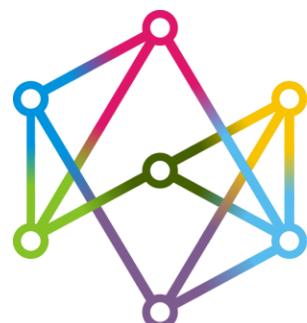


# Towards Scalable Online Decision Making Under Uncertainty in Partially Observable Environments

Vadim Indelman



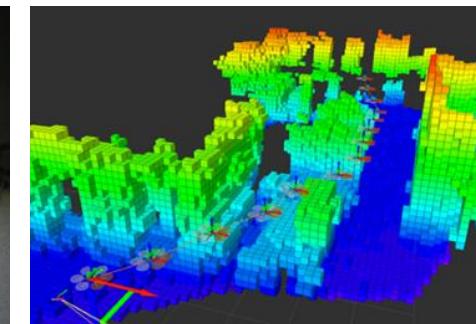
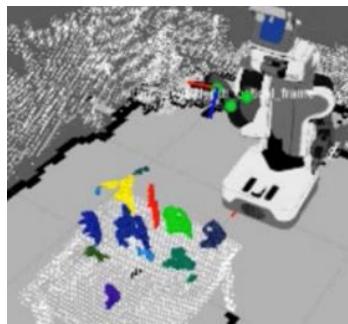
**TECHNION**  
Israel Institute  
of Technology



**ANPL**  
Autonomous Navigation and  
Perception Lab

# Advanced Autonomy

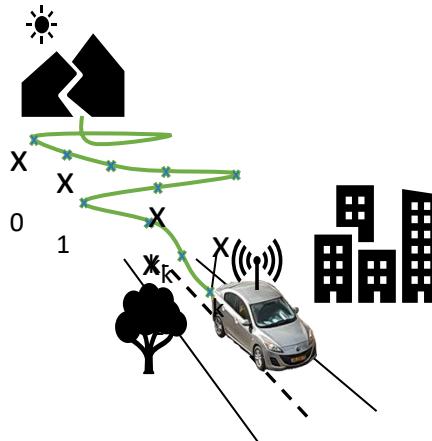
Involves autonomous navigation, active SLAM, informative gathering, active sensing, etc.



# Advanced Autonomy

## Perception and Inference

Where am I? What is the surrounding environment?



## Decision-Making Under Uncertainty

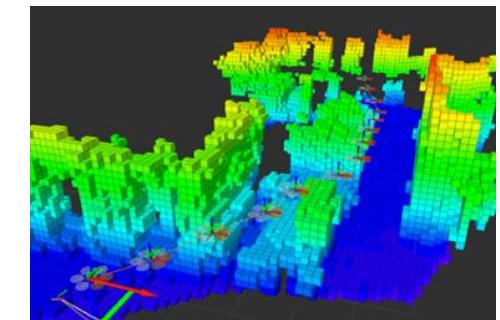
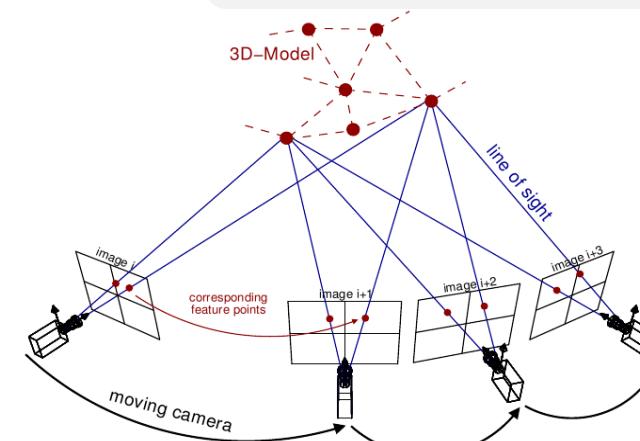
What should I be doing next?

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

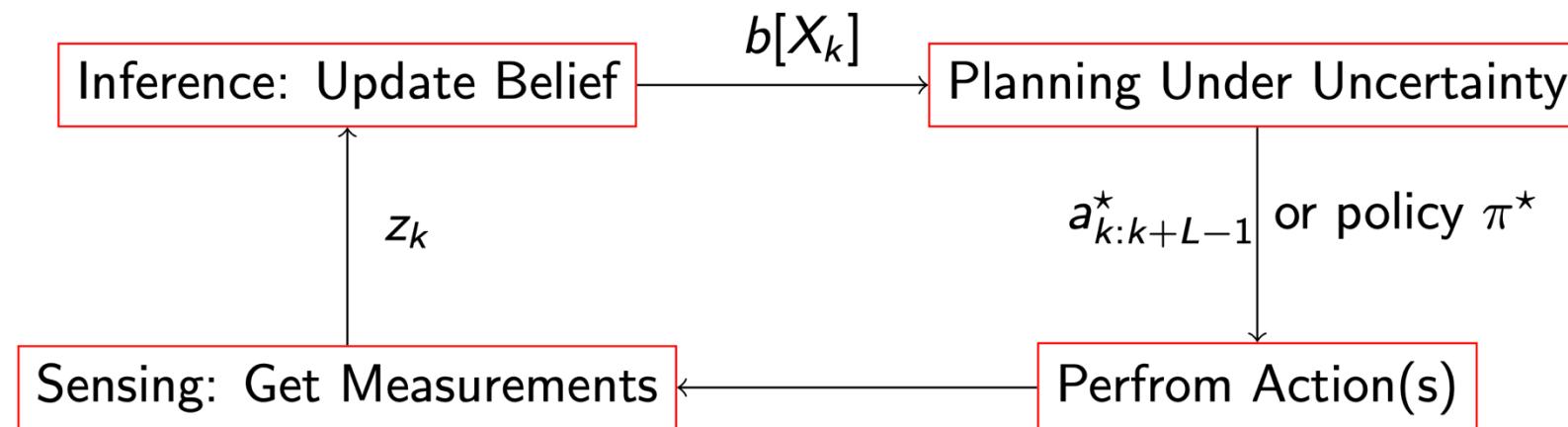
### Perception and Inference



### Decision-Making Under Uncertainty



# Autonomy Loop

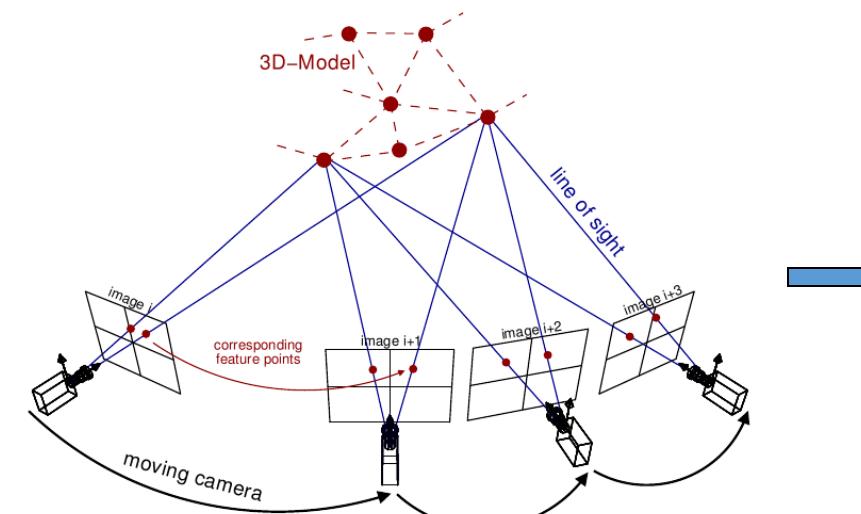
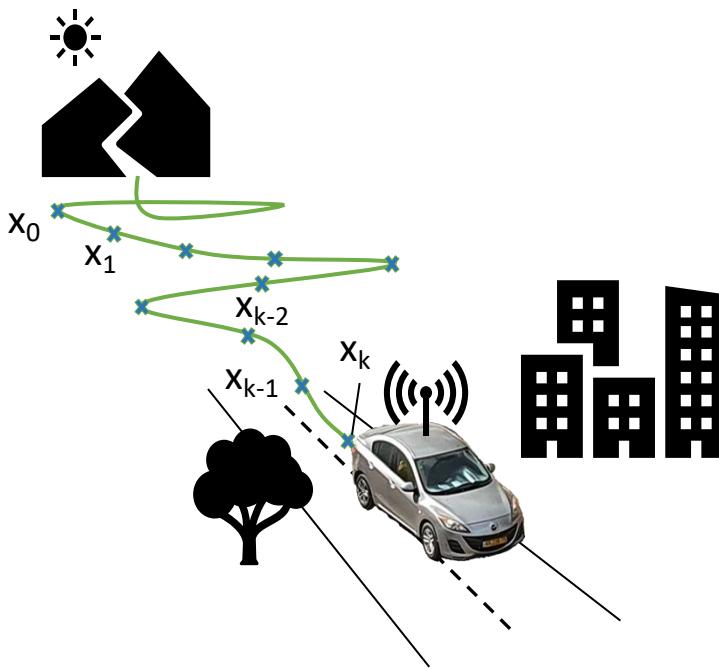


# Perception and Inference

- Posterior belief at time k:

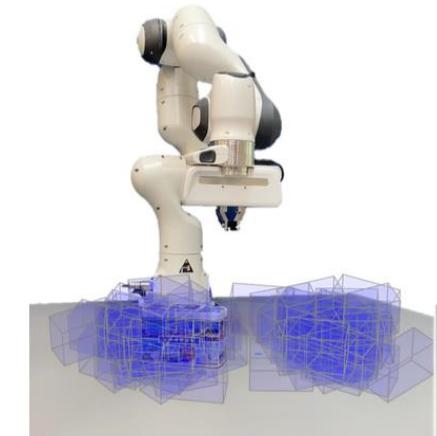
$$b_k \triangleq b[X_k] = \mathbb{P}(X_k \mid \underbrace{\mathcal{H}_k}_{\substack{\text{state/variables at} \\ \text{time instant } k}}, \underbrace{a_{0:k-1}}_{\text{actions}}, \underbrace{z_{1:k}}_{\text{observations}})$$

- Example:

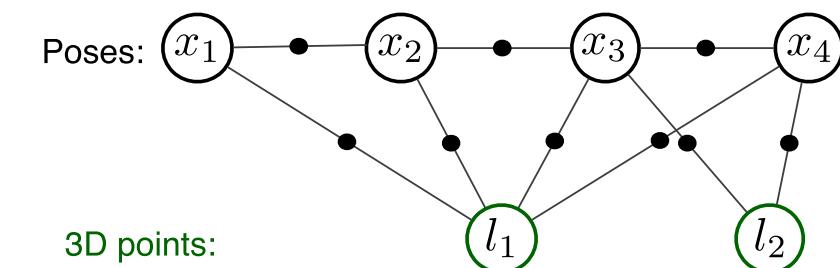


$$X_k \doteq \{x_0, \dots, x_k, L_k\}$$

Past & current robot states
Environment representation, e.g. Landmarks



Can be represented with graphical models, e.g. a Factor Graph



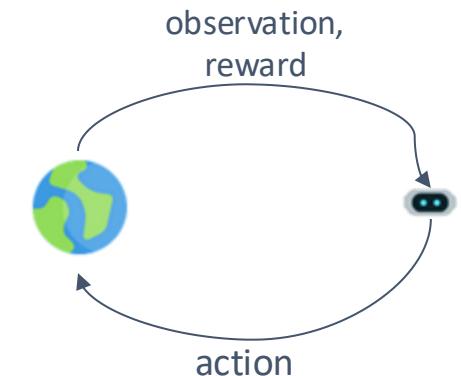
# Partially Observable Markov Decision Process (POMDP)

---

- POMDP tuple:

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

— state, observation, and action spaces  
— transition and observation models  
— Belief-dependent reward function  
— Belief at planning time instant k



- Value function

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

Belief-dependent reward function

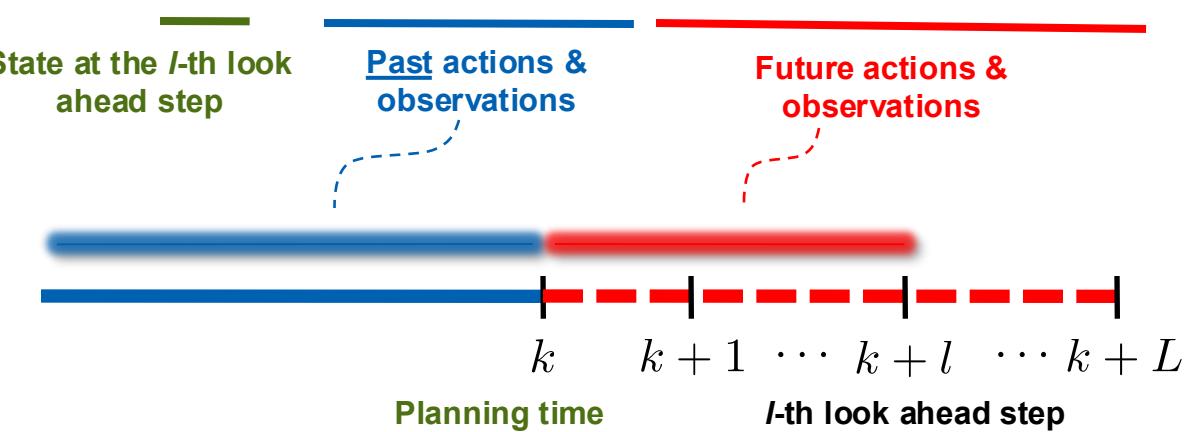
# Partially Observable Markov Decision Process (POMDP)

- Value function

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

Belief-dependent reward function

- Belief at the  $\ell$ -th look-ahead step:  $b_{k+\ell} \triangleq b[X_{k+\ell}] = \mathbb{P}(X_{k+\ell} \mid a_{0:k-1}, z_{0:k}, a_{k:k+\ell-1}, z_{k+1:k+\ell})$



- Examples for reward function  $\rho(b, a)$ :

- Expected distance to goal (navigate to a goal)
- Information theoretic reward (reduce uncertainty)
- ...

# Challenge

## Probabilistic Inference

Maintain a distribution over the state given data

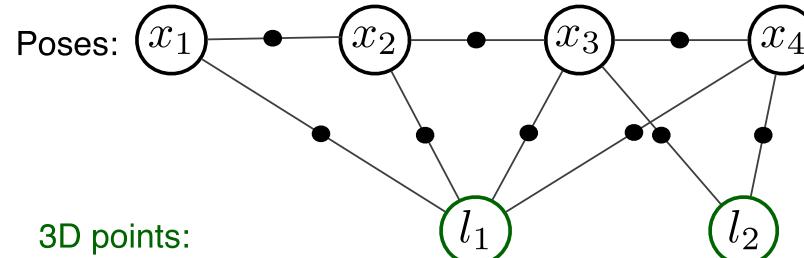
$$b_k \triangleq b[X_k] = \mathbb{P}(X_k \mid \underbrace{a_{0:k-1}}_{\text{state}}, \underbrace{z_{1:k}}_{\text{actions observations}})$$

## Decision-making under uncertainty

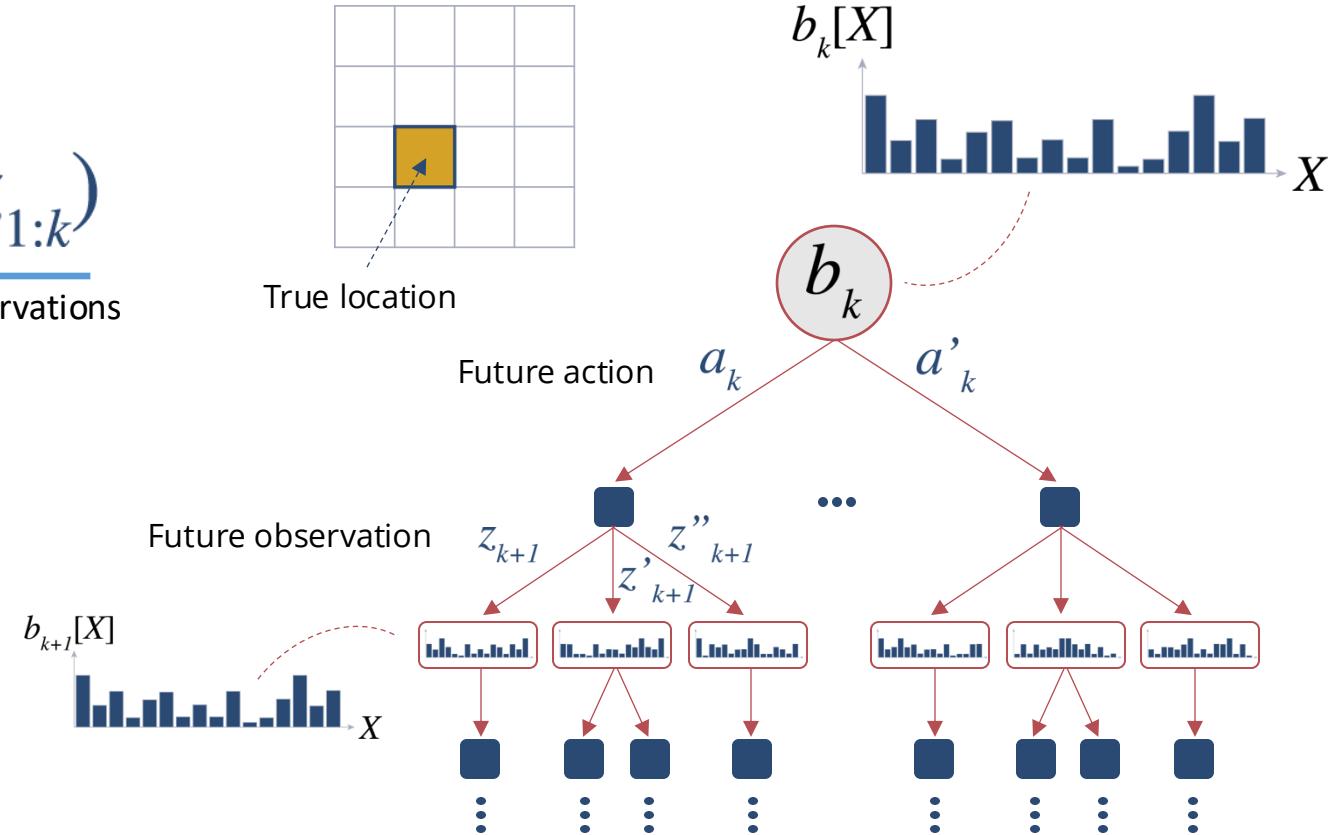
Involves reasoning about the entire observation and action spaces along planning horizon

## Computationally intractable

More so, in high dimensional settings



## Example - grid world



How can we act autonomously online and efficiently complete 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

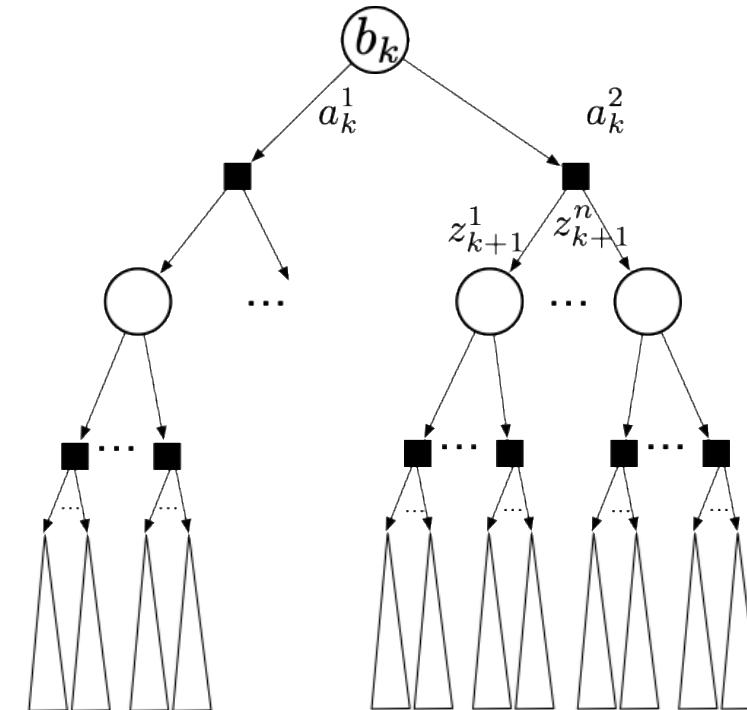
POMDP Planning with Hybrid Beliefs

Simplification of POMDP with Formal Guarantees

Multi-agent POMDP Planning with Inconsistent Beliefs

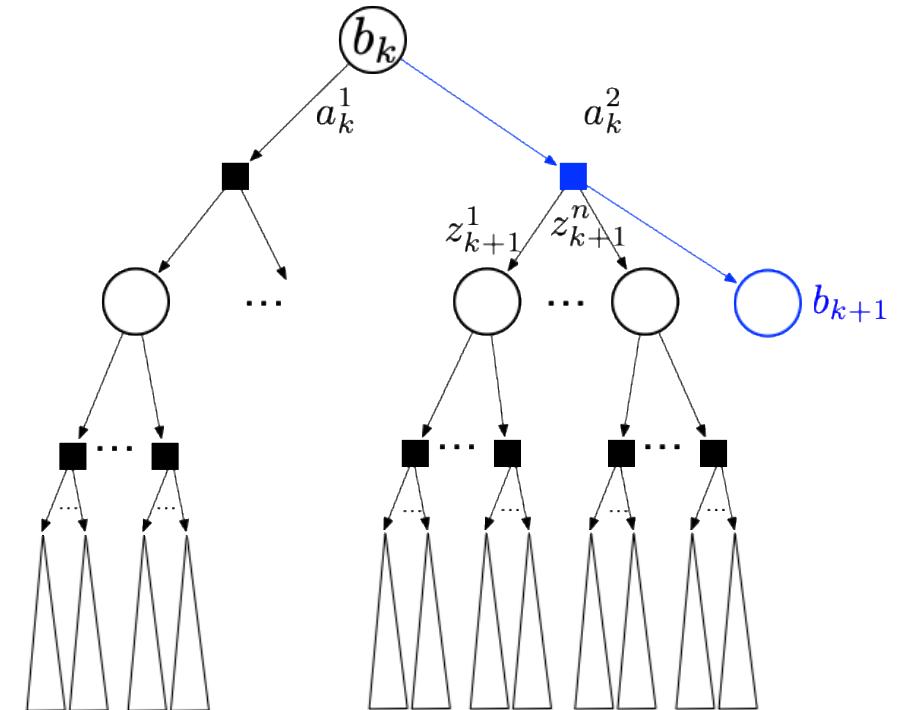
# Experience Reuse in POMDP Planning

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



# Experience Reuse in POMDP Planning

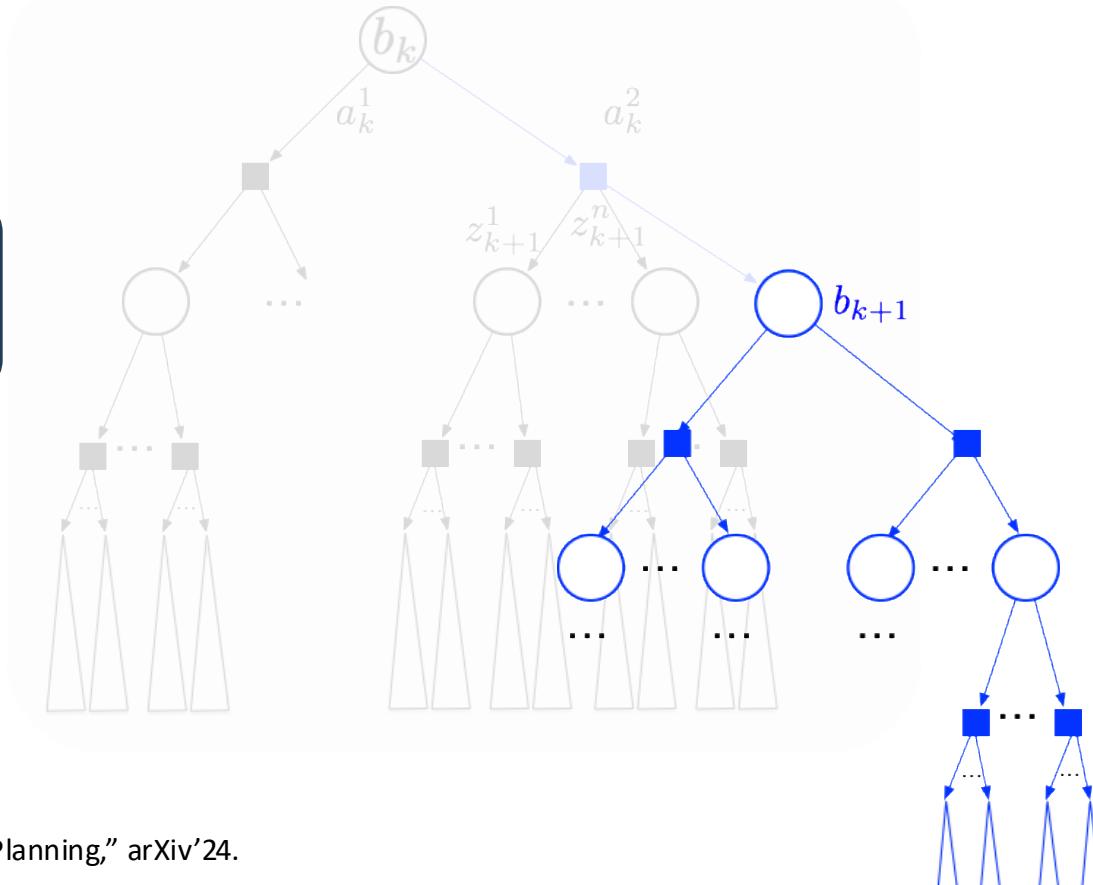
- Consider POMDPs with continuous state, action, and observation spaces
- The probability of sampling the same belief/observation twice is zero



# 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

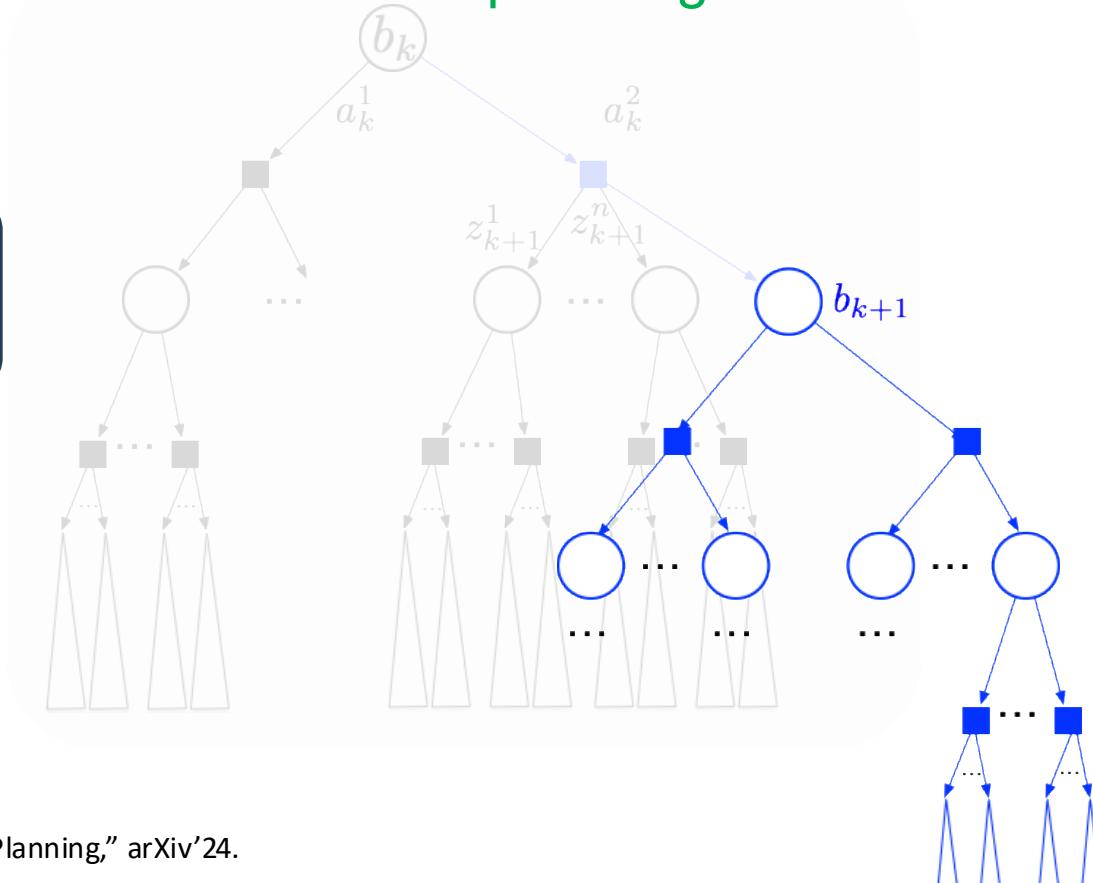
Online SOTA POMDP solvers typically perform calculations from **scratch at each planning session**



# 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**



# 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

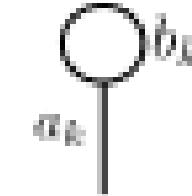
**Key idea:** Reuse previous trajectories/calculations to get an efficient estimation of

