$\texttt{iX-BSP}$: Incremental Belief Space Planning with Selective Resampling
Supplementary Material

Technical Report ANPL-2019-02

Elad I. Farhi and Vadim Indelman

This document provides supplementary material to the paper [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].

This report covers the formulation of Multiple Importance Sampling - Appendix A, and a toy example for our novel paradigm iX-BSP- Appendix B.

Appendix A: Multiple Importance Sampling Formulation

Let us assume we wish to express expectation over some function $f(x)$ with respect to distribution $p(x)$, by sampling x from a different distribution $q(x)$,

$$\mathbb{E}_p f(x) = \int f(x) \cdot p(x) dx = \int \frac{f(x) \cdot p(x)}{q(x)} q(x) dx = \mathbb{E}_q \left(\frac{f(x) \cdot p(x)}{q(x)} \right). \tag{10}$$

Eq. (10) presents the basic importance sampling problem, where $\mathbb{E}_q$ denotes expectation for $x \sim q(x)$. The probability ratio between the nominal distribution p and the importance sampling distribution q is usually referred to as the likelihood ratio. Our problem is more complex, since our samples are potentially taken from M different distributions while $M \in [1, (n_x \cdot n_z)^L]$, i.e. a multiple importance sampling problem

$$\mathbb{E}_p f(x) = \tilde{\mu}(x) \sim \sum_{m=1}^{M} \frac{1}{n_m} \sum_{i=1}^{n_m} w_m(x_{im}) \frac{f(x_{im}) p(x_{im})}{q_m(x_{im})}, \tag{11}$$

where $w_m(.)$ are weight functions satisfying $\Sigma_{m=1}^M w_m(x) = 1$. For $q_m(x) > 0$ whenever $w_m(x)p(x)f(x) \neq 0$, Eq. (11) forms an unbiased estimator

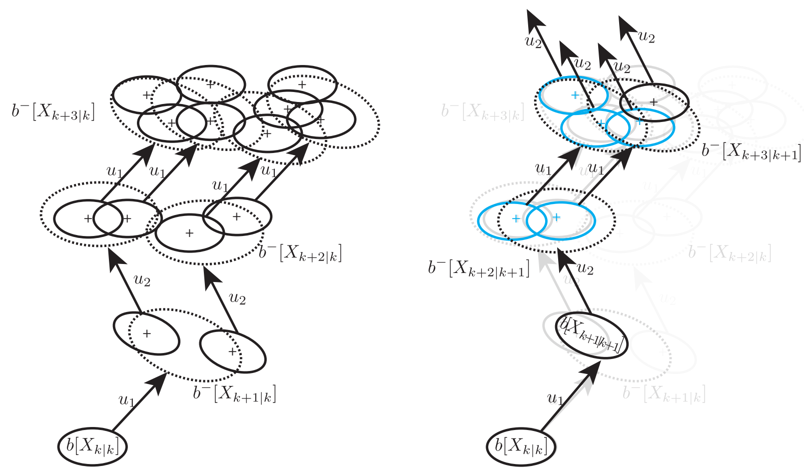
$$\mathbb{E}\left[\tilde{\mu}(x)\right] = \sum_{m=1}^{M} \mathbb{E}_{q_m} \left[\frac{1}{n_m} \sum_{i=1}^{n_m} w_m(x_{im}) \frac{f(x_{im}) p(x_{im})}{q_m(x_{im})} \right] = \tilde{\mu}(x). \tag{12}$$

Although there are numerous options for weight functions satisfying $\Sigma_{m=1}^M w_m(x) = 1$, we chose to consider the Balance Heuristic [2], considered to be nearly optimal in the sense of estimation variance,

$$w_m(x) = w_m^{BH}(x) = \frac{n_m q_m(x)}{\Sigma_{s=1}^M n_s q_s(x)}. \tag{13}$$

Using (13) in (11) produces the multiple importance sampling with the balance heuristic

$$\mathbb{E}_p f(x) \sim \frac{1}{n} \sum_{m=1}^{M} \sum_{i=1}^{n_m} \frac{p(x_{im})}{\Sigma_{s=1}^M \frac{n_s}{n} q_s(x_{im})} f(x_{im}). \tag{14}$$



