

Data Association Aware POMDP Planning with Hypothesis Pruning Performance Guarantees Supplementary Material

Moran Barenboim¹, Idan Lev-Yehudi¹ and Vadim Indelman^{2‡}

This document provides supplementary material to [1]. Therefore, it should not be considered a self-contained document, but instead regarded as an appendix of [1]. Throughout this report, all notations and definitions are with compliance to the ones presented in [1].

Contents

1	Theoretical analysis	1
1.1	Theorem 1	1
1.2	Corollary 1.1	7
1.3	Self Normalized Importance Sampling Estimator	8
1.4	Theorem 2	9
1.5	Corollary 2.1	11
1.6	Corollary 2.2	12
1.7	Corollary 2.3	12

1 Theoretical analysis

1.1 Theorem 1

Theorem 1. *Let time-step $t = 0$ denote the root of the planning tree. Then, the expected reward for the pruned POMDP, \bar{M} , is bounded with respect to the full POMDP,*

^{*1}Moran Barenboim and Idan Lev-Yehudi are with the Technion Autonomous Systems Program (TASP), Technion - Israel Institute of Technology, Haifa 32000, Israel, {moranbar, idanlev}@campus.technion.ac.il

^{†2} Vadim Indelman is with the Department of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel. vadim.indelman@technion.ac.il

[‡]This work was partially funded by US NSF/US-Israel BSF.

M , through the factor of the pruned weight values, and the maximum immediate reward,

$$\left| \mathbb{E}[r(b_t, a_t)] - \mathbb{E}[r(\bar{b}_t, a_t)] \right| \leq \mathcal{R}_{max} \left[\delta_0^\beta + \sum_{\tau=1}^{t-1} \mathbb{E}_{z_{1:\tau}}[\delta_\tau^\beta] \right], \quad (1)$$

where $\delta_\tau^\beta \triangleq \sum_{\beta_\tau \in D_\tau \setminus \bar{D}_\tau} \mathbb{P}(\beta_\tau | H_\tau)$, i.e. the sum of pruned hypotheses weights at time-step τ .

Proof. Denote D_t as the total number of new associations at time t , and \bar{D}_t as a subset thereof. By definition of the expected future reward,

$$\left| \mathbb{E}[r(b_t)] - \mathbb{E}[r(\bar{b}_t)] \right| \quad (2)$$

$$= \left| \int \int_{z_{1:t} x_{0:t}} r(x_t) \cdot [b_0 \prod_{\tau=1}^t \sum_{\beta_\tau}^{|D_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1})] \right| \quad (3)$$

$$- \bar{b}_0 \prod_{\tau=1}^t \sum_{\beta_\tau}^{|\bar{D}_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \Big| \quad (4)$$

$$(5)$$

by marginalizing out the variables β_t, z_t ,

$$\left| \int \int_{z_{1:t} x_{0:t}} r(x_t) \cdot [b_0 \mathbb{P}(x_t | x_{t-1}, \pi_{t-1})] \right| \quad (6)$$

$$\cdot \prod_{\tau=1}^{t-1} \sum_{\beta_\tau}^{|D_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \quad (7)$$

$$- \bar{b}_0 \mathbb{P}(x_t | x_{t-1}, \pi_{t-1}) \quad (8)$$

$$\cdot \prod_{\tau=1}^{t-1} \sum_{\beta_\tau}^{|\bar{D}_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \Big| \quad (9)$$

$$\leq \int \int_{z_{1:t} x_{0:t}} \left| r(x_t) \cdot [b_0 \mathbb{P}(x_t | x_{t-1}, \pi_{t-1})] \cdot \right. \quad (10)$$

$$\left. \prod_{\tau=1}^{t-1} \sum_{\beta_\tau}^{|D_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \right. \quad (11)$$

$$\left. - \bar{b}_0 \mathbb{P}(x_t | x_{t-1}, \pi_{t-1}) \right) \cdot \quad (12)$$

$$\left. \prod_{\tau=1}^{t-1} \sum_{\beta_\tau}^{|\bar{D}_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \right| \quad (13)$$

from Holder's inequality,

$$\leq \mathcal{R}_{max} \int \int_{z_{1:t} x_{0:t}} \left| \mathbb{P}(x_t | x_{t-1}, \pi_{t-1}) \right| \quad (14)$$

$$\left[b_0 \cdot \prod_{\tau=1}^{t-1} \sum_{\beta_\tau}^{|D_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \right] \quad (15)$$