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

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

# Experience Reuse in POMDP Planning

- Consider a planning session at time instant k

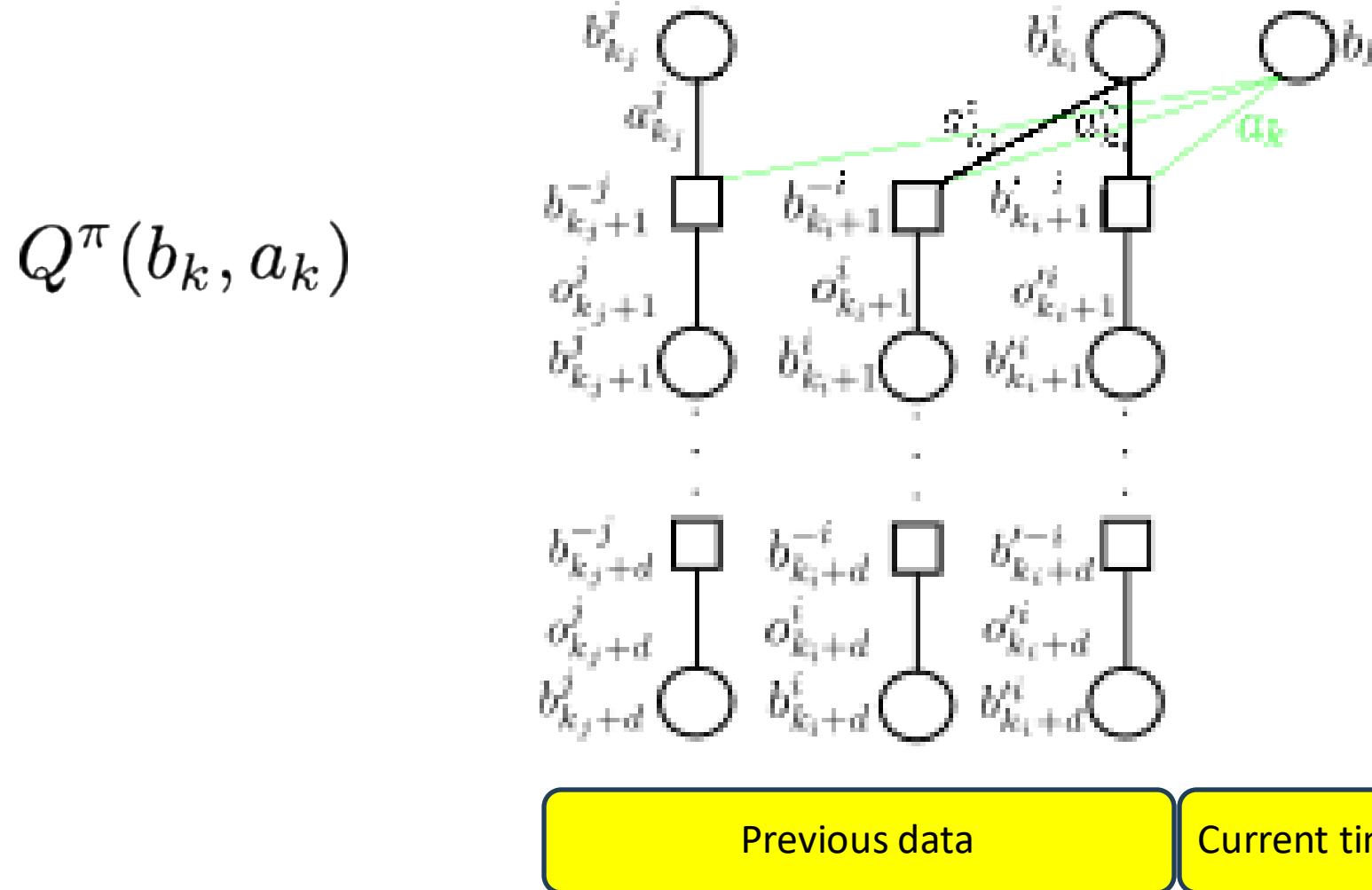


$$Q^\pi(b_k, a_k)$$

Current time

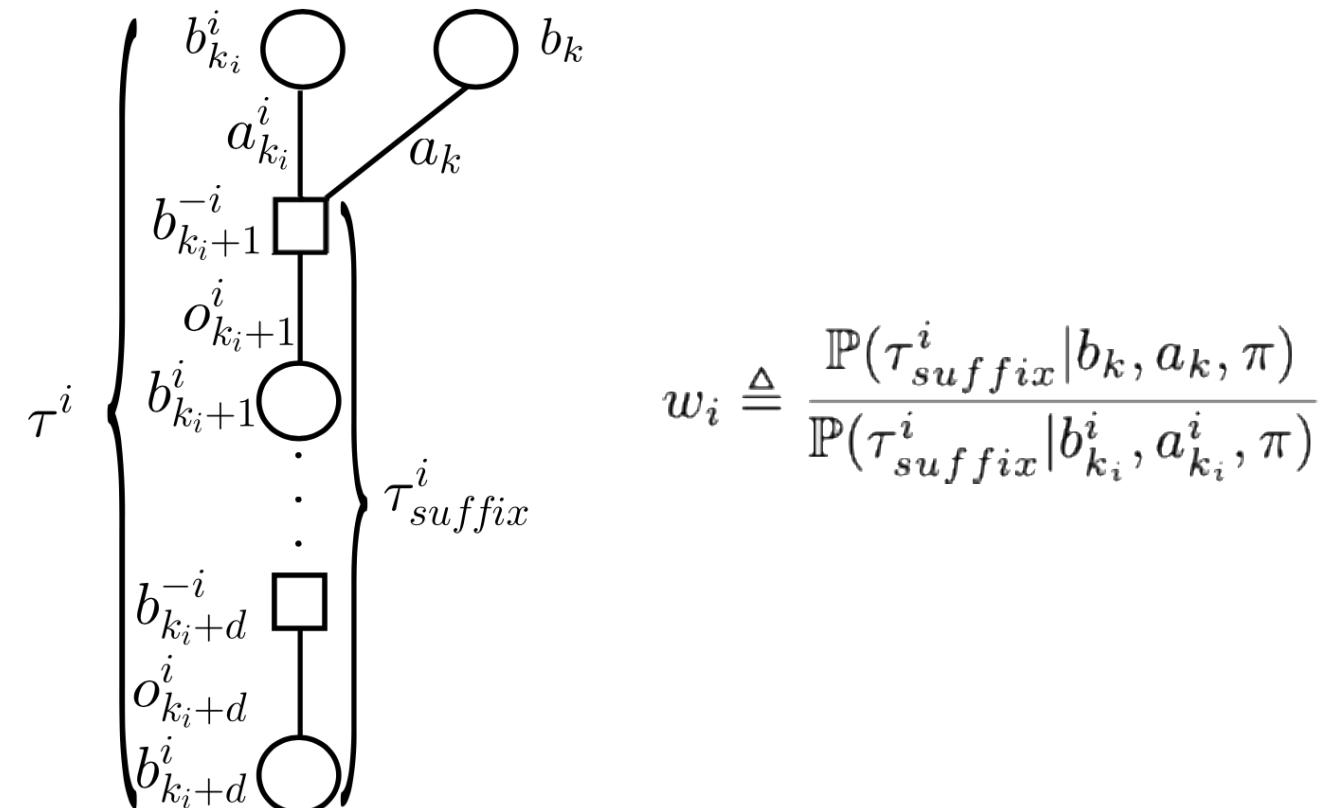
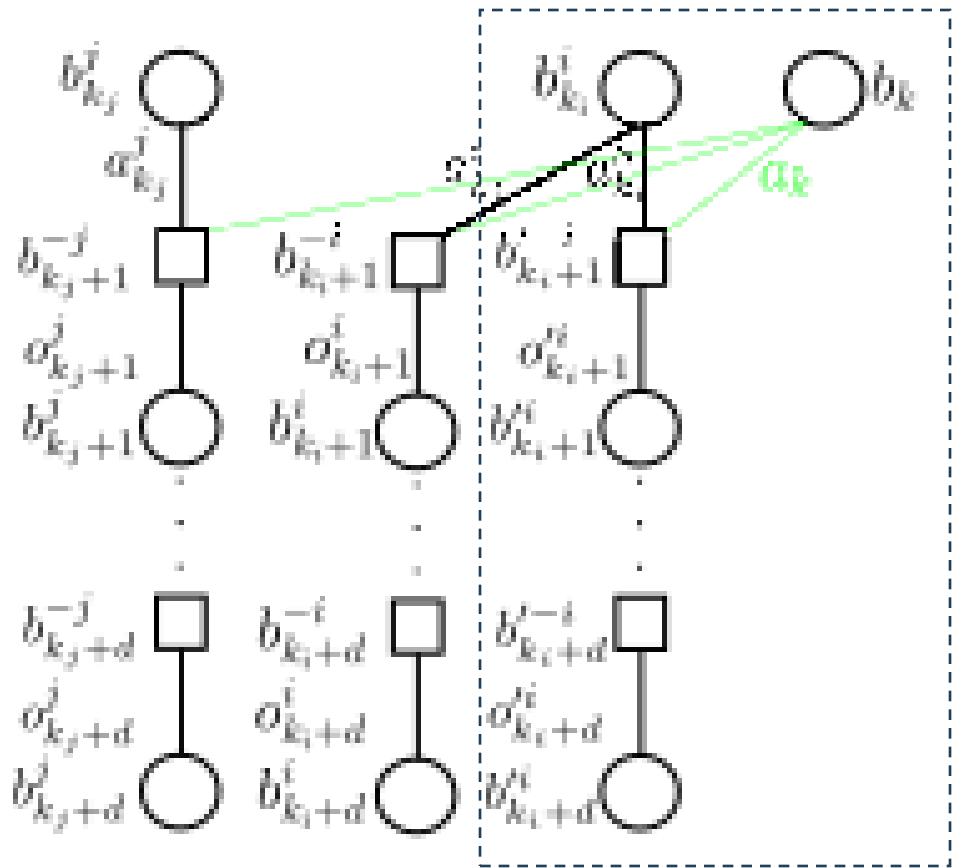
# Experience Reuse in POMDP Planning

- Consider a planning session at time instant  $k$



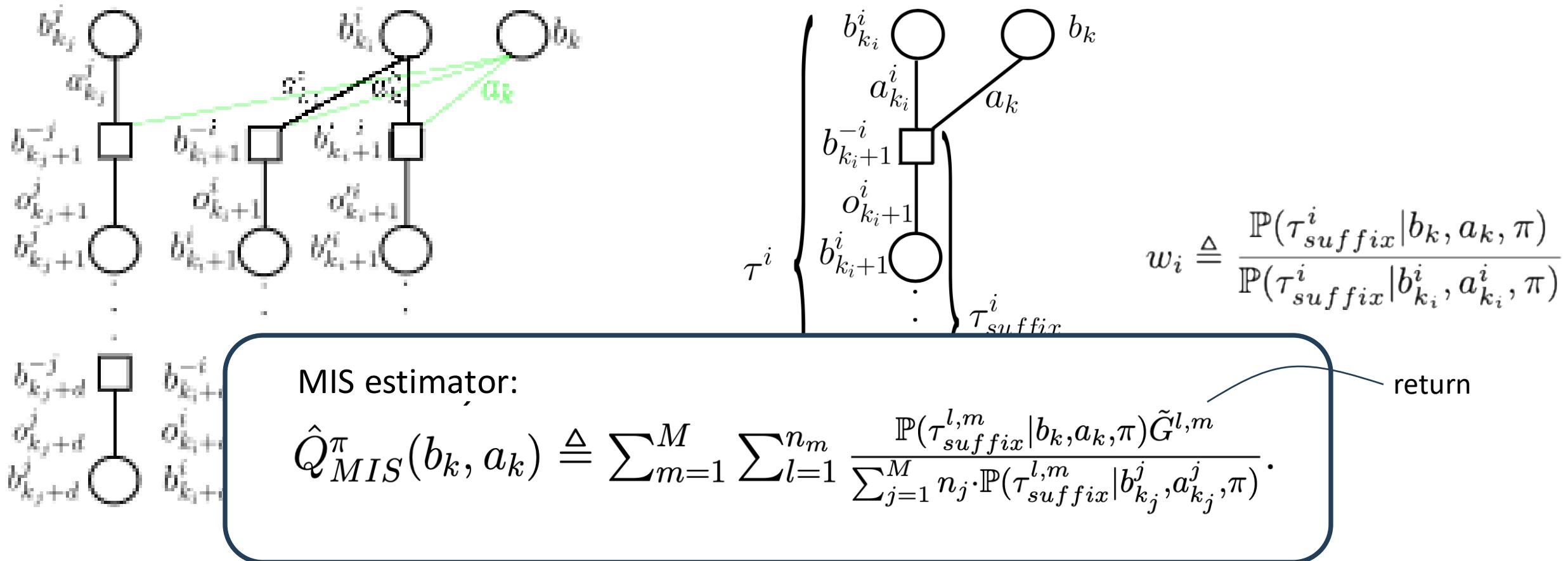
# Experience Reuse in POMDP Planning

- Key idea: multiple importance sampling (MIS) estimator



# Experience Reuse in POMDP Planning

- Key idea: multiple importance sampling (MIS) estimator



# Experience-Based Value Function Estimation

MIS estimator:

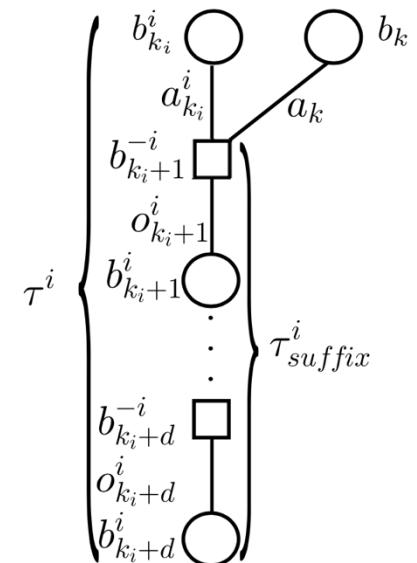
$$\hat{Q}_{MIS}^{\pi}(b_k, a_k) \triangleq \sum_{m=1}^M \sum_{l=1}^{n_m} \frac{\mathbb{P}(\tau_{suffix}^{l,m} | b_k, a_k, \pi) \tilde{G}^{l,m}}{\sum_{j=1}^M n_j \cdot \mathbb{P}(\tau_{suffix}^{l,m} | b_{k_j}^j, a_{k_j}^j, \pi)}.$$

## Theorem 1

$$\frac{\mathbb{P}(\tau_{suffix}^i | b_k, a_k, \pi)}{\mathbb{P}(\tau_{suffix}^i | b_{k_i}^i, a_{k_i}^i, \pi)} = \frac{\mathbb{P}(b_{k_i+1}^{-i} | b_k, a_k)}{\mathbb{P}(b_{k_i+1}^{-i} | b_{k_i}^i, a_{k_i}^i)}$$

## Proof.

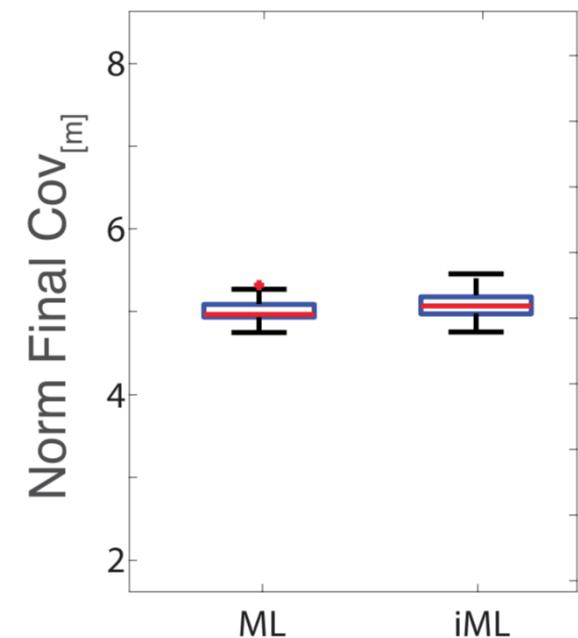
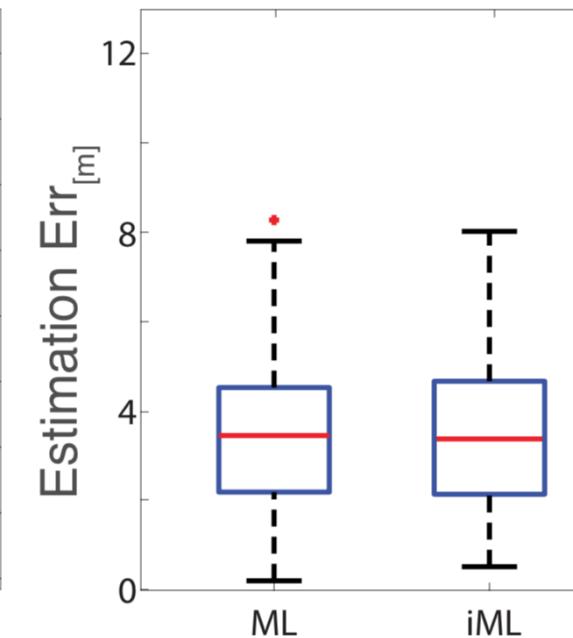
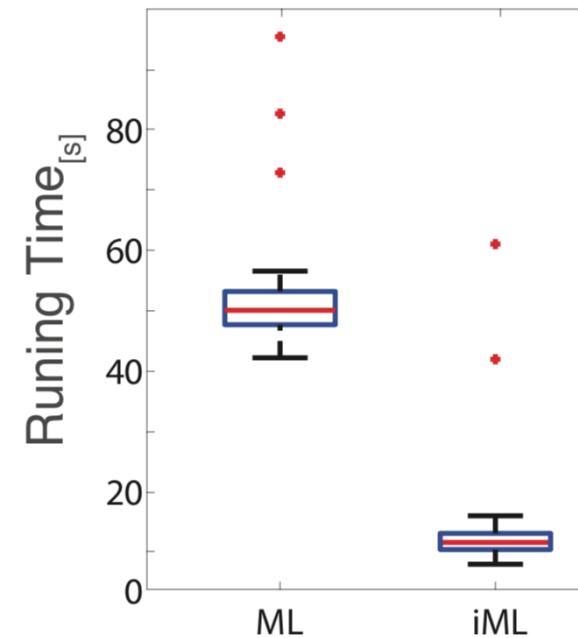
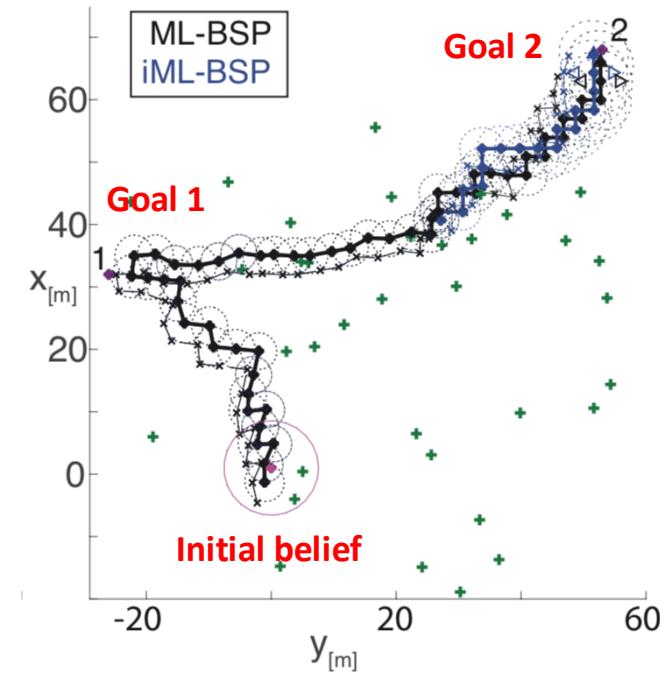
$$\begin{aligned} \frac{\mathbb{P}(\tau_{suffix}^i | b_k, a_k, \pi)}{\mathbb{P}(\tau_{suffix}^i | b_{k_i}^i, a_{k_i}^i, \pi)} &= \frac{\mathbb{P}(b_{k_i+1}^{-i}, o_{k_i+1}^i, \dots, b_{k_i+L}^i | b_k, a_k, \pi)}{\mathbb{P}(b_{k_i+1}^{-i}, o_{k_i+1}^i, \dots, b_{k_i+L}^i | b_{k_i}^i, a_{k_i}^i, \pi)} = \\ \frac{\mathbb{P}(b_{k_i+1}^{-i} | b_k, a_k)}{\mathbb{P}(b_{k_i+1}^{-i} | b_{k_i}^i, a_{k_i}^i)} \cdot \frac{\mathbb{P}(o_{k_i+1}^i, \dots, b_{k_i+L}^i | b_{k_i+1}^{-i}, \pi)}{\cancel{\mathbb{P}(o_{k_i+1}^i, \dots, b_{k_i+L}^i | b_{k_i+1}^{-i}, \pi)}} &= \frac{\mathbb{P}(b_{k_i+1}^{-i} | b_k, a_k)}{\mathbb{P}(b_{k_i+1}^{-i} | b_{k_i}^i, a_{k_i}^i)} \quad \square \end{aligned}$$



# Incremental Belief Space Planning

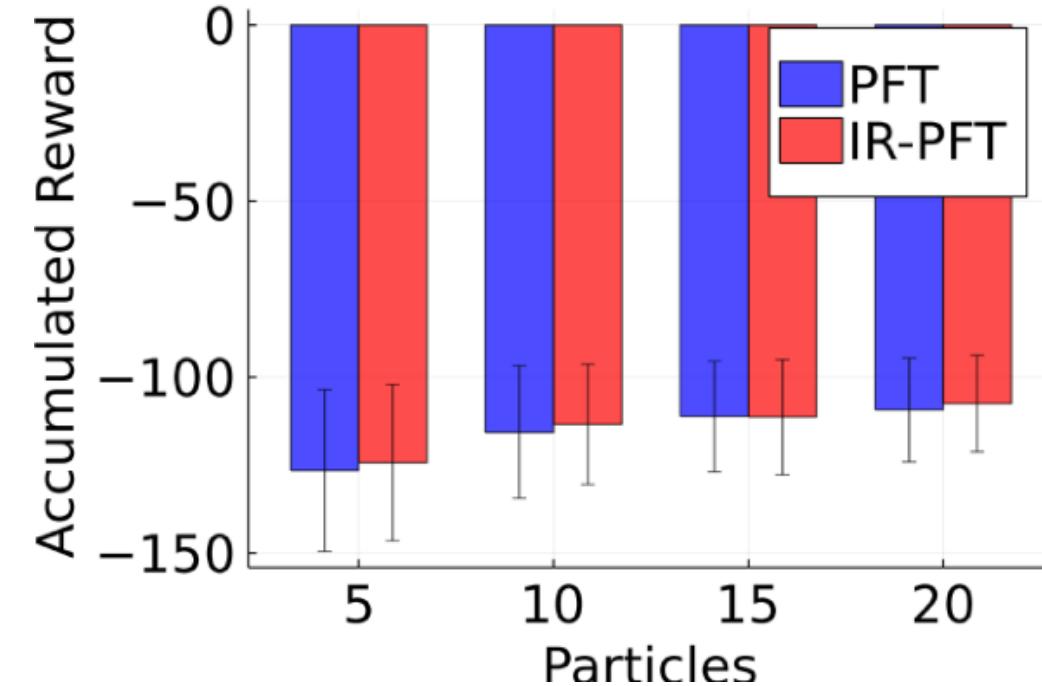
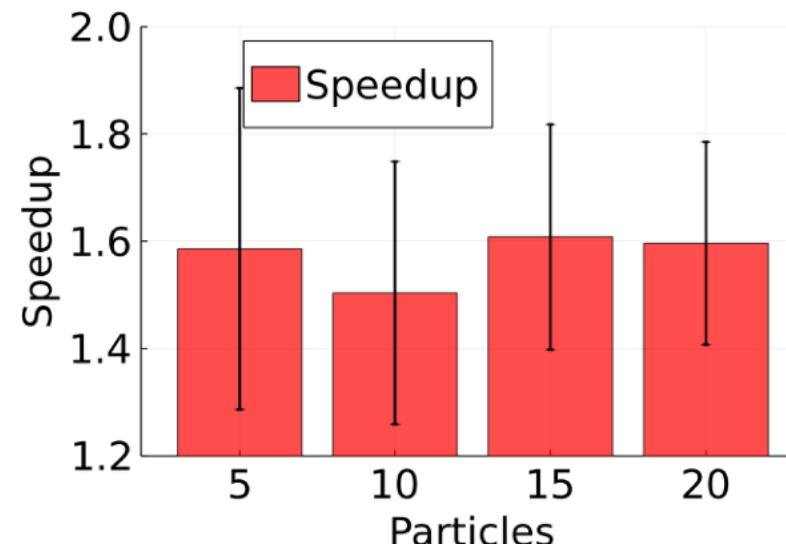
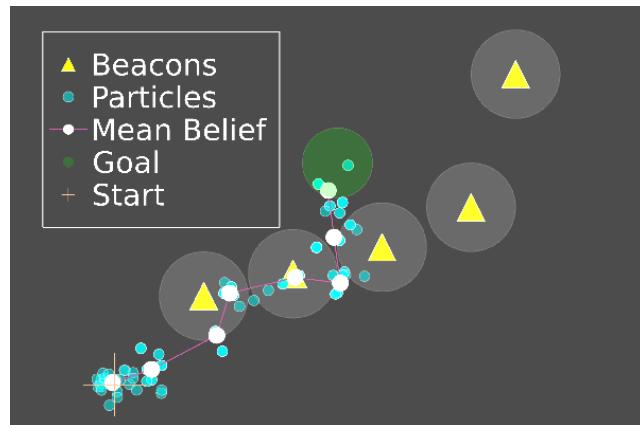
Basic simulation – autonomous navigation in unknown environments:

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



# Incremental Reuse Particle Filter Tree (IR-PFT)

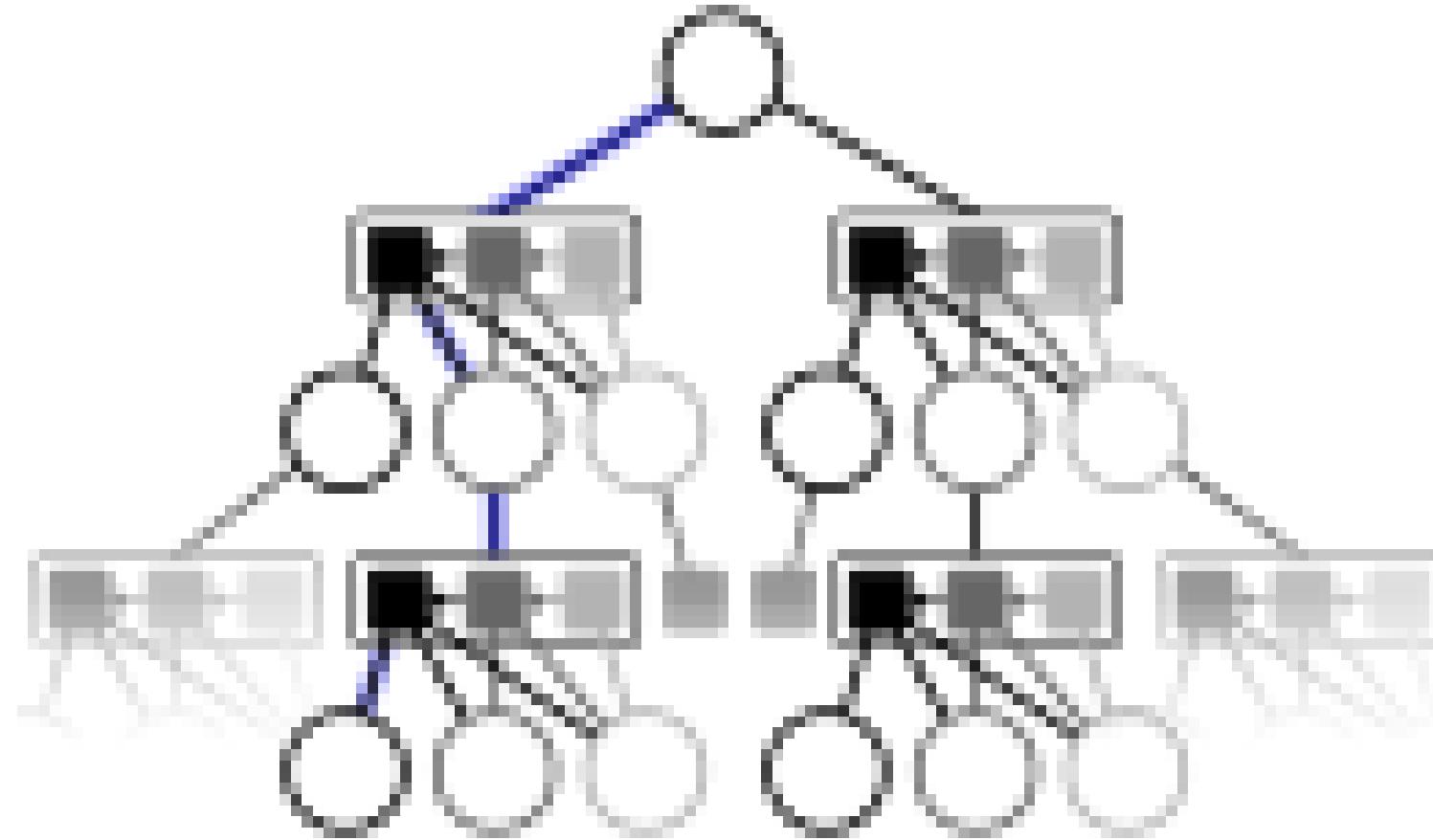
- Extend PFT-DPW<sup>1</sup>, incorporating trajectories from previous planning sessions for fast estimation of  $Q(b_k, a_k)$



<sup>1</sup>Z. Sunberg and M. Kochenderfer. "Online algorithms for POMDPs with continuous state, action, and observation spaces." ICAPS, 2018.

# Value Gradients with Action Adaptive Search Trees in Continuous (PO)MDPs

---



# Agenda

---

Experience Reuse in POMDP Planning

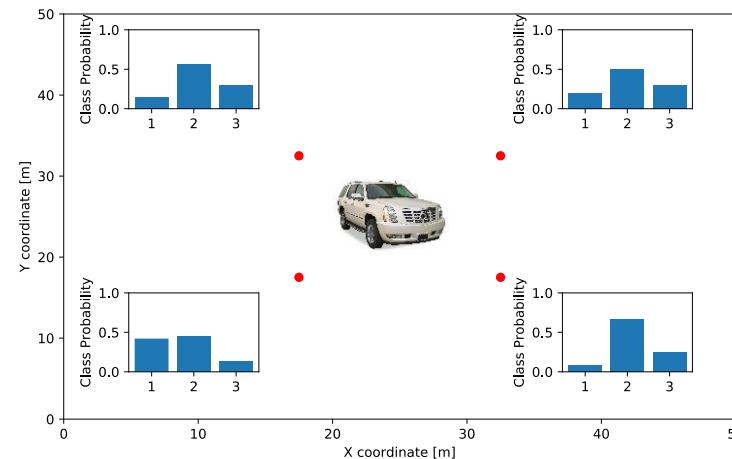
POMDP Planning with Hybrid Beliefs

Simplification of POMDP with Formal Guarantees

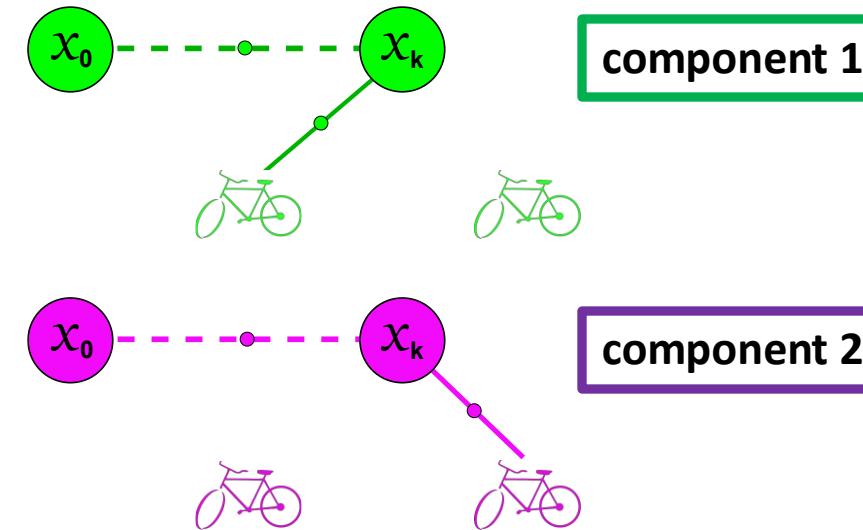
Multi-agent POMDP Planning with Inconsistent Beliefs

# Autonomous Semantic Perception & Ambiguous Environments

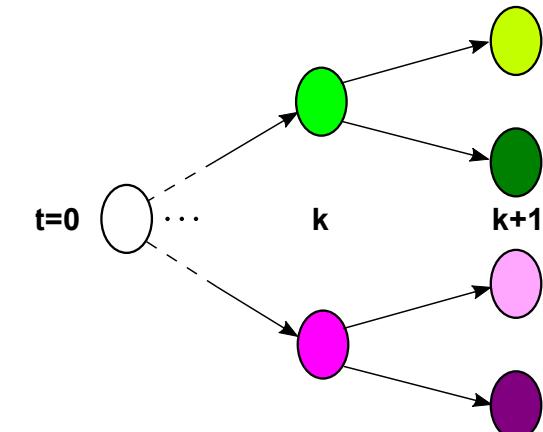
Viewpoint dependent semantic models



Data association hypotheses

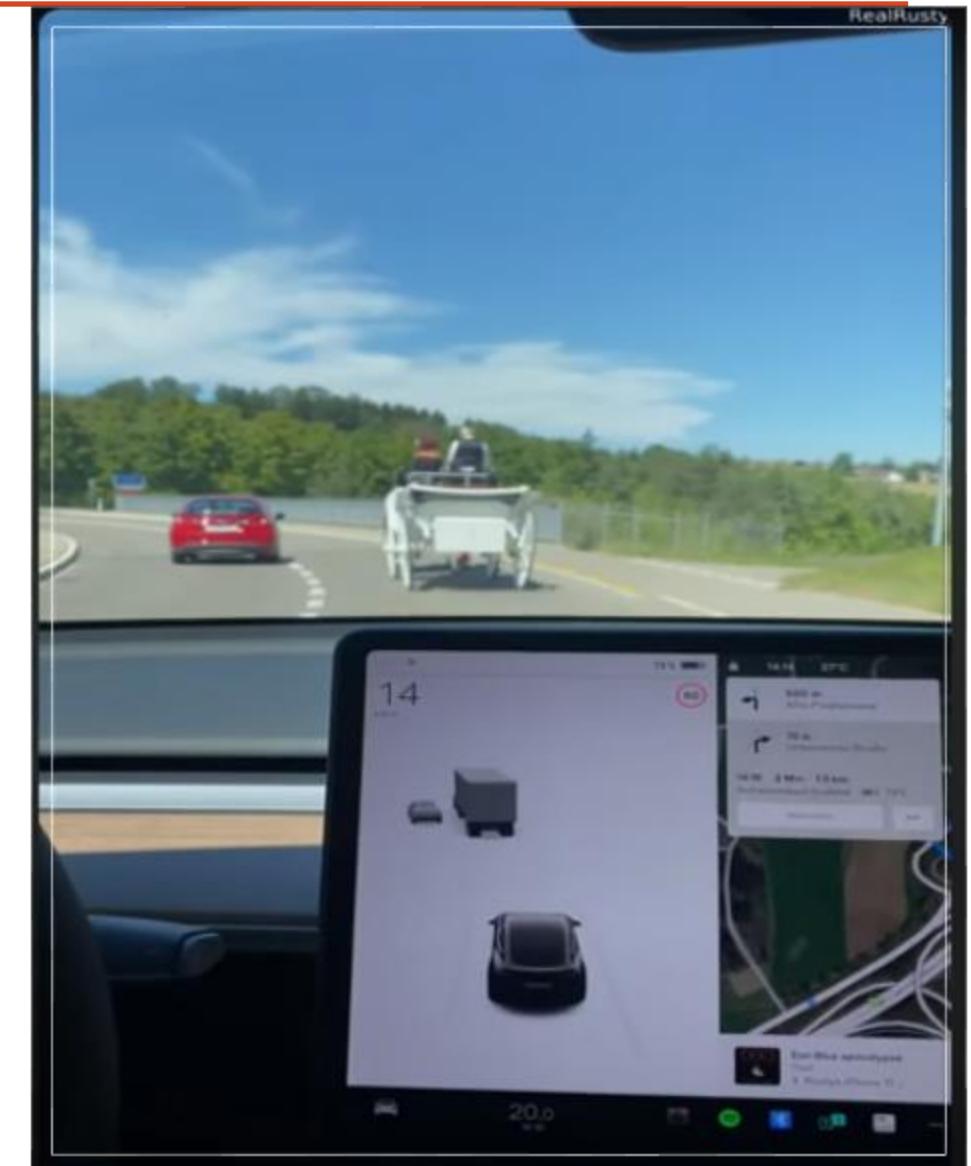


- Hybrid beliefs (over continuous and discrete RVs)
- The number of hypotheses can grow exponentially
- How do we do probabilistic inference and POMDP planning?



