Anytime Probabilistically Constrained Provably Convergent Belief Space Planning: Supplementary Material

Andrey Zhitnikov[1] and Vadim Indelman[2,3]

[1] Technion Autonomous Systems Program (TASP)
Technion - Israel Institute of Technology, Haifa 32000, Israel
andreyz@campus.technion.ac.il
[2] Department of Aerospace Engineering
[3] Department of Data and Decision Science
Technion - Israel Institute of Technology, Haifa 32000, Israel
vadim.indelman@technion.ac.il

This supplementary document contains:

1. VaR and CVaR as safety cost operators 1;
2. Proofs 2;
3. Additional Simulations 3.

1 VaR and CVaR as Safety Cost Operators.

Suppose we have particle represented belief and the obstacle of the circular form Fig. 1. We have two robots teal and blue. Each particle of the belief is a concatenated position of each robot such that if x is a particle, the $x[1{:}2]$ corresponds to the first robot and $x[3:4]$ corresponds to the second robot. We shall check such a constraint for each robot separately. Let x denote the position of the one of the robots. The map $\mathcal{M}$ is given. We first define a distance from the safe space $\mathcal{Y}\subseteq\mathcal{M}$ as $\text{dist}(x,\mathcal{Y}) = \min_{y\in\mathcal{Y}}\|x-y\|_2$. We then define Value at Risk (VaR) as

$$\theta(b)\triangleq\text{VaR}^b_\alpha[\text{dist}(x,\mathcal{Y})]= \min\{\xi|\text{P}(\text{dist}(x,\mathcal{Y})\leq\xi)\geq 1-\alpha\}. \tag{1}$$

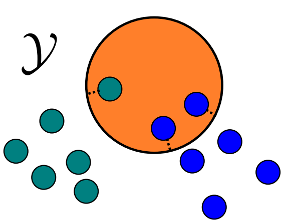


Fig. 1: Illustration of complex safety operators in multirobot setting.

The Conditional Value at Risk (CVaR) is specified as

$$\theta(b)\triangleq\text{CVaR}^b_\alpha[\text{dist}(x,\mathcal{Y})]= \mathbb{E}[\text{dist}(x,\mathcal{Y})|\{x:\text{dist}(x,\mathcal{Y})\geq \text{VaR}^b_\alpha[\text{dist}(x,\mathcal{Y})]\}]. \tag{2}$$

Both of these operators are **cost** operators.

2 Proofs

2.1 Proof of Lemma 1 (Representation of the Value function).

Before we begin, let us clarify that when we write $\{\mathbb{P}^{\pi}_{\ell}(a_{\ell}|b_{\ell})\}_{\ell=1}^{L-1}$, the a_{ℓ} and b_{ℓ} inside the $\{\mathbb{P}_{\ell}(a_{\ell}|b_{\ell})\}_{\ell=1}^{L-1}$ can be a random variables for all relevant ℓ or corresponding realizations. However, $\{\mathbb{P}^{\pi}_{\ell}\}_{\ell=1}^{L-1}$ is the series of distributions of length $L-1$ and corresponding actions and beliefs are unknown. In addition, we remind to the reader that $\pi_{\ell}(a_{\ell},b_{\ell})=\mathbb{P}^{\pi}_{\ell}(a_{\ell}|b_{\ell}) \quad \forall \ell \in 1:L-1$ and $\pi=\{\mathbb{P}^{\pi}_{\ell}\}_{\ell=1}^{L-1}$.

$$\begin{gathered}\mathbb{E}\big[\textstyle\sum_{\ell=0}^{L-1}\gamma^{\ell+1}\rho_{\ell+1}(b_{\ell},a_{\ell},b_{\ell+1})\big|b_0,\pi\big]=\\ \textstyle\sum_{\ell=0}^{L-1}\gamma^{\ell+1}\mathbb{E}\big[\rho_{\ell+1}\big|b_0,\pi\big]=\sum_{\ell=0}^{L-1}\gamma^{\ell+1}\mathbb{E}\big[\rho_{\ell+1}\big|b_0,\pi\big]\end{gathered}\tag{3}$$

