Simplified Continuous High Dimensional Belief Space Planning with Adaptive Probabilistic Belief-dependent Constraints

Andrey Zhitnikov and Vadim Indelman

IEEE Transactions on Robotics (T-RO), 2024

TECHNION Israel Institute of Technology

ICRA2024 May 13-17, Yokohama, Japan

ANPL Autonomous Navigation and Perception Lab

May 22, 2024

Belief Dependent POMDP

Continuous State space

Candidate action sequences

Continuous Obs. space

Discount factor

Prior

$$\langle \mathcal{X}, \mathcal{A}, \mathcal{Z}, T, O, \rho, \gamma, b_0 \rangle$$

$$\mathbb{P}_T(x'|x,a)$$

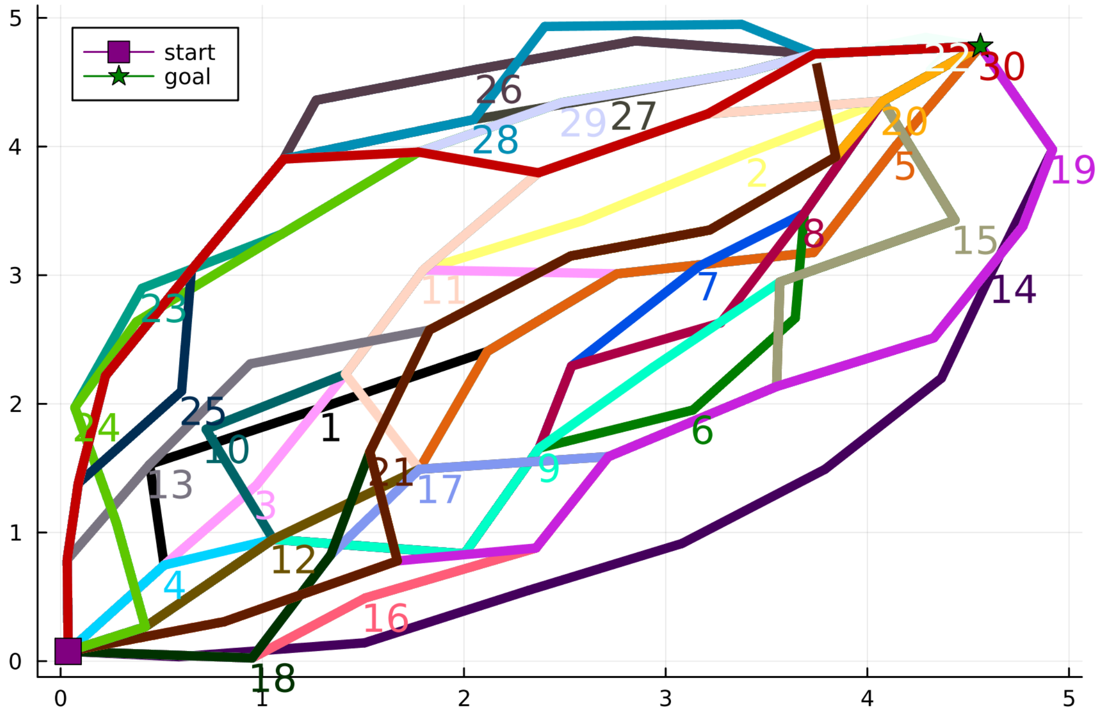
$$\mathbb{P}_Z(z|x)$$

Belief-dependent reward

$$b_k(x_k) \triangleq \mathbb{P}(x_k|b_0, a_{0:k-1}, z_{1:k})$$

Candidate action sequences

$$\mathcal{A} = \{a^i_{k:k+L^i-1}\}_{i=1}^{|\mathcal{A}|}$$



Belief-dependent rewards

Information-theoretic rewards

- Minus Differential entropy $\rho(b) = \mathbb{E}_{x\sim b(x)}\left[\log b(x)\right]$
- Minus Trace of the Covariance Matrix $\rho(b) = -\text{Trace}(\Sigma(b))$