# Semantic Perception & SLAM

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



# Class- and Viewpoint-Dependency

---

- Is it a floor or a roof?
- Depending on the viewpoint of the viewer!
  - Looking on the people below - it's a floor
  - Looking on the people above - it's a roof
- How do we know the viewpoint?

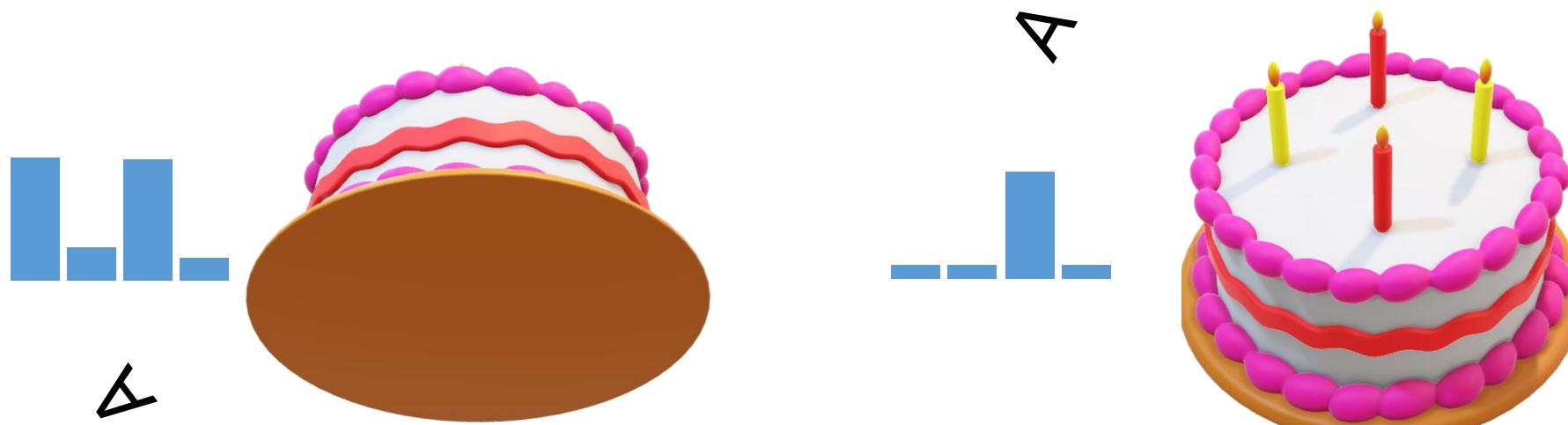


<http://www.puzzlesworld.com>

# Class- and Viewpoint-Dependency

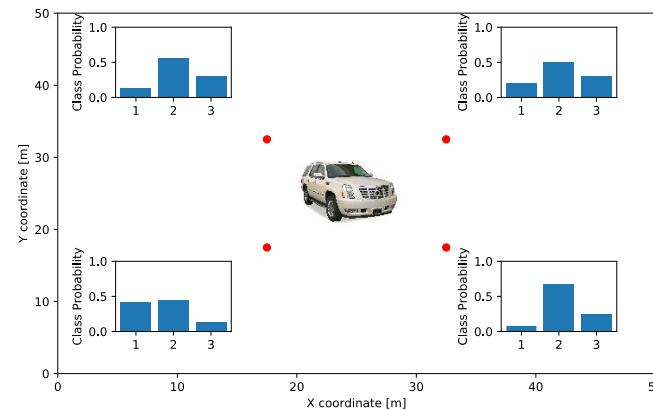
---

- Another example:



# Coupled Models

- View-dependent semantic observation model:



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

Semantic observation  
(from a classifier)

Object class

Agent's viewpoint  
relative to object

Find the cake.



- Class and poses can be coupled via learned prior probabilities.
- Reward/constraint can depend on both classes and poses (e.g., object search)

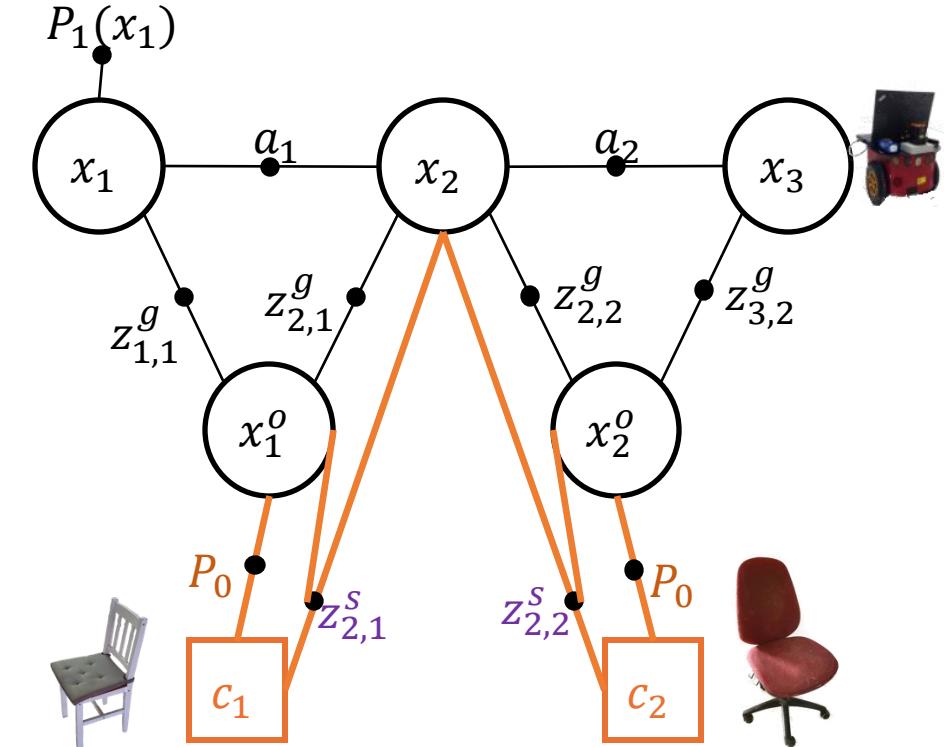
# Hybrid Belief

- **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 dependent
- Classes of different objects are dependent
- 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. Tchuiiev, Y. Feldman, and V. Indelman, "Data Association Aware Semantic Mapping and Localization via a Viewpoint Dependent Classifier Model," IROS'19.

V. Tchuiiev 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

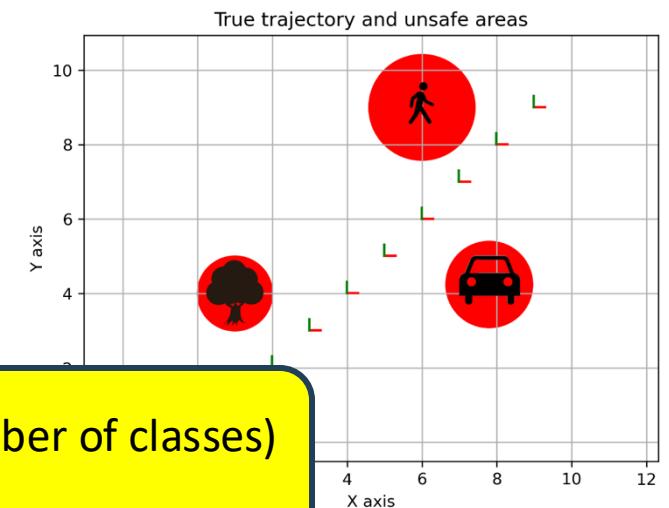
$$V^\pi(b_k) = \mathbb{E}_{z_{k+1:k+L}} \left[ \sum_{l=k}^{k+L-1} \rho(b_l, \pi_l(b_l), b_{l+1}) \right]$$

- Semantic Risk Awareness

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

Objects' classes      Objects' poses

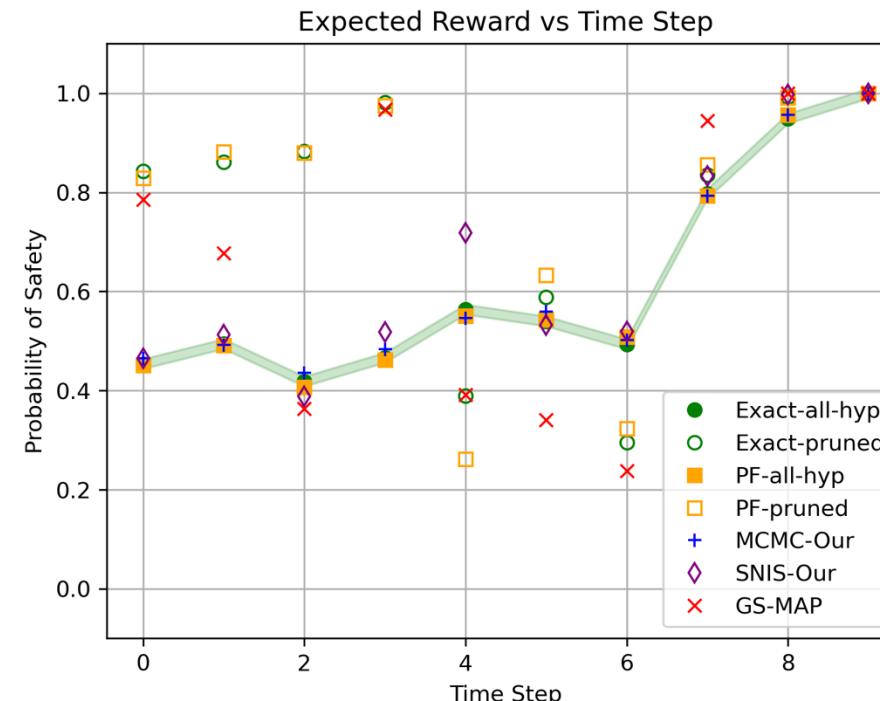
The number of classification hypotheses is  $M^N$  (N: number of objects, M: number of classes)  
 How to sample w/o pruning hypotheses? How to estimate  $\mathbb{P}_{safe}$ ?



# POMDP Planning with Hybrid Semantic-Geometric Beliefs

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

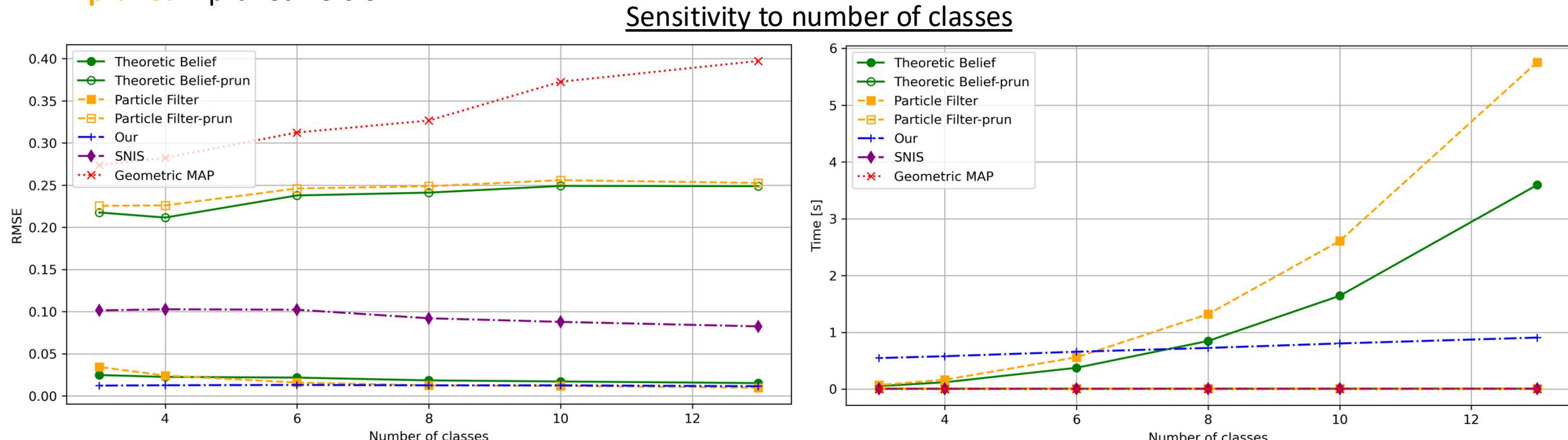
- **Exact-all-hyp** – belief computed exactly
  - **Exact-pruned** – pruned version
  - **PF-all-hyp** – Particle filter
  - **PF-pruned** – pruned version
- Our methods
- **MCMC-Our** – MCMC samples
  - **SNIS-Our** – self-normalized importance sampling
  - **GS-MAP** – separate semantic and geometric



# 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



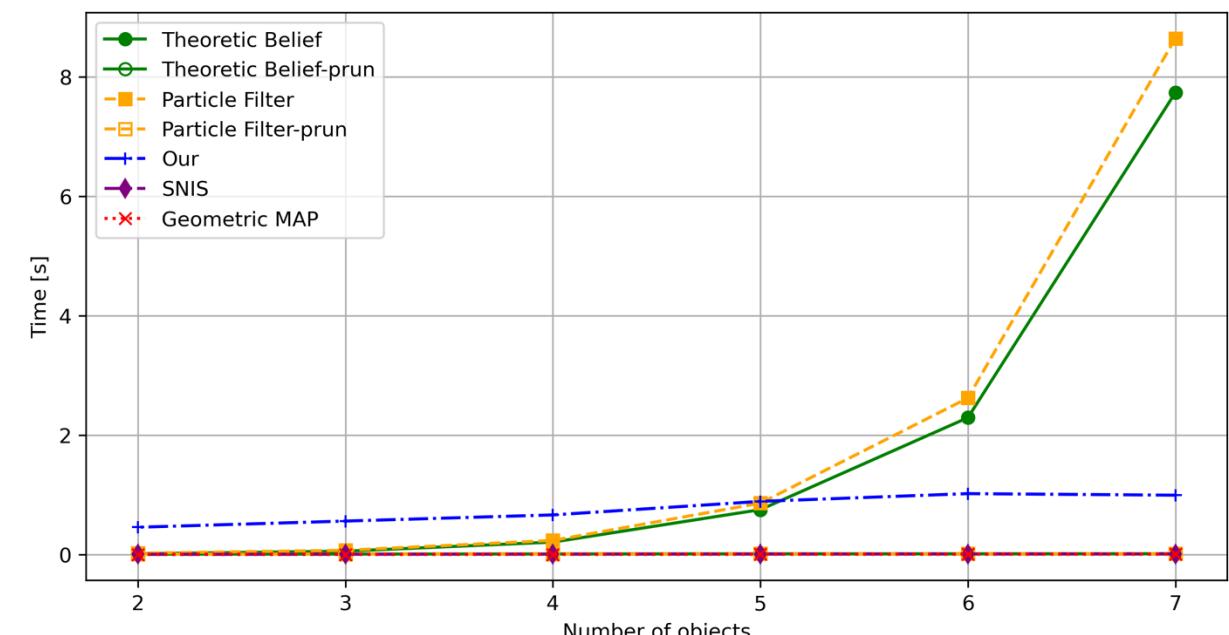
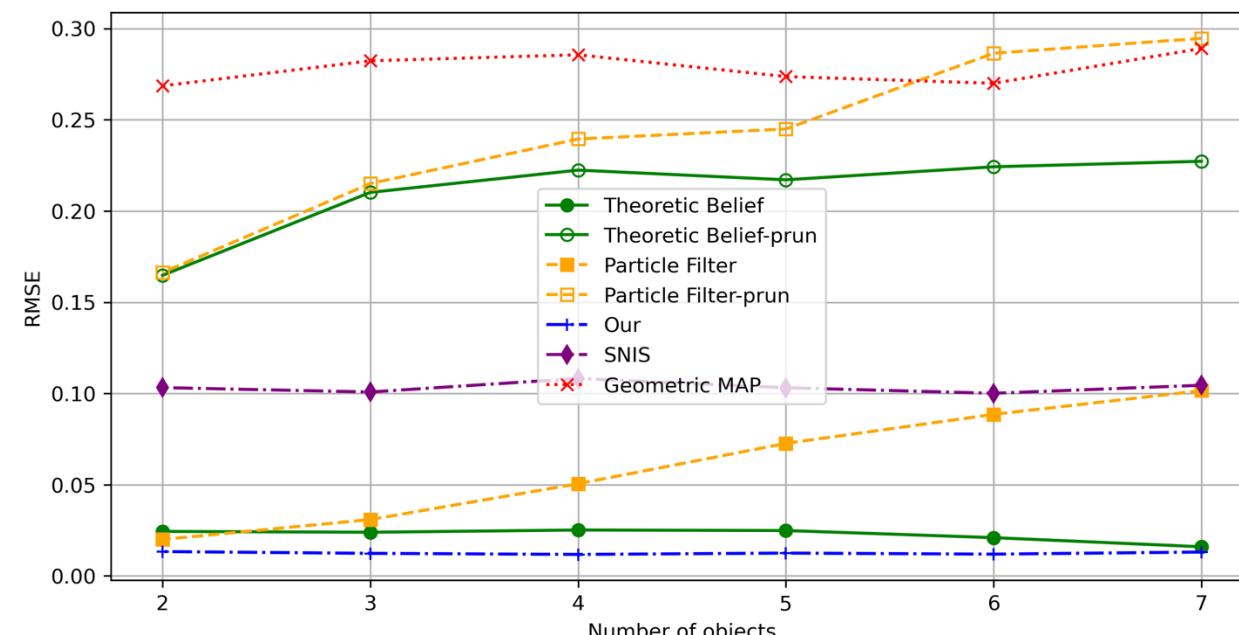
# POMDP Planning with Hybrid Semantic-Geometric Beliefs

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

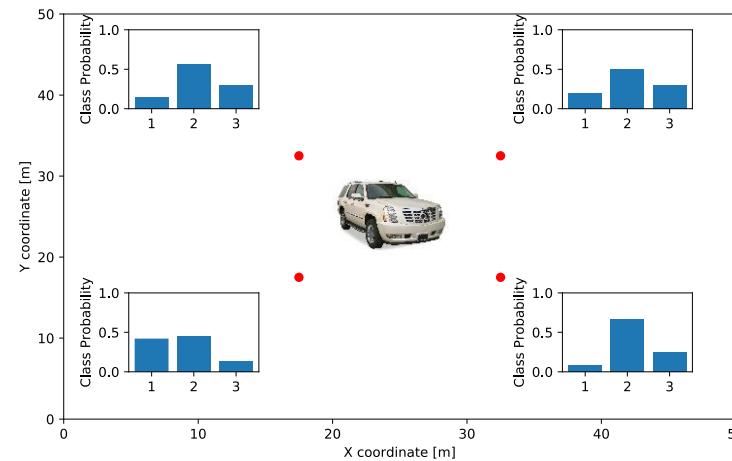
Our methods

Sensitivity to number of objects

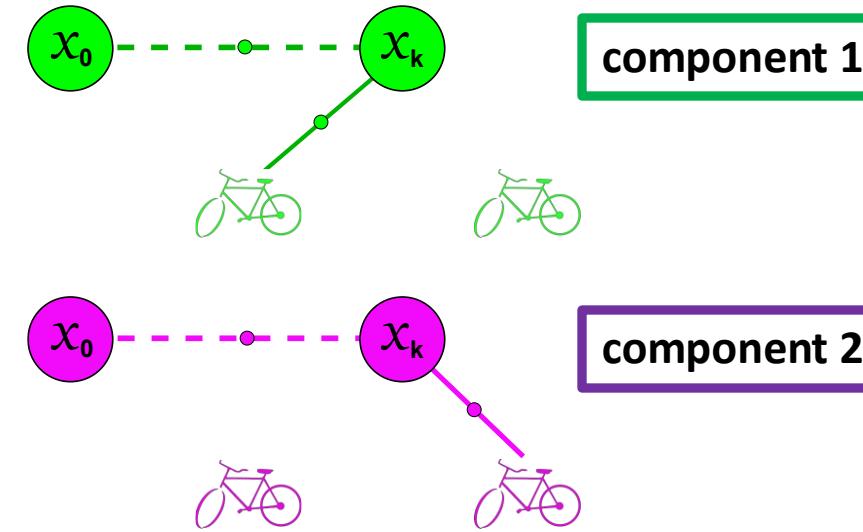


# Autonomous Semantic Perception & Ambiguous Environments

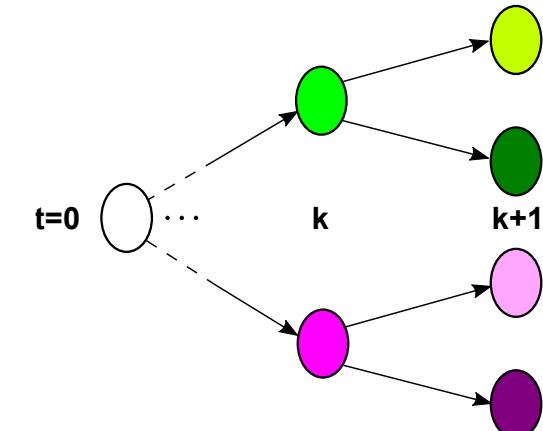
Viewpoint dependent semantic models



Data association hypotheses



- Hybrid beliefs (over continuous and discrete RVs)
- The number of hypotheses can grow exponentially
- How do we do probabilistic inference and POMDP planning?

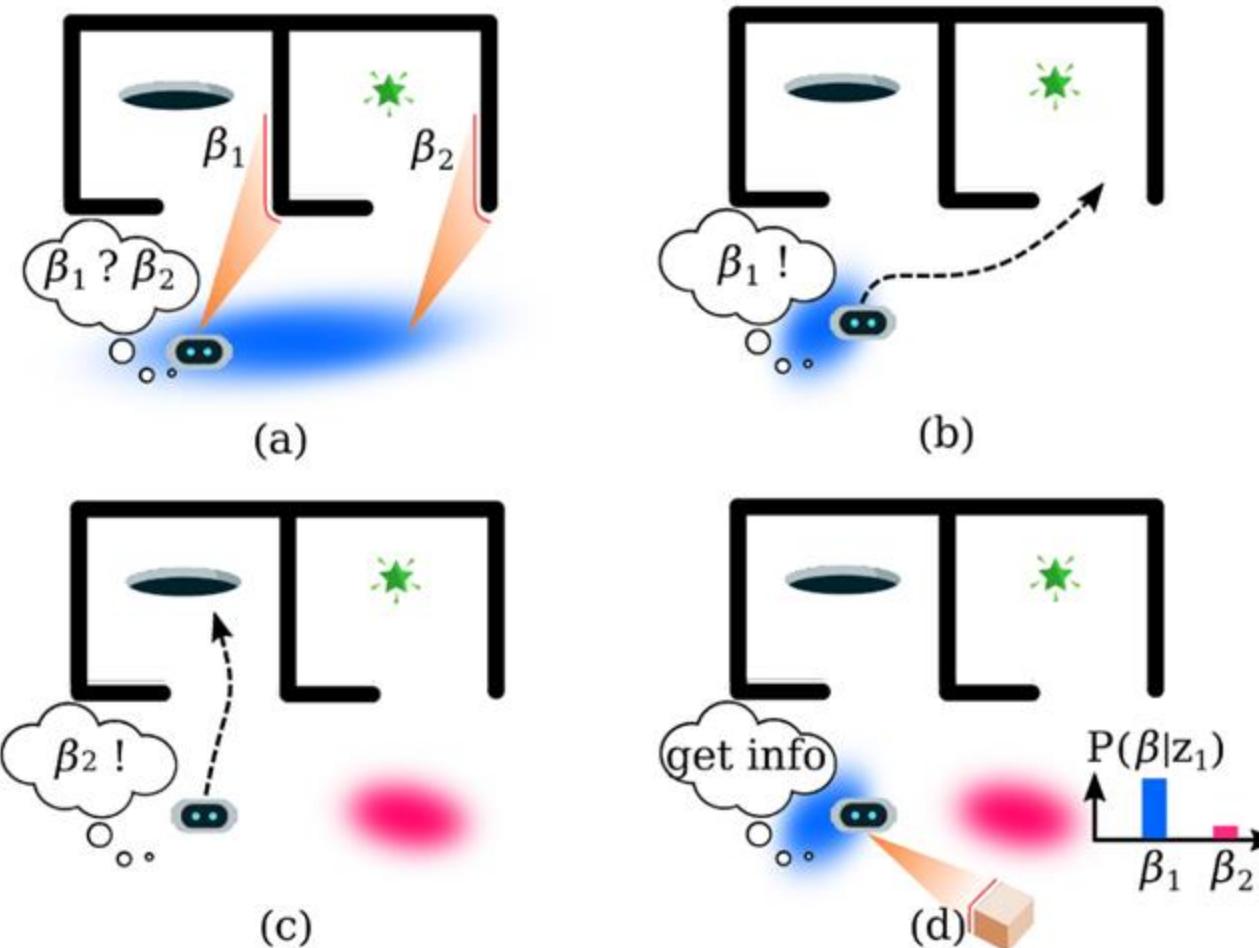


# Ambiguous Scenarios

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

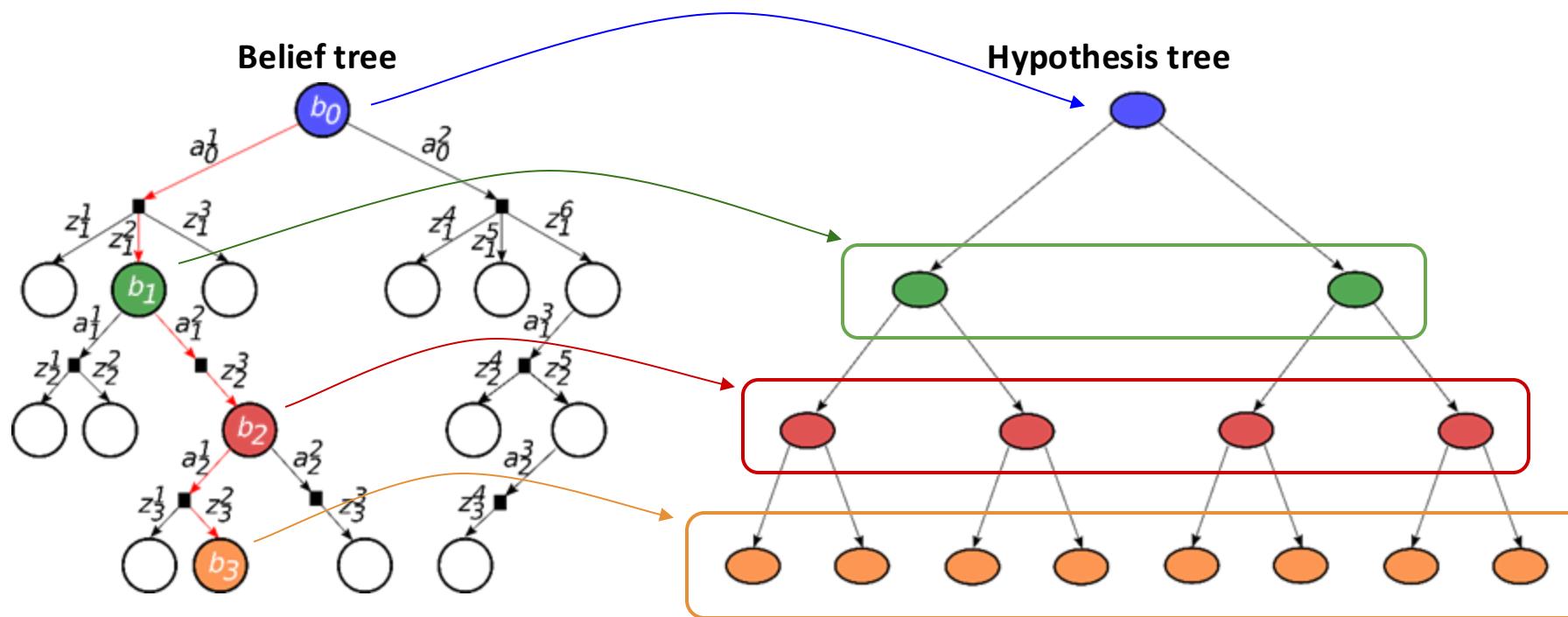
An observation:  
(e.g. LIDAR)

How should the agent act?