$$\left. - \bar{b}_0 \cdot \prod_{\tau=1}^{t-1} \sum_{\beta_\tau}^{|\bar{D}_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \right] \quad (16)$$

since the transition model is positive, we take out of the absolute operator and marginalize it out,

$$\leq \mathcal{R}_{max} \int \int_{z_{1:t} x_{0:t}} \left| b_0 \cdot \prod_{\tau=1}^{t-1} \sum_{\beta_\tau}^{|D_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \right| \quad (17)$$

$$\left. - \bar{b}_0 \cdot \prod_{\tau=1}^{t-1} \sum_{\beta_\tau}^{|\bar{D}_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \right| \quad (18)$$

to avoid clutter convenience, we denote

$$\tilde{b}_t \triangleq b_0 \prod_{\tau=1}^t \sum_{\beta_\tau}^{|D_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \quad (19)$$

$$\tilde{\bar{b}}_t \triangleq \bar{b}_0 \prod_{\tau=1}^t \sum_{\beta_\tau}^{|\bar{D}_\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}). \quad (20)$$

Then we can rewrite it as,

$$\begin{aligned} &= \mathcal{R}_{max} \int \int_{z_{1:t-1} x_{0:t-1}} \quad (21) \\ &\left| \tilde{b}_{t-2} \sum_{\beta_{t-1}}^{|D_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right. \\ &\left. - \tilde{\bar{b}}_{t-2} \sum_{\beta_{t-1}}^{|\bar{D}_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right| \end{aligned}$$

note how this expression can also be written as, $\mathcal{R}_{max} \int \int_{z_{1:t-1} x_{0:t-1}} \left| \tilde{b}_t - \tilde{\bar{b}}_t \right|$. This will be useful for a recursive structure to be discussed later.

We now add and subtract,

$$\begin{aligned}
&= \mathcal{R}_{max} \int_{z_{1:t-1}} \int_{x_{0:t-1}} \tag{22} \\
&\quad \left| \tilde{b}_{t-2} \sum_{\beta_{t-1}}^{|D_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right. \\
&\quad - \tilde{b}_{t-2} \sum_{\beta_{t-1}}^{|D_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \\
&\quad + \tilde{b}_{t-2} \sum_{\beta_{t-1}}^{|D_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \\
&\quad \left. - \tilde{b}_{t-2} \sum_{\beta_{t-1}}^{|\bar{D}_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right|
\end{aligned}$$

grouping terms and applying triangle inequality,

$$\begin{aligned}
&= \mathcal{R}_{max} \int_{z_{1:t-1}} \int_{x_{0:t-1}} \tag{23} \\
&\quad \left| \left[\tilde{b}_{t-2} - \tilde{b}_{t-2} \right] \sum_{\beta_{t-1}}^{|D_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right| \tag{24} \\
&\quad + \mathcal{R}_{max} \int_{z_{1:t-1}} \int_{x_{0:t-1}} \\
&\quad \left| \tilde{b}_{t-2} \cdot \left[\sum_{\beta_{t-1}}^{|D_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right. \right. \\
&\quad \left. \left. - \sum_{\beta_{t-1}}^{|\bar{D}_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right] \right|
\end{aligned}$$

The first summand describes the loss due to pruning in past time steps. The second summand describes the loss due to pruning at the latest time step. Focusing on the

second summand, recall that $\bar{D}t \subseteq D_t$, thus,

$$\mathcal{R}_{max} \int_{z_{1:t-1}} \int_{x_{0:t-1}} \quad (25)$$

$$\left| \tilde{b}_{t-2} \cdot \left[\sum_{\beta_{t-1}}^{|D_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right) \right. \quad (26)$$

$$\left. - \sum_{\beta_{t-1}}^{|\bar{D}t-1|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right] \Big| \quad (27)$$

$$= \mathcal{R}_{max} \int_{z_{1:t-1}} \int_{x_{0:t-1}} \tilde{b}_{t-2}. \quad (28)$$

$$\left| \sum_{\beta_{t-1} \in \bar{D}t-1} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right) \quad (29)$$

$$+ \sum_{\beta_{t-1} \in D_t \setminus \bar{D}t-1} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \quad (30)$$

$$\left. - \sum_{\beta_{t-1}}^{|\bar{D}t-1|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right) \Big| \quad (31)$$

$$= \mathcal{R}_{max} \int_{z_{1:t-1}} \int_{x_{0:t-1}} \tilde{b}_{t-2}. \quad (32)$$

