When is Particle Filtering Efficient for Planning in Partially Observed Linear Dynamical Systems? Supplementary material

Simon S. Du[1] **Wei Hu**[2] **Zhiyuan Li**[2] **Ruoqi Shen**[1] **Zhao Song**[2] **Jiajun Wu**[3]

[1]University of Washington
[2]Princeton University
[3]Stanford University

A PROBABILITY TOOLS

Lemma A.1 (Matrix Bernstein, Theorem 6.1.1 in Tropp [2015])**.** *Consider a finite sequence* $\{X_1, \cdots, X_m\} \subset \mathbb{R}^{n_1 \times n_2}$ *of independent, random matrices with common dimension* $n_1 \times n_2$*. Assume that*

$$\mathbb{E}[X_i] = 0, \forall i \in [m] \quad \text{and} \quad \|X_i\| \le M, \forall i \in [m].$$

Let $Z = \sum_{i=1}^m X_i$*. Let* $\mathrm{Var}[Z]$ *be the matrix variance statistic of sum:*

$$\mathrm{Var}[Z] = \max \left\{ \Big\| \sum_{i=1}^m \mathbb{E}[X_i X_i^\top] \Big\|, \Big\| \sum_{i=1}^m \mathbb{E}[X_i^\top X_i] \Big\| \right\}.$$

Then

$$\mathbb{E}[\|Z\|] \le (2\mathrm{Var}[Z] \cdot \log(n_1 + n_2))^{1/2} + M \cdot \log(n_1 + n_2)/3.$$

Furthermore, for all $t \ge 0$*,*

$$\Pr[\|Z\| \ge t] \le (n_1 + n_2) \cdot \exp\left(- \frac{t^2/2}{\mathrm{Var}[Z] + Mt/3} \right).$$

B PROOF OF MAIN RESULT

We state the complete version of the proofs shown in Section 3.4 in this section. Parts of Section B.1, B.2 and B.3 have been stated in Section 3.4, we restate here for completeness. In Section B.1, we give a concentration bound on the particle approximation of the latent state. In Section B.2, we study how the error of inference in each round accumulates through the sequential planning process. In Section B.3, we put the pieces together to give the upper bound on the number of particles needed so that the long-run rewards of the two processes are close. We show the proofs of most of the lemmas in this section to Section B.4.

B.1 PARTICLE CONCENTRATION

We first note that at time t, since we know the initial state x_0, the transition matrices $A_{0:t-1}$ and $B_{0:t-1}$ and the past actions $\widehat{u}_0, ..., \widehat{u}_{t-1}$, estimating the state x_t is equivalent to estimating $\xi_0, ..., \xi_{t-1}$. We show in Lemma 3.9 that we can write the states as a function of the initial state, past transformation noise and actions, which follows straightly from our definitions of the processes.

Lemma B.1 (Lemma 3.9)**.** *For any* $t \in [T]$*, we can write the state* x_t *as*

$$x_t = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot (\xi_s + B_s \cdot \widehat{u}_s) + \prod_{s=0}^{t-1} A_s \cdot x_0, \tag{1}$$

the state x_t^* as

$$x_t^* = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot (\xi_s + B_s \cdot u_s^*) + \prod_{s=0}^{t-1} A_s \cdot x_0, \tag{2}$$

and for particle $i \in [N]$,

$$x_t^{(i)} = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\xi_s^{(i)} + B_s \cdot \widehat{u}_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0. \tag{3}$$

Recall from Section 2 that the estimation $\widehat{y}_t$ is given by a weighted average of the states of the simulated particles,

$$\widehat{y}_t = \frac{\sum_{i=1}^N w_t^{(i)} x_t^{(i)}}{\sum_{i=1}^N w_t^{(i)}}. \tag{4}$$

and the estimation $\widetilde{y}_t$ is given by the posterior mean of x_t^* given observations $o_{0:t}$,

$$\widetilde{y}_t = \frac{\int_{x'_{1:t} \in \mathcal{X}^t} \prod_{s=1}^t \mathbb{P}\left[o_s^* \mid x'_s\right] x'_t \mathrm{d}\rho_t(x'_{1:t})}{\int_{x'_{1:t} \in \mathcal{X}^t} \prod_{s=1}^t \mathbb{P}\left[o_s^* \mid x'_s\right] \mathrm{d}\rho_t(x'_{1:t})}. \tag{5}$$

By Lemma 3.9, we know that to estimate x_t and x_t^*, it is enough to estimate $\xi_{0:t-1}$. Surprisingly, we can further show that the estimators $\widehat{y}_t$ and $\widetilde{y}_t$ can be written as a function of estimators $\widehat{\xi}_{t,0:t-1}$ and $\widetilde{\xi}_{t,0:t-1}$, past actions $\widehat{u}_{0:t-1}$ and $u_{0:t-1}^*$, and the initial state x_0. The estimator $\widehat{\xi}_{t,0:t-1}$ is given by a weighted average of the noise of the particles, $\xi_{0:t-1}^{(1)}, ..., \xi_{0:t-1}^{(N)}$, similar to (4). The estimator $\widetilde{\xi}_{t,0:t-1}$ is given by the posterior mean of the noise given observations, similar to (5). We show this formally in Lemma B.2.