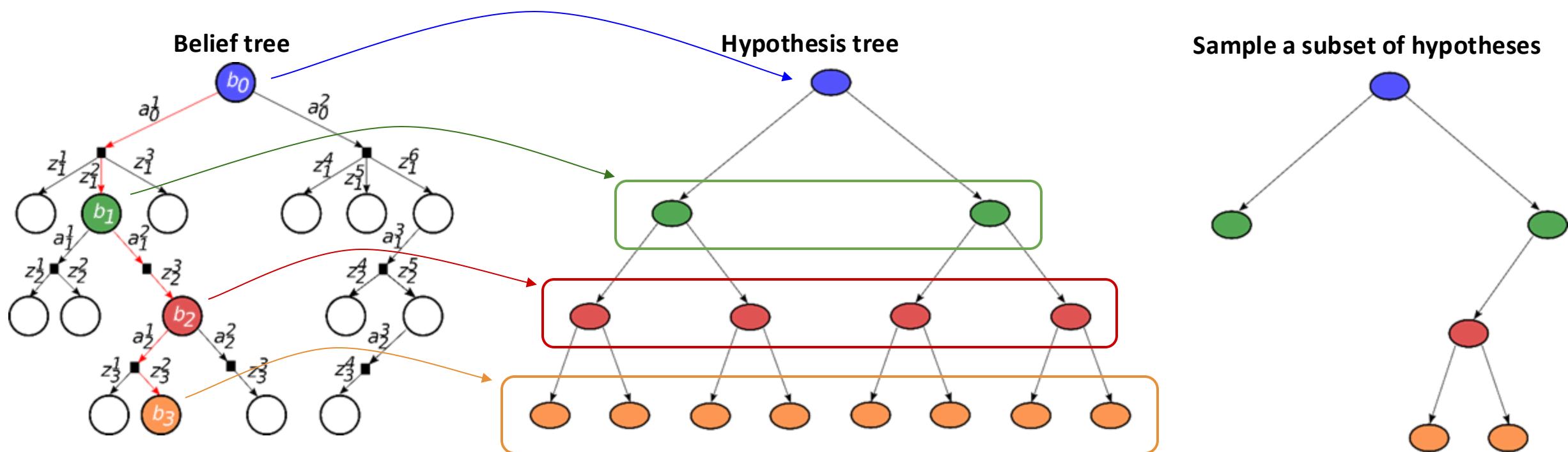
# Continuous-Discrete State Spaces - the Challenge

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



# Continuous-Discrete State Spaces - the Challenge

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



Impact on decision making?

# Agenda

---

Experience Reuse in POMDP Planning

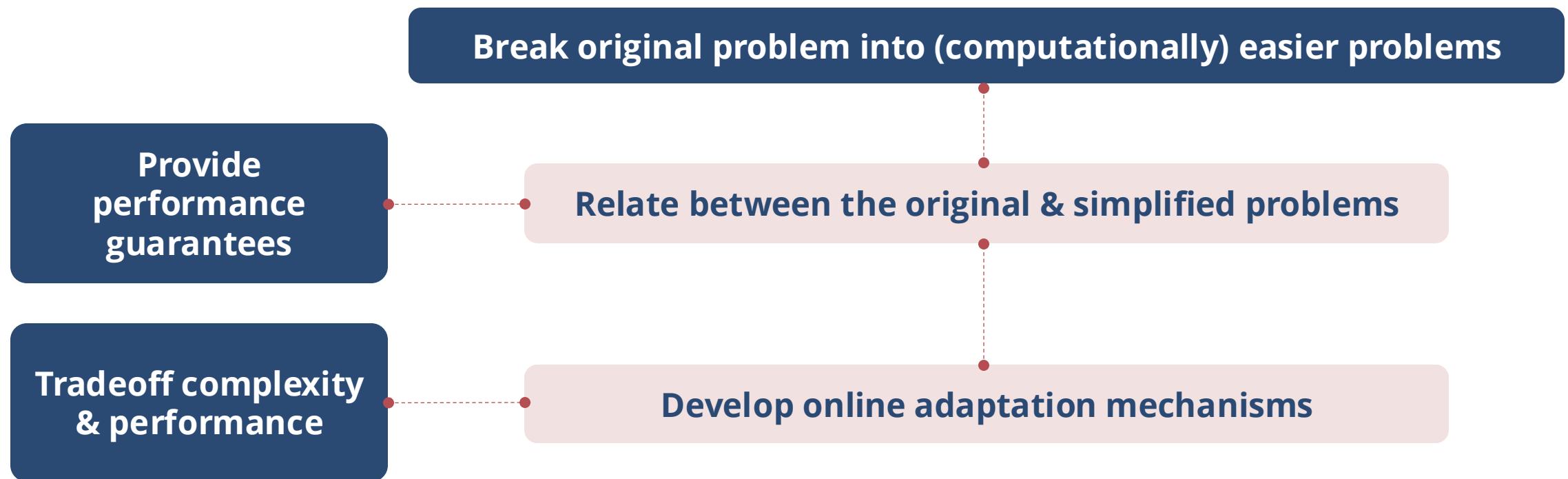
POMDP Planning with Hybrid Beliefs

Simplification of POMDP with Formal Guarantees

Multi-agent POMDP Planning with Inconsistent Beliefs

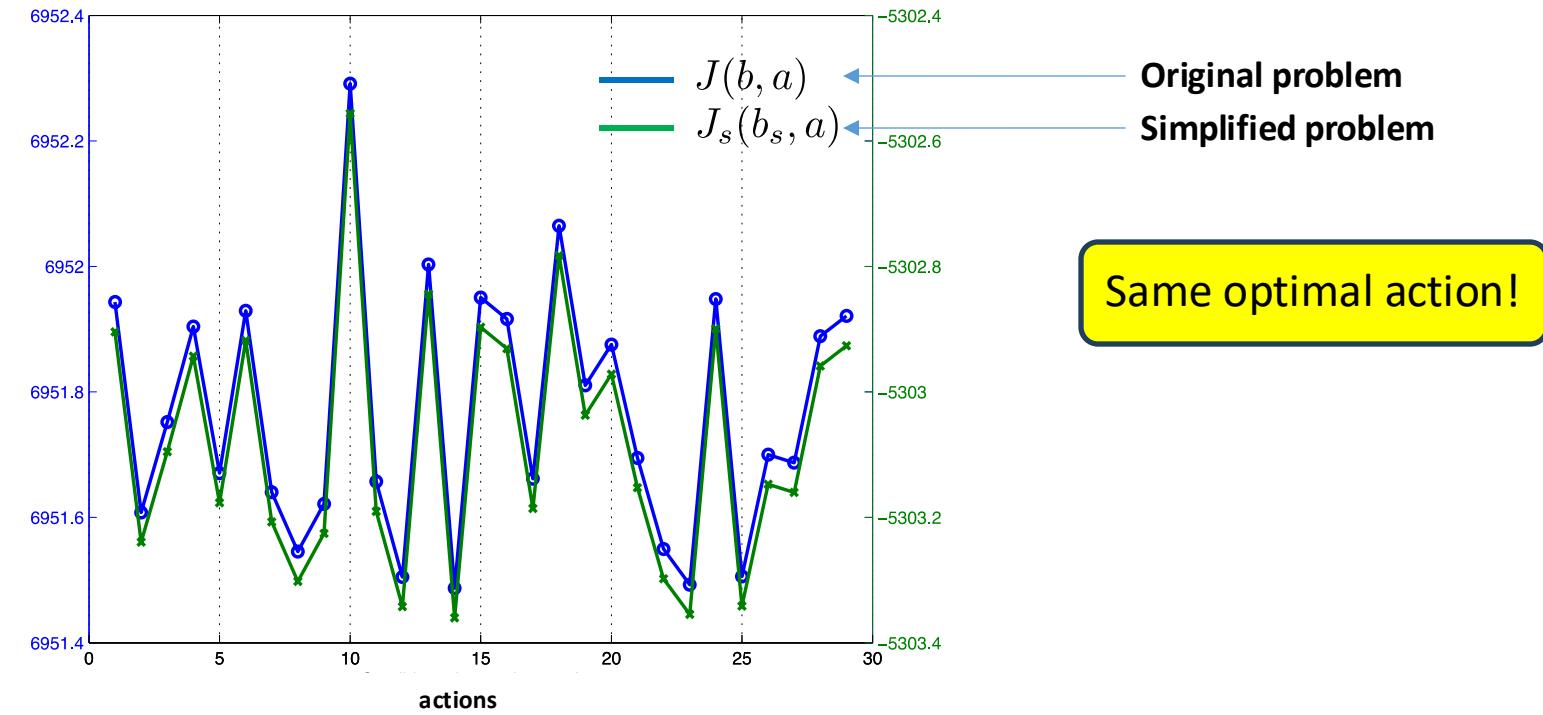
# Simplification Framework

**Accelerate decision making** by adaptive simplification while providing performance guarantees



# Simplification of Decision-Making Problems

- Each element of the decision-making problem can be simplified
- **Action-consistent simplification preserves order** between actions w.r.t. original problem



# Simplification of Decision-Making Problems

---

$$\mathcal{LB}(b, a) \leq Q(b, a) \leq \mathcal{UB}(b, a)$$

Computationally cheap(er)  
bounds

# Simplification of Decision-Making Problems

---

## Concept:

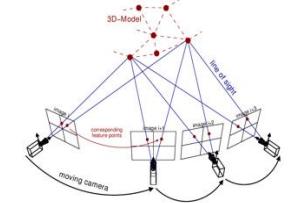
- Identify and solve a **simplified (computationally) easier** decision-making problem
- Provide performance guarantees

## 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

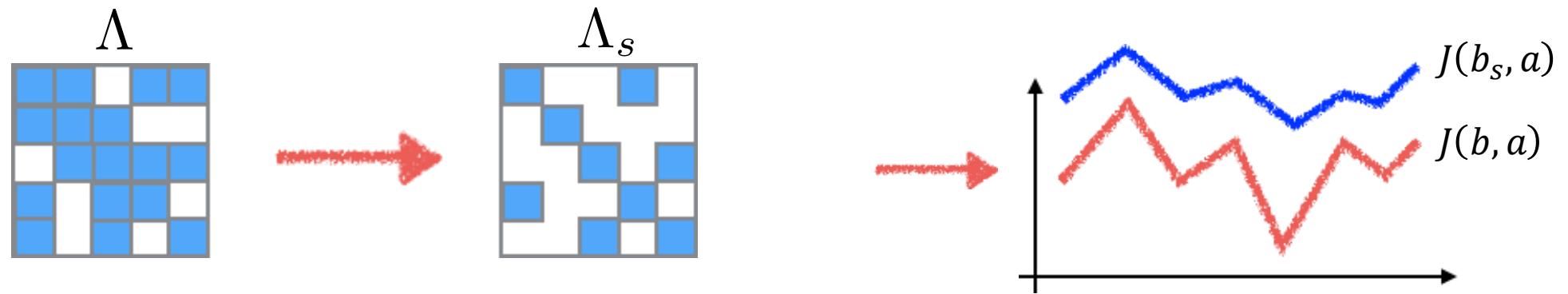
# Belief Sparsification for Gaussian BSP

- Find an appropriate **sparsified** (square root) information matrix
- Perform decision making using that, rather the original, information matrix



**Setting:**

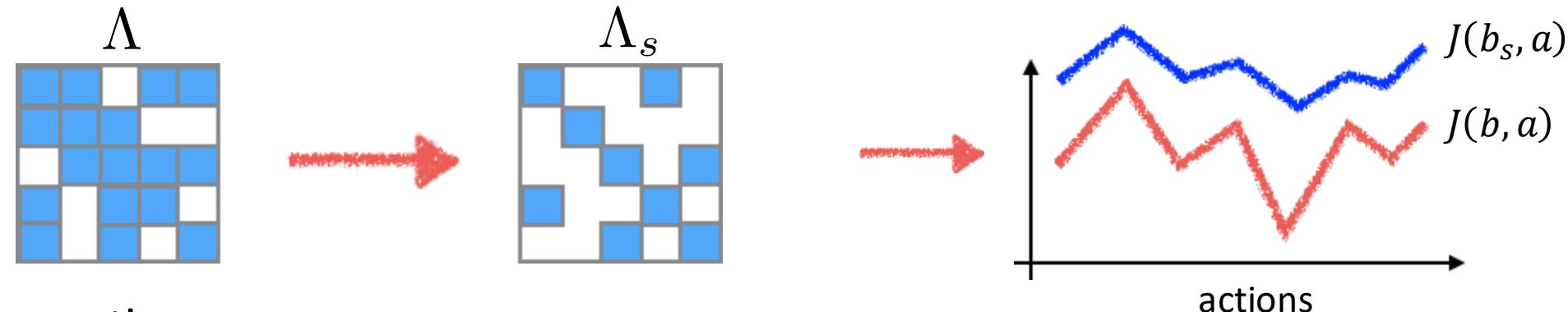
- Gaussian belief over high dim. state  $X \in \mathbb{R}^n$ :  $b[X] = \mathcal{N}(X^\star, \Lambda^{-1}) = \mathcal{N}(X^\star, (R^T R)^{-1})$
- Information-theoretic reward (entropy):  $H[X] = \frac{1}{2} \log((2\pi e)^n |\Lambda|^{-1})$



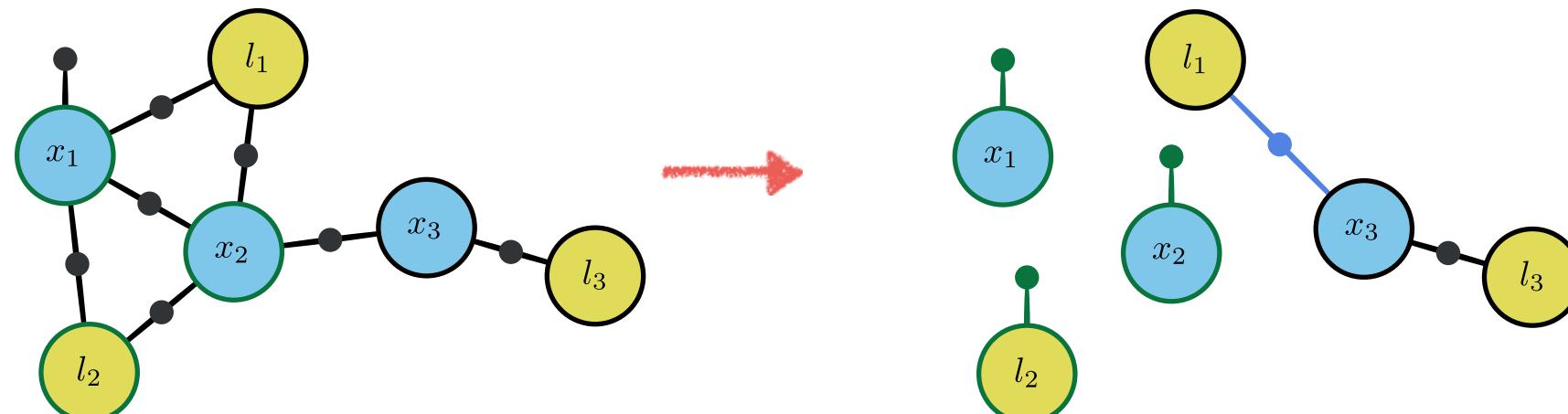
- Do we get the same performance (decisions), i.e. is it action consistent?

# Belief Sparsification for Gaussian BSP

- Sparsification of (square root) information matrix



- Graphical models perspective:

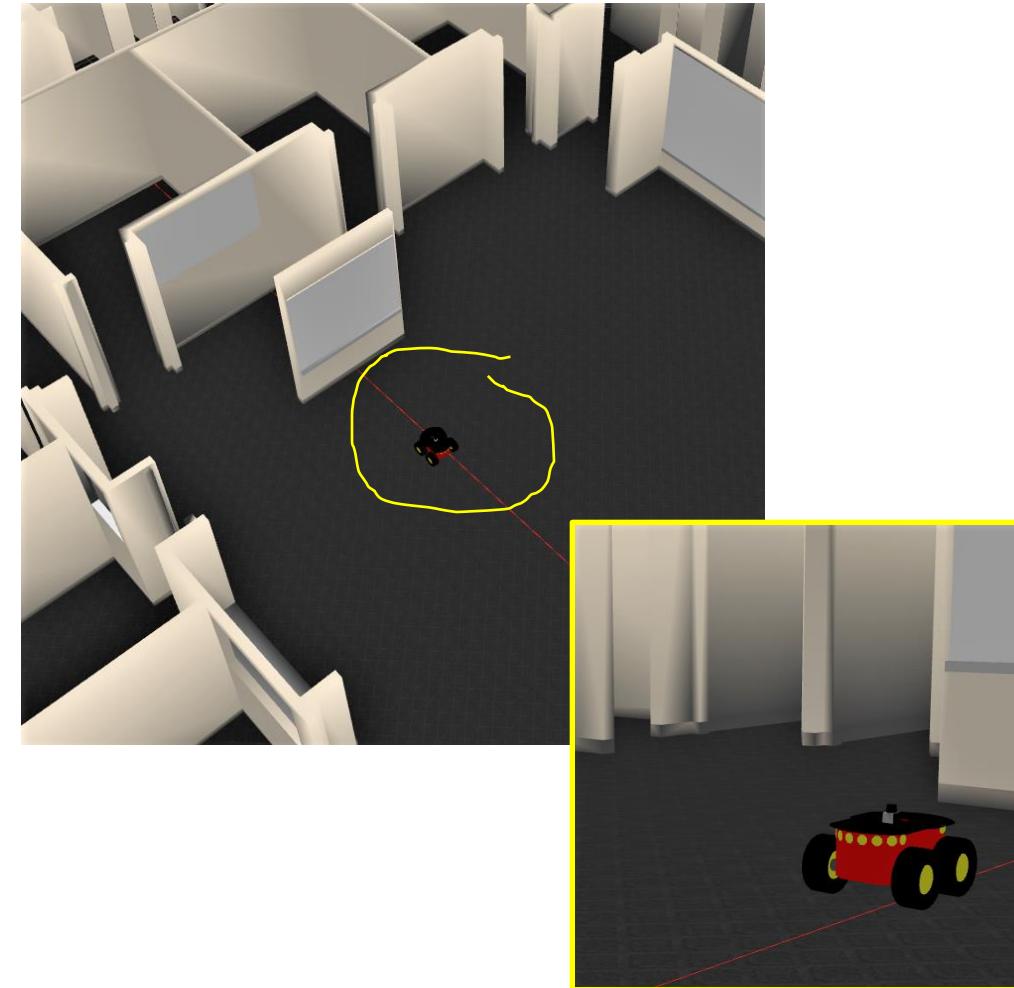
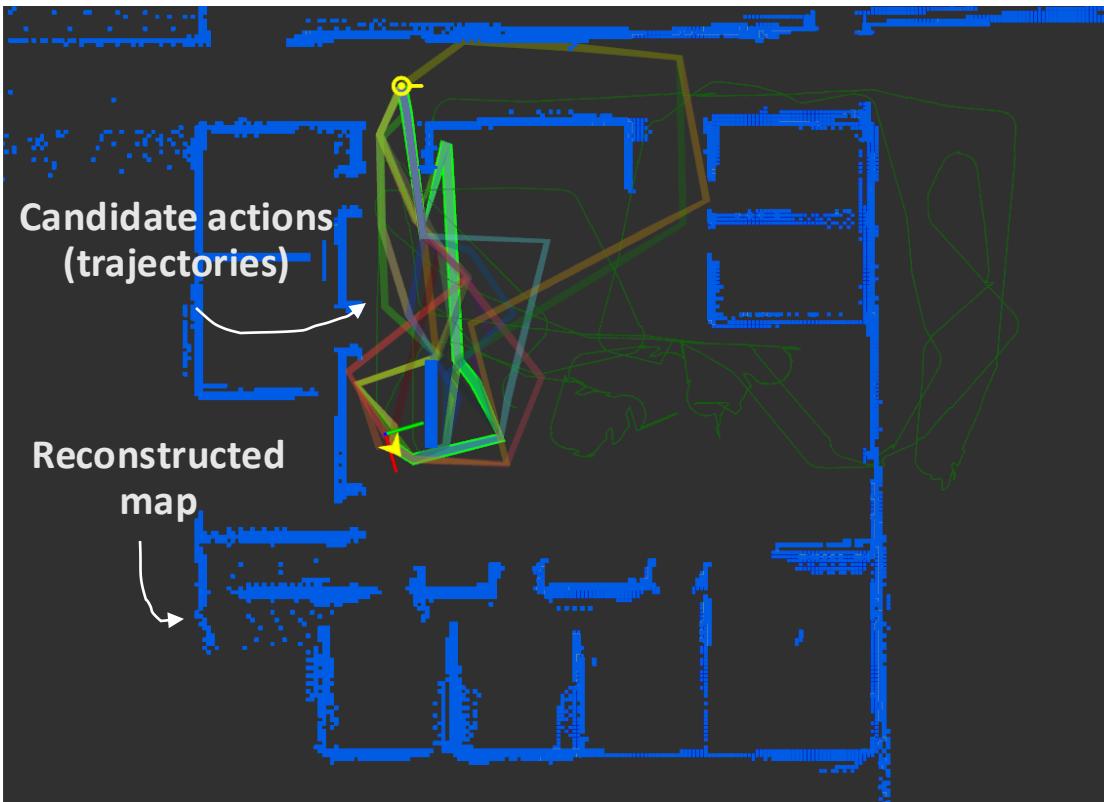


# Belief Sparsification for Gaussian BSP

- Agent performs simultaneous localization and mapping
- Maintains a multivariate Gaussian belief

$$b[X] = \mathcal{N}(X^*, (R^T R)^{-1})$$

- Task: reach a goal with **minimum uncertainty**

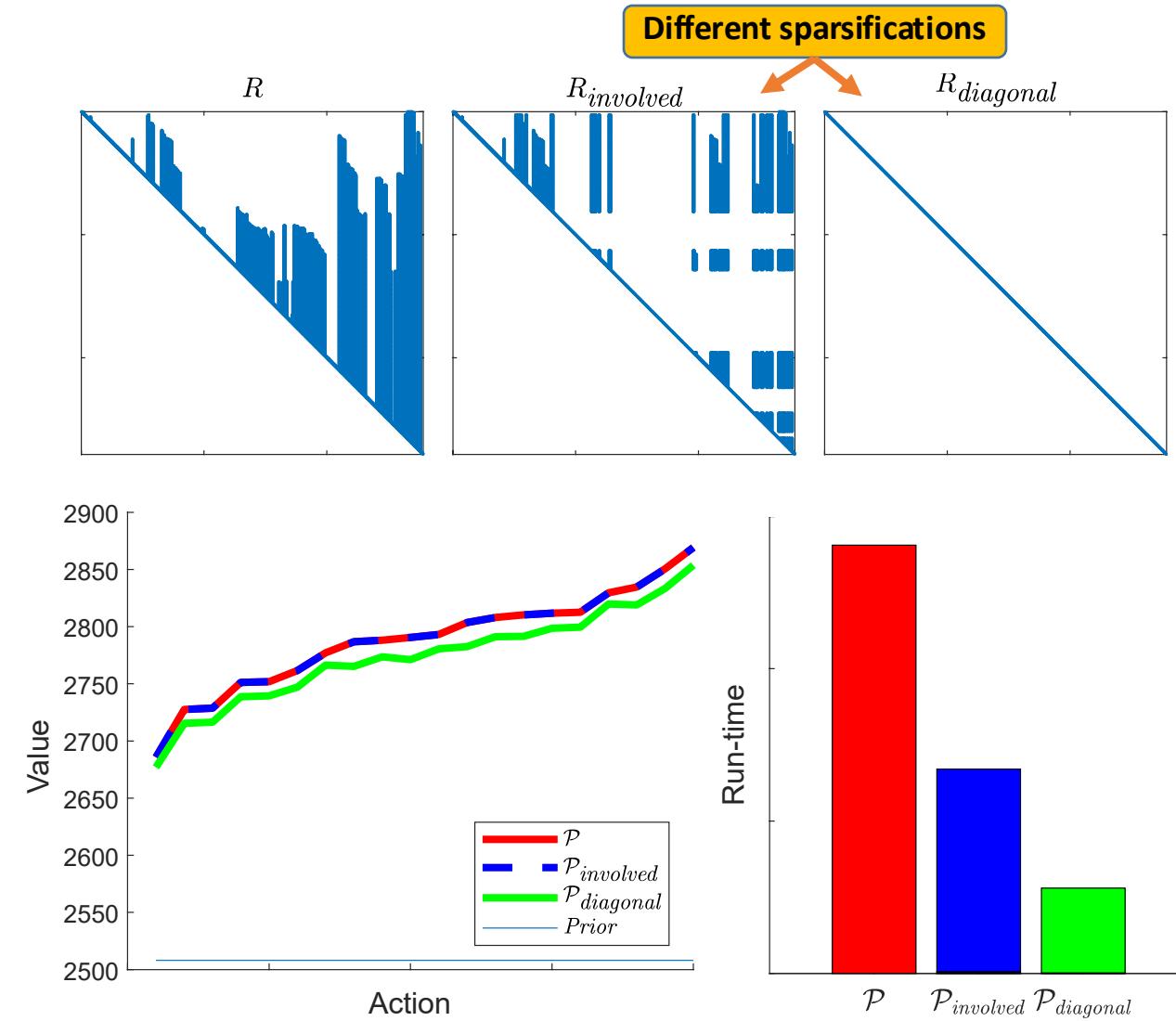
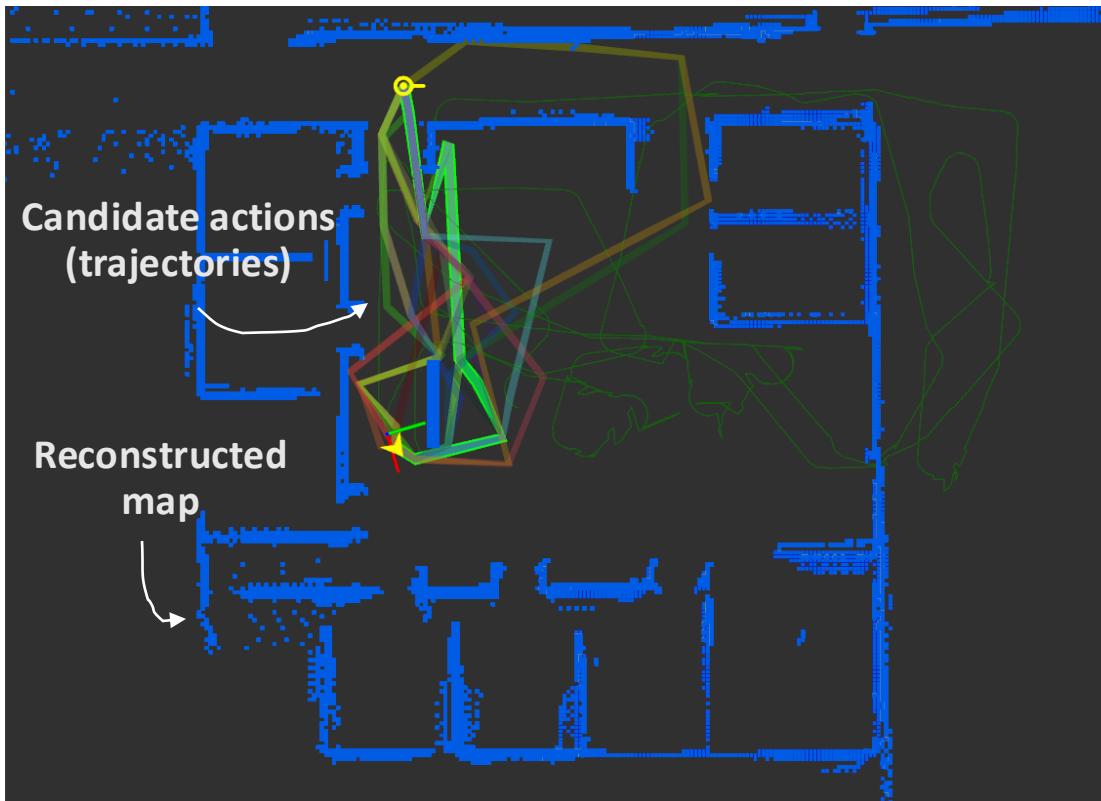


# Belief Sparsification for Gaussian BSP

- Agent performs simultaneous localization and mapping
- Maintains a multivariate Gaussian belief

$$b[X] = \mathcal{N}(X^*, (R^T R)^{-1})$$

- Task: reach a goal with **minimum uncertainty**



# Simplification of Decision-Making Problems

---

## Concept:

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

## Specific simplifications include:

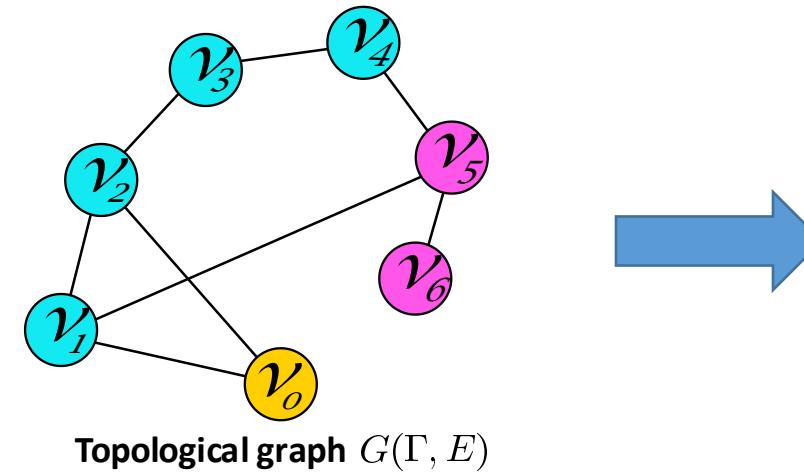
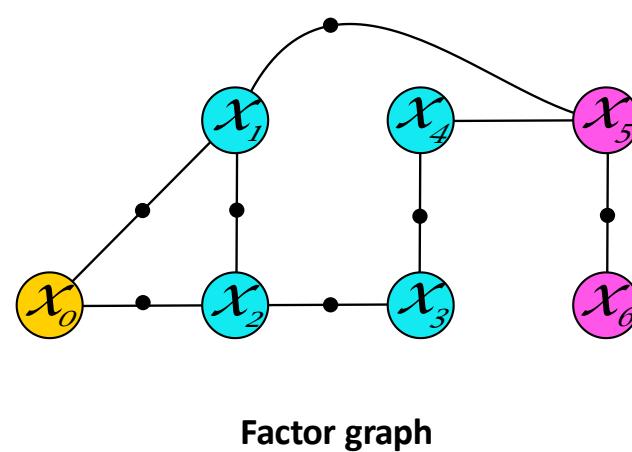
- Sparsification of Gaussian beliefs (high dim. state)
- Topological metric for Gaussian beliefs (high dim. state)
- Utilize a subset of samples (nonparametric beliefs)
- Utilize a subset of hypotheses (hybrid beliefs)
- Simplified models and spaces
- Simplification of Risk-Averse POMDP Planning
- Simplification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

# Topological Gaussian Belief Space Planning (t-BSP)

- Topological properties of factor graphs dominantly determine estimation accuracy<sup>1</sup>

**Key idea:**

- Design a metric of factor graph topology that is strongly correlated with entropy
- Determine best action using that topological metric (instead of entropy)
- **Does not require explicit inference, nor partial state covariance recovery**



topological  
metric  $s(G)$

graph signature

<sup>1</sup>K. Khosoussi, et al. "Reliable graphs for SLAM", IJRR'19.