$$\begin{gathered}\mathbb{E}\big[\rho_{\ell+1}\big|b_0,\pi\big]=\int_{\rho_{\ell+1}}\rho_{\ell+1}\mathbb{P}(\rho_{\ell+1}|b_0,\{\mathbb{P}^{\pi}_{i}\}_{i=0}^{L-1})\mathrm{d}\rho_{\ell+1}=\\ \int\limits_{\rho_{\ell+1}}\rho_{\ell+1}\int\limits_{\substack{b_{1:\ell}\\ a_{0:\ell}\in\times_{i=1}^{\ell}\mathcal{A}}}\mathbb{P}(\rho_{\ell+1},b_{1:\ell},a_{1:\ell}|b_0,\{\mathbb{P}^{\pi}_{i}\}_{i=0}^{L-1})\mathrm{d}b_{1:\ell}\mathrm{d}a_{0:\ell}\mathrm{d}\rho_{\ell+1}=\\ \int_{\rho_{\ell+1}}\rho_{\ell+1}\int\limits_{\substack{b_{1:\ell}\\ a_{0:\ell}\in\times_{i=1}^{\ell}\mathcal{A}}}\mathbb{P}(\rho_{\ell+1}|b_{0:\ell},a_{0:\ell})\\ \mathbb{P}(b_{1:\ell},a_{0:\ell}|b_0,\{\mathbb{P}^{\pi}_{i}\}_{i=0}^{L-1})\mathrm{d}b_{1:\ell}\mathrm{d}a_{0:\ell}\mathrm{d}\rho_{\ell+1}\end{gathered}\tag{4}$$

Previous equation equals to

$$\begin{gathered}\int\limits_{\substack{b_{1:\ell}\\ a_{0:\ell}\in\times_{i=1}^{\ell}\mathcal{A}}}\big(\int_{\rho_{\ell+1}}\rho_{\ell+1}\mathbb{P}(\rho_{\ell+1}|b_{0:\ell},a_{0:\ell})\mathrm{d}\rho_{\ell+1}\big)\\ \mathbb{P}(b_{1:\ell},a_{0:\ell}|b_0,\{\mathbb{P}^{\pi}_{i}\}_{i=0}^{L-1})\mathrm{d}b_{1:\ell}\mathrm{d}a_{0:\ell}\end{gathered}\tag{5}$$

We now use a chain rule from the future time back on $\mathbb{P}(b_{1:\ell},a_{0:\ell}|b_0,\{\mathbb{P}^{\pi}_{i}\}_{i=0}^{L-1})$ an got

$$\begin{gathered}\mathbb{E}\big[\rho_{\ell+1}\big|b_0,\pi\big]=\mathop{\mathbb{E}}_{a_0}\Big[\mathop{\mathbb{E}}_{b_1}\Big[\mathop{\mathbb{E}}_{a_1}\Big[\mathop{\mathbb{E}}_{b_2}\Big[\dots\\ \mathop{\mathbb{E}}_{a_{\ell}}\big[\mathbb{E}\big[\rho_{\ell+1}|b_{\ell},a_{\ell}\big]\big|b_{\ell},\pi_{\ell}\big]\dots\big|b_1,a_1\big]\big|b_1,\pi_1\big]\big|b_0,a_0\big]\big|b_0,\pi_0\big].\end{gathered}\tag{6}$$

■

2.2 Proof of Theorem 1 (Necessary condition for theoretical posteriors to be safe)

For the necessary condition we prove the inverse implication. Suppose that $\forall z_{\ell}\in\mathcal{Z}$ it holds that $\mathrm{P}\big(\{x_{\ell}\in\mathcal{X}^{\mathrm{safe}}_{\ell}\}\big|h^{-}_{\ell},z_{\ell}\big)\geq\delta$. We arrive at

$$(\int_{z_{\ell}\in\mathcal{Z}}\mathrm{P}\big(\{x_{\ell}\in\mathcal{X}^{\mathrm{safe}}_{\ell}\}\big|h^{-}_{\ell},z_{\ell}\big)\mathbb{P}(z_{\ell}|h^{-}_{\ell})\mathrm{d}z_{\ell})\geq\delta.\tag{7}$$

