A. Algorithm 1 Proofs

A.1. The Good Event

We begin with an explicit statement of the probabilistic events that comprise $\mathcal{E}_A$. Recall that

$$A_t = \arg\min_A \sum_{s=1}^{t-1} \|x_{s+1} - B_\star u_s - A x_s\|^2 + \lambda \|A\|_F^2,$$

and denote $\Delta_{A_t} = A_t - A_\star$, $V_t^x = \lambda I + \sum_{s=1}^{t-1} x_t x_t^T$. Now, define the following events

$$\mathcal{E}_{A_{ols}} = \left\{ \mathrm{Tr}\left(\Delta_{A_t}^T V_t^x \Delta_{A_t}\right) \le 4\sigma^2 d \log\left(3T^3 \frac{\det\left(V_t^x\right)}{\det\left(V_1^x\right)}\right) + 2\lambda d \vartheta^2, \text{ for all } t \ge 1 \right\}, \tag{9}$$

$$\mathcal{E}_{A_x} = \left\{ \sum_{t=1}^{\tau_i - 1} x_t x_t^T \succeq \frac{(\tau_i - 1)\sigma^2}{40} I, \text{ for all } 0 \le i \le n_T \right\}, \tag{10}$$

$$\mathcal{E}_{A_w} = \left\{ \max_{1 \le t \le T} \|w_t\| \le \sigma \sqrt{15 d \log 3T} \right\}, \tag{11}$$

Then we have the following lemma.

Lemma 17. *Let $\mathcal{E}_A = \mathcal{E}_{A_{ols}} \cap \mathcal{E}_{A_x} \cap \mathcal{E}_{A_w}$, and suppose that $T \ge 600 d \log 36T$. Then we have that $\mathbb{P}\left(\mathcal{E}_A\right) \ge 1 - T^{-2}$.*

Proof. First, we describe the parameter estimation error in terms of Lemma 6. To that end, let $z_t = x_t$, $y_{t+1} = x_{t+1} - B_\star u_t$, $V_t^x = \lambda I + \sum_{s=1}^{t-1} x_t x_t^T$, and $\Delta_{A_t} = A_t - A_\star$ Indeed, we have $y_{t+1} = A_\star x_t + w_t$, $w_t \sim \mathcal{N}(0, \sigma^2 I)$, and $\|A_\star\|_F^2 \le d\|A_\star\|^2 \le d\vartheta^2$ and so taking Lemma 6 with $\delta = \frac{1}{3}T^{-2}$, recalling that $T \ge d$, and simplifying, we get that $\mathbb{P}\left(\mathcal{E}_{A_{ols}}\right) \ge 1 - \frac{1}{3}T^{-2}$.

Next, for $\mathcal{E}_{A_x}$, we apply Lemma 36 to the sequence x_t with the filtration $\mathcal{F}_t = \sigma(x_1, u_1, \ldots, x_t, u_t)$. Notice that given x_{t-1}, u_{t-1} we have $x_t \sim \mathcal{N}(A_\star x_{t-1} + B_\star u_{t-1}, \sigma^2 I)$ and hence we also get

$$\mathbb{E}\left[x_t x_t^T \mid \mathcal{F}_{t-1}\right] \succeq (A_\star x_{t-1} + B_\star u_{t-1})(A_\star x_{t-1} + B_\star u_{t-1})^T + \sigma^2 I \succeq \sigma^2 I.$$

Finally, our choice of τ_0 ensures the minimal sum size assumption. We thus apply Lemma 36 $n_T + 1$ times with $\delta = \frac{1}{3}T^{-3}$ and apply a union bound. Since $n_T + 1 \le T$ we conclude that $\mathbb{P}\left(\mathcal{E}_{A_x}\right) \ge 1 - \frac{1}{3}T^{-2}$.

Finally, for $\mathcal{E}_{A_w}$ we apply Lemma 34 with $\delta = \frac{1}{3}T^{-2}$ to get $\mathbb{P}\left(\mathcal{E}_{A_w}\right) \ge 1 - \frac{1}{3}T^{-2}$. The final result is obtained by taking a union bound over the three events. ■

A.2. Proof of Lemma 7

We first need the following two lemmas.