# Topological Gaussian Belief Space Planning (t-BSP)

## Metric Space

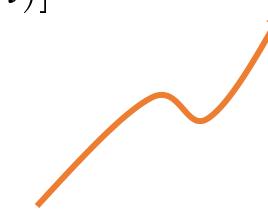
$$J(\mathcal{U}) = \frac{n}{2} \ln(2\pi e) + \frac{1}{2} \ln |\Sigma(X_{k+L})|$$

$$\mathcal{U}^* = \arg \min_{\mathcal{U}} J(\mathcal{U})$$

## Topological space

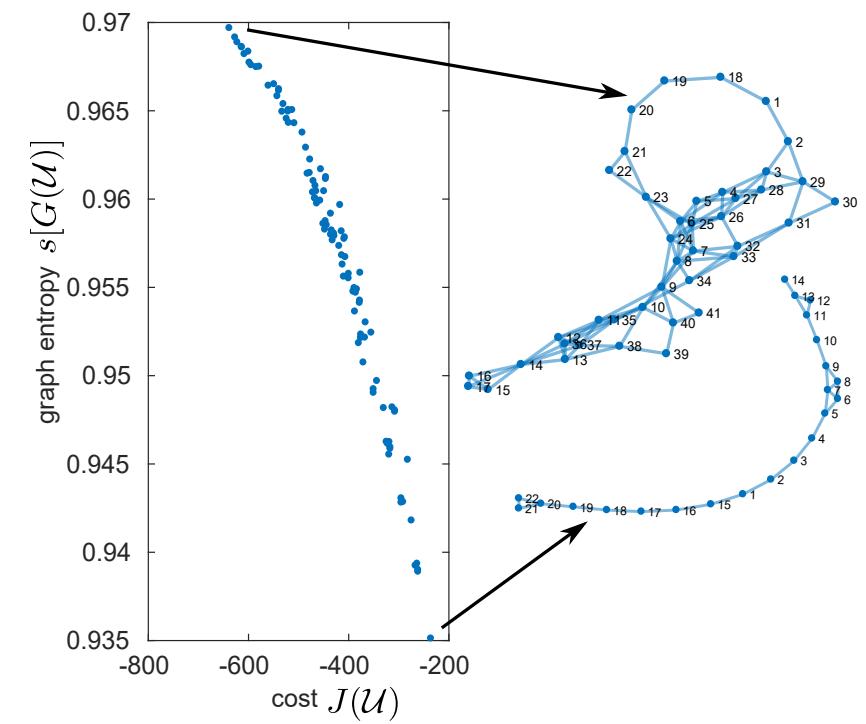
$$s(G) = H_{VN}(G) \approx 1 - \frac{1}{|\Gamma|} - \frac{1}{|\Gamma|^2} \sum_{(i,j) \in E} \frac{1}{d(i)d(j)}$$

$$\hat{\mathcal{U}}^* = \arg \max_{\mathcal{U}} s[G(\mathcal{U})]$$

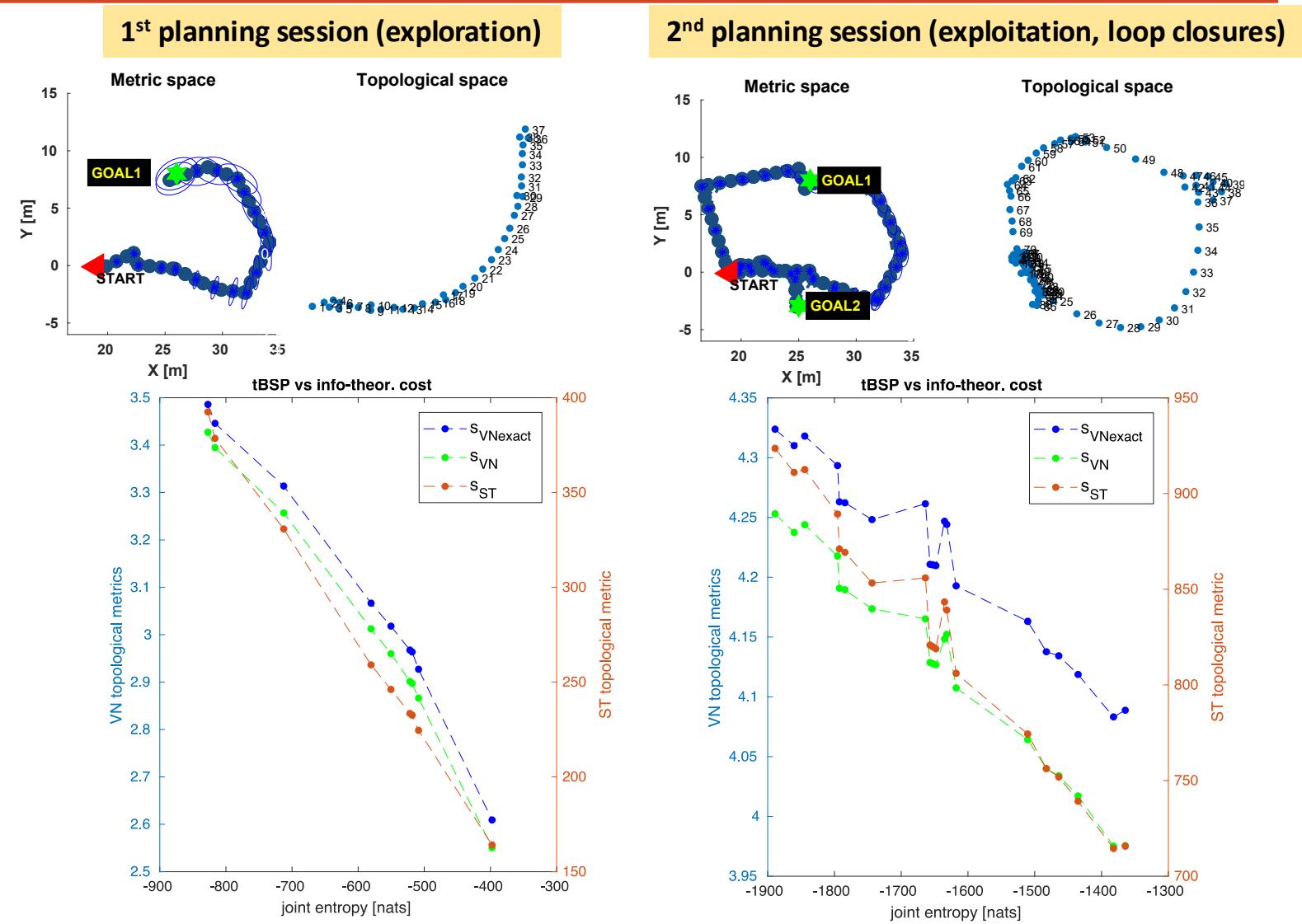
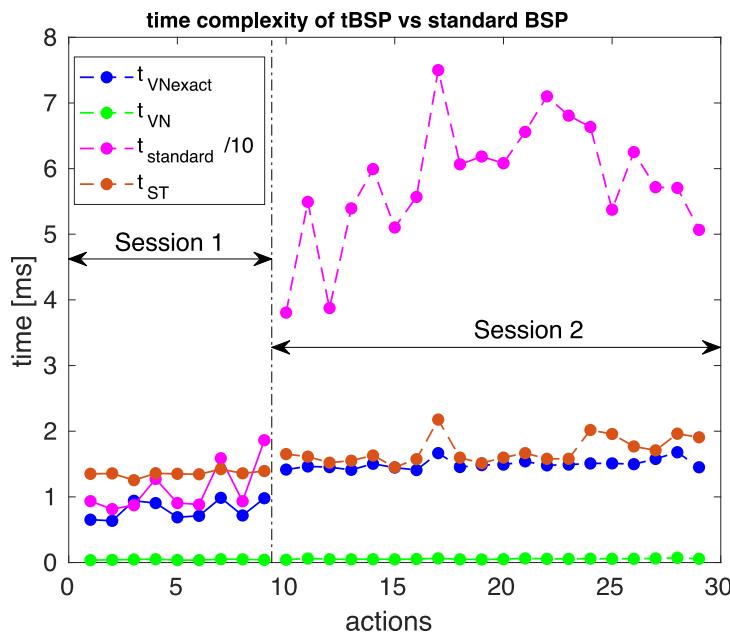
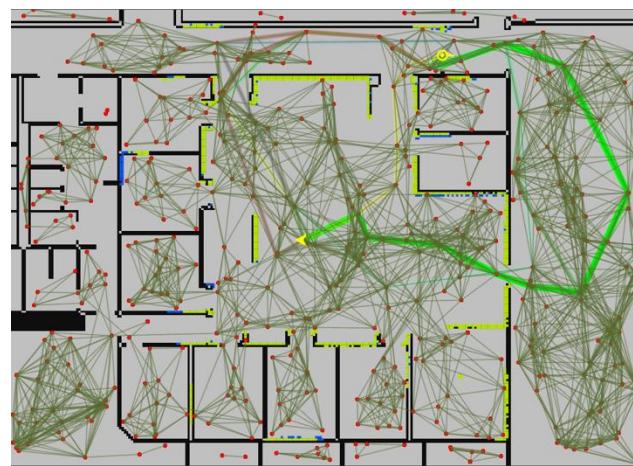


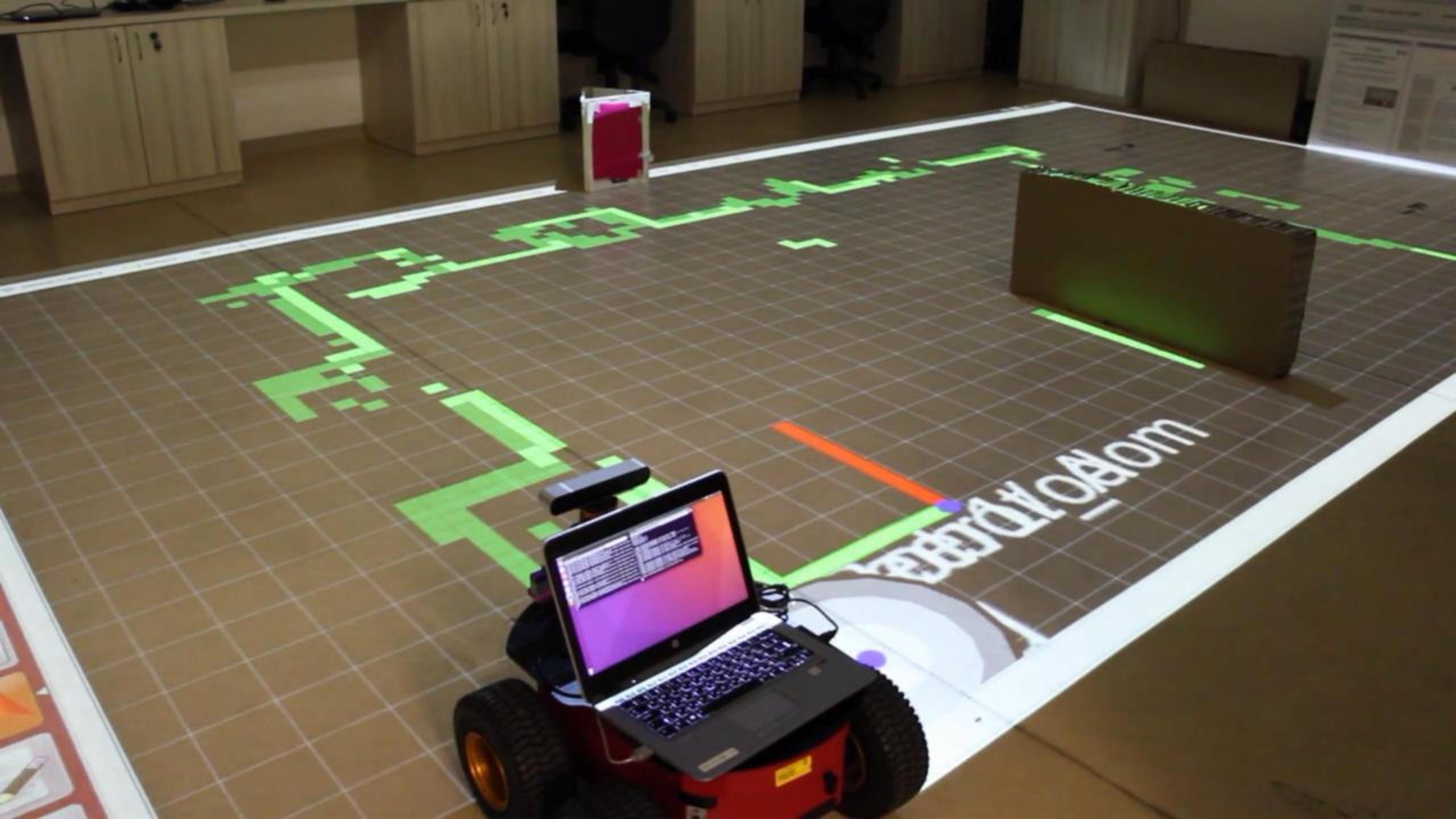
- Cheap to calculate, only a function of node degrees
- Supports incremental calculations
- Provided bounds on the error/loss  $|J_k(\hat{\mathcal{U}}) - J_k(\hat{\mathcal{U}}^*)|$

Topological and info-theoretic metrics are strongly correlated



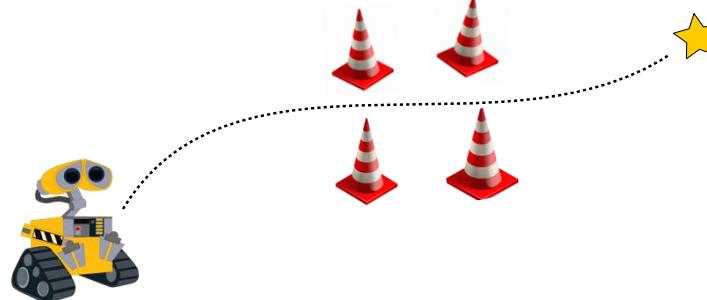
# t-BSP: Gazebo Results





# Focused Topological Gaussian Belief Space Planning (ft-BSP)

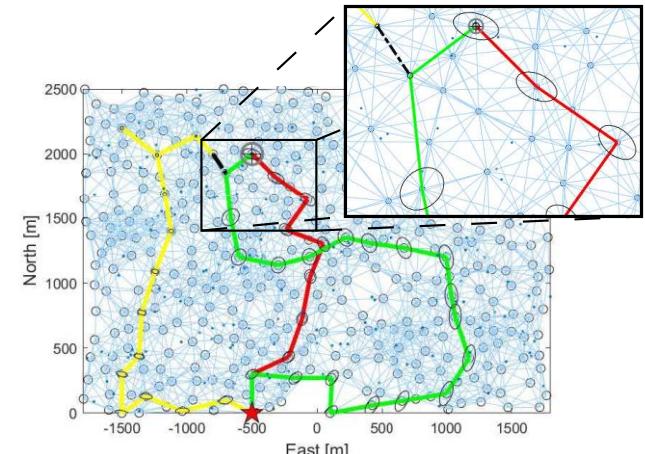
- **Unfocused** BSP – reduce uncertainty over all variables
- **Focused** BSP – reduce uncertainty over a predefined subset of variables (focused variables)



collision avoidance



*focused* reconstruction task



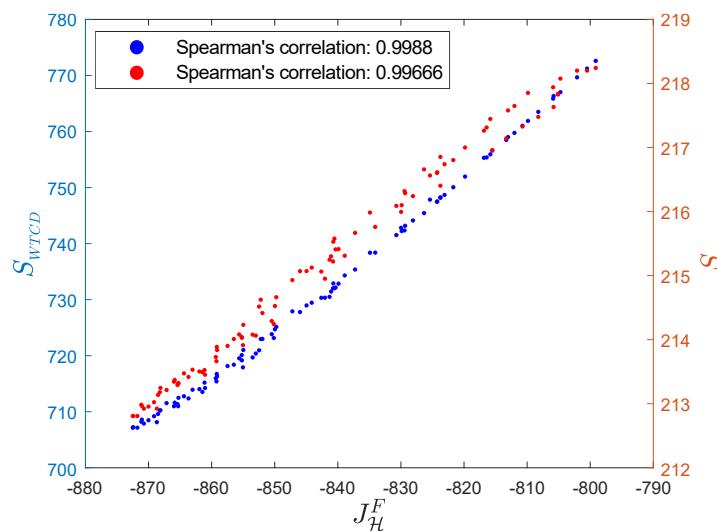
*focused* vs *unfocused*

$$J_{\mathcal{H}}^F(\mathcal{U}) = \frac{n^F}{2} \log(2\pi e) - \frac{1}{2} \log |\Lambda_{k+L}| + \frac{1}{2} \log |\Lambda_{k+L}^U| \quad \text{Expensive!}$$

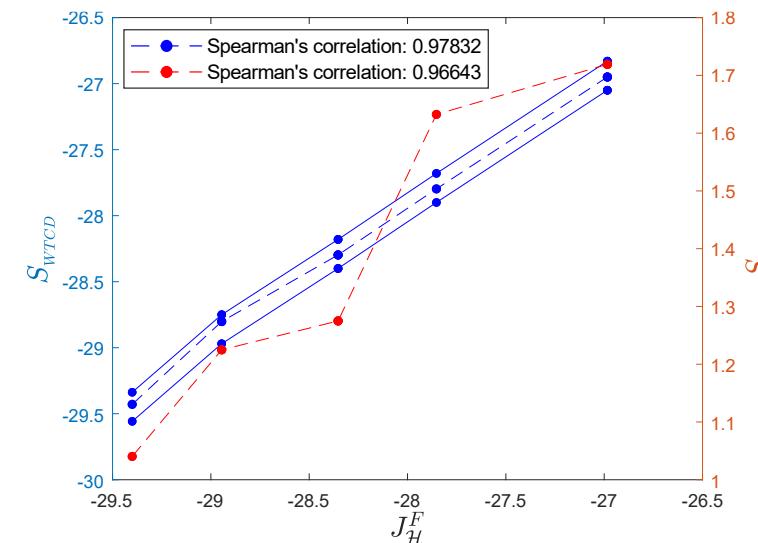
Topological signatures:  $\begin{matrix} S_{WTCD} \\ S_{VND} \end{matrix}$

# Focused Topological Gaussian Belief Space Planning (ft-BSP)

- **Unfocused BSP** – reduce uncertainty over all variables
- **Focused BSP** – reduce uncertainty over a predefined subset of variables (focused variables)



Measurement Selection



Active 2D Pose SLAM

signature	measurement selection	active SLAM
$S_{WTCD}$	18.88	1.21
$S_{VND}$	<b>12.02</b>	<b>0.14</b>
$J_H^F$	146.24	6.34

Average running time experiments in ms

# Simplification of Decision-Making Problems

---

## Concept:

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

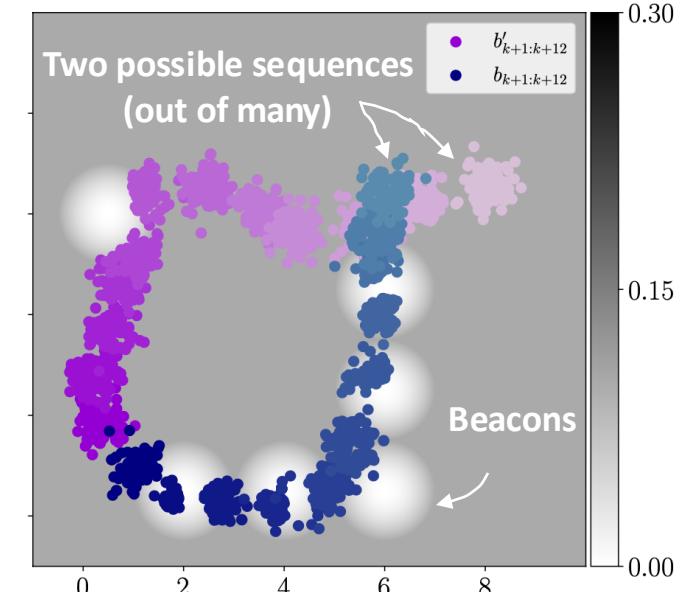
## Specific simplifications include:

- 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

# Simplification of POMDPs with Nonparametric Beliefs

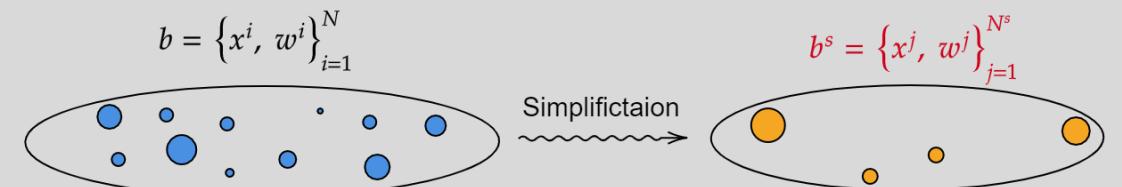
- Value function

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



## Simplification:

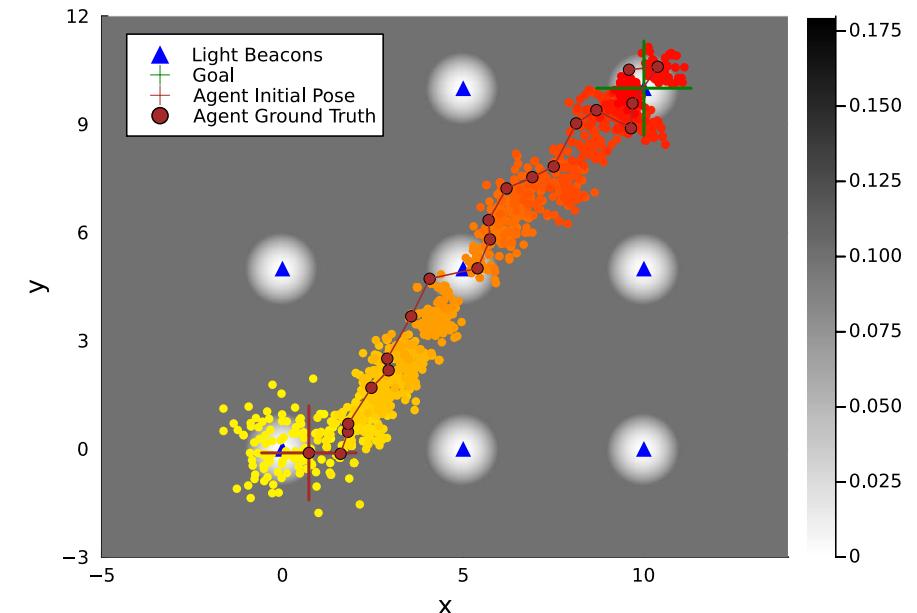
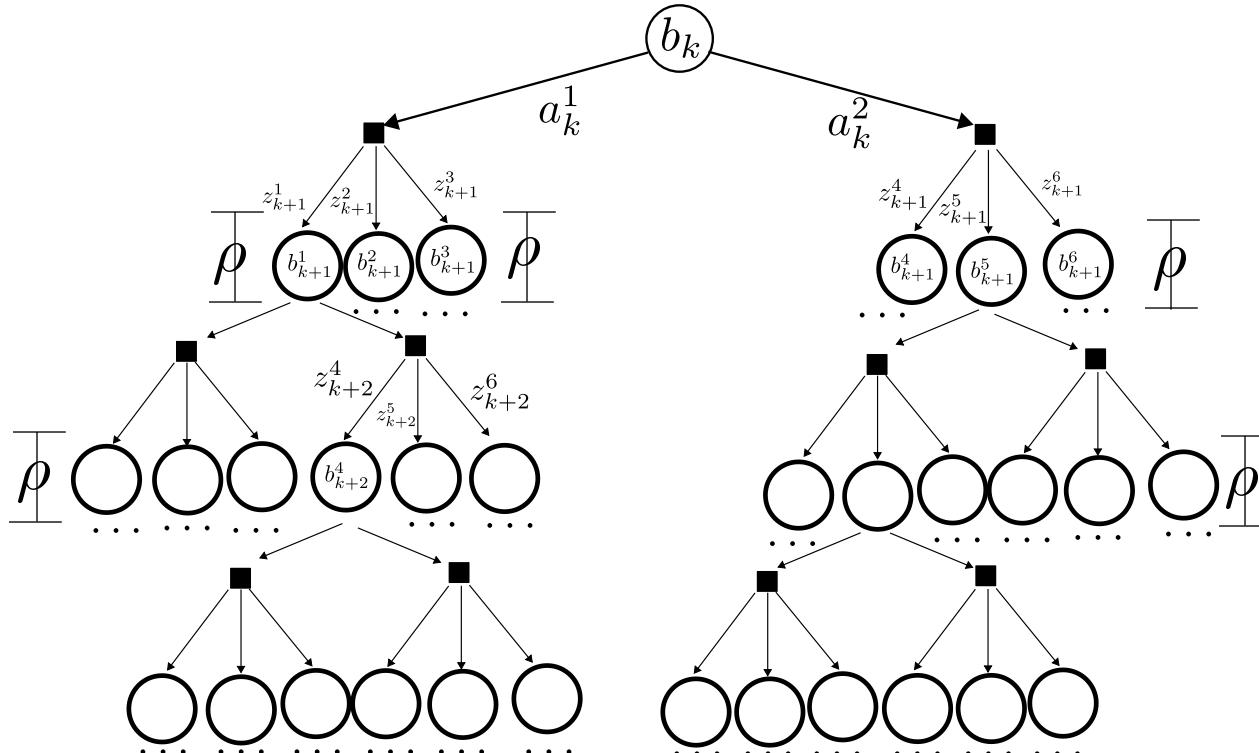
- Utilize a **subset** of samples for planning
- Information-theoretic reward (entropy)
- Analytical (**cheaper**) bounds over the reward



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

# Simplification of POMDPs with Nonparametric Beliefs

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

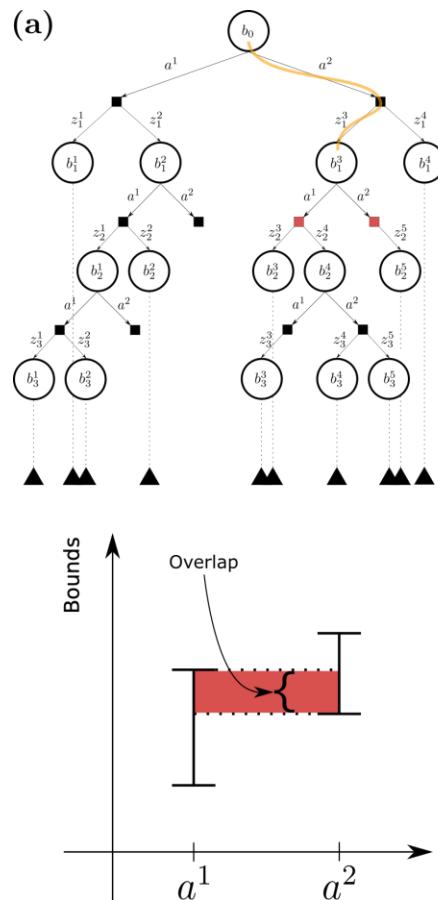


Typical speedup of 20% - 50%,  
Same performance!

# Simplification of POMDPs with Nonparametric Beliefs

---

- Adaptive multi-level simplification in an MCTS setting:



# Simplification of Decision-Making Problems

---

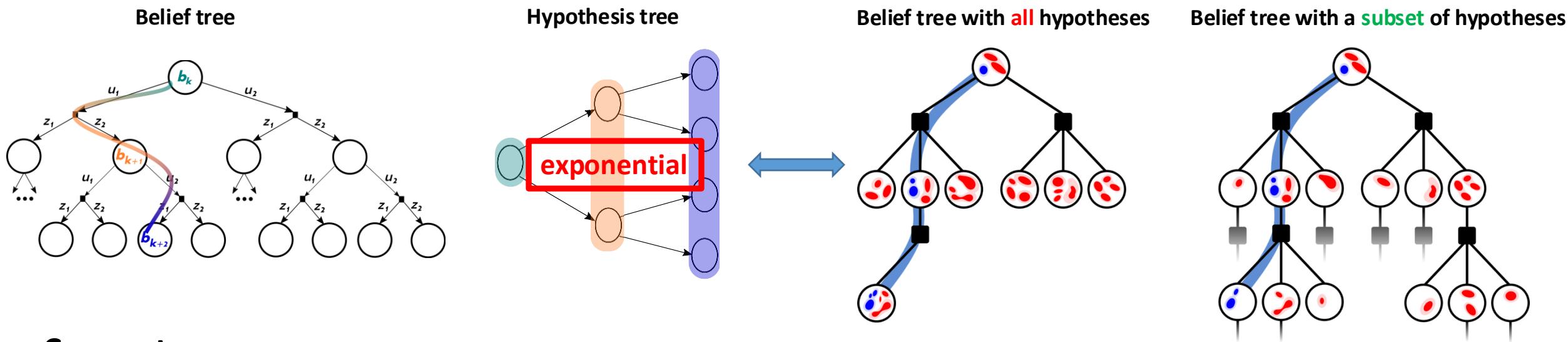
## Concept:

- 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

# Simplification of BSP/POMDP with Hybrid Beliefs



## Concept:

- Instead, utilize only a **subset** of hypotheses
- Derive reward bounds, given planning task (reward)
  - Disambiguate between hypotheses
  - Navigate to a goal
  - ..

$$\mathcal{LB}(b_k, \pi) \leq V^\pi(b_k) \leq \mathcal{UB}(b_k, \pi)$$

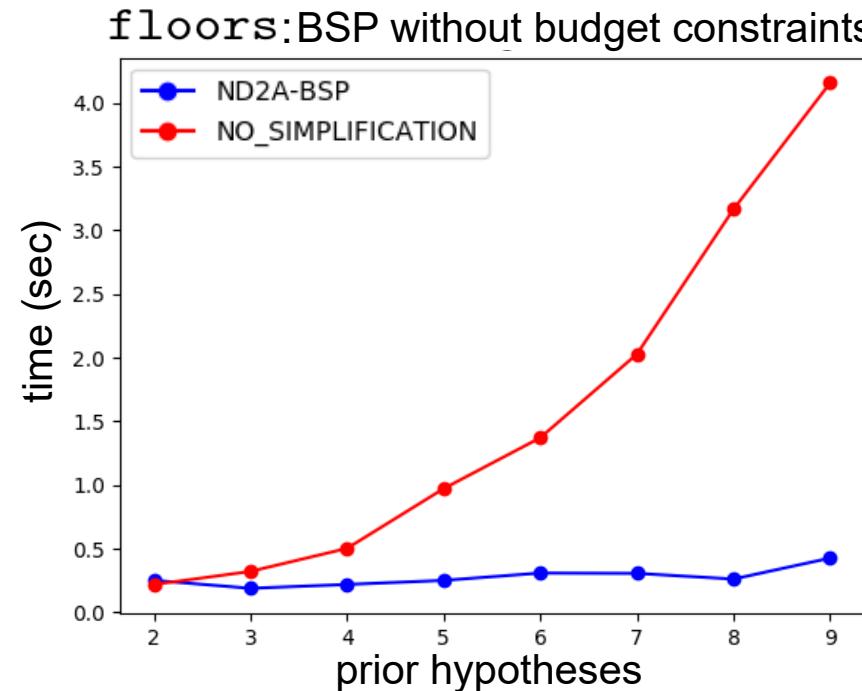
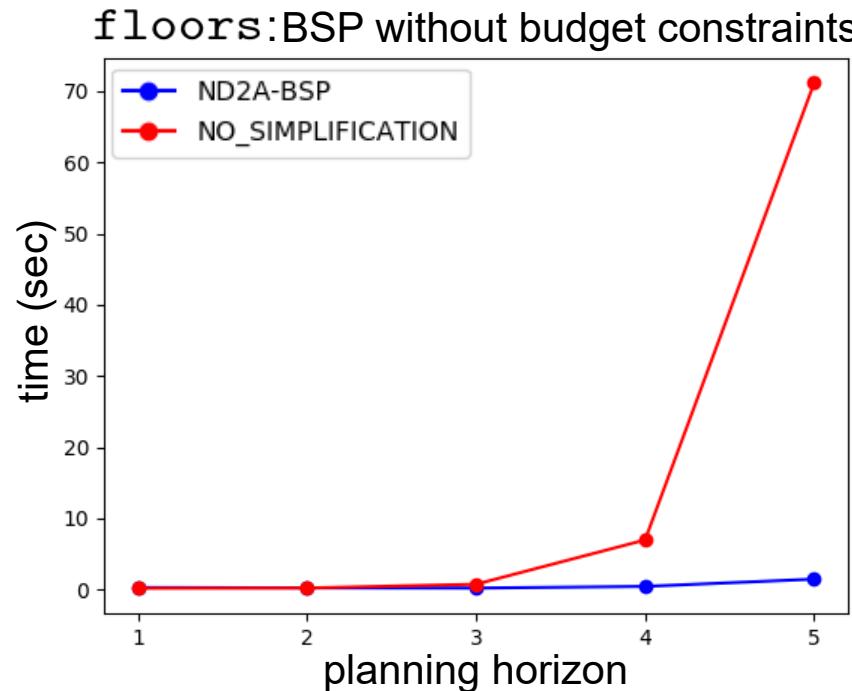
M. Shienman and V. Indelman, "D2A-BSP: Distilled Data Association Belief Space Planning with Performance Guarantees Under Budget Constraints," ICRA'22, **Outstanding Paper Award Finalist**.