■

2.3 Proof of Theorem 2(Representation of Our Outer Constraint).

Before we begin, let us clarify that when we write

$$\mathrm{P}((\mathbf{1}_{\Phi_0^\delta}(b_0)\prod_{\ell=1}^{L}\mathbf{1}_{\Phi_\ell^\delta}(b_\ell)){=}1|b_0,a_0,\{\mathbb{P}_\ell^\pi\}_{\ell=1}^{L-1}),$$

the actions a_ℓ and the beliefs b_ℓ inside $\{\mathbb{P}_\ell^\pi\}_{\ell=1}^{L-1}$ are unknown random quantities. In addition, we remind to the reader that $\pi_\ell(a_\ell,b_\ell){=}\mathbb{P}_\ell^\pi(a_\ell|b_\ell)\quad\forall\ell{\in}1{:}L{-}1$ and $\pi{=}\{\mathbb{P}_\ell^\pi\}_{\ell=1}^{L-1}$. Moreover, in this paper each posterior belief is associated with corresponding propagated belief. Therefore we can rescind the explicit dependence of the indicator on propagated belief.

$$\begin{gathered}\mathbb{E}\Big[\mathbf{1}_{\Phi_0^\delta}(b_0)\prod_{\ell=1}^{L}\mathbf{1}_{\Phi_\ell^\delta}(b_\ell)|b_0,a_0,\{\mathbb{P}_\ell^\pi\}_{\ell=1}^{L-1}\Big]=\\ \int\limits_{\substack{b_{1:L}\\ a_{1:L-1}\in\times_{\ell=1}^{L-1}\mathcal{A}}}\mathbf{1}_{\Phi_0^\delta}(b_0)\prod_{\ell=1}^{L}\mathbf{1}_{\Phi_\ell^\delta}(b_\ell)\\ \mathbb{P}(b_{1:L},a_{1:L-1}|b_0,a_0,\{\mathbb{P}_\ell^\pi\}_{\ell=1}^{L-1})\mathrm{d}b_{1:L}\mathrm{d}a_{1:L-1}.\end{gathered}\tag{8}$$

Now, we need to handle $\mathbb{P}(b_{1:L},a_{0:L-1}|b_0,a_0,\{\mathbb{P}_\ell^\pi\}_{\ell=1}^{L-1})$. It holds that

