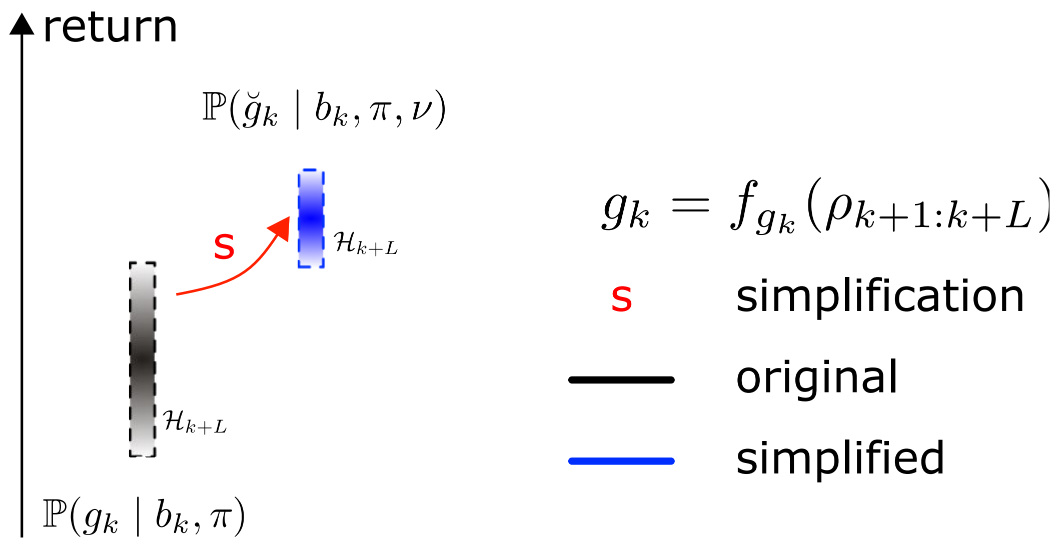
Specific simplifications include:

- Sparsification of Gaussian beliefs (high dim. state)
- Topological metric for Gaussian beliefs (high dim. state)
- Utilize a subset of samples (nonparametric beliefs)
- Utilize a subset of hypotheses (hybrid beliefs)
- Simplified models and spaces
- Simplification of Risk-Averse POMDP Planning
- Simplification in a multi-agent setting

[Indelman RAL'16; Elimelech & Indelman IJRR'22; Sztyglic & Indelman IROS'22, Zhitnikov & Indelman AIJ'22, TRO'24; Shienman & Indelman ICRA'22; Kitanov & Indelman IJRR'24; Zhitnikov et al. IJRR'24, Barenboim & Indelman NIPS'23; Lev-Yehudi, Barenboim & Indelman AAAI'24, Kong & Indelman ISRR 2024]

Simplification of Risk Averse POMDP Planning

- Impact of simplification on **distribution** over returns/rewards
- Simplified **risk aware** decision making with belief-dependent rewards



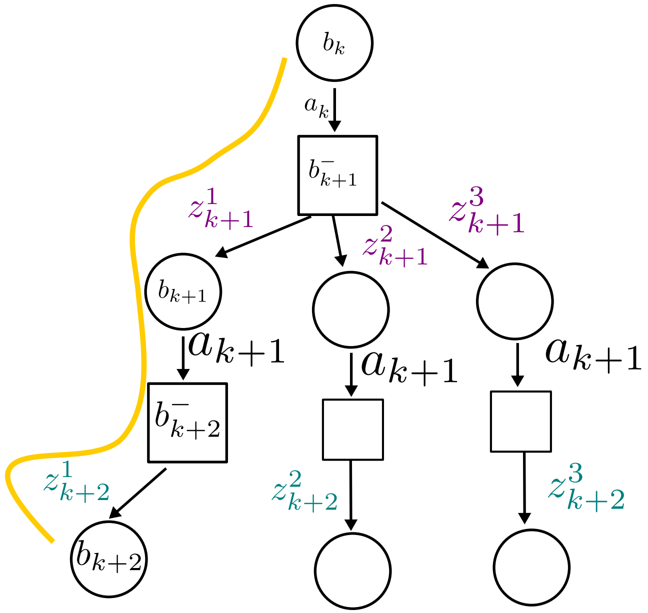
$$V^{\pi}(b_k) = \varphi\Big(\mathbb{P}(\rho_{k+1:k+L} | b_k, \pi_{k:k+L-1}), g_k\Big)$$

Y. Pariente and V. Indelman, "Simplification of Risk Averse POMDPs with Performance Guarantees,", arXiv'24.
A. Zhitnikov and V. Indelman, "Simplified Risk Aware Decision Making with Belief Dependent Rewards in Partially Observable Domains," Artificial Intelligence, 2022.