Lemma B.2. *At time $t \in [T]$, for any $s = 0, .., t-1$, if we estimate ξ_s as $\widehat{\xi}_{t,s}$, given by,*

$$\widehat{\xi}_{t,s} = \frac{\sum_{i=1}^N w_t^{(i)} \xi_s^{(i)}}{\sum_{i=1}^N w_t^{(i)}},$$

$\widehat{y}_t$ can be written as

$$\widehat{y}_t = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widehat{\xi}_{t,s} + B_s \cdot \widehat{u}_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0.$$

Moreover, let $\pi_t(\xi_{0:t-1})$ be the distribution of $\xi_{0:t-1}$, given by the density $\prod_{s=0}^{t-1} \eta_s(\xi_s)$. If we estimate ξ_s as $\widetilde{\xi}_{t,s}$, given by,

$$\widetilde{\xi}_{t,s} = \frac{\int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t}^* \mid \xi'_{0:t-1}, u_{0:t-1}^*, x_0\right] \xi'_s \mathrm{d}\pi_t(\xi'_{0:t-1})}{\int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t}^* \mid \xi'_{0:t-1}, u_{0:t-1}^*, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1})},$$

$\widetilde{y}_t$ can be written as

$$\widetilde{y}_t = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widetilde{\xi}_{t,s} + B_s \cdot u_s^*\right) + \prod_{s=0}^{t-1} A_s \cdot x_0.$$

By Lemma B.2, since $\widehat{u}_{0:t-1}$ and $u_{0:t-1}$ are determined by $\widehat{y}_{0:t-1}$ and $\widetilde{y}_{t-1}$, to show that $\widehat{y}_t$ is close to $\widetilde{y}_t$, it is enough to show that $\widehat{\xi}_{s,0:s-1}$ is close to $\widetilde{\xi}_{s,0:s-1}$ in all rounds $s = 1, ..., t-1$. In this section, we focus on showing how accurately $\xi_{t,0:t-1}$ can approximate $\widetilde{\xi}_{t,0:t-1}$ in one time step. We postpone the discussion of how the error of this approximation accumulates through the process to Section B.2.

To see how $\widehat{\xi}_{t,s}$ can approximate $\widetilde{\xi}_{t,s}$, we study the numerator and the denominator of $\widehat{\xi}_{t,s}$ and $\widetilde{\xi}_{t,s}$ seperately. To simplify our notation, we make the following definition.

Definition B.3. *We define the scalar* $\gamma_t \in R$ *and vector* $\Gamma_{t,s} \in \mathbb{R}^d$ *for any time* $0 \leq s < t \leq T$ *as follows:*

$$\gamma_t = \int_{\xi'_{0:t-1}} \mathbb{P}\left[o^*_{1:t} \mid \xi'_{0:t-1}, u^*_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1}),$$
$$\Gamma_{t,s} = \int_{\xi'_{0:t-1}} \mathbb{P}\left[o^*_{1:t} \mid \xi'_{0:t-1}, u^*_{0:t-1}, x_0\right] \xi'_s \mathrm{d}\pi_t(\xi'_{0:t-1}).$$

We can further show that since we couple our two processes using the same noise, the posterior mean of the noise given observations, $o_{1:t}$, and actions $\widehat{u}_{0:t-1}$ in the approximate process is the same as that given observations, $o^*_{1:t}$, and actions $u^*_{0:t-1}$ in the ideal process.

Lemma B.4. *For any time* $0 \leq s < t \leq T$ *,*

$$\gamma_t = \int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1}),$$
$$\Gamma_{t,s} = \int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \xi'_s \mathrm{d}\pi_t(\xi'_{0:t-1}).$$

Then, to show that the particle approximation, $\widehat{\xi}_{t,s}$, is close to $\widetilde{\xi}_{t,s}$, it is enough to show that $\widehat{\xi}_{t,s}$ concentrates around the posterior mean of $\xi_{t,s}$. We show the relationship between the accuracy of particle approximation and the number of particles, N, in the following lemma.

Lemma B.5 (Lemma 3.10)**.** *Let* $M := \sqrt{\frac{d}{m}(1 + 2\sqrt{\log \beta'/d} + 2\log \beta'/d)}$ *for some* $\beta' > 1$. *At time* $t \in [T]$, *for each* $s = 0, .., t-1$, *we have for any* $\beta \leq \frac{1}{2}$,

$$\|\widehat{\xi}_{t,s} - \widetilde{\xi}_{t,s}\| \leq 4\beta M,$$

holds with probability at least

$$1 - (d+1)\exp(-N\beta^2\gamma_t/3) - N\exp(-\beta').$$

We defer the proof of Lemma B.2, Lemma B.4 and Lemma 3.10 to Appendix B.4.

B.2 ERROR ACCUMULATION

In Section B.1, we studied how the particle approximation concentrates in one time step. In this Section, we discuss how the error of approximation in one time step can affect the actions in the future and further affect the long-run reward of the process.

Lemma 3.9 shows that the states of the two processes, x_t and x^*_t, at time step t, can be written as

$$x_t = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot (\xi_s + B_s \cdot \widehat{u}_s) + \prod_{s=0}^{t-1} A_s \cdot x_0,$$
$$x^*_t = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot (\xi_s + B_s \cdot u^*_s) + \prod_{s=0}^{t-1} A_s \cdot x_0.$$

It is easy to see that the distance between x_t and x^*_t is determined by the distance between actions in the past time steps, $\widehat{u}_{0:t-1}$ and $u^*_{0:t-1}$.