$$\left| \sum_{\beta_{t-1} \in D_{t-1} \setminus \bar{D}t-1} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right) \Big| \quad (33)$$

since all terms within the absolute operator are positive, we can now drop it entirely, we then marginalize the observation at time $t - 1$,

$$= \mathcal{R}_{max} \int_{z_{1:t-2}} \int_{x_{0:t-1}} \bar{b}_0 \cdot \prod_{\tau=1}^{t-2} \sum_{\beta_\tau}^{|\bar{D}\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1}) \quad (34)$$

$$\sum_{\beta_{t-1} \in \bar{D}t-1} \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \quad (35)$$

by introducing back the normalizer of the pruned belief, $\mathbb{P}(z_\tau | H_\tau^-)$, we get,

$$= \mathcal{R}_{max} \int_{z_{1:t-2}} \prod_{k=1}^{t-2} \bar{\mathbb{P}}(z_k | H_k^-) \int_{x_{0:t-1}} \quad (36)$$

$$\bar{b}_0 \cdot \prod_{\tau=1}^{t-2} \left[\frac{\sum_{\beta_\tau}^{|\bar{D}\tau|} \mathbb{P}(z_\tau | x_\tau, \beta_\tau) \mathbb{P}(\beta_\tau | x_\tau) \mathbb{P}(x_\tau | x_{\tau-1}, \pi_{\tau-1})}{\bar{\mathbb{P}}(z_\tau | H_\tau^-)} \right] \quad (37)$$

$$\sum_{\beta_{t-1} \in \neg \bar{D}t-1} \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \quad (38)$$

$$= \mathcal{R}_{max} \int_{z_{1:t-2}} \prod_{k=1}^{t-2} \bar{\mathbb{P}}(z_k | H_k^-) \int_{x_{t-2:t-1}} \bar{b}_{t-2} \sum_{\beta_{t-1} \in \neg \bar{D}t-1} \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \quad (39)$$

or, equivalently,

$$= \mathcal{R}_{max} \bar{\mathbb{E}}_{z_{1:t-2}} \left[\int_{x_{t-2:t-1}} \bar{b}_{t-2} \sum_{\beta_{t-1} \in \neg \bar{D}t-1} \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right]. \quad (40)$$

Crucially, note how the following term depends only on the survived hypotheses (no access to the pruned hypotheses is required). Finally, by rearranging and marginalizing state variables, we get,

$$= \mathcal{R}_{max} \bar{\mathbb{E}}_{z_{1:t-2}} \left[\int_{x_{t-1}} \sum_{\beta_{t-1} \in \neg \bar{D}t-1} \mathbb{P}(\beta_{t-1} | x_{t-1}) \bar{\mathbb{P}}(x_{t-1} | H_{t-1}^-) \right] \quad (41)$$

$$= \mathcal{R}_{max} \bar{\mathbb{E}}_{z_{1:t-2}} \left[\sum_{\beta_{t-1} \in \neg \bar{D}t-1} \bar{\mathbb{P}}(\beta_{t-1} | H_{t-1}^-) \right] \quad (42)$$

$$= \mathcal{R}_{max} \bar{\mathbb{E}}_{z_{1:t-2}} \left[\sum_{\beta_{t-1} \in \neg \bar{D}t-1} \int_{z_{t-1}} \bar{\mathbb{P}}(\beta_{t-1} | H_{t-1}) \bar{\mathbb{P}}(z_{t-1} | H_{t-1}^-) \right] \quad (43)$$

$$= \mathcal{R}_{max} \bar{\mathbb{E}}_{z_{1:t-1}} \left[\sum_{\beta_{t-1} \in \neg \bar{D}t-1} \bar{\mathbb{P}}(\beta_{t-1} | H_{t-1}) \right] \quad (44)$$

Going back to the first summand from equation (22) and applying triangle inequality,

we have that,

$$\mathcal{R}_{max} \int_{z_{1:t-1}} \int_{x_{0:t-1}} \quad (45)$$

$$\left| \left[\tilde{b}_{t-2} - \tilde{\tilde{b}}_{t-2} \right] \sum_{\beta_{t-1}}^{|D_{t-1}|} \mathbb{P}(z_{t-1} | x_{t-1}, \beta_{t-1}) \mathbb{P}(\beta_{t-1} | x_{t-1}) \mathbb{P}(x_{t-1} | x_{t-2}, \pi_{t-2}) \right| \quad (46)$$