State-dependent rewards (sometimes also reduce uncertainty) $\rho(b) = -\mathbb{E}_{x\sim b}\left[\underbrace{\|x - x^{\text{g}}\|_2^2}_{r(x)}\right]$

$$\boxed{b_k(x_k) \triangleq \mathbb{P}(x_k|b_0, a_{0:k-1}, z_{1:k})}$$

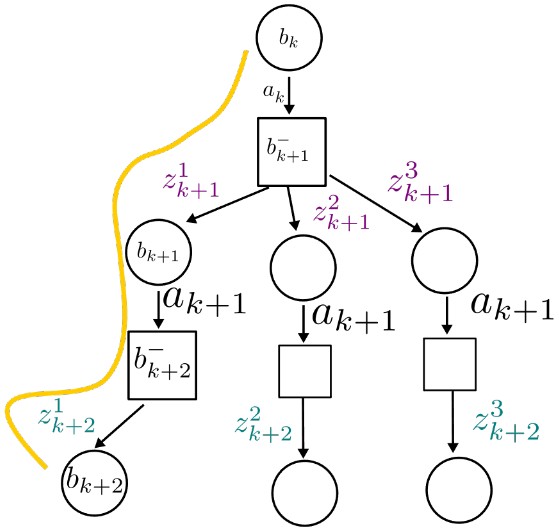
Existing Approaches and the gap

Information-theoretic payoff $\phi(b_\ell, b_{\ell+1}) = -\mathrm{h}(b_{\ell+1}) + \mathrm{h}(b_\ell)$

Differential entropy, uncertain map

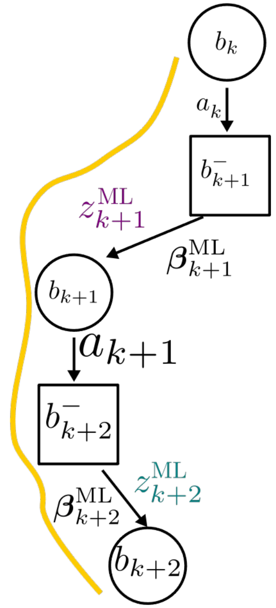
Our:

$$\mathbb{E}_{z_{k+1:k+L}}\left[\sum_{\ell=k}^{k+L-1} \phi(b_\ell, b_{\ell+1}) \middle| b_k, a_{k+}\right] \geq \delta$$



Maximum likely observation:

$$\left(\sum_{\ell=k}^{k+L-1} \phi(b^{\mathrm{ML}}_\ell, b^{\mathrm{ML}}_{\ell+1})\right) \geq \delta$$



Contributions of this work

- We utilize our Probabilistically Constrained POMDP in the context of information-theoretic constraint;
- Maximize Value at Risk adaptively;
- We rigorously derive a theory of the simplification.

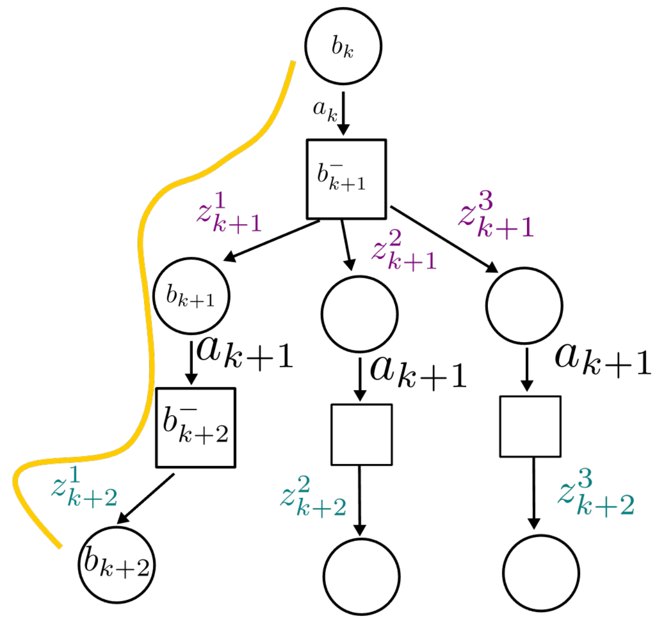
Our Probabilistic Belief Dependent Constraint

$$\max_{a_{k+} \in \mathcal{A}} \mathbb{E}\left[\sum_{t=k}^{k+L-1} \rho_{t+1} \middle| b_k, a_{k+} \right]$$

Probabilistic constraint sample approximation

$$\hat{\mathrm{P}}^{(m)}(c(b_{k:k+L};\phi,\delta)=1|b_k,a_{k+})=\frac{1}{m}\sum_{l=1}^{m}c(b^l_{k:k+L};\phi,\delta)$$