$$\mathbb{P}(b_{1:L},a_{1:L-1}|b_0,a_0,\{\mathbb{P}_\ell^\pi\}_{\ell=1}^{L-1})$$

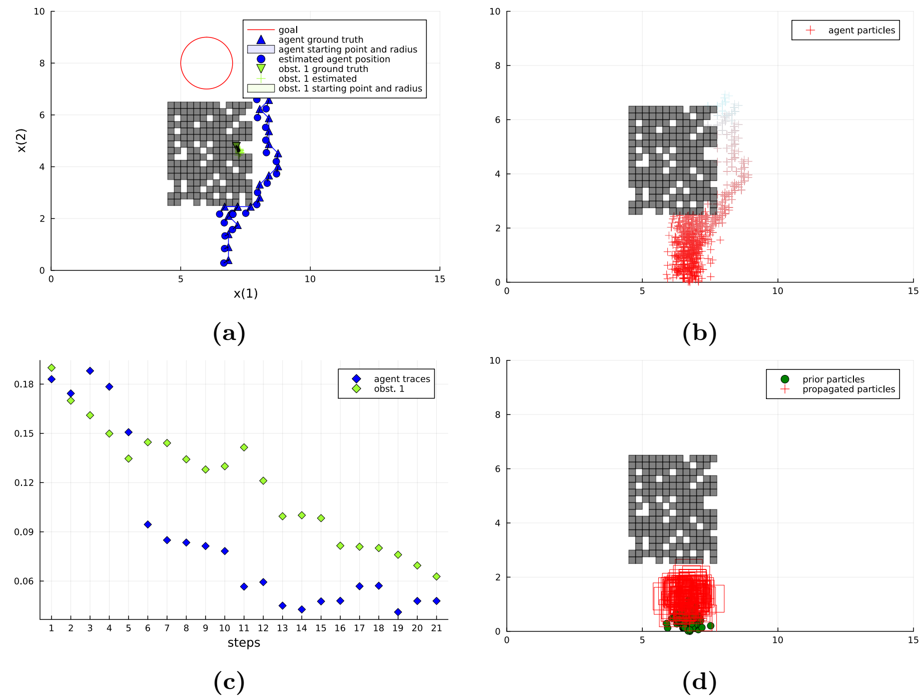
equals to

$$\begin{gathered}\mathbb{P}(b_{2:L},a_{2:L-1}|b_0,a_0,b_1,a_1,\mathbb{P}_1^\pi(a_1|b_1),\{\mathbb{P}_\ell^\pi\}_{\ell=2}^{L-1})\\ \mathbb{P}(b_1,a_1|b_0,a_0,\{\mathbb{P}_\ell\}_{\ell=1}^{L-1})=\\ \mathbb{P}(b_{2:L},a_{2:L-1}|b_1,a_1,\{\mathbb{P}_\ell^\pi\}_{\ell=2}^{L-1})\mathbb{P}_1^\pi(a_1|b_1)\mathbb{P}(b_1|b_0,a_0)=\\ \mathbb{P}(b_L|b_{L-1},a_{L-1})\prod_{\ell=1}^{L-1}\mathbb{P}_\ell^\pi(a_\ell|b_\ell)\mathbb{P}(b_\ell|b_{\ell-1},a_{\ell-1}).\end{gathered}\tag{9}$$

We now merge (8) and (9), and land at the desired result

$$\begin{gathered}\mathbf{1}_{\Phi_0^\delta}(b_0)\int\limits_{\substack{b_{1:L}\\ a_{1:L-1}\in\times_{i=1}^{L-1}\mathcal{A}}}\mathbf{1}_{\Phi_L^\delta}(b_L)\mathbb{P}(b_L|b_{L-1},a_{L-1})\\ \prod_{\ell=1}^{L-1}(\mathbb{P}_\ell^\pi(a_\ell|b_\ell)\mathbb{P}(b_\ell|b_{\ell-1},a_{\ell-1})\mathbf{1}_{\Phi_\ell^\delta}(b_\ell))\mathrm{d}b_{1:L}\mathrm{d}a_{1:L-1}=\\ \mathbf{1}_{\Phi_0^\delta}(b_0)\int_{b_1}\mathbb{P}(b_1|b_0,a_0)\mathbf{1}_{\Phi_1^\delta}(b_1)\int_{a_1}\mathbb{P}_1^\pi(a_1|b_1)(\ldots\\ \int_{b_L}\mathbf{1}_{\Phi_L^\delta}(b_L)\mathbb{P}(b_L|b_{L-1},a_{L-1})\mathrm{d}b_L\ldots)\mathrm{d}a_1\mathrm{d}b_1=\\ \mathbf{1}_{\Phi_0^\delta}(b_0)\mathop{\mathbb{E}}_{b_1}[\mathbf{1}_{\Phi_1^\delta}(b_1)\mathop{\mathbb{E}}_{a_1}[\mathop{\mathbb{E}}_{b_2}[\mathbf{1}_{\Phi_2^\delta}(b_2)\ldots\\ \mathbb{E}[\mathbf{1}_{\Phi_L^\delta}(b_L)|b_{L-1},a_{L-1}]\ldots|b_1,a_1]|b_1,\pi_1]|b_0,a_0]=\\ \mathbf{1}_{\Phi_0^\delta}(b_0)\mathop{\mathbb{E}}_{b_1}[\mathop{\mathbb{E}}_{a_1\sim\mathbb{P}_1^\pi(a_1|b_1)}[\\ \mathrm{P}((\prod_{\ell=1}^{L}\mathbf{1}_{\Phi_\ell^\delta}(b_\ell)){=}1|b_1,a_1,\pi)|b_1,\pi_1]|b_0,a_0].\end{gathered}$$

■



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

Fig. 2: This simulation setup is associated with Table 1. In this figure we plot one of the trials shown in Table 1. We nullify unsafe part of the belief in planning and run PC-SB-PFT-DPW. **(a)** Here, we plot the goal, agent ground truth, estimated agent positions and the obstacles; **(b)** Belief particles, where the colors symbolize the time instance; **(c)** Traces of the agent and the landmark (obstacle); **(d)** Visualization of the truncation. Here we move each particle of b_0 with action selected by the agent and plot the truncation region of the stochastic motion model.