$$\leq \mathcal{R}_{max} \int_{z_{1:t-2}} \int_{x_{0:t-2}} \left| \tilde{b}_{t-2} - \tilde{\tilde{b}}_{t-2} \right| \quad (47)$$

recall the recursive structure from equation (21), thus,

$$\left| \mathbb{E}[r(b_t)] - \mathbb{E}[r(\tilde{b}_t)] \right| \leq \quad (48)$$

$$\mathcal{R}_{max} \mathbb{E}_{z_{1:t-1}} \left[\sum_{\beta_{t-1} \in \bar{D}_{t-1}} \bar{\mathbb{P}}(\beta_{t-1} | H_{t-1}) \right] + \mathcal{R}_{max} \int_{z_{1:t-2}} \int_{x_{0:t-2}} \left| \tilde{b}_{t-2} - \tilde{\tilde{b}}_{t-2} \right| \quad (49)$$

$$\leq \mathcal{R}_{max} \left(\mathbb{E}_{z_{1:t-1}} \left[\sum_{\beta_{t-1} \in \bar{D}_{t-1}} \bar{\mathbb{P}}(\beta_{t-1} | H_{t-1}) \right] \right) \quad (50)$$

$$+ \mathbb{E}_{z_{1:t-2}} \left[\sum_{\beta_{t-2} \in \bar{D}_{t-2}} \bar{\mathbb{P}}(\beta_{t-2} | H_{t-2}) \right] + \int_{z_{1:t-3}} \int_{x_{0:t-3}} \left| \tilde{b}_{t-3} - \tilde{\tilde{b}}_{t-3} \right| \leq \dots \quad (51)$$

$$\leq \mathcal{R}_{max} \left(\sum_{\tau=1}^{t-1} \mathbb{E}_{z_{1:\tau}} \left[\sum_{\beta_{\tau} \in \bar{D}_{\tau}} \bar{\mathbb{P}}(\beta_{\tau} | H_{\tau}) \right] + \int_{x_0} |b_0 - \tilde{b}_0| dx_0 \right) \quad (52)$$

$$\equiv \mathcal{R}_{max} \left(\sum_{\tau=1}^{t-1} \mathbb{E}_{z_{1:\tau}} [\delta^{\beta}(H_{\tau})] + \delta_0^{\beta} \right) \quad (53)$$

which concludes our derivation. \square

1.2 Corollary 1.1

Corollary 1.1. *Without loss of generality, assume that the time step at the root node of the planning tree is $t = 0$. Then, for any policy π , the following holds,*

$$\left| V^{\pi}(b_0) - \bar{V}^{\pi}(\tilde{b}_0) \right| \leq \mathcal{R}_{max} \left[T \cdot \delta_0^{\beta} + \sum_{k=1}^T \sum_{\tau=1}^k \mathbb{E}_{z_{1:\tau}} [\delta_{\tau}^{\beta}] \right]. \quad (54)$$

Proof. The proof is a direct consequence of the linearity of expectation. \square

1.3 Self Normalized Importance Sampling Estimator

In this subsection we will derive the SN estimator. The theoretical expected reward at time step t may be written as,

$$\mathbb{E}_{z_{1:t}}[\rho(b_t)] = \int_{z_{1:t}} \prod_{\tau=1}^t \mathbb{P}(z_\tau | H_\tau^-) \sum_{\beta_{0:t} \in D_{0:t}} \mathbb{P}(\beta_{0:t} | H_t) \int_{x_t} \mathbb{P}(x_t | \beta_{0:t}, H_t) r(x_t) \quad (55)$$

where $D_{0:t}$ is the set of all hypotheses at time step t . Applying Bayes rule followed by a chain rule on $\mathbb{P}(\beta_{0:t} | H_t)$,

$$\int_{z_{1:t}} \prod_{\tau=1}^t \mathbb{P}(z_\tau | H_\tau^-) \sum_{\beta_{0:t} \in D_{0:t}} \frac{\mathbb{P}(z_t, \beta_t | \beta_{0:t-1}, H_t^-)}{\mathbb{P}(z_t | H_t^-)} \mathbb{P}(\beta_{0:t-1} | H_{t-1}) \int_{x_t} \mathbb{P}(x_t | \beta_{0:t}, H_t) r(x_t) \quad (56)$$

applying this step repeatedly on $\mathbb{P}(\beta_{0:\tau} | H_\tau) \forall \tau \in [1, t-1]$ results in,