Probabilistically Constrained Belief Space Planning

$$\max_{\pi_{k+}} \mathbb{E}\left[\sum_{\ell=k}^{k+L-1} \rho_{\ell+1} \middle| b_k, \pi_{k+}\right]$$

$$\text{subject to } \mathrm{P}(c(b_{k:k+L};\phi,\delta)=1|b_k,\pi_{k+}) \geq 1-\epsilon$$



Information gain[1]:

$$c(b_{k:k+L};\phi,\delta) \triangleq \mathbf{1}_{\left\{\left(\sum_{\ell=k}^{k+L-1}\phi(b_t,b_{t+1})\right)\geq\delta\right\}}(b_{k:k+L})$$

Safety[2]:

$$c(b_{k:k+L};\phi,\delta) \triangleq \prod_{\ell=k}^{k+L} \mathbf{1}_{\{b_\ell:\phi(b_\ell)\geq\delta\}}(b_\ell)$$

[1]A. Zhitnikov and V. Indelman, "Simplified Continuous High Dimensional Belief Space Planning with Adaptive Probabilistic Belief-dependent Constraints," T-RO'24.
[2]A. Zhitnikov and V. Indelman, "Anytime Probabilistically Constrained Provably Convergent Online Belief Space Planning," arXiv'24.

Agenda

Experience Reuse in POMDP Planning

POMDP Planning with Hybrid Beliefs

Simplification of POMDP with Formal Guarantees

Multi-agent POMDP Planning with Inconsistent Beliefs

Simplification of Decision-Making Problems

Concept:

- Identify and solve a **simplified (computationally) easier** decision-making problem
- Provide performance guarantees

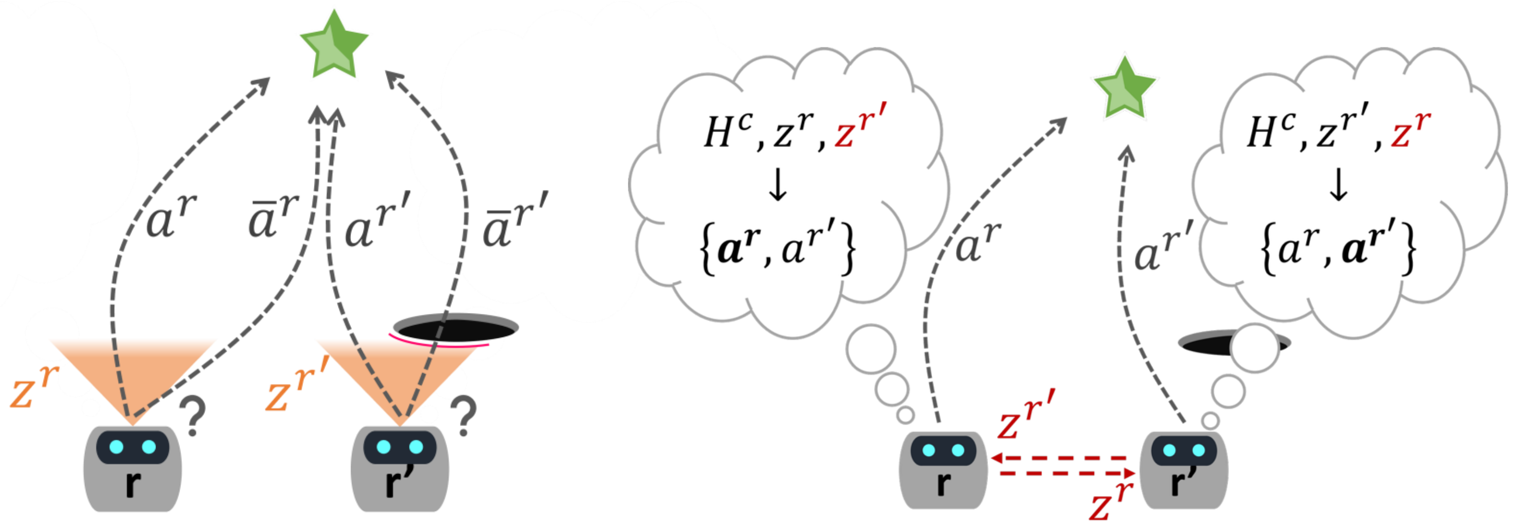
Specific simplifications include:

- Sparsification of Gaussian beliefs (high dim. state)
- Topological metric for Gaussian beliefs (high dim. state)
- Utilize a subset of samples (nonparametric beliefs)
- Utilize a subset of hypotheses (hybrid beliefs)
- Simplified models and spaces
- Simplification of Risk-Averse POMDP Planning
- Simplification in a multi-agent setting

[Indelman RAL'16; Elimelech & Indelman IJRR'22; Sztyglic & Indelman IROS'22, Zhitnikov & Indelman AIJ'22, TRO'24; Shienman & Indelman ICRA'22; Kitanov & Indelman IJRR'24; Zhitnikov et al. IJRR'24, Barenboim & Indelman NIPS'23; Lev-Yehudi, Barenboim & Indelman AAAI'24, Kong & Indelman ISRR 2024]

Multi-Robot Belief Space Planning

- **A common assumption**: Beliefs of different robots are consistent at planning time
- Requires prohibitively frequent data-sharing capabilities!



TECHNION Israel Institute of Technology

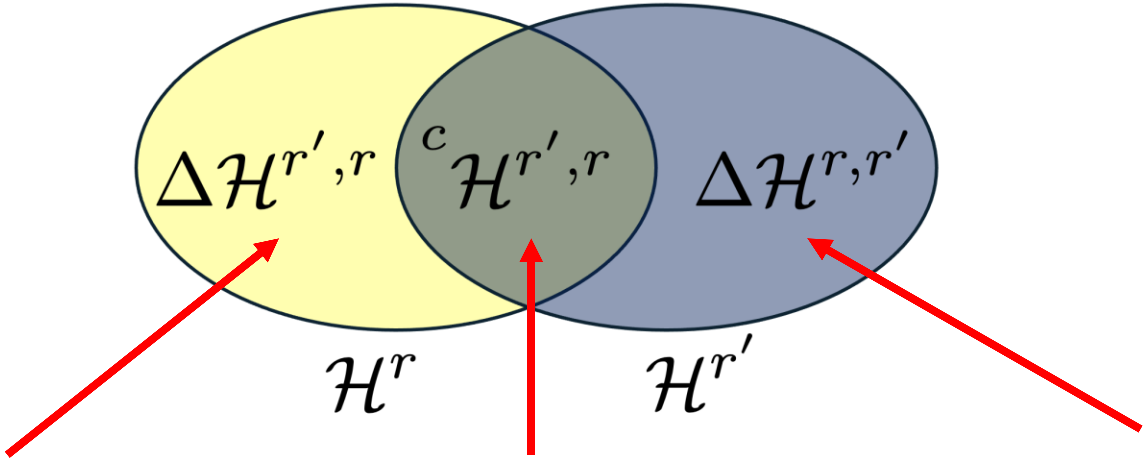
ANPL Autonomous Navigation and Perception Lab

Multi-Robot Cooperative BSP with Inconsistent Beliefs

What happens when data-sharing capabilities between the robots are limited?