3 Additional Simulations

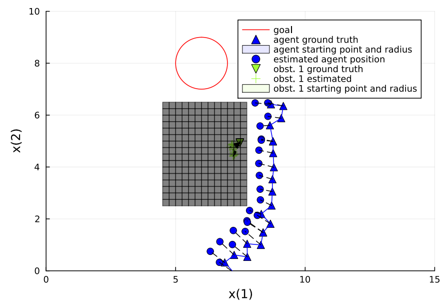
We now describe additional simulations we have done. We remind to the reader that in PC-PFT-DPW the operator ϕ comply with

$$\phi(b_\ell){=}\mathrm{P}\big(\{x_\ell{\in}\mathcal{X}^{\text{safe}}_\ell\}\big|b_\ell\big){=}\mathrm{P}\big(\{x_\ell{\in}\mathcal{X}^{\text{safe}}_\ell\}\big|b_0,a_{0:\ell-1},z_{1:\ell}\big) \tag{10}$$

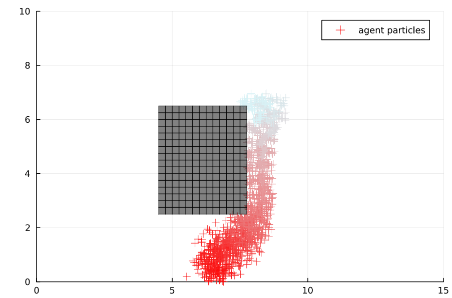
$$\phi(b^-_\ell){=}\mathrm{P}\big(\{x_\ell{\in}\mathcal{X}^{\text{safe}}_\ell\}\big|b^-_\ell\big){=}\mathrm{P}\big(\{x_\ell{\in}\mathcal{X}^{\text{safe}}_\ell\}\big|b_0,a_{0:\ell-1},z_{1:\ell-1}\big). \tag{11}$$

and in PC-SB-PFT-DPW if follows

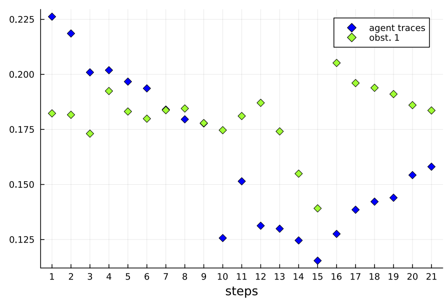
$$\begin{gathered}\phi(\bar{b}_\ell){=}\mathrm{P}\big(\{x_\ell{\in}\mathcal{X}^{\text{safe}}_\ell\}\big|\bar{b}_\ell\big){=}\\ \mathrm{P}\big(\{x_\ell{\in}\mathcal{X}^{\text{safe}}_\ell\}\big|b_0,a_{0:\ell-1},z_{1:\ell},\textstyle\bigcap_{i=0}^{\ell-1}\{x_i{\in}\mathcal{X}^{\text{safe}}_i\}\big)\end{gathered} \tag{12}$$



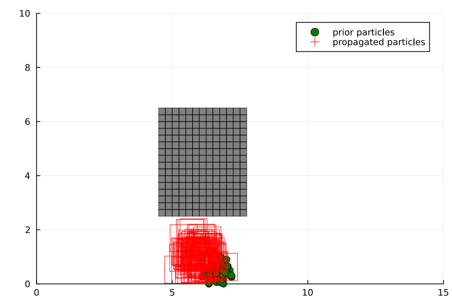
(a)



(b)



(c)



(d)

Fig. 3: This simulation setup is associated with Table 1 columns related to PC-SB-PFT-DPW and here we show one of the trials. In this figure we nullify unsafe part of the belief in planning. **(a)** Here, we plot the goal, agent ground truth, estimated agent positions and the obstacles; **(b)** Belief particles, where the colors symbolize the time instance; **(c)** Traces of the agent and the landmark (obstacle); **(d)** Visualization of the truncation. Here we move each particle of b_0 with action selected by the agent and plot the truncation region of the stochastic motion model.