$$\int_{z_{1:t}} \prod_{\tau=1}^t \mathbb{P}(z_\tau | H_\tau^-) \sum_{\beta_{0:t} \in D_{0:t}} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \frac{\mathbb{P}(z_\tau, \beta_\tau | \beta_{0:\tau-1}, H_\tau^-)}{\mathbb{P}(z_\tau | H_\tau^-)} \int_{x_t} \mathbb{P}(x_t | \beta_{0:t}, H_t) r(x_t) \quad (57)$$

$$= \int_{z_{1:t}} \sum_{\beta_{0:t} \in D_{0:t}} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \mathbb{P}(z_\tau, \beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \int_{x_t} \mathbb{P}(x_t | \beta_{0:t}, H_t) r(x_t) \quad (58)$$

$$= \sum_{\beta_0 \in D_0} \mathbb{P}(\beta_0) \sum_{\beta_1 \in D_1} \mathbb{P}(\beta_1 | \beta_0, H_1^-) \int_{z_1} \mathbb{P}(z_1 | \beta_{0:1}, H_1^-) \cdots \quad (59)$$

$$\cdots \sum_{\beta_t \in D_t} \mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) \int_{z_t} \mathbb{P}(z_t | \beta_{0:t}, H_t^-) \int_{x_t} \mathbb{P}(x_t | \beta_{0:t}, H_t) r(x_t)$$

where the second equality is due to chain rule on $\mathbb{P}(z_\tau, \beta_\tau | \beta_{0:\tau-1}, H_\tau^-)$ and rearranging terms.

According to equation (59) we define a self-normalized importance sampling estimator for the expected reward, at time step t , where both the observations and states are sampled,

$$\hat{\mathbb{E}}_{z_{1:t}}[\rho(\hat{b}_t)] \triangleq \sum_{\beta_0 \in D_0} \sum_{\beta_1 \in D_1} \sum_{c_1} \cdots \sum_{\beta_t \in D_t} \sum_{z_\tau^c} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r}(b_t^\beta) \quad (60)$$

$$= \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^k} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r}(b_t^\beta) \quad (61)$$

where $\omega(z_\tau) = \frac{\mathbb{P}(z_\tau | \beta_{0:\tau}, H_\tau^-)}{Q(z_\tau | H_\tau^-)}$ and $Q(\cdot)$ is the proposal distribution according to which the sampling-based estimator generates observations. Similarly, we define the *pruned*

estimator, where the only difference is the summation over a pruned subset of the hypotheses, denoted \bar{D} ,

$$\hat{\mathbb{E}}_{z_{1:t}} \left[\rho \left(\hat{b}_t \right) \right] = \sum_{z_{1:t}^c} \sum_{\beta_{0:t} \in \bar{D}_{0:t}} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r} \left(b_t^\beta \right). \quad (62)$$

1.4 Theorem 2

Theorem 2. *Let π be the policy, then the expected reward for the estimated pruned POMDP, \hat{M} , is bounded with respect to the estimated full POMDP, \hat{M} , as follows,*

$$\left| \hat{\mathbb{E}}_{z_{1:t}}^\pi [\rho(\hat{b}_t)] - \hat{\mathbb{E}}_{z_{1:t}}^\pi \left[\rho \left(\hat{b}_t \right) \right] \right| \leq \mathcal{R}_{max} \left[\hat{\delta}_0^\beta + \sum_{\tau=1}^t \hat{\delta}_\tau^\beta \right]. \quad (63)$$

where, $\hat{\delta}_\tau^\beta = \hat{\mathbb{E}}_{z_{1:t}^c} \bar{\mathbb{E}}_{\beta_{0:t-1}} \sum_{\beta_t \in D_t \setminus \bar{D}_t} \mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-)$ for all $\tau \in [1, t]$ represents the expected sum of conditional hypotheses' weights which are myopically pruned and $\hat{\delta}_0^\beta = \sum_{\beta_0 \in D_0 \setminus \bar{D}_0} \mathbb{P}(\beta_0 | H_t^-)$.