M. Shienman and V. Indelman, "Nonmyopic Distilled Data Association Belief Space Planning Under Budget Constraints," ISRR'22.

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

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

# Simplification of BSP/POMDP with Hybrid Beliefs



- Significant speed-up in planning
- Same planning performance is **guaranteed** (no overlap between bounds)

# Simplification of BSP/POMDP with Hybrid Beliefs

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

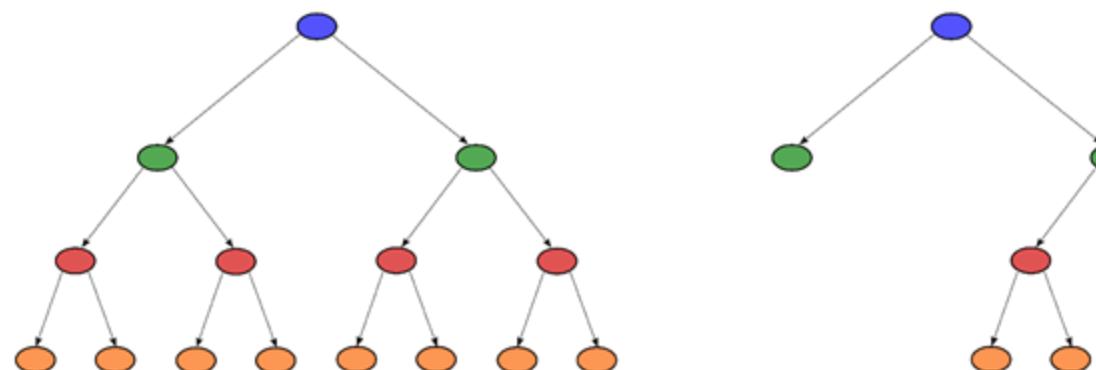
## Corollary

For any policy  $\pi$ , and selection of hypotheses set  $\{\beta_{0:\tau}^i\}_{i=0}^{|\mathcal{B}|}$  the following holds,

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

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

# Simplification of Decision-Making Problems

---

## Concept:

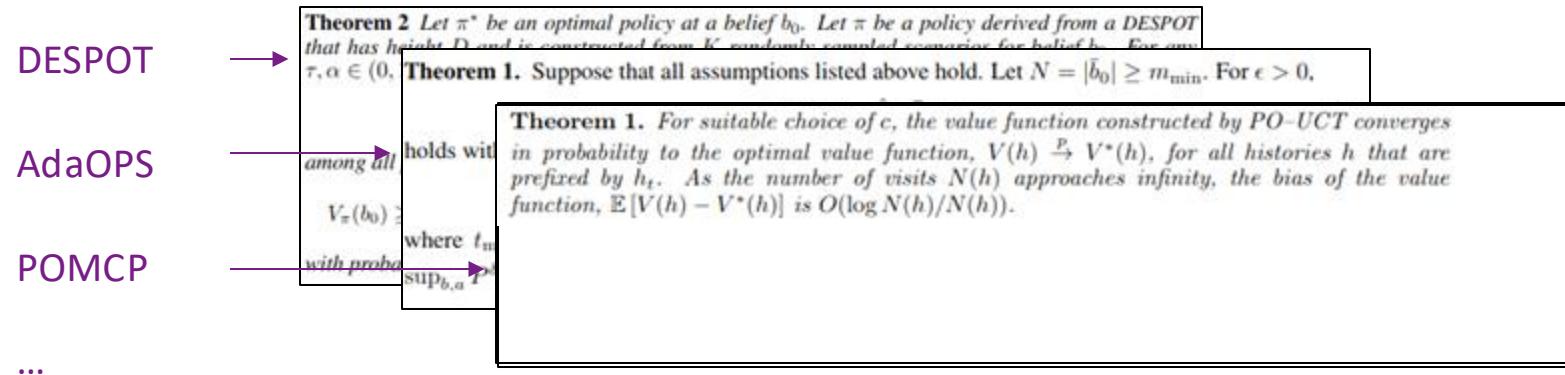
- Identify and solve a **simplified (computationally) easier** decision-making problem
- Provide performance guarantees

## 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

# POMDPs with Deterministic Guarantees

SOTA sampling based approaches come with probabilistic theoretical guarantees



Can we get deterministic guarantees?

We show that deterministic guarantees are indeed possible!

# Online POMDP Planning with Anytime Deterministic Guarantees

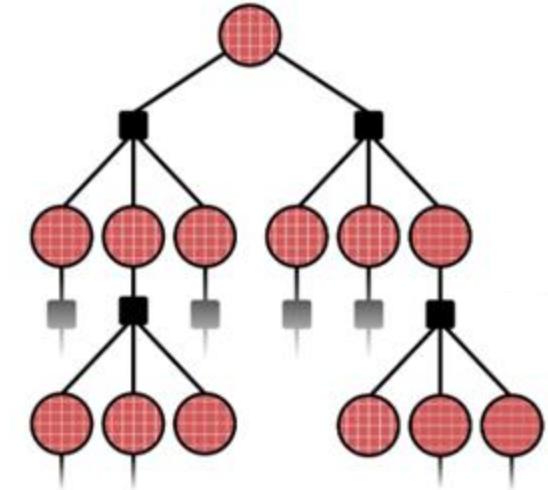
---

## Concept:

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

$$\mathcal{M} \xrightarrow{\text{wavy arrow}} \bar{\mathcal{M}}$$

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



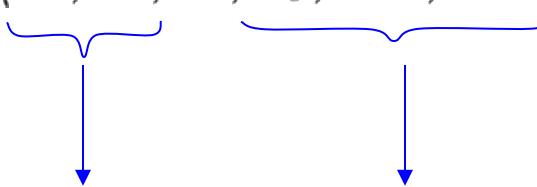
$$V(b) \xleftarrow{\text{wavy arrow}} ? \xrightarrow{\text{wavy arrow}} \bar{V}(b)$$

# 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,

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

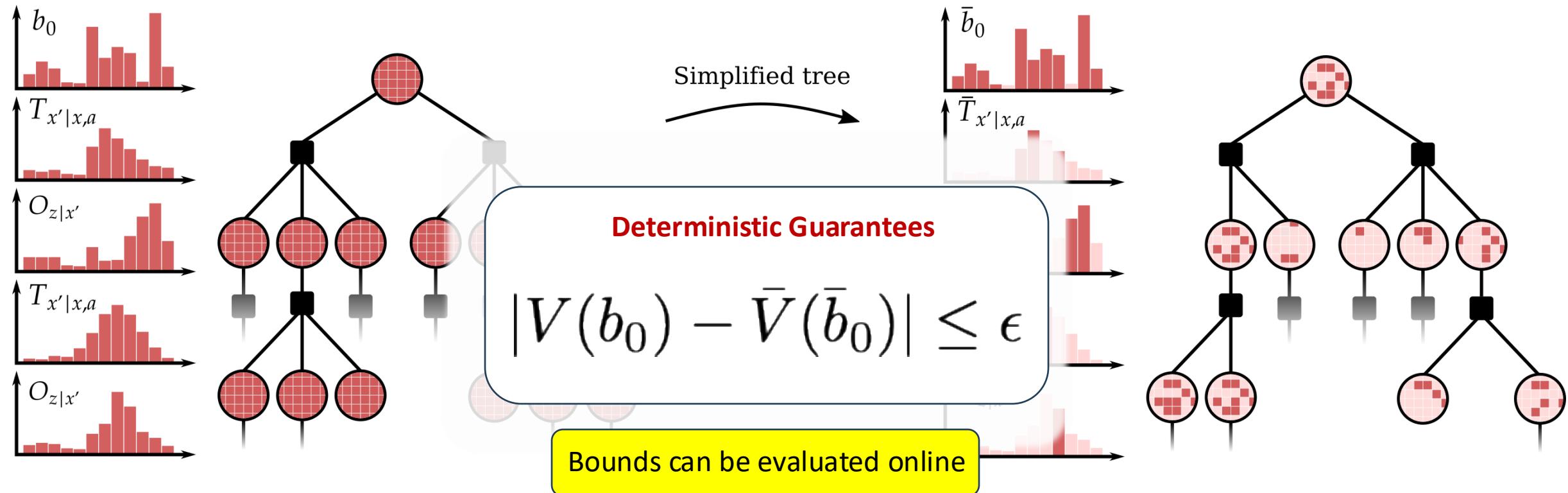

  
 $\bar{\mathcal{X}}(H_t) \subset \mathcal{X}$        $\bar{b}_0(x) \triangleq \begin{cases} b_0(x) & , x \in \bar{\mathcal{X}}_0 \\ 0 & , otherwise \end{cases}$   
 $\bar{\mathcal{Z}}(H_t) \subset \mathcal{Z}$        $\bar{\mathbb{P}}(x_{t+1} | x_t, a_t) \triangleq \begin{cases} \mathbb{P}(x_{t+1} | x_t, a_t) & , x_{t+1} \in \bar{\mathcal{X}}(H_{t+1}^-) \\ 0 & , otherwise \end{cases}$   
 $\bar{\mathbb{P}}(z_t | x_t) \triangleq \begin{cases} \mathbb{P}(z_t | x_t) & , z_t \in \bar{\mathcal{Z}}(H_t) \\ 0 & , otherwise \end{cases}$

- Simplified value function

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

# Online POMDP Planning with Anytime Deterministic Guarantees

- Deterministic guarantees (assuming discrete spaces)



# Online POMDP Planning with Anytime Deterministic Guarantees

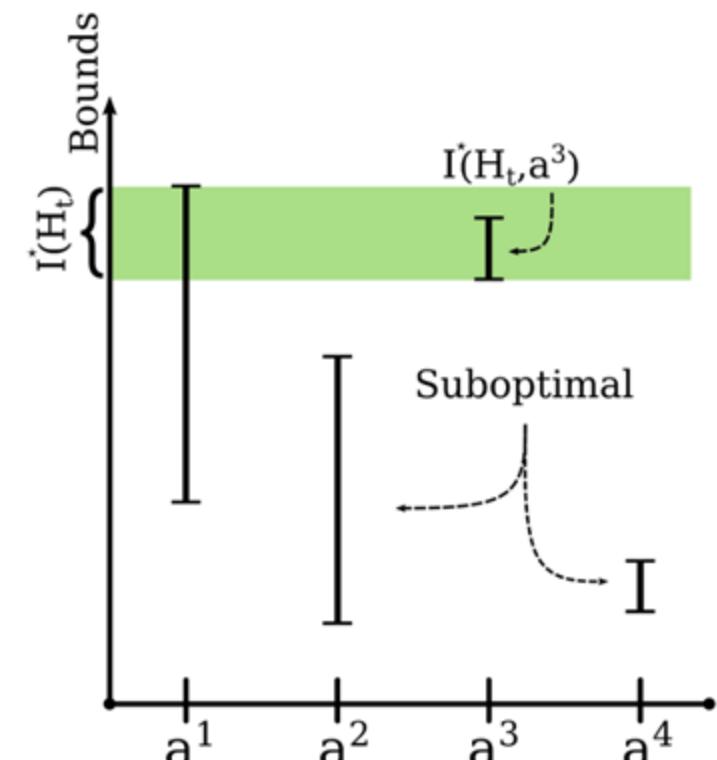
Importantly, the bounds can be calculated during planning.

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**

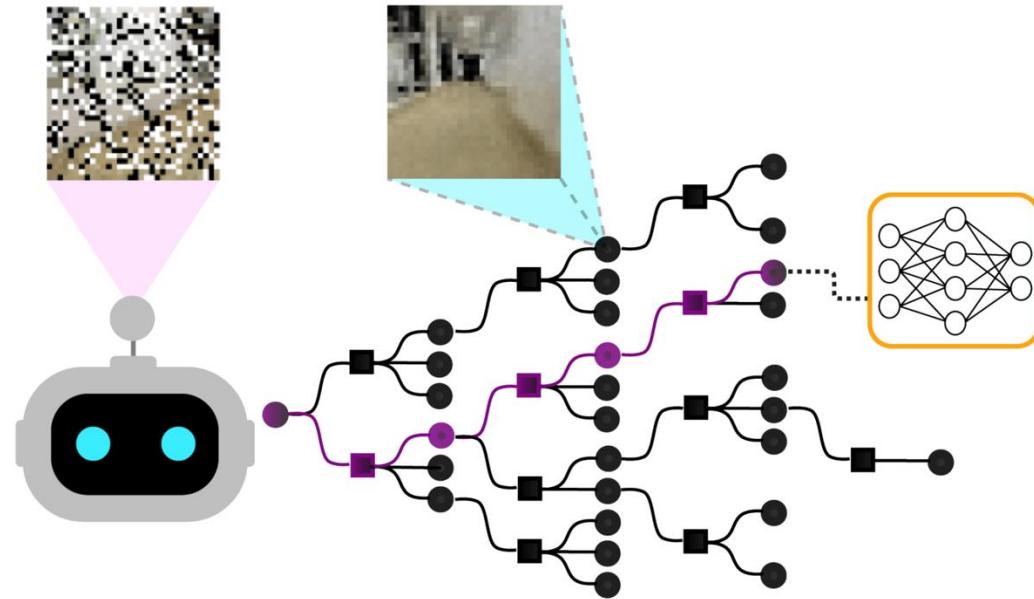
$$|V(b_0) - \bar{V}(\bar{b}_0)| \leq \epsilon$$



# Simplifying Complex Observation Models with Probabilistic Guarantees

## Visual POMDP planning

- Visual observations are complex to model in planning<sup>1,2</sup>
- Learned observation models are (often) impractical for solving POMDPs in real time



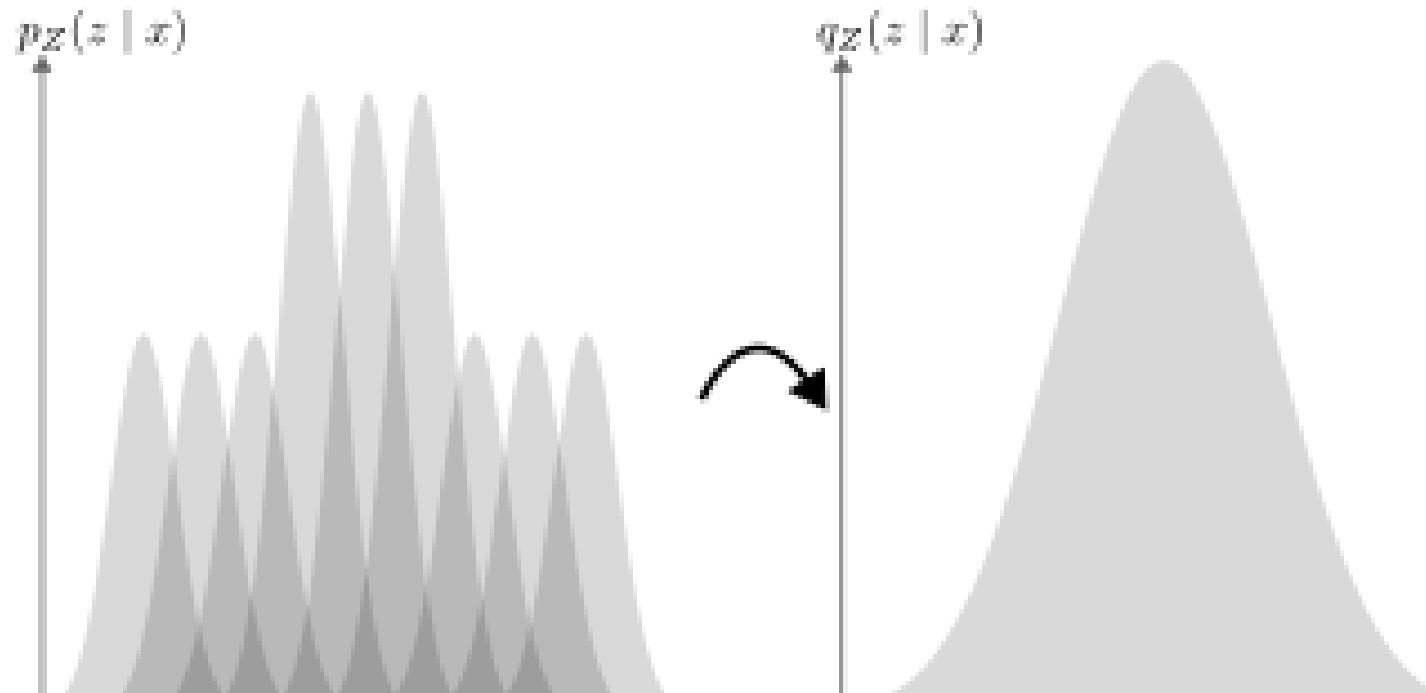
**Can we simplify the learned models?  
What is the impact on planning performance?**

<sup>1</sup>Wang et al., “DualSMC: Tunneling Differentiable Filtering and Planning under Continuous POMDPs”.

<sup>2</sup>Deglurkar et al., “Compositional Learning-based Planning for Vision POMDPs”.

# 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:



**Simplified models**

$$p_\theta(z | x)$$

Original, **expensive**

$$q_\phi(z | x)$$

Simplified, **cheap**

# 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:  $Q_P^{q_Z}$

## Corollary 3

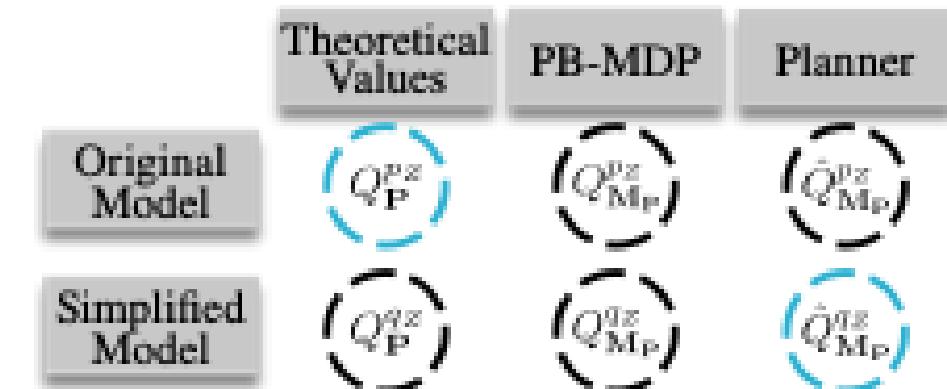
For arbitrary  $\varepsilon, \delta > 0$  there exists a number of particles for which

$$\underline{|Q_P^{p_Z}(b_t, a) - \hat{Q}_{M_P}^{q_Z}(\bar{b}_t, a)|} \leq \Phi_{M_P}(\bar{b}_t, a) + \varepsilon$$

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

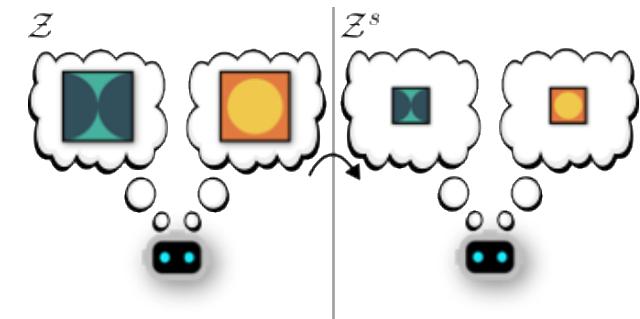
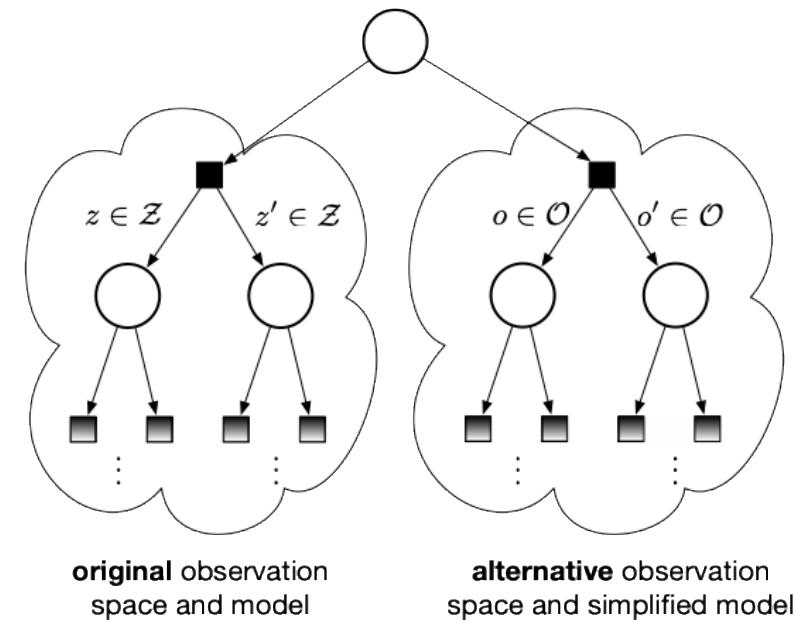
# Simplified POMDP Planning with an Alternative Observation Space

- 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



# Simplified POMDP Planning with an Alternative Observation Space

- Switch to an alternative observation space and model

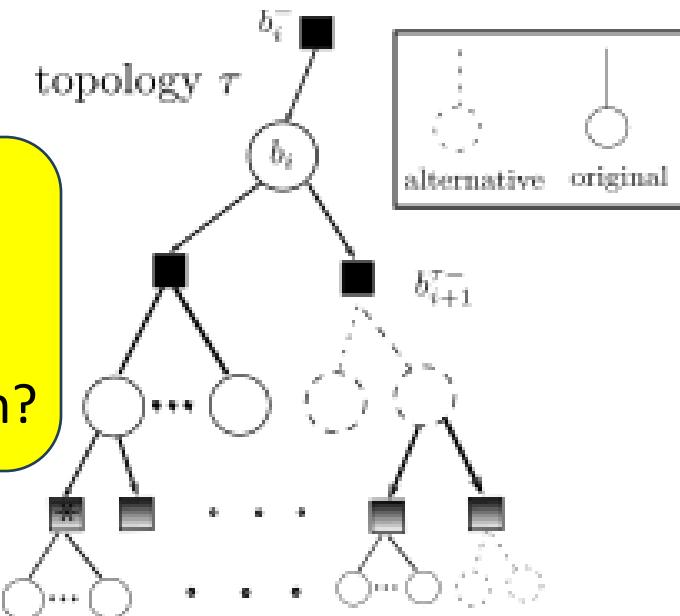
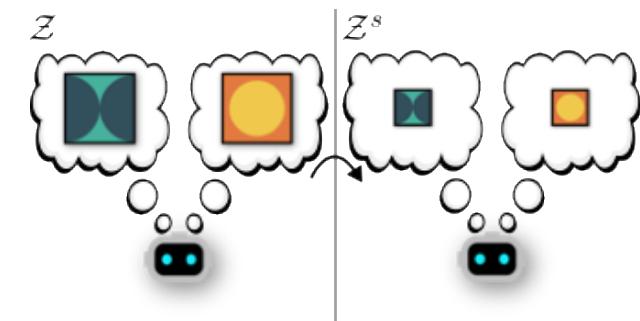
## Model Definition

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?



# 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

# Partitioning of a Multivariate Observation Space

---

- Value function:  $V^\pi(b_k) = R(b_k, \pi_k(b_k)) + \mathbb{E}_{z_{k+1:k+\ell}} \left[ \sum_{i=k+1}^{k+\ell} R(b_i, \pi_i(b_i)) \right]$

- Belief-dependent reward: entropy

$$R(b, \pi(b)) \triangleq -\mathcal{H}(X) \equiv \mathbb{E}_{X \sim b} (\log b[X])$$

- The expected reward at each  $i$ th look-ahead step:

$$\mathbb{E}_{Z_{k+1:i}} [R(b_i, a_{i-1})] = -\mathcal{H}(X_i | Z_{k+1:i})$$

- Future observations are drawn from the distribution  $\mathbb{P}(Z_{k+1:i} | b_k, \pi)$

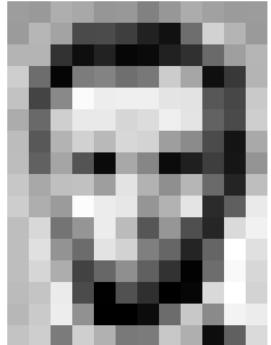
# Partitioning of a Multivariate Observation Space

- Consider a multivariate random variable  $Z \in \mathcal{Z}$ , that represents future observations:

$$Z = (Z^1, Z^2, \dots, Z^m)$$

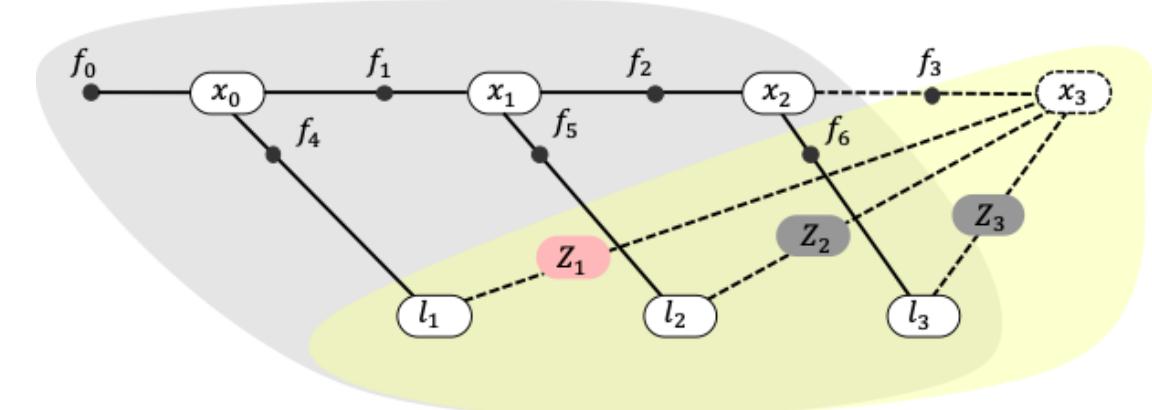
- Examples:

Raw measurement of an image sensor



167	153	174	168	150	152	129	151	172	161	155	156
185	182	163	74	59	62	33	17	110	210	180	154
180	180	60	14	54	6	10	33	48	106	159	181
205	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	228	227	87	71	201	
172	105	207	233	258	214	220	239	228	96	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	158	134	11	51	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	235	187	85	150	79	36	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	56	153	143	96	50	2	109	249	215
187	196	236	75	1	81	47	0	6	217	25	211
183	202	237	145	0	9	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Factor graph



# Partitioning of a Multivariate Observation Space

---

- Consider a multivariate random variable  $Z \in \mathcal{Z}$ , that represents future observations:

$$Z = (Z^1, Z^2, \dots, Z^m)$$

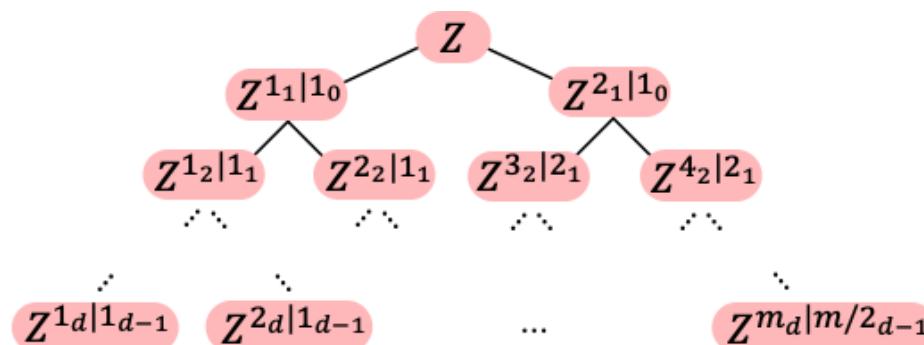
- We can partition  $Z \in \mathcal{Z}$  into different subsets/components, e.g.

$$Z^s = \{Z^1, Z^2, \dots, Z^n\}$$

$$Z^{\bar{s}} = \{Z^{n+1}, Z^{n+2}, \dots, Z^m\}$$

$$Z = Z^s \cup Z^{\bar{s}}$$

- Hierarchical Partitioning:

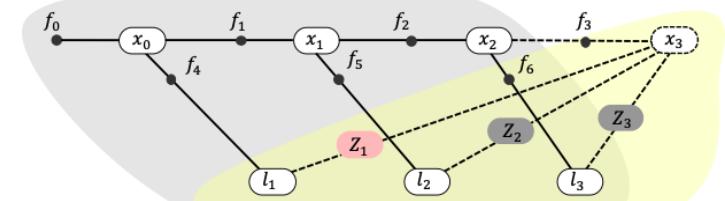
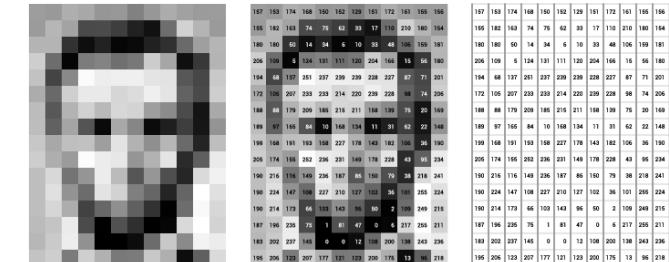


# Partitioning of a Multivariate Observation Space

---

But why is this a good idea?