Lemma B.6 (Lemma 3.11)**.** *At time* t,

$$x_t - x^*_t = \sum_{s=0}^{t-1} \left(\prod_{s'=s+1}^{t-1} A_{s'} \right) B_s \left(\widehat{u}_s - u^*_s\right).$$

The actions $\widehat{u}_s$ and u_s^* at $s = 1, ..., t-1$ is determined by the state estimations, $\widehat{y}_s$ and $\widetilde{y}_s$,

$$\widehat{u}_s = g(\widehat{y}_t) \quad \text{and} \quad u_s^* = g(\widetilde{y}_s).$$

Moreover, by Lemma B.2, at time step t,

$$\widehat{y}_t = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widehat{\xi}_{t,s} + B_s \cdot \widehat{u}_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0,$$
$$\widetilde{y}_t = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widetilde{\xi}_{t,s} + B_s \cdot u_s^*\right) + \prod_{s=0}^{t-1} A_s \cdot x_0.$$

which shows that $\widehat{y}_t$ and $\widetilde{y}_s$ are in turn determined by the past actions. Thus, the key step of bounding the error accumulation is bounding the distance between the actions in the two processes. We show the upper bound on the action distance in the following lemma.

Lemma B.7 (Lemma 3.12). *Assume that* $\max_{0 \le s < t \le T} \|\widehat{\xi}_{t,s} - \widetilde{\xi}_{t,s}\| = \epsilon$. *At time t, we can show the following bounds on* $\|\widehat{u}_t - u_t^*\|$.

- *Under Assumptions 3.1, 3.2 and 3.3, for* $t \in [T]$, *let* $\Sigma_a^{(t)} = 1 + C_a \sum_{s=0}^{t-2} \rho_a^s$ *and* $\Sigma_{ab}^{(t-1)} = \sum_{s=0}^{t-2} (C_a + C_b L_g)^s$. *Then, we have*

$$\|\widehat{u}_t - u_t^*\| \le L_g \Sigma_a^{(t)} \left(1 + L_g C_b \Sigma_{ab}^{(t-1)}\right) \cdot \epsilon.$$

- *Under Assumptions 3.1, 3.2 and 3.6, for* $t \in [T]$, *let* $\Sigma_a^{(t)} = 1 + C_a \sum_{s=0}^{t-2} \rho_a^s$ *and* $\bar{\Sigma}_{ab}^{(t-1)} = 1 + C_{ab} \sum_{s=0}^{t-3} \rho_{ab}^s$. *Then, we have*

$$\|\widehat{u}_t - u_t^*\| \le L_g \Sigma_a^{(t)} \left(1 + C_{bg} \bar{\Sigma}_{ab}^{(t-1)}\right) \cdot \epsilon.$$

We defer the proof of Lemma 3.12 to Appendix B.4. Lemma 3.11 and Lemma 3.12 together show that we can bound the distance between the states and the action of the two processes in terms of the accuracy of the particle approximation of transformation noise ξ_t.

B.3 BOUND ON REWARD DIFFERENCE

In this section, we combine the results from Section B.1 and Section B.2 to show an upper bound on the number of particle needed so that the rewards of the two processes are close. Lemma 3.10 upper bounds the number of particles needed so that the particle approximation of the noise ξ_t is accurate. Lemma 3.11 and Lemma 3.12 show that the actions and the states of the two processes are close if the particle approximation is accurate. Then, for reward function that depends on states and actions, we can combine these results to upper bound the number of particles that can guarantee the rewards of the two processes are close. We state our main result in Theorem 3.7 and show the proof below.

Proof of Theorem 3.4 and Theorem 3.7. We state the proof for the Lipschitz g case here. The proof for linear g follows the same steps. We first show the number of particles needed so that the estimation of the noise, ξ_t, in a single round is accurate. If

$$N = \Omega(\beta^{-2} p^{-1} \log(dT/\delta)), \tag{6}$$

then

$$(d+1) \cdot \exp(-N\beta^2 \gamma_t / 3) \le (d+1) \cdot \exp(-N\beta^2 p/3) \le \delta/(2T^2),$$

where the first inequality follows from

$$\gamma_t = \mathbb{P}_{O_{1:t}^*}[o_{1:t}^* | u_{0:t}^*, x_0] = \mathbb{P}_{O_{1:t}}[o_{1:t} | \widehat{u}_{0:t}, x_0] \ge p.$$

Let $M := \sqrt{\frac{d}{m}(1+2\sqrt{\log\beta'/d}+2\log\beta'/d)}$. If we choose $\beta' = \log(2T^2N/\delta)$ and $\beta = \epsilon/(4MT)$, by Lemma 3.10, with success probability at least

$$1 - \sum_{t=1}^{T}\sum_{s=0}^{t-1}\delta/(2T^2) - \sum_{t=1}^{T}\sum_{s=0}^{t-1}\delta/(2T^2) \geq 1-\delta,$$

we have for all time step $t = 1, ..., T$ and $s = 0, .., t-1$,

$$\|\widehat{\xi}_{t,s} - \widetilde{\xi}_{t,s}\| \leq 4\beta M = \epsilon/T.$$

Next, we bound the distance between actions in the two processes. By Lemma 3.12, for any $t = 1, ..., T$,

$$\|\widehat{u}_t - u_t^*\| \leq L_g \Sigma_a^{(t)}\left(1 + L_g C_b \Sigma_{ab}^{(t-1)}\right)\cdot\frac{\epsilon}{T}. \tag{7}$$