(a) (b)

Figure 4: Consider BSP presented in a spatial belief propagation diagram, with propagated beliefs denoted by dashed line ellipse, samples denoted by $+$, actions denoted by arrows, beliefs denoted by solid line ellipse and with $n_x = 2$ and $n_z = 1$. (a) Presents `X-BSP` at planning time k for the action sequence $u_1 \to u_2 \to u_1$, as such all measurements were sampled from their original distributions. Assuming action u_1 have been executed, (b) presents `iX-BSP` for the succeeding planning time $k+1$, for the action sequence $u_2 \to u_1 \to u_2$, where re-used samples and corresponding re-used beliefs are denoted in blue, and freshly sampled measurements and corresponding freshly calculated beliefs are denoted in black.

Appendix B: Toy example for `iX-BSP`

To better understand the problem let us consider a toy example. We have access to all calculations from planning time k, in-which we performed `X-BSP` (or `iX-BSP`) for a horizon of 3 steps, with $n_x = 2$ and $n_z = 1$, while considering 2 candidate actions u_1 and u_2. Figure 4a illustrates a specific action sequence, $u_1 \to u_2 \to u_1$, considered as part of planning at time k. We are currently at time k+1, performing planning using `iX-BSP` with the same horizon length and number of samples per action, for the action sequence $u_2 \to u_1 \to u_2$, see Figure 4b.

Following Alg. 1 line 1, out of the two available beliefs from planning time k, $\{b[X_{k+1|k}]\}_1^2$, the left one is closer to $b[X_{k+1|k+1}]$, so we consider all its descendants as the set $\mathcal{B}_{k+1|k}$, and denote the difference between $b[X_{k+1|k}]$ and $b[X_{k+1|k+1}]$ as $\{\delta\}$. We continue with re-using the beliefs in the set $\mathcal{B}_{k+1|k}$ (Alg. 1 line 3). First we check whether the two available samples from planning time k constitute an adequate representation for $b^-[X_{k+2|k+1}]$ (Alg. 2 line 7); since they are, we flag them both for re-use, and update $\{b[X_{k+2|k}]\}_1^2$ into $\{b[X_{k+2|k+1}]\}_1^2$ (Alg. 2 line 11). Now for the next future time step, we propagate $\{b[X_{k+2|k+1}]\}_1^2$ with action u_1 to obtain $\{b^-[X_{k+3|k+1}]\}_1^2$ (Alg. 2 line 3), and check whether the four available samples from planning time k constitute an adequate representation for $\{b^-[X_{k+3|k+1}]\}_1^2$ (Alg. 2 line 7); since only three of them are, we flag them for re-use and sample the forth one (black colored belief at $k+3|k+1$ in Figure 4b) from $b^-[X_{k+3|k+1}]$. We then update $\{b[X_{k+3|k}]\}_1^3$ into $\{b[X_{k+3|k+1}]\}_1^3$, and $b^-[X_{k+3|k+1}]$ into $b^4[X_{k+3|k+1}]$ using the newly sampled measurement (Alg. 2 line 11). The last step of the horizon $k+4|k+1$ is calculated using `X-BSP` (Alg. 1 line 4).

At this point we have all inference results for all beliefs along the action sequence $u_2 \to u_1 \to u_2$, so we can calculate all reward(cost) values for this action sequence for planning at time k+1. For the look ahead at time $k+2$ of planning session at time $k+1$, i.e. $k+2|k+1$, we have two reward(cost) values, $\{r_{k+2|k+1}(b[X_{k+2|k+1}], u_2)\}_1^2$, each calculated with a different belief $b[X_{k+2|k+1}]$ considering a different sample $z_{k+2|k}$. Calculating the expected reward(cost) value for future time step $k+2|k+1$ would mean in this case, using measurements sampled from $\mathbb{P}(z_{k+2|k}|H_{k+1|k}, u_2)$ rather then from $\mathbb{P}(z_{k+2|k+1}|H_{k+1|k+1}, u_2)$. This problem, of performing estimation using forced samples is called importance sampling. Since for a single time step we might have samples from multiple different distributions, e.g. future time $k+3|k+1$ in Figure 4b, our problem falls within the special case of Multiple Importance Sampling (see Appendix A). Using the formulation of multiple importance sampling using the balance heuristic (14)