- Apply partitioning to a raw image measurement of size 20x20 binary pixels
- Consider all of the different permutations for each pixel,  $2^{400}$  in total
- If we partition  $Z^s \triangleq \{Z^{x,y} \mid y \leq 10\}$  and  $Z^{\bar{s}} \triangleq \{Z^{x,y} \mid y > 10\}$ 
  - Need to consider  $2^{200}$  permutations for each
  - Overall,  $2^{201}$  vs  $2^{400}$  permutations



# Partitioning of a Multivariate Observation Space

---

## Lemma 2

Given two sets of expected measurements  $(Z^s, Z^{\bar{s}})$ , the conditional Entropy can be factorized as

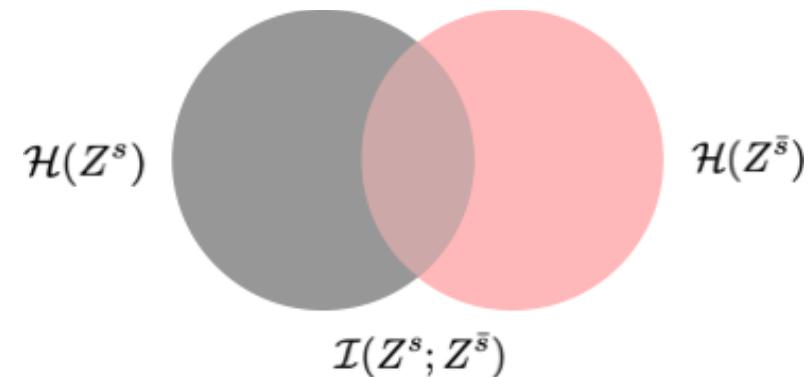
$$\mathcal{H}(X|Z) = \mathcal{H}(Z^s|X) + \mathcal{H}(Z^{\bar{s}}|X) - \mathcal{H}(Z^s, Z^{\bar{s}}) + \mathcal{H}(X)$$

$$\mathcal{H}(Z^s, Z^{\bar{s}}) = \mathcal{H}(Z^s) + \mathcal{H}(Z^{\bar{s}}) - \mathcal{I}(Z^s; Z^{\bar{s}})$$

$$\mathcal{LB} \leq \mathcal{H}(X|Z) \leq \mathcal{UB}$$

$$\mathcal{LB} \triangleq \mathcal{H}(Z^s | X) + \mathcal{H}(Z^{\bar{s}}|X) - \mathcal{H}(Z^s) - \mathcal{H}(Z^{\bar{s}}) + \mathcal{H}(X)$$

$$\mathcal{UB} \triangleq \mathcal{H}(Z^s|X) + \mathcal{H}(X) - \mathcal{H}(Z^s)$$



# Partitioning of a Multivariate Observation Space

## Multivariate Gaussian Belief

- Posterior information matrix:

$$\Lambda_i = \Lambda_k^{\text{Aug}} + A_i^T A_i$$

- Prior work<sup>1,2</sup> - application of the matrix determinant lemma:

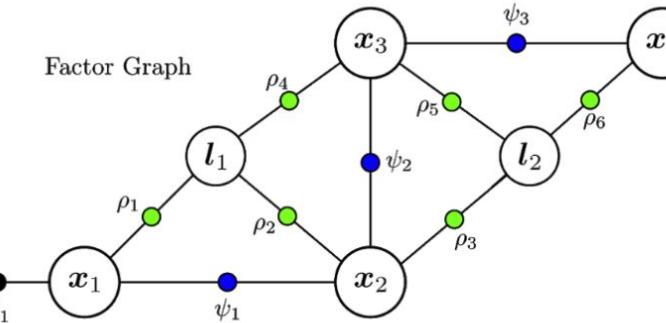
entropy  $\propto |\Lambda_k + A^T \cdot A| = |\Lambda_k| \cdot |I_m + A \cdot \Sigma_k \cdot A^T|$

posterior info matrix

For information gain:

$$\ln \frac{|\Lambda_k + A^T A|}{|\Lambda_k|} = \ln |I_m + A \cdot \Sigma_k \cdot A^T| = \ln |I_m + {}^I A \cdot \Sigma_k^{M, {}^I X} \cdot ({}^I A)^T|$$

One-time calculations for all candidate actions  
 Recover entries only for the involved variables in any of the actions



Measurement Jacobian

$l_1$	$l_2$	$x_1$	$x_2$	$x_3$	$x_4$	
X	X	X	X	X	X	$\rho_1$
X	X	X	X	X	X	$\rho_2$
X	X	X	X	X	X	$\rho_4$
X	X	X	X	X	X	$\rho_3$
X	X	X	X	X	X	$\rho_5$
X	X	X	X	X	X	$\rho_6$
						$\phi_1$
						$\psi_1$
						$\psi_2$
						$\psi_3$

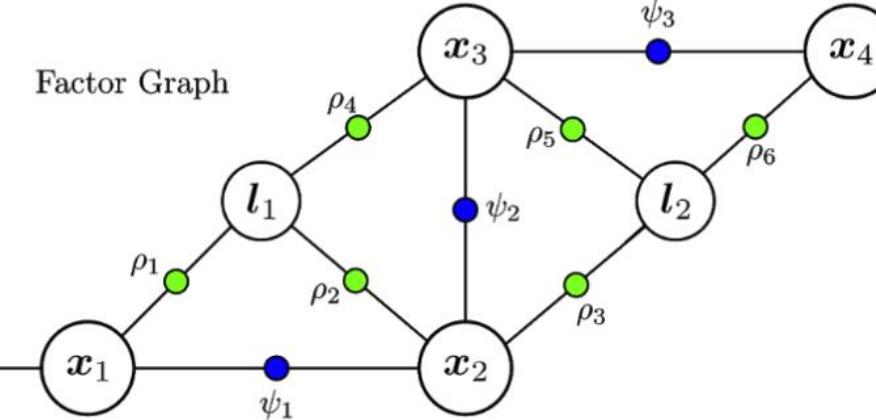
<sup>1</sup>D. Kopitkov and V. Indelman, "No belief propagation required: belief space planning in high-dimensional state spaces via factor graphs, the matrix determinant lemma, and re-use of calculation", IJRR'17.

<sup>2</sup>D. Kopitkov and V. Indelman, "General-purpose incremental covariance update and efficient belief space planning via a factor-graph propagation action tree", IJRR'19.

# Partitioning of a Multivariate Observation Space

## Partitioning of a Gaussian Belief

$$\Lambda_i = \Lambda_k^{\text{Aug}} + A_i^T A_i$$

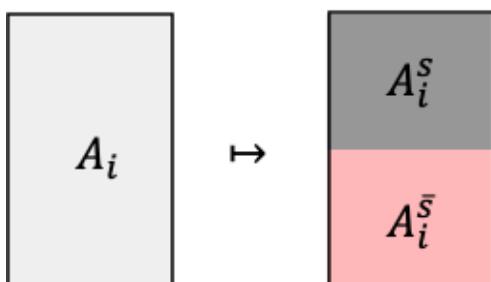


Measurement Jacobian

	$l_1$	$l_2$	$x_1$	$x_2$	$x_3$	$x_4$	
$\rho_1$	X						
$\rho_2$	X						
$\rho_3$	X						
$\rho_4$		X					
$\rho_5$		X					
$\rho_6$		X					
$\phi_1$			X				
$\psi_1$			X				
$\psi_2$			X				
$\psi_3$			X				

**Observation partitioning** corresponds to splitting the Jacobian into blocks

$$Z_i \mapsto (Z_i^s, Z_i^{\bar{s}})$$



Bounds<sup>1</sup>:

$$\mathcal{LB} = C - \frac{1}{2Z_{k+1:i}} \mathbb{E} \left[ \ln \frac{f \left( \Lambda_k^{\text{Aug}-}, A_i^s \right) \cdot f \left( \Lambda_k^{\text{Aug}-}, A_i^{\bar{s}} \right)}{|\Lambda_k^{\text{Aug}-}|} \right]$$

$$\mathcal{UB} = C - \frac{1}{2Z_{k+1:i}} \mathbb{E} \left[ \ln f \left( \Lambda_k^{\text{Aug}-}, A_i^s \right) \right],$$

$$f(\Lambda, A) \triangleq |\Lambda + A^T A|$$

Reduced complexity wrt rAMDL<sup>2</sup>:  
 $O\left(\frac{m^3}{4}\right)$  vs  $O(m^3)$

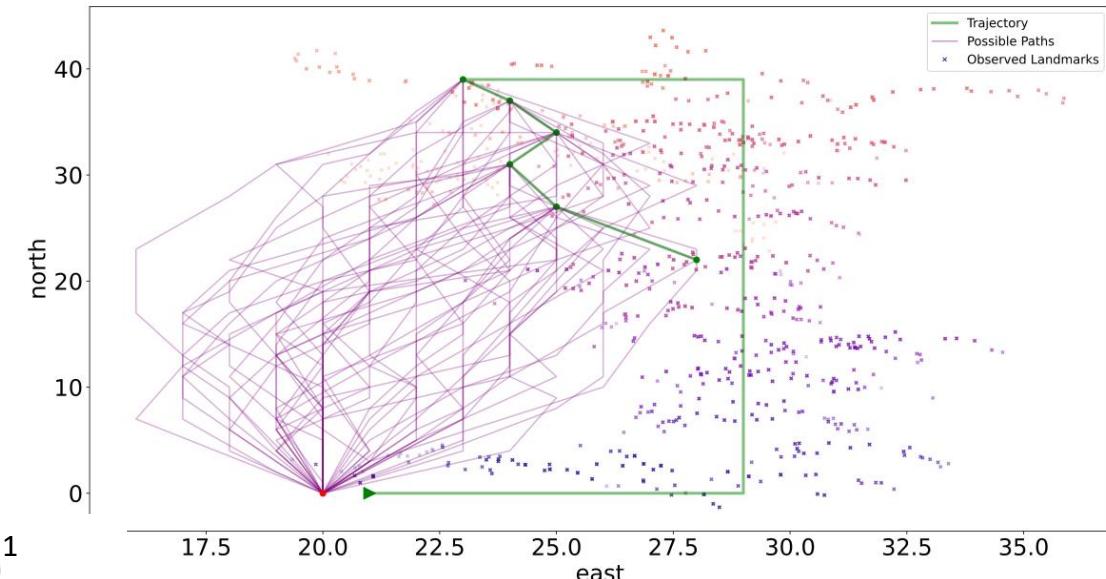
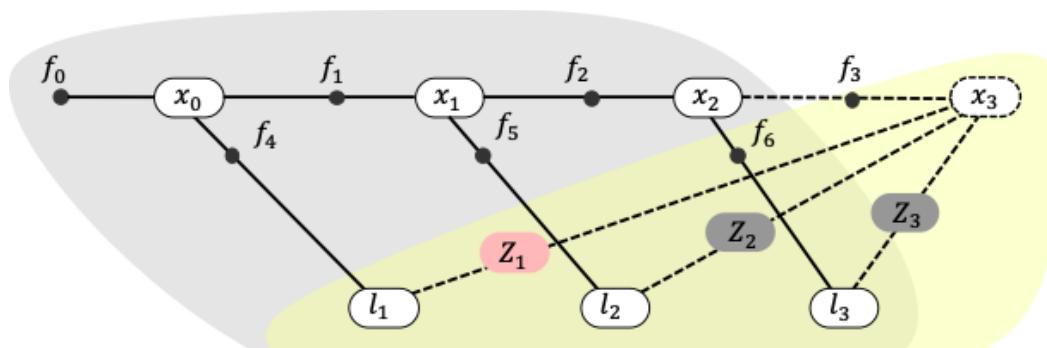
<sup>1</sup>T. Yotam and V. Indelman, "Measurement Simplification in p-POMDP with Performance Guarantees," IEEE T-RO'24.

<sup>2</sup>D. Kopitkov and V. Indelman, "No Belief Propagation Required: Belief Space Planning in High-Dimensional State Spaces via Factor Graphs, Matrix Determinant Lemma and Re-use of Calculation," IJRR'17.

<sup>2</sup>D. Kopitkov and V. Indelman, "General-purpose incremental covariance update and efficient belief space planning via a factor-graph propagation action tree", IJRR'19.

# Partitioning of a Multivariate Observation Space

## Application to Active SLAM



# Paths	# Factors	RP	rAMDL <sup>2</sup>	MP (ours) <sup>1</sup>
100	2956	No	$11.521 \pm 0.537$	<b><math>6.888 \pm 0.155</math></b>
100	2956	Yes	$24.636 \pm 1.381$	<b><math>11.758 \pm 0.372</math></b>
100	5904	Yes	$84.376 \pm 14.458$	<b><math>32.069 \pm 4.913</math></b>

Table: Total planning time in seconds (lower is better)

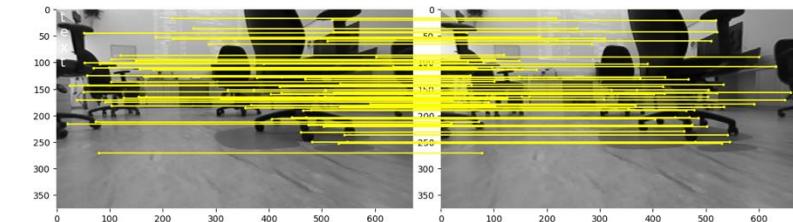
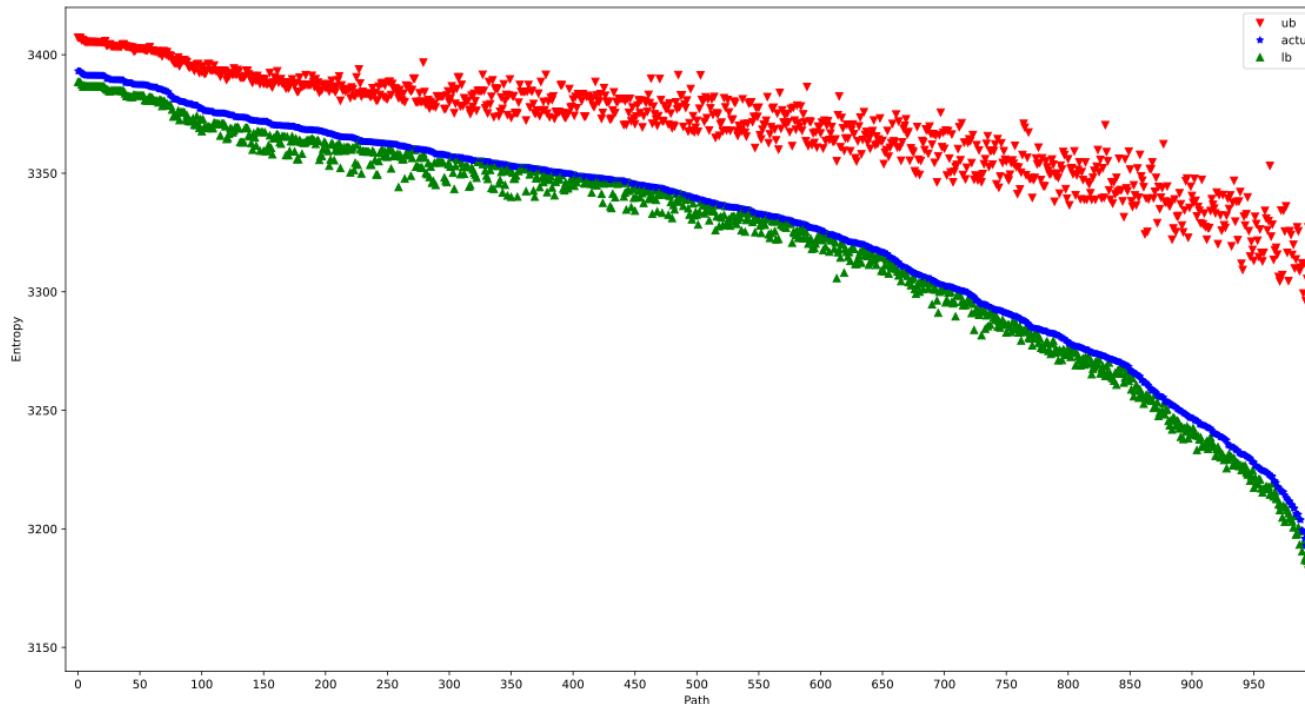
<sup>1</sup>T. Yotam and V. Indelman, "Measurement Simplification in p-POMDP with Performance Guarantees," IEEE T-RO'24.

<sup>2</sup>D. Kopitkov and V. Indelman, "No Belief Propagation Required: Belief Space Planning in High-Dimensional State Spaces via Factor Graphs, Matrix Determinant Lemma and Re-use of Calculation," IJRR'17.

<sup>2</sup>D. Kopitkov and V. Indelman, "General-purpose incremental covariance update and efficient belief space planning via a factor-graph propagation action tree", IJRR'19.

# Partitioning of a Multivariate Observation Space

## Application to Active SLAM



Method	time [sec]
MP (ours) <sup>1</sup>	<b>585.507 ± 27.153</b>
rAMD <sup>2</sup>	802.545 ± 25.651
iSAM2 <sup>3</sup>	1764.835 ± 26.521

Table: Total planning time in seconds (lower is better)

<sup>1</sup>T. Yotam and V. Indelman, "Measurement Simplification in p-POMDP with Performance Guarantees," IEEE T-RO'24.

<sup>2</sup>D. Kopitkov and V. Indelman, "No Belief Propagation Required: Belief Space Planning in High-Dimensional State Spaces via Factor Graphs, Matrix Determinant Lemma and Re-use of Calculation," IJRR'17.

<sup>2</sup>D. Kopitkov and V. Indelman, "General-purpose incremental covariance update and efficient belief space planning via a factor-graph propagation action tree", IJRR'19.

<sup>3</sup>M. Kaess, et al., "iSAM2: Incremental smoothing and mapping using the Bayes tree," IJRR'12.

# Simplification of Decision-Making Problems

---

## Concept:

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

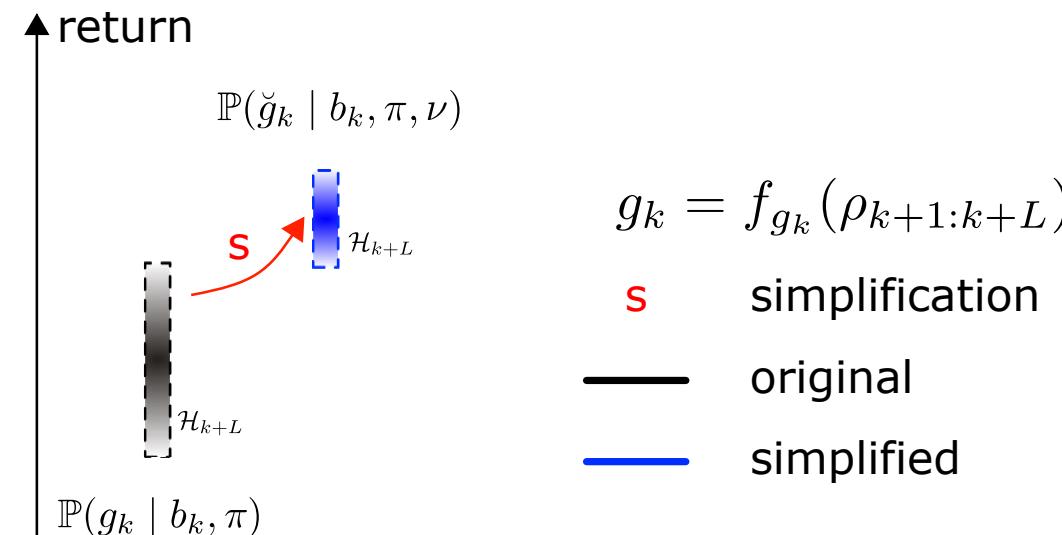
## Specific simplifications include:

- 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

# Simplification of Risk Averse POMDP Planning

---

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



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

# Probabilistically Constrained Belief Space Planning

---

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

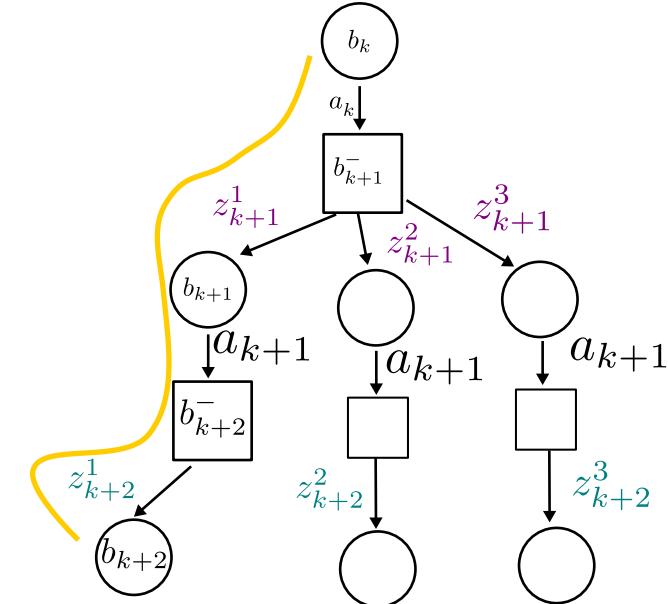
subject to  $P(c(b_{k:k+L}; \phi, \delta) = 1 | b_k, \pi_{k+}) \geq 1 - \epsilon$

Information gain<sup>1</sup>:

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

Safety<sup>2</sup>:

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



<sup>1</sup>A. Zhitnikov and V. Indelman, "Simplified Continuous High Dimensional Belief Space Planning with Adaptive Probabilistic Belief-dependent Constraints," T-RO'24.

<sup>2</sup>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:

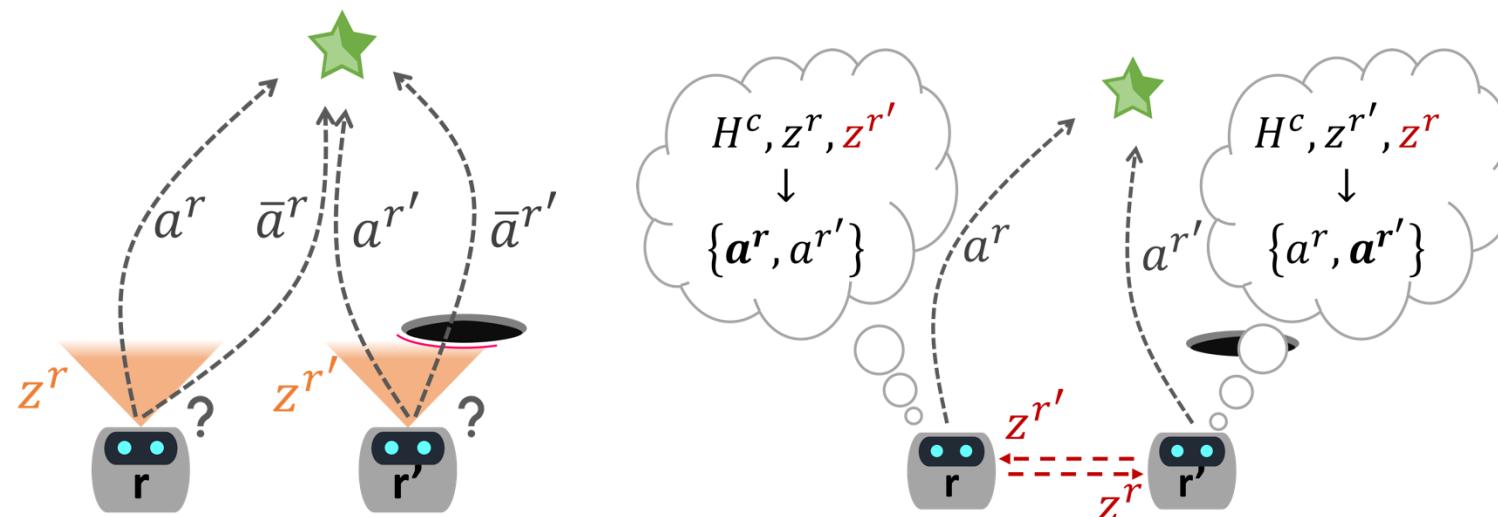
- Identify and solve a **simplified (computationally) easier** decision-making problem
- Provide performance guarantees

## 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

# Multi-Robot Belief Space Planning

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



# Multi-Robot Cooperative BSP with Inconsistent Beliefs

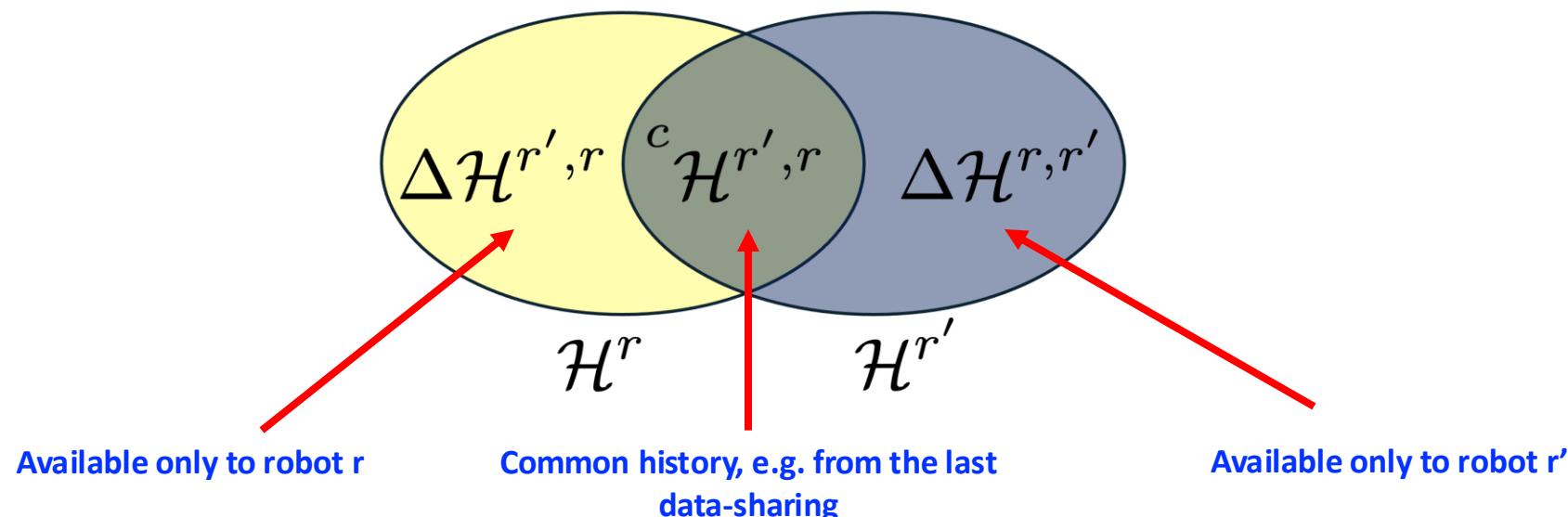
**What happens when data-sharing capabilities between the robots are limited?**

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

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

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

$$\mathcal{H}_k^r \neq \mathcal{H}_k^{r'}$$



# 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

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

- 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:

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

# Multi-Robot Cooperative BSP with Inconsistent Beliefs

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

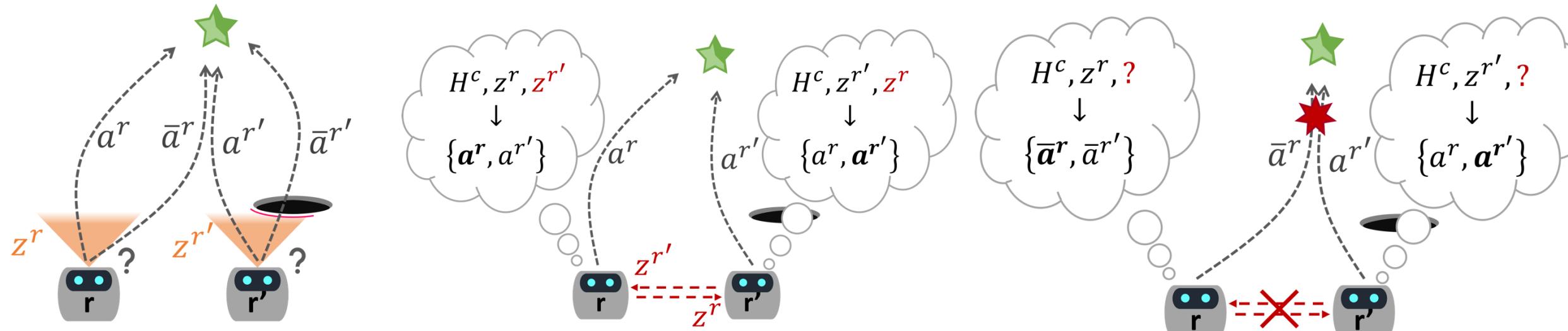
- Histories & beliefs of the robots may **differ** due to limited data-sharing capabilities

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

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

$$\mathcal{H}_k^r \neq \mathcal{H}_k^{r'}$$

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



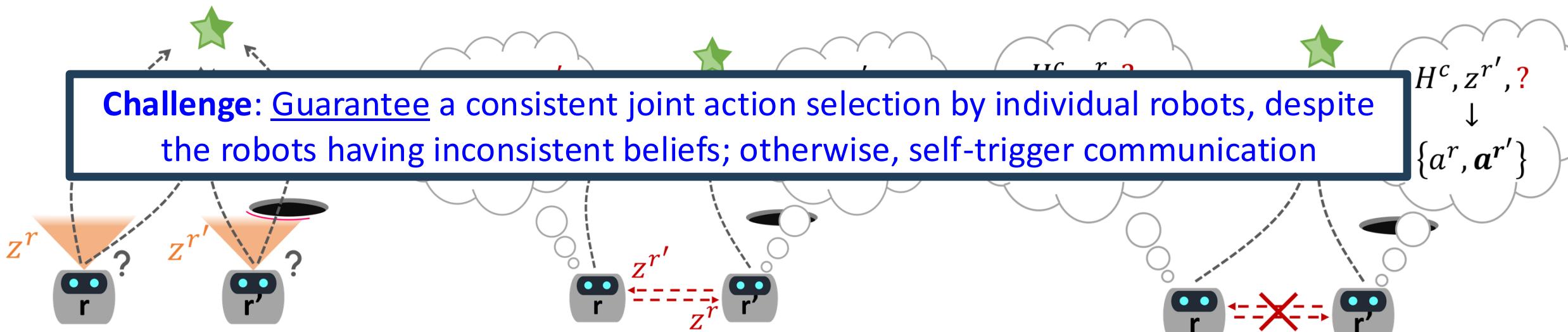
# Multi-Robot Cooperative BSP with Inconsistent Beliefs

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

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

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

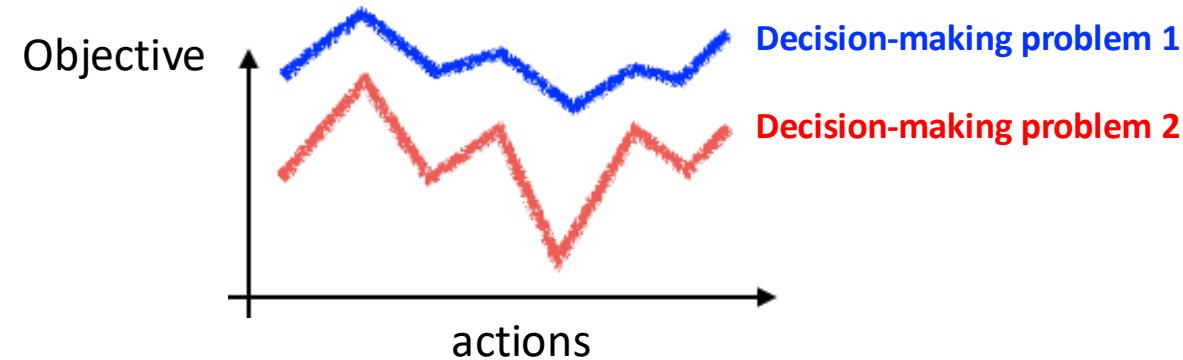
- **Can lead to a lack of coordination and unsafe and sub-optimal actions**



# Action Consistency

---

- If two decision-making problems have the **same action preference**, this implies both have the **same best action regardless of the actual objective/value function values**

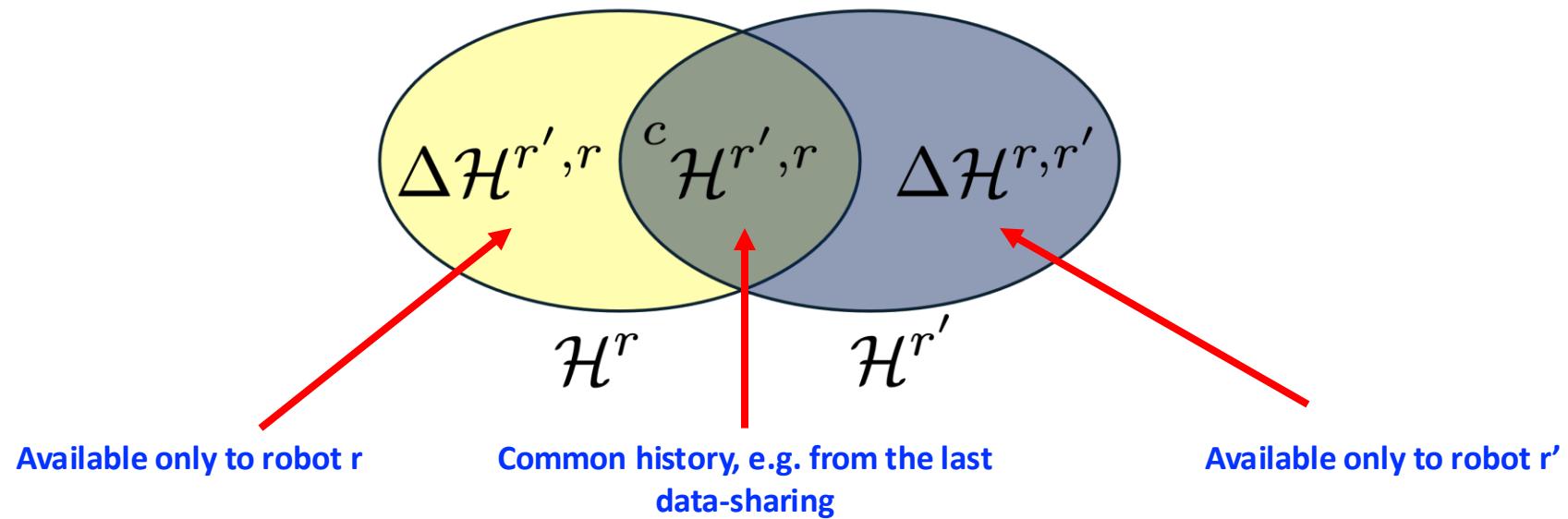


- **Key idea:** to guarantee consistent multi-robot decision-making, each robot
  - reasons about its own and other robots' action preferences while accounting for the **missing information** between the robots
  - checks if for all these realizations, we get the same best joint action

# Decentralized Verification of Multi-Robot Action Consistency (MR-AC)

---

- From the perspective of robot  $r$ , MR-AC holds if the selected joint actions are the same based on:
  - Its local information
  - What it perceives about the reasoning of the other robot  $r'$
  - What it perceives about the reasoning of itself perceived by the other robot  $r'$



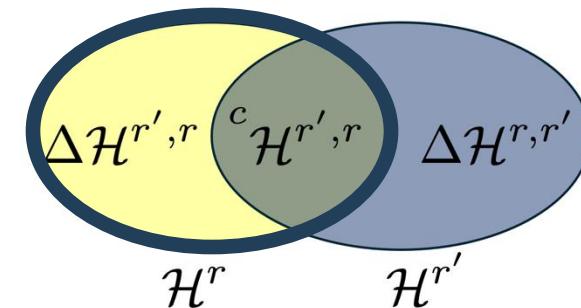
T. Kundu, M. Rafaeli, and V. Indelman, "Multi-Robot Communication-Aware Cooperative Belief Space Planning with Inconsistent Beliefs: An Action-Consistent Approach," IROS'24.