The second inequality follows from our assumption. By Lemma 3.11 and our assumptions, we can further bound the distance between the states of the two processes as

$$\begin{aligned}\|x_t - x_t^*\| = \left\|\sum_{s=0}^{t-1}\left(\prod_{s'=s+1}^{t-1}A_{s'}\right)B_s(\widehat{u}_s - u_s^*)\right\| &\leq\ C_b\left(\left\|\widehat{u}_{t-1} - u_{t-1}^*\right\| + C_a\sum_{s=0}^{t-2}\rho_a^s\left\|\widehat{u}_s - u_s^*\right\|\right)\\ &\leq C_b\Sigma_a^{(t)}\cdot L_g\Sigma_a^{(t)}\left(1 + L_g C_b\Sigma_{ab}^{(t-1)}\right)\cdot\frac{\epsilon}{T},\end{aligned} \tag{8}$$

Thus, combining (7) and (8), we can get for any L_r-Lipschitz reward function r_T,

$$\begin{aligned}r_T(x_{1:T}, \widehat{u}_{0:T-1}) - r_T(x_{1:T}^*, u_{0:T-1}^*) &\leq \sum_{t=1}^{T}L_r\|x_t - x_t^*\| + \sum_{t=1}^{T}L_r\|\widehat{u}_{t-1} - u_{t-1}^*\|\\ \leq L_r L_g\Sigma_a^{(T)}\left(1 + C_b\Sigma_a^{(T)}\right)\left(1 + L_g C_b\Sigma_{ab}^{(T-1)}\right)\epsilon,\end{aligned}$$

where the first step follows from r_T is L_r-Lipschitz and the second step follows from (7) and (8).

Plugging $\beta^2 = \epsilon^2/(16T^2M^2) = \widetilde{\Theta}(\epsilon^2T^{-2}d^{-1}m)$ into (6), the number of particles needed is

$$N = \widetilde{O}(\beta^{-2}p^{-1}) = \widetilde{O}(T^2dm^{-1}\epsilon^{-2}p^{-1}),$$

which completes the proof. Similarly, we can also show that the number of particles needed for linear g so that

$$r_T(x_{1:T}, \widehat{u}_{0:T-1}) - r_T(x_{1:T}^*, u_{0:T-1}^*) \leq\ L_r L_g\Sigma_a^{(T)}\left(1 + C_b\Sigma_a^{(T)}\right)\left(1 + C_{bg}\bar{\Sigma}_{ab}^{(T-1)}\right)\epsilon$$

is

$$N = \widetilde{O}(T^2dm^{-1}\epsilon^{-2}p^{-1}).$$

□

B.4 DEFERRED PROOFS

Lemma B.2. *At time $t \in [T]$, for any $s = 0, .., t-1$, if we estimate ξ_s as $\widehat{\xi}_{t,s}$, given by,*

$$\widehat{\xi}_{t,s} = \frac{\sum_{i=1}^N w_t^{(i)}\xi_s^{(i)}}{\sum_{i=1}^N w_t^{(i)}},$$

$\widehat{y}_t$ can be written as

$$\widehat{y}_t = \sum_{s=0}^{t-1}\prod_{s'=s+1}^{t-1}A_{s'}\cdot\left(\widehat{\xi}_{t,s} + B_s\cdot\widehat{u}_s\right) + \prod_{s=0}^{t-1}A_s\cdot x_0.$$

Moreover, let $\pi_t(\xi_{0:t-1})$ be the distribution of $\xi_{0:t-1}$, given by the density $\prod_{s=0}^{t-1} \eta_s(\xi_s)$. If we estimate ξ_s as $\widetilde{\xi}_{t,s}$, given by,

$$\widetilde{\xi}_{t,s} = \frac{\int_{\xi'_{0:t-1}} \mathbb{P}\left[o^*_{1:t} \mid \xi'_{0:t-1}, u^*_{0:t-1}, x_0\right] \xi'_s \mathrm{d}\pi_t(\xi'_{0:t-1})}{\int_{\xi'_{0:t-1}} \mathbb{P}\left[o^*_{1:t} \mid \xi'_{0:t-1}, u^*_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1})},$$

$\widetilde{y}_t$ can be written as

$$\widetilde{y}_t = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widetilde{\xi}_{t,s} + B_s \cdot u^*_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0.$$

Proof. By Lemma 3.9, for every particle $i \in [N]$,

$$x_t^{(i)} = \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \left(\xi_s^{(i)} + B_s \cdot \widehat{u}_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0.$$