- Histories & beliefs of the robots may **differ** due to limited data-sharing capabilities

$$b_k^r = \mathbb{P}(x_k \mid \mathcal{H}_k^r) \qquad b_k^{r'} = \mathbb{P}(x_k \mid \mathcal{H}_k^{r'}) \qquad \mathcal{H}_k^r \neq \mathcal{H}_k^{r'}$$



Available only to robot r

Common history, e.g. from the last data-sharing

Available only to robot r’

T. Kundu, M. Rafaeli, and V. Indelman, “Multi-Robot Communication-Aware Cooperative Belief Space Planning with Inconsistent Beliefs: An Action-Consistent Approach,” IROS’24.
T. Kundu, M. Rafaeli, A. Gulyaev, and V. Indelman, “Action-Consistent Decentralized Belief Space Planning with Inconsistent Beliefs and Limited Data Sharing: Framework and Simplification Algorithms with Formal Guarantees,” arXiv’25.

Multi-Robot Cooperative BSP with Inconsistent Beliefs

What happens when data-sharing capabilities between the robots are limited?

- Histories & beliefs of the robots may **differ** due to limited data-sharing capabilities

$$b_k^r = \mathbb{P}(x_k \mid \mathcal{H}_k^r) \qquad b_k^{r'} = \mathbb{P}(x_k \mid \mathcal{H}_k^{r'}) \qquad \mathcal{H}_k^r \neq \mathcal{H}_k^{r'}$$

- Decentralized POMDP tuple from the perspective of robot r:

$$\langle \mathcal{X}, \mathcal{Z}, \mathcal{A}, T, O, \rho, b_k^r \rangle$$

- Objective function:

$$J(b_k^r, a_{k+}) = \mathop{\mathbb{E}}_{z_{k+1:k+L}} \left[\sum_{l=0}^{L-1} \rho(b_{k+l}^r, a_{k+l}) + \rho(b_{k+L}^r)\right]$$

T. Kundu, M. Rafaeli, and V. Indelman, "Multi-Robot Communication-Aware Cooperative Belief Space Planning with Inconsistent Beliefs: An Action-Consistent Approach," IROS'24.
T. Kundu, M. Rafaeli, A. Gulyaev, and V. Indelman, "Action-Consistent Decentralized Belief Space Planning with Inconsistent Beliefs and Limited Data Sharing: Framework and Simplification Algorithms with Formal Guarantees," arXiv'25.

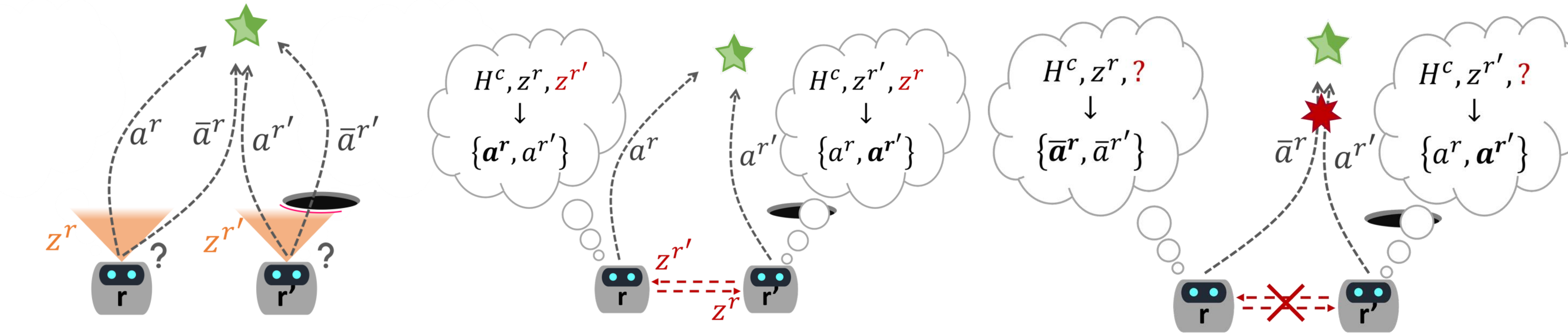
Multi-Robot Cooperative BSP with Inconsistent Beliefs

What happens when data-sharing capabilities between the robots are limited?