Proof. Hereon forward, we assume that conditioned the same hypothesis, $\beta_{0:\tau}$, the same observations and states are sampled. This is required in order to obtain a deterministic bound and can be achieved in practice by fixing some seed number. Additionally, we also define a pruned conditionals,

$$\bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \triangleq \begin{cases} \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) & , \beta_{0:\tau} \in \bar{D}_{0:\tau} \\ 0 & , \text{otherwise} \end{cases}. \quad (64)$$

Then,

$$\left| \hat{\mathbb{E}}_{z_{1:t}} [\rho(\hat{b}_t)] - \hat{\mathbb{E}}_{z_{1:t}} \left[\rho \left(\hat{b}_t \right) \right] \right| \quad (65)$$

$$= \left| \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r} \left(b_t^\beta \right) \right| \quad (66)$$

$$- \sum_{\beta_0 \in \bar{D}_0} \prod_{k=1}^t \sum_{\beta_k \in \bar{D}_k} \sum_{z_\tau^c} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r} \left(b_t^\beta \right) |$$

$$= \left| \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r} \left(b_t^\beta \right) \right| \quad (67)$$

$$- \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^t \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r} \left(b_t^\beta \right) |$$

add and subtract,

$$\begin{aligned}
& \left| \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r}(b_t^\beta) \right. \\
& - \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) \frac{\omega_t^c}{\sum_{c_t'} \omega_t^{c_t'}} \hat{r}(b_t^\beta) \\
& + \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) \frac{\omega_t^c}{\sum_{c_t'} \omega_t^{c_t'}} \hat{r}(b_t^\beta) \\
& \left. - \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^t \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r}(b_t^\beta) \right| \quad (68)
\end{aligned}$$

applying triangle inequality then focusing on the second pair of terms,

$$\begin{aligned}
& \left| \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) \frac{\omega(z_\tau^c)}{\sum_{z_t^k} \omega(z_t^k)} \hat{r}(b_t^\beta) \right. \\
& \left. - \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^t \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r}(b_t^\beta) \right| \quad (69)
\end{aligned}$$

$$\begin{aligned}
& = \left| \sum_{\beta_0 \in D_0} \prod_{k=1}^{t-1} \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \right. \\
& \left. \sum_{\beta_t \in D_t} [\mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) - \bar{\mathbb{P}}(\beta_t | \beta_{0:t-1}, H_t^-)] \sum_{z_t^c} \frac{\omega(z_t^c)}{\sum_{z_t^k} \omega(z_t^k)} \hat{r}(b_t^\beta) \right| \quad (70)
\end{aligned}$$

$$\begin{aligned}
& \leq \mathcal{R}_{max} \sum_{\beta_0 \in D_0} \prod_{k=1}^{t-1} \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \quad (71)
\end{aligned}$$

applying again triangle inequality followed by Holder inequality,

$$\begin{aligned}
& \left| \sum_{\beta_t \in D_t} [\mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) - \bar{\mathbb{P}}(\beta_t | \beta_{0:t-1}, H_t^-)] \right| \\
& = \mathcal{R}_{max} \sum_{\beta_0 \in D_0} \prod_{k=1}^{t-1} \sum_{\beta_k \in D_k} \sum_{z_\tau^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \quad (72)
\end{aligned}$$

$$\begin{aligned}
& \sum_{\beta_t \in D_t \setminus \bar{D}_t} \mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) \\
& \triangleq \mathcal{R}_{max} \hat{\delta}_t^\beta
\end{aligned}$$

where $\hat{\delta}_t^\beta$ is the empirical expected weight of all the pruned hypotheses at time step t . Crucially, its value depends only on past pruned hypotheses, which are known to us. Now focusing on the first pair of terms from equation (68),

$$\left| \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_k^c} \mathbb{P}(\beta_0) \prod_{\tau=1}^t \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \hat{r}(b_t^\beta) \right. \quad (73)$$

$$\begin{aligned} & - \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_k^c} \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) \frac{\omega(z_t^c)}{\sum_{z_t^k} \omega(z_t^k)} \hat{r}(b_t^\beta) \Big| \\ & = \left| \sum_{\beta_0 \in D_0} \prod_{k=1}^t \sum_{\beta_k \in D_k} \sum_{z_k^c} \left[\mathbb{P}(\beta_0) \prod_{\tau=1}^{t-1} \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} - \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \right] \right. \\ & \quad \left. \mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) \frac{\omega(z_t^c)}{\sum_{z_t^k} \omega(z_t^k)} \hat{r}(b_t^\beta) \right| \end{aligned} \quad (74)$$

$$\mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) \frac{\omega(z_t^c)}{\sum_{z_t^k} \omega(z_t^k)} \hat{r}(b_t^\beta) \Big|$$

triangle and Holder inequalities,

$$\leq \sum_{\beta_0 \in D_0} \prod_{k=1}^{t-1} \sum_{\beta_k \in D_k} \sum_{z_k^c} \left| \mathbb{P}(\beta_0) \prod_{\tau=1}^{t-1} \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \right. \quad (75)$$