Then,

$$\begin{aligned}
\widehat{y}_t &= \frac{\sum_{i=1}^N w_t^{(i)} x_t^{(i)}}{\sum_{i=1}^N w_t^{(i)}}, \\
&= \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\frac{\sum_{i=1}^N w_t^{(i)} \xi_t^{(i)}}{\sum_{i=1}^N w_t^{(i)}} + B_s \cdot \widehat{u}_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0 \\
&= \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widehat{\xi}_s + B_s \cdot \widehat{u}_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0.
\end{aligned}$$

Similarly, by Lemma 3.9 and the definition of ρ_t,

$$\begin{aligned}
\widetilde{y}_t &= \frac{\int_{x'_{1:t} \in \mathcal{X}^t} \prod_{s=0}^t \mathbb{P}\left[o^*_s \mid x'_s\right] x'_t \mathrm{d}\rho_t(x'_{1:t})}{\int_{x'_{1:t} \in \mathcal{X}^t} \prod_{s=0}^t \mathbb{P}\left[o^*_s \mid x'_s\right] \mathrm{d}\rho_t(x'_{1:t})} \\
&= \frac{\int_{\xi'_{0:t-1}} \mathbb{P}\left[o^*_{1:t} \mid \xi'_{0:t-1}, u^*_{0:t-1}, x_0\right]}{\int_{\xi'_{0:t-1}} \mathbb{P}\left[o^*_{1:t} \mid \xi'_{0:t-1}, u^*_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1})} \left[\sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\xi'_s + B_s \cdot u^*_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1}) \\
&= \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widetilde{\xi}_{t,s} + B_s \cdot u^*_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0.
\end{aligned}$$

□

Lemma B.4. *For any time $0 \le s < t \le T$,*

$$\begin{aligned}
\gamma_t &= \int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1}), \\
\Gamma_{t,s} &= \int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \xi'_s \mathrm{d}\pi_t(\xi'_{0:t-1}).
\end{aligned}$$

Proof. For any $t \in [T]$, we have

$$
\begin{aligned}
&\mathbb{P}_{O^*_{1:t}}\left[o^*_{1:t} \mid \xi'_{0:t-1}, u^*_{0:t-1}, x_0\right] \\
&= \prod_{t'=1}^{t} \mathbb{P}_{O^*_{1:t}}\left[o^*_{t'} \mid \xi'_{0:t'-1}, u^*_{0:t'-1}, x_0\right] \\
&= \prod_{t'=1}^{t} \eta_{t'}\left(\left[\sum_{s=0}^{t'-1}\prod_{s'=s+1}^{t'-1} A_{s'}\cdot(\xi_s + B_s\cdot\widehat{u}_s) + \prod_{s=0}^{t'-1} A_s\cdot x_0 + \zeta_{t'}\right] - \left[\sum_{s=0}^{t'-1}\prod_{s'=s+1}^{t'-1} A_{s'}\cdot\left(\widehat{\xi'}_s + B_s\cdot\widehat{u}_s\right) + \prod_{s=0}^{t'-1} A_s\cdot x_0\right]\right) \\
&= \prod_{t'=1}^{t} \eta_{t'}\left(\sum_{s=0}^{t'-1}\prod_{s'=s+1}^{t'-1} A_{s'}\cdot(\xi_s - \xi'_s) + \zeta_{t'}\right) \\
&= \mathbb{P}_{O_{1:t}}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right],
\end{aligned}
$$

where the third step follows from Lemma 3.9, so

$$
\int_{\xi'_{0:t-1}} \mathbb{P}_{O_{1:t}}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1}) = \int_{\xi'_{0:t-1}} \mathbb{P}_{O^*_{1:t}}\left[o^*_{1:t} \mid \xi'_{0:t-1}, u^*_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1}),
$$

$$
\int_{\xi'_{0:t-1}} \mathbb{P}_{O_{1:t}}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \xi'_s \mathrm{d}\pi_t(\xi'_{0:t-1}) = \int_{\xi'_{0:t-1}} \mathbb{P}_{O^*_{1:t}}\left[o^*_{1:t} \mid \xi'_{0:t-1}, u^*_{0:t-1}, x_0\right] \xi'_s \mathrm{d}\pi_t(\xi'_{0:t-1}).
$$

□

Now, we state the proof of Lemma 3.10.

Proof of Lemma 3.10. We first consider the random variables

$$
\mathbb{P}\left[o_{1:t} \mid \xi^{(i)}_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] = \prod_{t'=1}^{t} \eta_{t'}\left(\sum_{s=0}^{t'-1}\prod_{s'=s+1}^{t'-1} A_{s'}\cdot(\xi_s - \xi'_s) + \zeta_{t'}\right),
$$

for $i = 1, ..., N$. By the way we generate $\xi^{(1)}_{0:t}, \xi^{(2)}_{0:t}, ..., \xi^{(N)}_{0:t}$,

$$
\mathbb{P}\left[o_{1:t} \mid \xi^{(1)}_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right], \mathbb{P}\left[o_{1:t} \mid \xi^{(2)}_{0:t}, \widehat{u}_{0:t}, x_0\right], ..., \mathbb{P}\left[o_{1:t} \mid \xi^{(N)}_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right]
$$

are independent. Also, for $i = 1, ..., N$, by Lemma B.4

$$
\mathbb{E}\left[\mathbb{P}\left[o_{1:t} \mid \xi^{(i)}_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right]\right] = \int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1}) = \gamma_t.
$$

and

$$
\mathbb{E}\left[\mathbb{P}\left[o_{1:t} \mid \xi^{(i)}_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \xi'_s\right] = \int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \xi'_s \mathrm{d}\pi_t(\xi'_{0:t-1}) = \Gamma_{t,s}.
$$