$$\begin{aligned}\phi(\bar{b}_\ell^-)=\mathrm{P}\big(\{x_\ell\in\mathcal{X}_\ell^{\text{safe}}\}\big|\bar{b}_\ell^-\big)=\\ \mathrm{P}\big(\{x_\ell\in\mathcal{X}_\ell^{\text{safe}}\}\big|b_0, a_{0:\ell-1}, z_{1:\ell-1}, \textstyle\bigcap_{i=0}^{\ell-1}\{x_i\in\mathcal{X}_i^{\text{safe}}\}\big).\end{aligned} \tag{13}$$

3.1 SLAM

In our second setup we fill the complete rectangle with tiny obstacles in a random manner as debated (Fig. 4) in the manuscript. We show our results in Table 1. We did not obtained a significant difference in two approaches. Interestingly, as we see the safety is much challenging in this problem due to challenging robot localization with simultaneous mapping of uncertain single landmark. Additionally we see that the reward is slightly higher in PC-PFT-DPW. We think that this maybe related to the fact that dropping unsafe particles helps to localize the robot. Thus

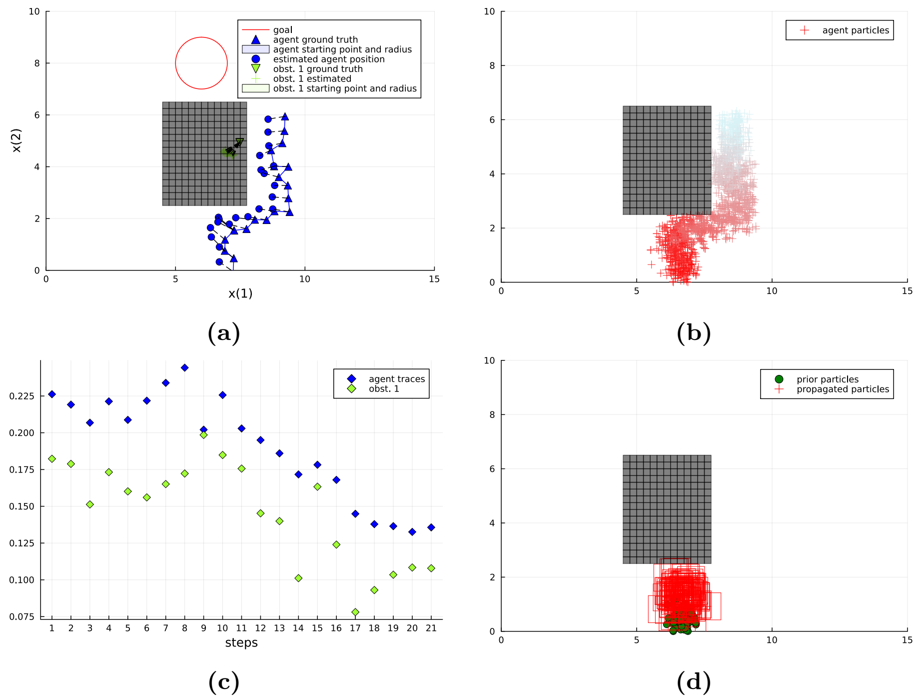


Fig. 4: This simulation setup is associated with Table 1 columns related to PC-PFT-DPW and here we show one of the trials. In this figure we **do not** nullify unsafe part of the belief in planning. **(a)** Here, we plot the goal, agent ground truth, estimated agent positions and the obstacles; **(b)** Belief particles, where the colors symbolize the time instance; **(c)** Traces of the agent and the landmark (obstacle); **(d)** Visualization of the truncation. Here we move each particle of b_0 with action selected by the agent and plot the truncation region of the stochastic motion model.