Maximal amount of expanded laces



$$c(b^l_{k:k+L};\phi,\delta)\sim\mathrm{P}(c|b_k,a_{k+})$$

$$c\in\{0,1\}$$

The inner constraint is violated

The inner constraint is satisfied

Probabilistic constraints, the bounds

$$\tilde{m} \leq m$$

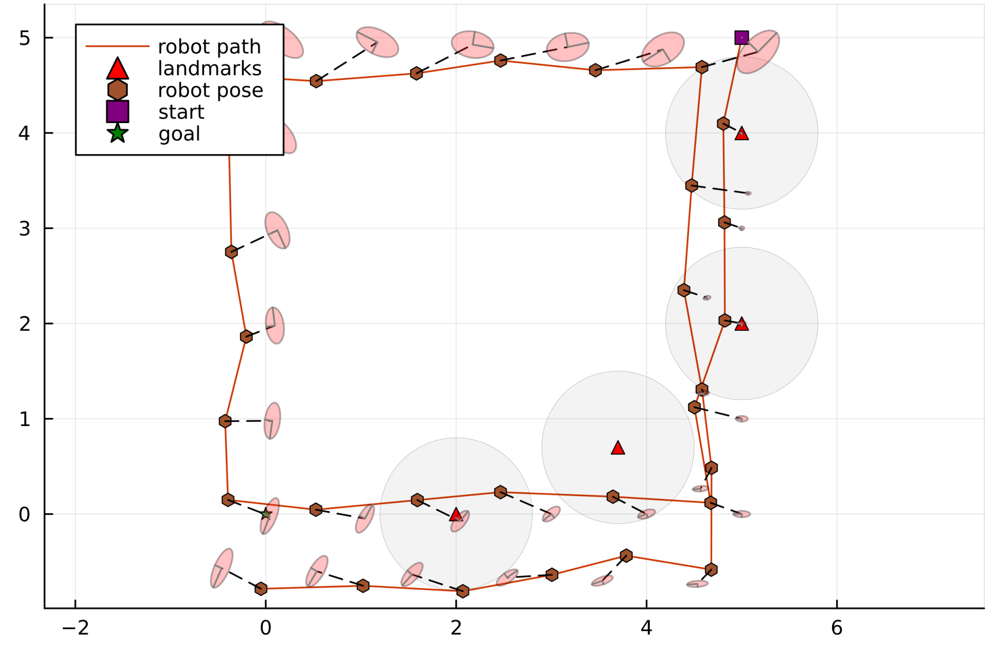
$$\text{lb} \ \leq \ \frac{1}{m}\sum_{l=1}^{m} c(b^{l}_{k:k+L};\phi,\delta) \ \leq \ \text{ub}$$