By Lemma A.1,

$$
\begin{aligned}
&\Pr\left[\left|\frac{1}{N}\sum_{i=1}^{N} \mathbb{P}\left[o_{1:t} \mid \xi^{(i)}_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] - \gamma_t\right| \geq \beta\gamma_t\right] \\
&\leq \exp\left(-\frac{N\beta^2\gamma_t^2}{2\mathrm{Var}\left[\mathbb{P}\left[o_{1:t} \mid \xi^{(i)}_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right]\right] + \frac{2}{3}\max\left|\mathbb{P}\left[o_{1:t} \mid \xi^{(i)}_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right]\right|\beta\gamma_t}\right) \\
&\leq \exp(-N\beta^2\gamma_t/3).
\end{aligned}
$$

where the third step follows from

$$\mathrm{Var}\left[\mathbb{P}\left[o_{1:t} \mid \xi_{0:t-1}^{(i)}, \widehat{u}_{0:t-1}, x_0\right]\right] \leq \mathbb{E}\left[\mathbb{P}^2\left[o_{1:t} \mid \xi_{0:t-1}^{(i)}, \widehat{u}_{0:t-1}, x_0\right]\right]$$
$$= \int_{\xi'_{0:t-1}} \mathbb{P}^2\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1})$$
$$\leq \gamma_t,$$

and

$$\max\left|\mathbb{P}\left[o_{1:t} \mid \xi_{0:t-1}^{(i)}, \widehat{u}_{0:t-1}, x_0\right]\right| \leq 1.$$

Without loss of generality, we assume the noise ξ has mean zero. Since the noise $\|\xi_s^{(i)}\|$ is sub-gaussian, with probability at least $1 - N\exp(-\beta')$, for all $i \in [N]$,

$$\|\xi_s^{(i)}\|^2 \leq M^2 = \frac{d}{m}(1 + 2\sqrt{\log\beta'/d} + 2\log\beta'/d).$$

Similarly, by Lemma A.1, since the noise $\|\xi_s^{(i)}\| \leq M$ for all $i \in [N]$,

$$\Pr\left[\left\|\frac{1}{N}\sum_{i=1}^{N}\mathbb{P}\left[o_{1:t} \mid \xi_{0:t-1}^{(i)}, \widehat{u}_{0:t-1}, x_0\right]\xi_s^{(i)} - \Gamma_t\right\| \geq \beta\gamma_t M\right]$$
$$\leq d\cdot\exp\left(-\frac{N\beta^2\gamma_t^2 M^2}{2\mathrm{Var}\left[\mathbb{P}\left[o_{1:t} \mid \xi_{0:t-1}^{(i)}, \widehat{u}_{0:t-1}, x_0\right]\xi_s^{(i)}\right] + \frac{2}{3}\max\left\|\mathbb{P}\left[o_{1:t} \mid \xi_{0:t-1}^{(i)}, \widehat{u}_{0:t-1}, x_0\right]\xi_s^{(i)}\right\|\beta\gamma_t M}\right)$$
$$\leq d\cdot\exp(-N\beta^2\gamma_t/3).$$

where the third step follows from

$$\mathrm{Var}\left[\mathbb{P}\left[o_{1:t} \mid \xi_{0:t-1}^{(i)}, \widehat{u}_{0:t-1}, x_0\right]\xi_s^{(i)}\right] \leq \gamma_t M^2,$$

and

$$\max\left\|\mathbb{P}\left[o_{1:t} \mid \xi_{0:t-1}^{(i)}, \widehat{u}_{0:t-1}, x_0\right]\xi_s^{(i)}\right\| \leq M.$$

Then, with probability at least

$$1 - (d+1)\exp(-N\beta^2\gamma_t/3) - N\exp(-\beta'),$$

we have

$$\|\widehat{\xi}_{t,s} - \widetilde{\xi}_{t,s}\| = \left\|\widehat{\xi}_{t,s} - \frac{\Gamma_{t,s}}{\gamma_t}\right\|$$
$$\leq \max\left\{\left\|\frac{1}{(1-\beta)\gamma_t}\sum_{i=1}^{N} w_t^{(i)}\xi_{t'}^{(i)} - \frac{\Gamma_{t,s}}{\gamma_t}\right\|, \left\|\frac{1}{(1+\beta)\gamma_t}\sum_{i=1}^{N} w_t^{(i)}\xi_{t'}^{(i)} - \frac{\Gamma_{t,s}}{\gamma_t}\right\|\right\}$$
$$\leq \max\left\{\frac{1}{1-\beta}\left\|\frac{\sum_{i=1}^{N} w_t^{(i)}\xi_{t'}^{(i)}}{\gamma_t} - \frac{\Gamma_{t,s}}{\gamma_t}\right\| + \left(\frac{1}{1-\beta} - 1\right)\left\|\frac{\Gamma_{t,s}}{\gamma_t}\right\|,\right.$$
$$\left.\frac{1}{1+\beta}\left\|\frac{\sum_{i=1}^{N} w_t^{(i)}\xi_{t'}^{(i)}}{\gamma_t} - \frac{\Gamma_{t,s}}{\gamma_t}\right\| + \left(1 - \frac{1}{1+\beta}\right)\left\|\frac{\Gamma_{t,s}}{\gamma_t}\right\|\right\}$$
$$\leq \max\left\{\frac{\beta M}{1-\beta} + \frac{\beta M}{1-\beta}, \frac{\beta M}{1+\beta} + \frac{\beta M}{1+\beta}\right\}$$
$$\leq 4\beta M$$