- Histories & beliefs of the robots may **differ** due to limited data-sharing capabilities

$$b_k^r = \mathbb{P}(x_k \mid \mathcal{H}_k^r) \qquad b_k^{r'} = \mathbb{P}(x_k \mid \mathcal{H}_k^{r'}) \qquad \mathcal{H}_k^r \neq \mathcal{H}_k^{r'}$$

- **Can lead to a lack of coordination and unsafe and sub-optimal actions**



T. Kundu, M. Rafaeli, and V. Indelman, "Multi-Robot Communication-Aware Cooperative Belief Space Planning with Inconsistent Beliefs: An Action-Consistent Approach," IROS'24.
T. Kundu, M. Rafaeli, A. Gulyaev, and V. Indelman, "Action-Consistent Decentralized Belief Space Planning with Inconsistent Beliefs and Limited Data Sharing: Framework and Simplification Algorithms with Formal Guarantees," arXiv'25.

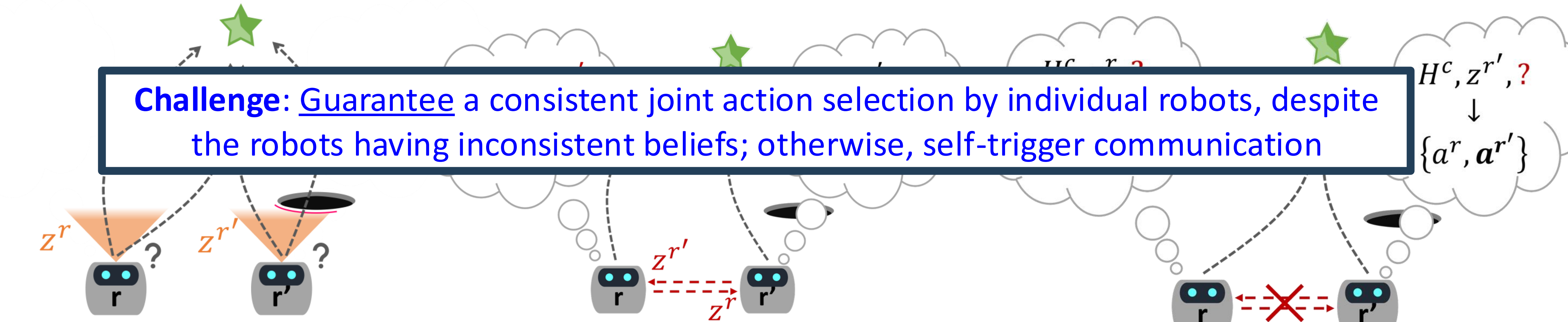
Multi-Robot Cooperative BSP with Inconsistent Beliefs

What happens when data-sharing capabilities between the robots are limited?

- Histories & beliefs of the robots may **differ** due to limited data-sharing capabilities

$$b_k^r = \mathbb{P}(x_k \mid \mathcal{H}_k^r) \qquad b_k^{r'} = \mathbb{P}(x_k \mid \mathcal{H}_k^{r'}) \qquad \mathcal{H}_k^r \neq \mathcal{H}_k^{r'}$$

- **Can lead to a lack of coordination and unsafe and sub-optimal actions**



Challenge: Guarantee a consistent joint action selection by individual robots, despite the robots having inconsistent beliefs; otherwise, self-trigger communication

T. Kundu, M. Rafaeli, and V. Indelman, "Multi-Robot Communication-Aware Cooperative Belief Space Planning with Inconsistent Beliefs: An Action-Consistent Approach," IROS'24.

T. Kundu, M. Rafaeli, A. Gulyaev, and V. Indelman, "Action-Consistent Decentralized Belief Space Planning with Inconsistent Beliefs and Limited Data Sharing: Framework and Simplification Algorithms with Formal Guarantees," arXiv'25.

Agenda

- Experience Reuse in POMDP Planning
- POMDP Planning with Hybrid Beliefs
- Simplification of POMDP with Formal Guarantees
- Multi-agent POMDP Planning with Inconsistent Beliefs

Semantic Risk Awareness

Ambiguous Environments

Thank You

TECHNION Israel Institute of Technology

ANPL Autonomous Navigation and Perception Lab

ISRAEL SCIENCE FOUNDATION

NSF

BSF United States – Israel Binational Science Foundation

inistry of Science and Technology

Meta

intel

THE ISRAELI SMART TRANSPORTATION RESEARCH CENTER