$$\left. - \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \right| \sum_{\beta_t \in D_t} \sum_{z_t^c} \mathbb{P}(\beta_t | \beta_{0:t-1}, H_t^-) \frac{\omega(z_t^c)}{\sum_{z_t^k} \omega(z_t^k)} \mathcal{R}_{max}$$

$$= \mathcal{R}_{max} \sum_{\beta_0 \in D_0} \prod_{k=1}^{t-1} \sum_{\beta_k \in D_k} \sum_{z_k^c} \left| \mathbb{P}(\beta_0) \prod_{\tau=1}^{t-1} \mathbb{P}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \right. \quad (76)$$

$$\left. - \bar{\mathbb{P}}(\beta_0) \prod_{\tau=1}^{t-1} \bar{\mathbb{P}}(\beta_\tau | \beta_{0:\tau-1}, H_\tau^-) \frac{\omega(z_\tau^c)}{\sum_{z_\tau^k} \omega(z_\tau^k)} \right|$$

then, applying similar steps recursively on the obtained term yields,

$$\left| \hat{\mathbb{E}}_{z_{1:t}}[\rho(\hat{b}_t)] - \hat{\mathbb{E}}_{z_{1:t}}[\rho(\hat{b}_t)] \right| \leq \mathcal{R}_{max} \sum_{\tau=0}^t \hat{\delta}_\tau^\beta \quad (77)$$

which concludes our derivation. \square

1.5 Corollary 2.1

Corollary 2.1. *The difference between the estimated value function of the full POMDP, \hat{M} , and the estimated value function of the pruned POMDP, $\hat{\bar{M}}$, is bounded by,*

$$|\hat{V}^\pi(\hat{b}_0) - \hat{\bar{V}}^\pi(\hat{b}_0)| \leq \mathcal{R}_{max} \left[\hat{\delta}_0^\beta + \sum_{k=1}^T \sum_{\tau=1}^k \hat{\delta}_\tau^\beta \right]. \quad (78)$$

Proof. The proof is a direct consequence of the linearity of expectation. \square

1.6 Corollary 2.2

Corollary 2.2. *Let π be a policy and let \mathcal{A} be a sampling-based estimator for the value function such that $|V^\pi(b_0) - \hat{V}^\pi(\hat{b}_0)| \leq \epsilon_{\mathcal{A}}$ with probability at least $1 - \delta_{\mathcal{A}}$. Then, the following corollary holds for the loss in the value function for the pruned hypotheses,*

$$|V^\pi(b_0) - \hat{\hat{V}}^\pi(\hat{\hat{b}}_0)| \leq \quad (79)$$

$$|V^\pi(b_0) - \hat{V}^\pi(\hat{b}_0)| + |\hat{V}^\pi(\hat{b}_0) - \hat{\hat{V}}^\pi(\hat{\hat{b}}_0)| \leq \epsilon_{\mathcal{A}} + \hat{\epsilon}_{\mathcal{D}}^{hs}, \quad (80)$$

hold with probability $1 - \delta_{\mathcal{A}}$. We use $\hat{\epsilon}_{\mathcal{D}}^{hs}$ as a shorthand for the bounds provided in corollary 2.1.

1.7 Corollary 2.3

Corollary 2.3. *Let π^* be the optimal policy for the full theoretical POMDP with a respective value function, $V(b_t)$. Let $\bar{\pi}$ be the optimal policy for the pruned POMDP and a value function, $\bar{V}(\hat{b}_t)$. Last, let $\hat{\hat{\pi}}$ be the optimal policy for the pruned, sampled-based POMDP with a value function, $\hat{\hat{V}}(\hat{\hat{b}}_t)$. Then, the following holds,*

$$\left| V^{\pi^*}(b_t) - \hat{\hat{V}}^{\hat{\hat{\pi}}}(\hat{\hat{b}}_t) \right| \leq 2(\epsilon_{\mathcal{A}} + \epsilon_{\mathcal{D}}). \quad (81)$$

Proof. For conciseness, we drop the explicit dependence on the belief at each value function.