where the first step follows from $|\sum_{i=1}^N w_t^{(i)} - \gamma_t| \le \beta\gamma_t$, the second step follows from triangle inequality, the third step follows from $\left\|\sum_{i=1}^N w_t^{(i)} \xi_{t'}^{(i)} - \Gamma_{t,s}\right\| \le \beta\gamma_t M$, and

$$\left\|\frac{\Gamma_{t,s}}{\gamma_t}\right\| = \left\|\frac{\int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \xi'_s \mathrm{d}\pi_t(\xi'_{0:t-1})}{\int_{\xi'_{0:t-1}} \mathbb{P}\left[o_{1:t} \mid \xi'_{0:t-1}, \widehat{u}_{0:t-1}, x_0\right] \mathrm{d}\pi_t(\xi'_{0:t-1})}\right\| \le M,$$

and the last step follows from $\beta \le \frac{1}{2}$. □

Finally, we state the proof of Lemma 3.12.

Proof of Lemma 3.12. When $t = 0$, we have $\widehat{u}_0 = u_0^* = g(x_0)$, so $\|\widehat{u}_0 - u_0^*\| = 0$. For $t > 0$, we study the two cases separately.

In the first case, for L_g-Lipschitz g, at time $t > 0$,

$$\begin{aligned}
&\|\widehat{u}_t - u_t^*\| = \|g(\widehat{y}_t) - g(\widetilde{y}_t)\| \\
&= \left\|g\left(\sum_{s=0}^{t-1}\prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widehat{\xi}_{t,s} + B_s \cdot \widehat{u}_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0\right) - g\left(\sum_{s=0}^{t-1}\prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widetilde{\xi}_{t,s} + B_s \cdot u_s^*\right) + \prod_{s=0}^{t-1} A_s \cdot x_0\right)\right\| \\
&\le L_g \cdot \left\|\left(\sum_{s=0}^{t-1}\prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widehat{\xi}_{t,s} + B_s \cdot \widehat{u}_s\right) + \prod_{s=0}^{t-1} A_s \cdot x_0\right) - \left(\sum_{s=0}^{t-1}\prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widetilde{\xi}_{t,s} + B_s \cdot u_s^*\right) + \prod_{s=0}^{t-1} A_s \cdot x_0\right)\right\| \\
&\le L_g \cdot \left\|\sum_{s=0}^{t-1}\prod_{s'=s+1}^{t-1} A_{s'} \cdot (\widehat{\xi}_{t,s} - \widetilde{\xi}_{t,s})\right\| + L_g \cdot \left\|\sum_{s=0}^{t-1}\prod_{s'=s+1}^{t-1} A_{s'} B_s \cdot (\widehat{u}_s - u_s^*)\right\|.
\end{aligned}$$

where the first step follows from definitions of $\widehat{u}_t$ and u_t^*, the second step follows from Lemma B.2, the third step follows from g is L_g-Lipschitz and the last step follows from triangle inequality.

We define f_t and h_t as follows:

$$f_t = \left\|\sum_{s=0}^{t-1}\prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widehat{\xi}_{t,s} - \widetilde{\xi}_{t,s}\right)\right\|,$$
$$h_t = \left\|\sum_{s=0}^{t-1}\prod_{s'=s+1}^{t-1} A_{s'} B_s \cdot (\widehat{u}_s - u_s^*)\right\|.$$

Then,

$$\begin{aligned}
h_t &\le C_a h_{t-1} + C_b\|\widehat{u}_{t-1} - u_{t-1}^*\| \\
&\le C_a h_{t-1} + C_b L_g (f_{t-1} + h_{t-1}) \\
&\le C_b L_g f_{t-1} + (C_a + C_b L_g)\left(C_a h_{t-2} + C_b L_g (f_{t-2} + h_{t-2})\right) \\
&\le \cdots \\
&\le C_b L_g \sum_{s=1}^{t-1} (C_a + C_b L_g)^{t-s-1} f_s,
\end{aligned}$$

The first step follows from definition. The second step follows from $\|\widehat{u}_{t-1} - u_{t-1}^*\| \le L_g(f_{t-1} + h_{t-1})$. The third step and the last step follow from induction. Thus, for L_g-Lipschitz g,

$$\begin{aligned}
\|\widehat{u}_t - u_t^*\| &\le L_g \cdot f_t + L_g \cdot C_b L_g \sum_{s=1}^{t-1} (C_a + C_b L_g)^{t-s-1} f_s \\
&\le L_g\left(1 + C_a \sum_{s=0}^{t-2} \rho_a^s\right)\epsilon + L_g \cdot C_b L_g \sum_{s=1}^{t-1} (C_a + C_b L_g)^{t-s-1} \cdot \left(1 + C_a \sum_{s'=0}^{s-2} \rho_a^{s'}\right) \cdot \epsilon \\
&\le L_g \Sigma_a^{(t)} \cdot \epsilon + L_g \cdot C_b L_g \Sigma_{ab}^{(t-1)} \cdot \Sigma_a^{(t)} \cdot \epsilon = L_g \Sigma_a^{(t)}\left(1 + L_g C_b \Sigma_{ab}^{(t-1)}\right) \cdot \epsilon.
\end{aligned}$$

The second step follows from our assumption and the last two steps follow from our definitions of $\Sigma_a^{(t)}$ and $\Sigma_{ab}^{(t)}$.