Elad I. Farhi is with the Technion Autonomous Systems Program (TASP), Technion - Israel Institute of Technology, Haifa 32000, Israel, `eladf@campus.technion.ac.il`. Vadim Indelman is with the Department of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel, `vadim.indelman@technion.ac.il`.

we can write down the estimation for the expected reward value at planning time $k+2|k+1$,

$$\mathbb{E}\left[r_{k+2|k+1}(.)\right] \sim \frac{1}{2}\frac{p_1(z^1_{k+2|k})}{\frac{2}{2}q_1(z^1_{k+2|k})}\cdot r^1_{k+2|k+1}(.)+\frac{1}{2}\frac{p_1(z^2_{k+2|k})}{\frac{2}{2}q_1(z^2_{k+2|k})}\cdot r^2_{k+2|k+1}(.), \tag{15}$$

where $p_1(.) \doteq \mathbb{P}(z_{k+2|k+1}|H_{k+1|k+1}, u_2)$ and $q_1(.) \doteq \mathbb{P}(z_{k+2|k}|H_{k+1|k}, u_2)$. In the same manner, following (14), we can also write down the estimation for the expected reward(cost) value at look ahead step $k+3$ from planning session at time $k+1$, i.e. $k+3|k+1$,

$$\begin{aligned}\mathbb{E}\left[r_{k+3|k+1}(.)\right] \sim{}& \frac{1}{4}\frac{p_2(z^1_{k+2:k+3|k})}{\frac{3}{4}q_2(z^1_{k+2:k+3|k})+\frac{1}{4}p_2(z^1_{k+2:k+3|k})}r^1_{k+3|k+1}(.)+\frac{1}{4}\frac{p_2(z^2_{k+2:k+3|k})}{\frac{3}{4}q_2(z^2_{k+2:k+3|k})+\frac{1}{4}p_2(z^2_{k+2:k+3|k})}r^2_{k+3|k+1}(.)+\\ &\frac{1}{4}\frac{p_2(z^3_{k+2:k+3|k})}{\frac{3}{4}q_2(z^3_{k+2:k+3|k})+\frac{1}{4}p_2(z^3_{k+2:k+3|k})}r^3_{k+3|k+1}(.)+\frac{1}{4}\frac{p_2(z^4_{k+2:k+3|k+1})}{\frac{3}{4}q_2(z^4_{k+2:k+3|k+1})+\frac{1}{4}p_2(z^4_{k+2:+3|k+1})}r^4_{k+3|k+1}(.),\end{aligned} \tag{16}$$

where $p_2(.) \doteq \mathbb{P}(z_{k+2:k+3|k+1}|H_{k+1|k+1}, u_2, u_1)$ and $q_2(.) \doteq \mathbb{P}(z_{k+2:k+3|k}|H_{k+1|k}, u_2, u_1)$. When considering

$$\mathbb{P}(z_{k+1:k+L|k}|H_{k|k}, u_{k:k+L-1}) = \prod_{i=k+1}^{k+L}\mathbb{P}(z_{i|k}|H^-_{i|k}) \tag{17}$$

we can re-write the measurement likelihood from (16) into a product of measurement likelihoods per look ahead step, e.g. $p_2(z^1_{k+2:k+3|k}) = p_1(z^1_{k+2|k})\tilde{p}_2(z^1_{k+3|k})$, when $p_1(.)$ need not be calculated at look ahead step $k+3$, since it is given from (15).

References

[1] E. Farhi and V. Indelman. ix-bsp: Incremental belief space planning with selective resampling. In *Proc. of the Intl. Symp. of Robotics Research (ISRR)*, October 2019. Submitted.

[2] Eric Veach and Leonidas J Guibas. Optimally combining sampling techniques for monte carlo rendering. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 419–428. ACM, 1995.