$$\left| V^{\pi^*} - \hat{\hat{V}}^{\hat{\hat{\pi}}} \right| \leq \underbrace{\left| V^{\pi^*} - \bar{V}^{\bar{\pi}} \right|}_{(a)} + \underbrace{\left| \bar{V}^{\bar{\pi}} - \hat{\hat{V}}^{\hat{\hat{\pi}}} \right|}_{(b)} \quad (82)$$

we split the derivation of term (a) into two cases, case 1a: if $V^{\bar{\pi}} \geq \bar{V}^{\pi^*}$, then, $V^{\pi^*} \geq V^{\bar{\pi}} \geq \bar{V}^{\pi^*}$

$$\begin{aligned} (a) &= \left| V^{\pi^*} - \bar{V}^{\bar{\pi}} \right| \leq \left| V^{\pi^*} - V^{\bar{\pi}} \right| + \left| V^{\bar{\pi}} - \bar{V}^{\bar{\pi}} \right| \\ &\leq \left| V^{\pi^*} - \bar{V}^{\pi^*} \right| + \left| V^{\bar{\pi}} - \bar{V}^{\bar{\pi}} \right| \leq 2\epsilon_{\mathcal{D}} \end{aligned} \quad (83)$$

where the last inequality is a result of Corollary 2.2. case 2a: $V^{\bar{\pi}} \leq \bar{V}^{\pi^*}$. Then, $V^{\bar{\pi}} \leq \bar{V}^{\pi^*} \leq \bar{V}^{\bar{\pi}}$. By triangle inequality,

$$(a) = \left| V^{\pi^*} - \bar{V}^{\bar{\pi}} \right| \leq \left| V^{\pi^*} - \bar{V}^{\pi^*} \right| + \left| \bar{V}^{\pi^*} - \bar{V}^{\bar{\pi}} \right| \quad (84)$$

$$\leq \left| V^{\pi^*} - \bar{V}^{\pi^*} \right| + \left| V^{\bar{\pi}} - \bar{V}^{\bar{\pi}} \right| \leq 2\epsilon_{\mathcal{D}} \quad (85)$$

Similarly, we split the handling of term (b) to two different cases, case 1b: if $\bar{V}^{\bar{\pi}} \geq \hat{\hat{V}}^{\hat{\hat{\pi}}}$, then, $\bar{V}^{\bar{\pi}} \geq \bar{V}^{\hat{\hat{\pi}}} \geq \hat{\hat{V}}^{\hat{\hat{\pi}}}$. By triangle inequality,

$$(b) = \left| \bar{V}^{\bar{\pi}} - \hat{\hat{V}}^{\hat{\hat{\pi}}} \right| \leq \left| \bar{V}^{\bar{\pi}} - \bar{V}^{\hat{\hat{\pi}}} \right| + \left| \bar{V}^{\hat{\hat{\pi}}} - \hat{\hat{V}}^{\hat{\hat{\pi}}} \right| \quad (86)$$

$$\leq \left| \bar{V}^{\bar{\pi}} - \hat{\hat{V}}^{\hat{\hat{\pi}}} \right| + \left| \bar{V}^{\hat{\hat{\pi}}} - \hat{\hat{V}}^{\hat{\hat{\pi}}} \right| \leq 2\epsilon_{\mathcal{A}} \quad (87)$$

case 2b: if $\bar{V}^{\hat{\pi}} \leq \hat{V}^{\bar{\pi}}$, then, $\bar{V}^{\hat{\pi}} \leq \hat{V}^{\bar{\pi}} \leq \hat{V}^{\hat{\pi}}$. From triangle inequality,

$$(b) = \left| \bar{V}^{\bar{\pi}} - \hat{V}^{\hat{\pi}} \right| \leq \left| \bar{V}^{\bar{\pi}} - \hat{V}^{\bar{\pi}} \right| + \left| \hat{V}^{\bar{\pi}} - \hat{V}^{\hat{\pi}} \right| \quad (88)$$

$$\leq \left| \bar{V}^{\bar{\pi}} - \hat{V}^{\bar{\pi}} \right| + \left| \bar{V}^{\hat{\pi}} - \hat{V}^{\hat{\pi}} \right| \leq 2\epsilon_{\mathcal{A}} \quad (89)$$

which covers all the cases and result in,

$$\left| V^{\pi^*}(b_t) - \hat{V}^{\hat{\pi}}(\hat{b}_t) \right| \leq 2(\epsilon_{\mathcal{A}} + \epsilon_{\bar{D}}). \quad (90)$$

□

References

- [1] M. Barenboim, I. Lev-Yehudi, and V. Indelman, “Data association aware pomdp planning with hypothesis pruning performance guarantees,” *arXiv preprint arXiv:2303.02139*, 2023.