In the second case, for linear $g = G$, similarly, we have

$$\begin{aligned} \widehat{u}_t - u_t^* &= g(\widehat{y}_t) - g(\widetilde{y}_t) \\ &= G \cdot \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot (\widehat{\xi}_{t,s} - \widetilde{\xi}_{t,s}) + G \cdot \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} B_s \cdot (\widehat{u}_s - u_s^*). \end{aligned}$$

We define f_t and h_t as follows:

$$\begin{aligned} f_t &= \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} \cdot \left(\widehat{\xi}_{t,s} - \widetilde{\xi}_{t,s}\right), \\ h_t &= \sum_{s=0}^{t-1} \prod_{s'=s+1}^{t-1} A_{s'} B_s \cdot (\widehat{u}_s - u_s^*). \end{aligned}$$

then by induction,

$$\begin{aligned} h_t &= A_{t-1} h_{t-1} + B_{t-1} \left(\widehat{u}_{t-1} - u_{t-1}^*\right) \\ &\le A_{t-1} h_{t-1} + B_{t-1} G (f_{t-1} + h_{t-1}) \\ &\le B_{t-1} G f_{t-1} + (A_{t-1} + B_{t-1} G) \left(A_{t-2} h_{t-2} + B_{t-2} G (f_{t-2} + h_{t-2})\right) \\ &\le \cdots \\ &\le \sum_{s=1}^{t-1} \prod_{s'=s+1}^{t-1} (A_{s'} + B_{s'} G) B_s G f_s, \end{aligned}$$

Thus, by our assumptions and the definitions of $\Sigma_a^{(t)}$ and $\bar{\Sigma}_{ab}^{(t)}$.

$$\begin{aligned} \|\widehat{u}_t - u_t^*\| &= \left\| G \cdot f_t + G \cdot \sum_{s=1}^{t-1} \prod_{s'=s+1}^{t-1} (A_{s'} + B_{s'} G) B_s G f_s \right\| \\ &\le L_g \left(1 + C_a \sum_{s=0}^{t-2} \rho_a^s\right) \cdot \epsilon + L_g \cdot C_{bg} \left(1 + C_{ab} \sum_{s=0}^{t-3} \rho_{ab}^s\right) \cdot \left(1 + C_a \sum_{s'=0}^{s-2} \rho_a^{s'}\right) \cdot \epsilon \\ &\le L_g \Sigma_a^{(t)} \left(1 + C_{bg} \bar{\Sigma}_{ab}^{(t-1)}\right) \cdot \epsilon. \end{aligned}$$

□

C EXPERIMENT

In this section, we use simulations to show the error of particle filtering can accumulate and be amplified through sequential planning. We run a process with a maximum time step $T = 40$ and $d = 1$. Since our bound shows the number of particles needed is insensitive to the dimension d, we mainly show how the number of time steps can affect the accuracy of particle filtering.

We consider the following process, for all $t \in [T]$,

$$x_t = x_{t-1} + u_{t-1} + \xi_{t-1}, \quad \text{and } o_t = x_t + \zeta_t.$$

The process can suffer a random shift of size 1, i.e., ξ_t follows a uniform distribution on set $\{0, 1\}$. ζ_t follows the standard normal distribution $\mathcal{N}(0, 1)$. The regret is defined as the average ℓ_1 norm of the states, i.e., $r(x_{1:t}) = \sum_{i=1}^t |x_t|/t$. The policy function is $g(x) = -x$. We show in Figure 1 the regret and its standard deviation of the estimation using different number of particles. The result shows the number of particles needed for an accurate estimation can increase fast as the number of time step increases due to error accumulation. This experiment corroborates the importance of our theoretical results.

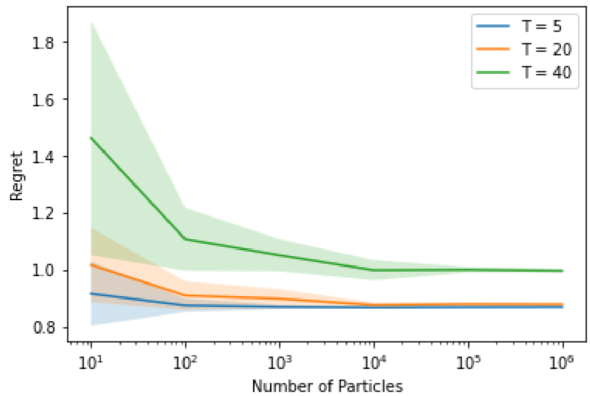

Figure 1: Relationship between the regret and the number of particles.

References

Joel A Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2): 1–230, 2015.