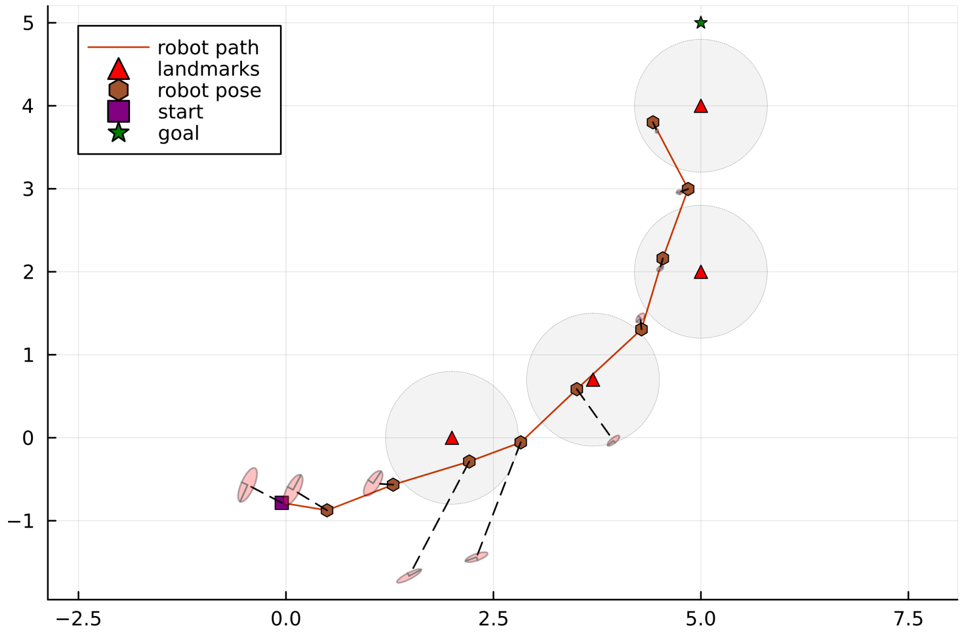
Cheaper to calculate

Expensive to calculate $\tilde{m}$ laces

Partial Results – Optimality under PC

active SLAM: 20% speedup



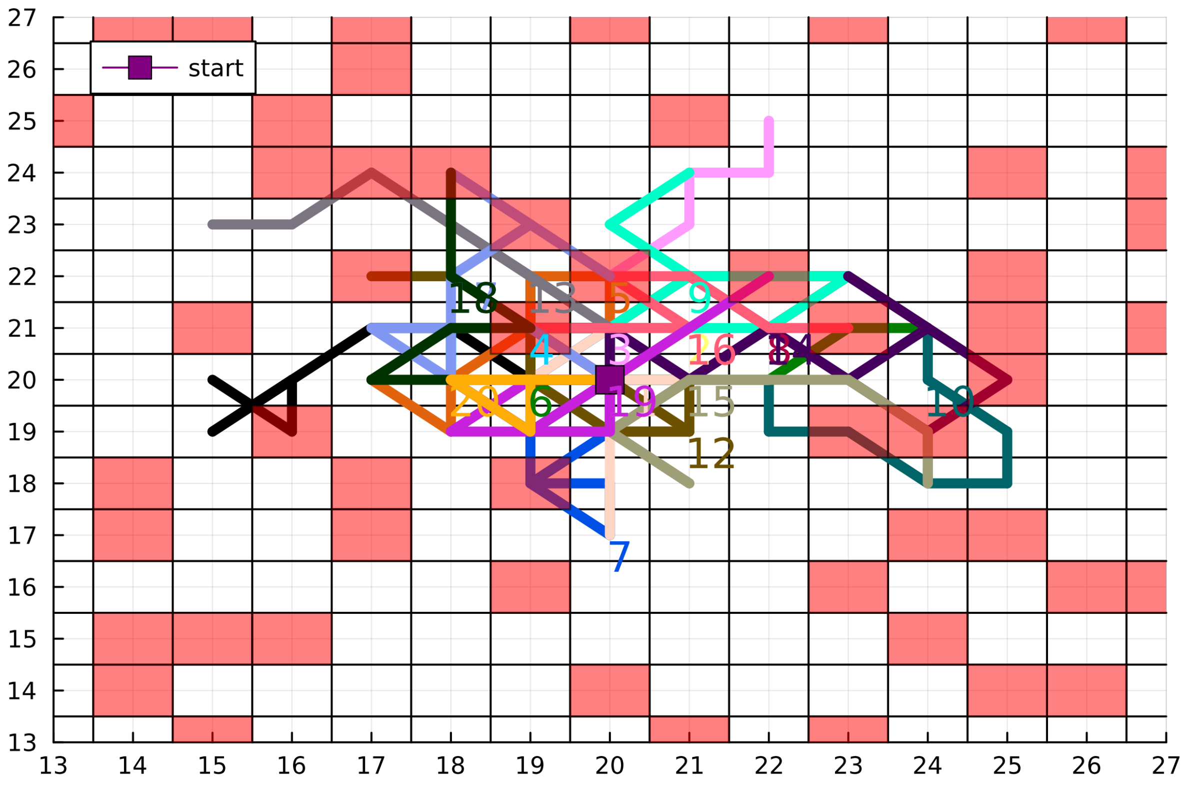


TECHNION Israel Institute of Technology

ANPL Autonomous Navigation and Perception Lab

Partial Results – Maximal Feasible Return

Sensor Deployment: sometimes 80% speedup



Summary

- We utilized our Probabilistically Constrained POMDP in the context of information-theoretic constraint;
- We provided the mechanism to stop exploration and accept profitable or discard unprofitable candidate action sequences early;
- We Maximized Value at Risk adaptively;
- We formulated simplification for the PC.

Thank you for your attention!