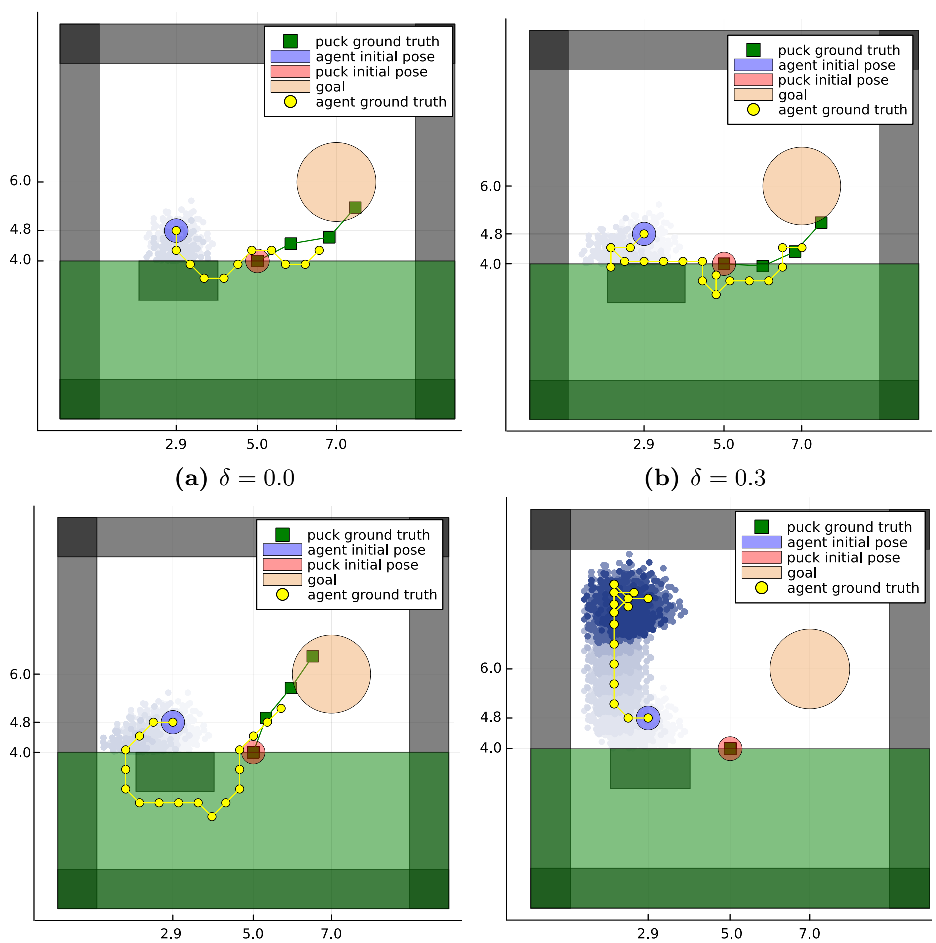
the PC-SB-PFT-DPW is less sensitive to the minimization of the trace of the covariance matrix of the belief.

3.2 PushBox2D

In Fig. 5 we show one of the trials with several values of δ.

Table 1: 50 Trials of at most 20 cycles of autonomy loop where planning sessions implemented by Alg. PC-SB-PFT-DPW versus PC-PFT-DPW. Same seed in both Alg. This problem is the **SLAM** described in the main manuscript. In **our second scenario** shown at Fig. 3 and Fig. 4. Here we study the probability of the safe trajectory while running autonomy loop, number of collisions and the reward value. The operator ϕ conforms to (12) and (13) in PC-SB-PFT-DPW and to (10) and (11) in PC-PFT-DPW. The inner threshold $\delta = 0.8$.

$\hat{\mathrm{P}}(S\|b_0)$		num coll.		mean cum. rew. $\pm$ std	
PC-SB-PFT-DPW	PC-PFT-DPW	PC-SB-PFT-DPW	PC-PFT-DPW	PC-SB-PFT-DPW	PC-PFT-DPW
0.6	0.6	28/70	28/70	-109.92 ± 11.55	-106.68 ± 12.77



(a) $\delta = 0.0$

(b) $\delta = 0.3$

(c) $\delta = 0.7$

(d) $\delta = 1.0$

Fig. 5: Visualization of actual PushBox2D simulation with several values of δ.