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.

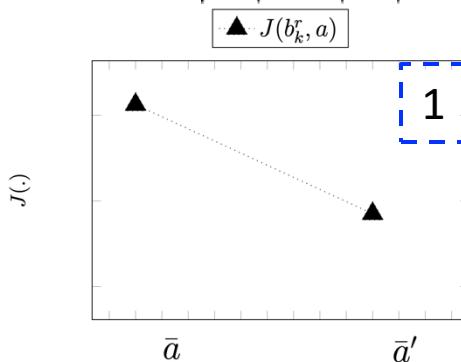
# Decentralized Verification of Multi-Robot Action Consistency (MR-AC)

---

- From the perspective of robot  $r$ , MR-AC holds if the selected joint actions are the same based on:
  - Its local information



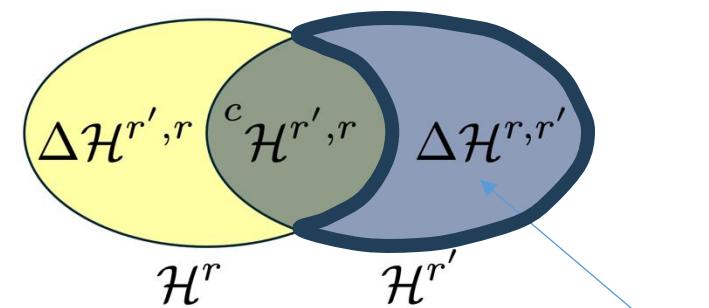
Toy example for  $|\mathcal{A}| = |\mathcal{Z}| = 2$ :



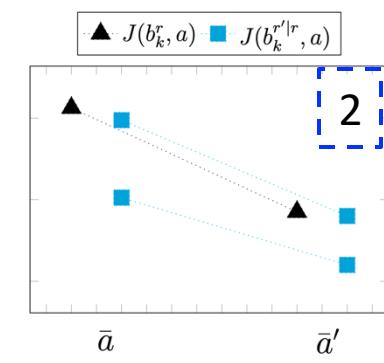
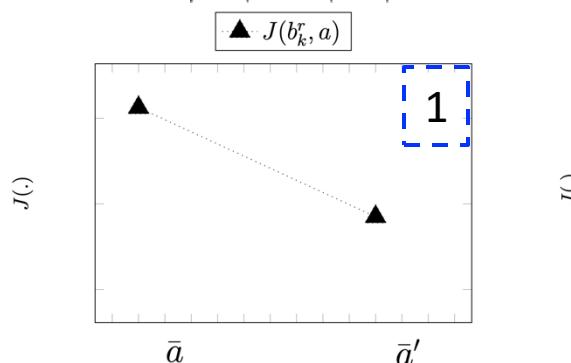
# Decentralized Verification of Multi-Robot Action Consistency (MR-AC)

---

- From the perspective of robot  $r$ , MR-AC holds if the selected joint actions are the same based on:
  - Its local information
  - What it perceives about the reasoning of the other robot  $r'$



Toy example for  $|\mathcal{A}| = |\mathcal{Z}| = 2$ :

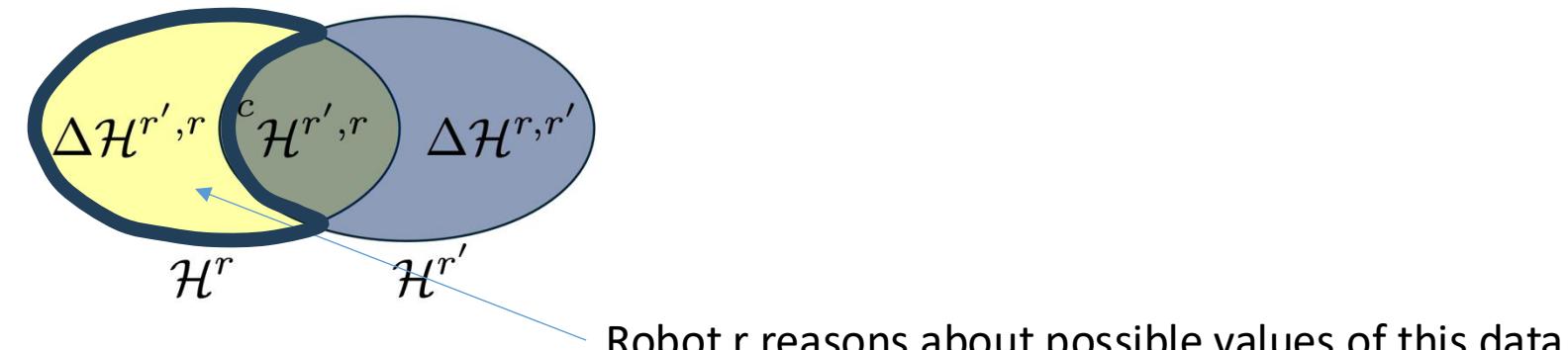


Robot  $r$  reasons about possible values of this data

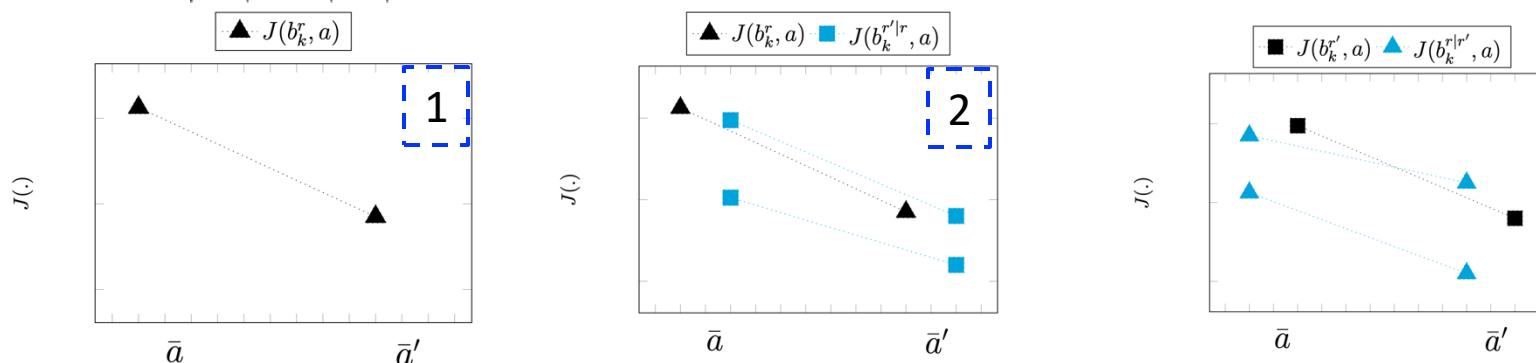
# Decentralized Verification of Multi-Robot Action Consistency (MR-AC)

---

- From the perspective of robot  $r$ , MR-AC holds if the selected joint actions are the same based on:
  - Its local information
  - What it perceives about the reasoning of the other robot  $r'$
  - What it perceives about the reasoning of itself perceived by the other robot  $r'$**



Toy example for  $|\mathcal{A}| = |\mathcal{Z}| = 2$ :



# Decentralized Verification of Multi-Robot Action Consistency (MR-AC)

- From the perspective of robot  $r$ , MR-AC holds if the selected joint actions are the same based on:
  - Its local information
  - What it perceives about the reasoning of the other robot  $r'$
  - What it perceives about the reasoning of itself perceived by the other robot  $r'$**

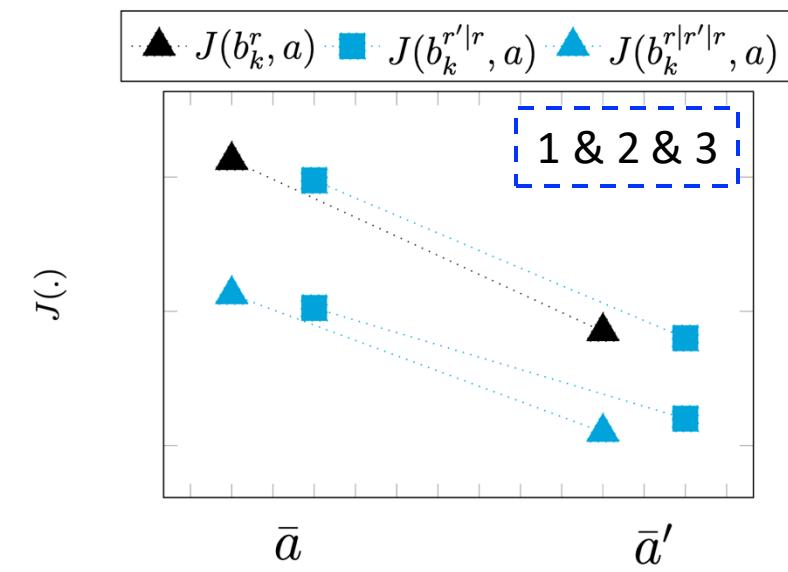
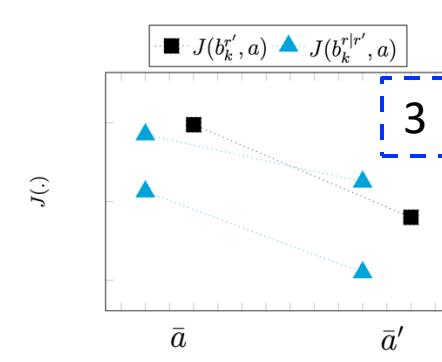
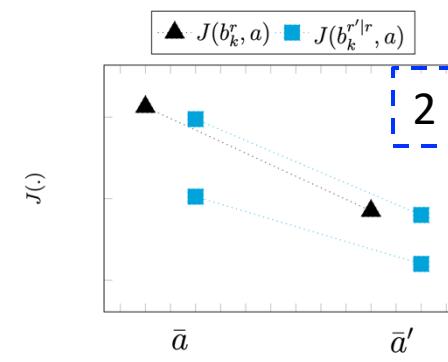
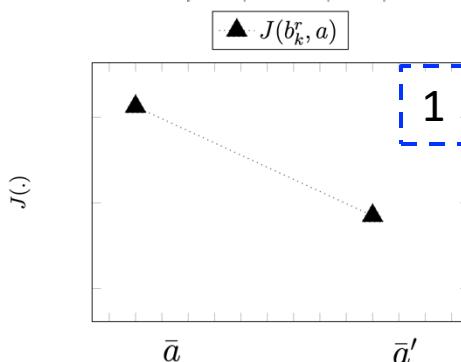
For each possible observation of  $r'$ ,  $\tilde{z}^r \in \Delta \mathcal{Z}_k^{r',r}$ , robot  $r$

constructs a plausible belief of robot  $r'$ :  $b_k^{r|r'|r}(\tilde{z}^r) \triangleq \mathbb{P}(x_k | {}^c\mathcal{H}_k^{r',r}, \tilde{z}^r)$

evaluates  $J(b_k^{r|r'|r}(\tilde{z}^r), a) \quad \forall a \in \mathcal{A}$

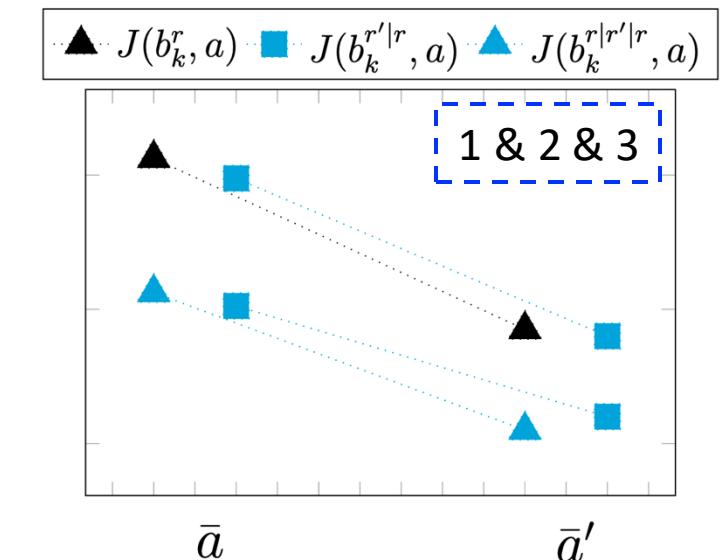
Checks if  $\bar{a}$  is selected

Toy example for  $|\mathcal{A}| = |\mathcal{Z}| = 2$ :



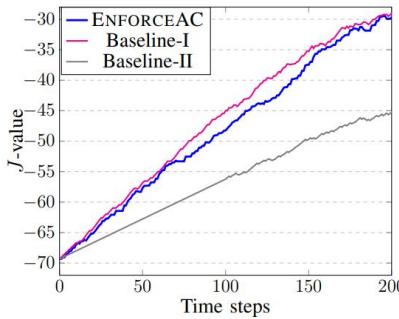
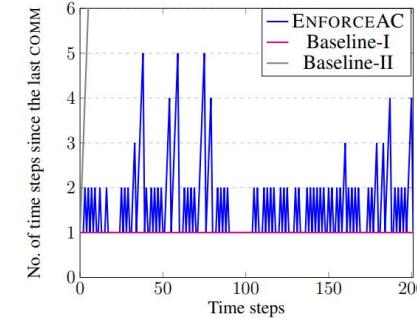
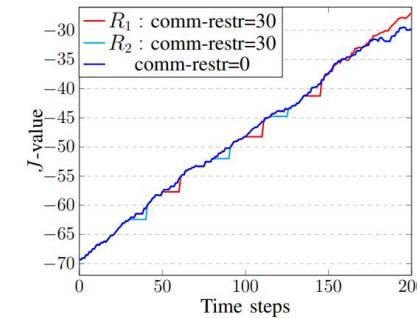
# Decentralized Verification of Multi-Robot Action Consistency (MR-AC)

- From the perspective of robot  $r$ , MR-AC holds if the selected joint actions are the same based on:
  1. Its local information
  2. What it perceives about the reasoning of the other robot  $r'$
  3. What it perceives about the reasoning of itself perceived by the other robot  $r'$
- Same best action in all cases?
  - Yes:** MR-AC is guaranteed to be satisfied
    - Robots are **guaranteed** to choose the same joint action
    - No further data sharing** is needed!
  - No:** self-trigger communication, share some data, repeat Steps 1-3



# Simulation Results (Search & Rescue Scenario)

- EnforceAC: our approach
- Baseline I: always communicate all data
- Baseline II: never communicate

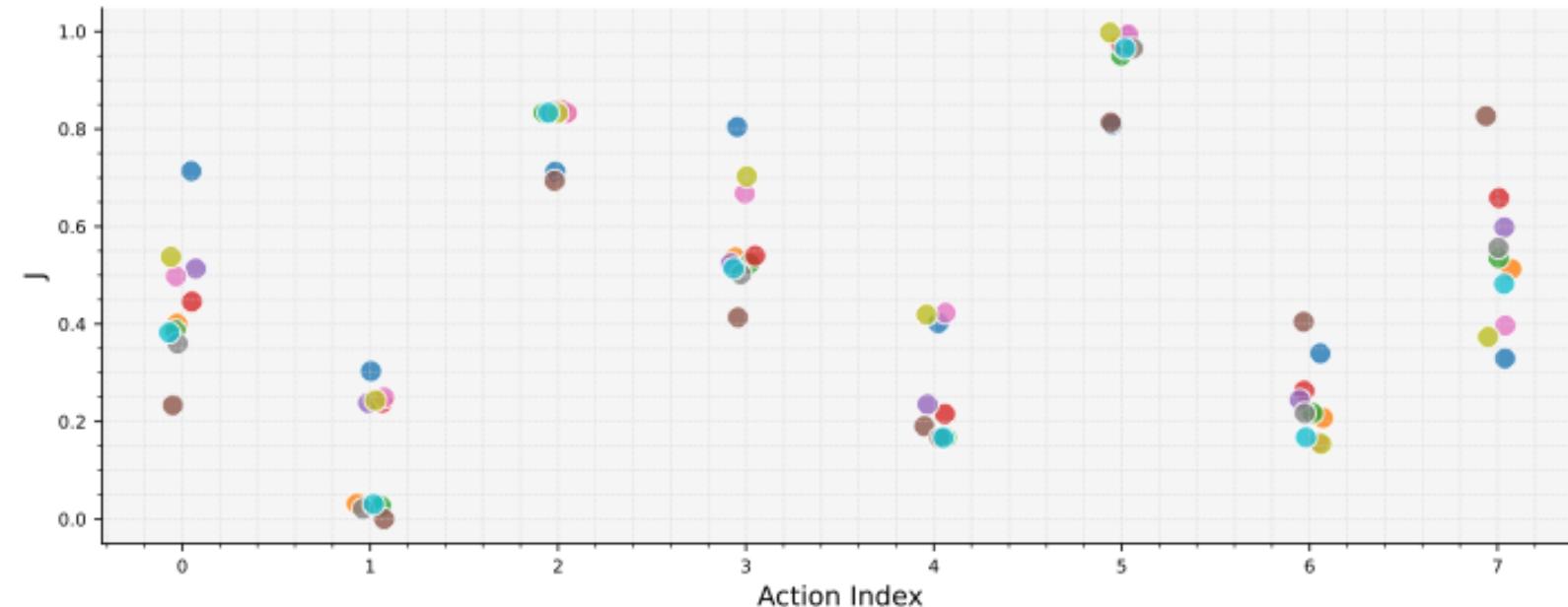
(a)  $comm\text{-}restr = 0$ (b)  $comm\text{-}restr = 0$ (c)  $comm\text{-}restr = 30$ 

NOT-AC (ACTION INCONSISTENCY), COMMS AND TIME FOR  $E = 200$ .

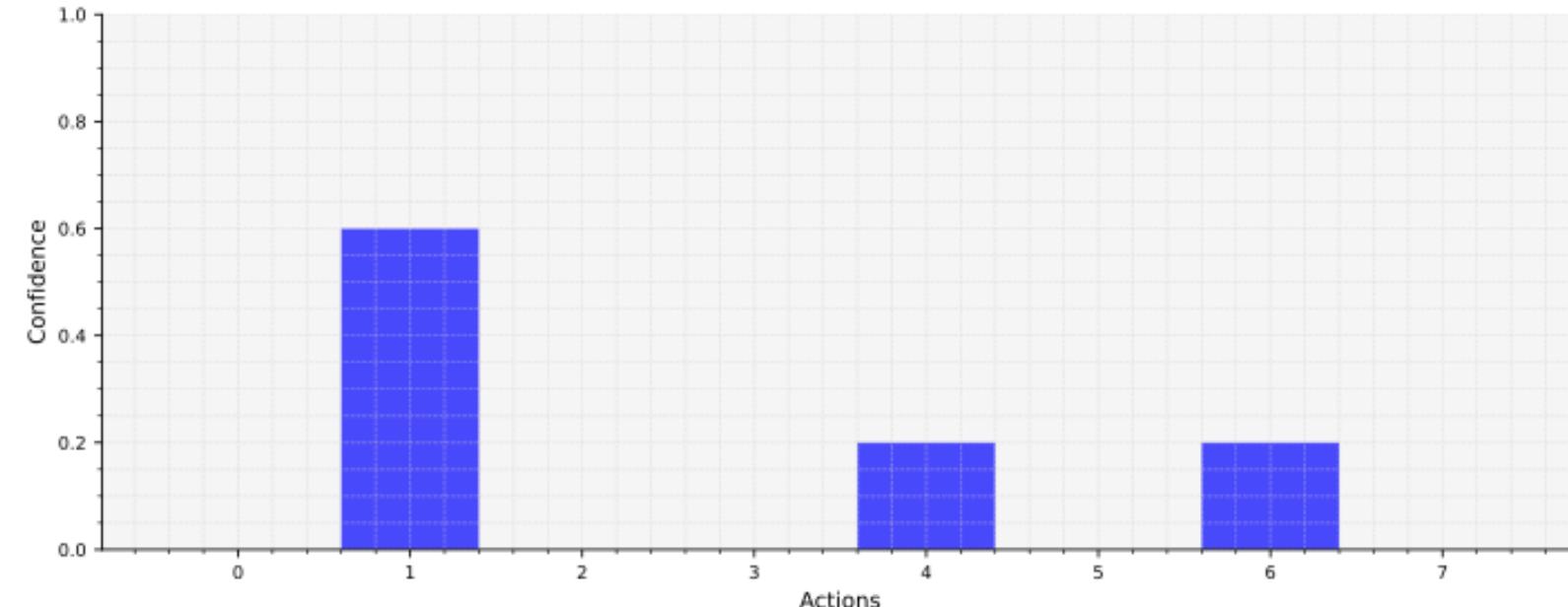
Input	Algorithm	Not-AC	COMM	Time
$comm\text{-}restr = 0$	Baseline-II	181	0	1.3s
$Motion\ prim. = 4$	Baseline-I	0	400	1.3s
$MaxEntropy\text{-}Init$	<b>ENFORCEAC</b>	0	238	12.4s
$comm\text{-}restr = 0$	Baseline-II	185	0	1.3s
$Motion\ prim. = 4$	Baseline-I	0	400	1.4s
$Entropy\text{-}Init$	<b>ENFORCEAC</b>	0	268	8.7s
$comm\text{-}restr = 0$	Baseline-II	194	0	3.6s
$Motion\ prim. = 8$	Baseline-I	0	400	3.5s
$MaxEntropy\text{-}Init$	<b>ENFORCEAC</b>	0	248	36.4s
$comm\text{-}restr = 0$	Baseline-II	188	0	3.6s
$Motion\ prim. = 8$	Baseline-I	0	400	3.6s
$Entropy\text{-}Init$	<b>ENFORCEAC</b>	0	278	31.1s
$comm\text{-}restr = 20$	Baseline-II	194	0	3.3s
$Motion\ prim. = 8$	Baseline-I	14	360	4.3s
$MaxEntropy\text{-}Init$	<b>ENFORCEAC</b>	13	224	94.9s
$comm\text{-}restr = 20$	Baseline-II	188	0	3.2s
$Motion\ prim. = 8$	Baseline-I	14	360	3.6s
$Entropy\text{-}Init$	<b>ENFORCEAC</b>	10	251	31.2s
$comm\text{-}restr = 30$	Baseline-II	188	0	3.4s
$Motion\ prim. = 8$	Baseline-I	22	340	4.0s
$MaxEntropy\text{-}Init$	<b>ENFORCEAC</b>	20	238	46.9s

## Probabilistic MRAC

**Step 2:**  
Sampled measurement realizations



MRAC probability for each action

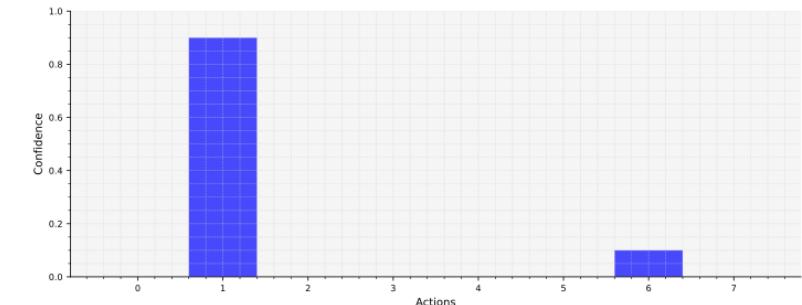
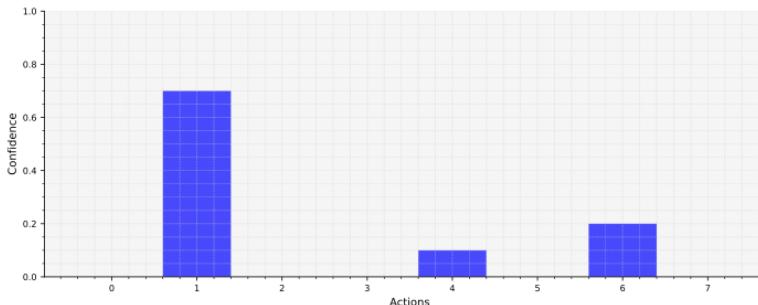
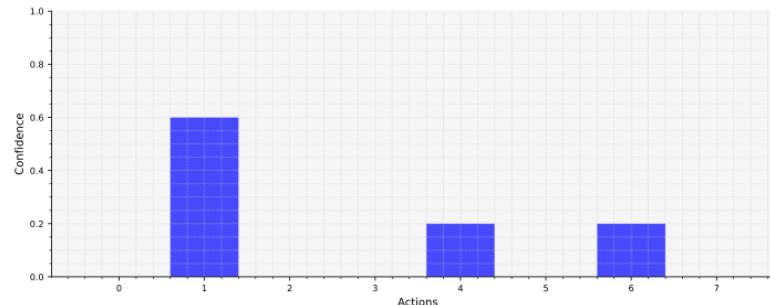
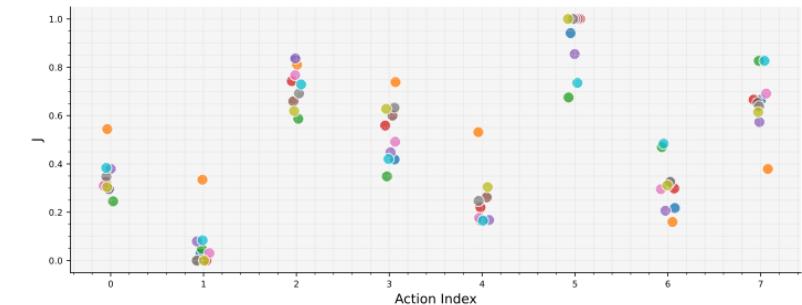
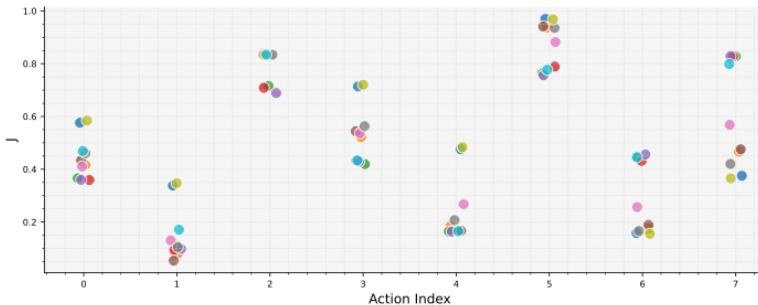
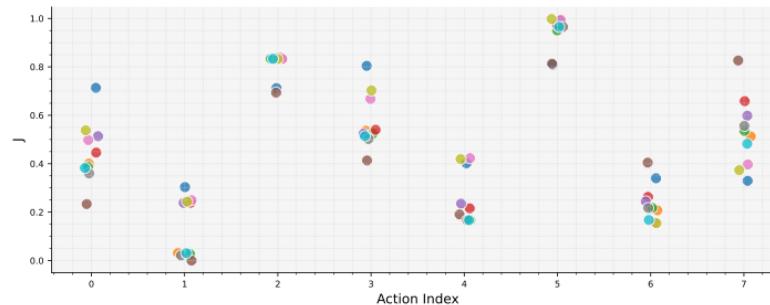


T. Kundu, M. Rafaeli, and V. Indelman, "Multi-Robot Communication-Aware Cooperative Belief Space Planning with Inconsistent Beliefs: An Action-Consistent Approach," IROS'24.

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.

## Probabilistic MRAC

data sharing



# Agenda

## Towards Scalable Online Decision Making Under Uncertainty in Partially Observable Environments

Experience Reuse in POMDP Planning

POMDP Planning with Hybrid Beliefs

Simplification of POMDP with Formal Guarantees

Multi-agent POMDP Planning with Inconsistent Beliefs

# Thank You

---