Lemma 18 (Bounded warm-up)**.** *On $\mathcal{E}_A$ we have that $\|x_t\| \le \sigma\kappa_0^3\sqrt{60 d \log 3T} \le \sqrt{x_b}$, for all $1 \le t \le \tau_0$.*

Proof. First, by Lemma 41, $J(K_0) \le \nu_0$ implies that K_0 is (κ_0, γ_0)–strongly stable with $\gamma_0^{-1} = 2\kappa_0^2$. So, applying Lemma 38 with $x_1 = 0$ we get that for all $1 \le t \le \tau_0$

$$\|x_t\| \le 2\kappa_0^3 \max_{1 \le t \le T} \|w_t\|,$$

and applying the noise bound in Eq. (11) we obtain the desired result. ■

Lemma 19 (Conditional parameter estimation)**.** *On $\mathcal{E}_A$ fix some i such that $0 \le i \le n_T$ and suppose that $\|x_t\|^2 \le x_b$ for all $1 \le t \le \tau_i$. Then we have that $\|\Delta_{A_{\tau_i}}\| \le \varepsilon_0 2^{-i}$.*

Proof. First, on $\mathcal{E}_A$ by Eq. (10) we have that

$$V_{\tau_i}^x = \lambda I + \sum_{t=1}^{\tau_i - 1} x_t x_t^T \succeq \left(\lambda + \frac{(\tau_i - 1)\sigma^2}{40}\right) I \succeq \frac{\tau_i \sigma^2}{40} I,$$

and so we conclude that

$$\mathrm{Tr}\Big(\Delta_{A_{\tau_i}}^T V_{\tau_i}^x \Delta_{A_{\tau_i}}\Big) \geq \mathrm{Tr}\Big(\Delta_{A_{\tau_i}}^T \Delta_{A_{\tau_i}}\Big)\frac{\tau_i\sigma^2}{40} \geq \|\Delta_{A_{\tau_i}}\|^2 \frac{\tau_i\sigma^2}{40}.$$

Rearranging and applying Eq. (9) we obtain

$$\|\Delta_{A_{\tau_i}}\|^2 \leq \frac{1}{\tau_i}\left(160d\log\left(3T^3\frac{\det\big(V_{\tau_i}^x\big)}{\det\big(V_1^x\big)}\right) + 80\frac{\lambda d\vartheta^2}{\sigma^2}\right).$$

Now, since we assumed $\|x_t\|^2 \leq x_b = \lambda$, we can apply Lemma 37 to conclude that

$$\log\frac{\det\big(V_{\tau_i}^x\big)}{\det\big(V_1^x\big)} \leq d\log T,$$

and plugging this into the above we get that

$$\|\Delta_{A_{\tau_i}}\|^2 \leq \frac{1}{\tau_i}\left(640d^2\log(3T) + 80\frac{\lambda d\vartheta^2}{\sigma^2}\right) \leq \frac{1}{\tau_i}\frac{80\lambda d\big(1+\vartheta^2\big)}{\sigma^2} \leq \frac{\varepsilon_0^2\tau_0}{\tau_i} \leq \varepsilon_0^2 4^{-i},$$

where all transitions are due to our choice of parameters. ■

Proof of Lemma 7. First recall that by Lemma 42, if $\|\Delta_{A_t}\| \leq \varepsilon_0$ then K_t is (κ,γ)–strongly stable. We now show by induction on n that for all $0 \leq i \leq n$, K_{τ_i} is (κ,γ)–strongly stable. Note that $0 \leq n \leq n_T$.

For the base case, $n = 0$, Lemma 18 shows that $\|x_t\|^2 \leq x_b$ for all $1 \leq t \leq \tau_0$, which in turn satisfies Lemma 19, i.e., $\|\Delta_{A_{\tau_0}}\| \leq \varepsilon_0$ and so the required strong stability of K_{τ_0} is obtained.

Now, suppose the induction holds up to $n-1$ and we show for n. By the strong stability of the controllers up to time $\tau_n - 1$, and since $\tau_0 \geq \frac{\log\kappa}{\gamma}$, we can apply Lemma 39 to conclude that

$$\|x_t\| \leq 3\kappa\max\left\{\frac{\|x_{\tau_0}\|}{2}, \frac{\kappa}{\gamma}\max_{1\leq t\leq T}\|w_t\|\right\}, \qquad \text{for all } \tau_0 \leq t \leq \tau_i.$$

recalling that $\gamma^{-1} = 2\kappa^2$, bounding the noise with Eq. (11), and bounding $\|x_{\tau_0}\|$ by Lemma 18 we get that

$$\|x_t\| \leq 3\kappa\max\left\{\frac{\sigma\kappa_0^3\sqrt{60d\log 3T}}{2}, 2\kappa^3\sigma\sqrt{15d\log 3T}\right\}, \leq \sigma\kappa\max\big\{\kappa_0^3, 2\kappa^3\big\}\sqrt{135d\log 3T} = \sqrt{x_b},$$

and as for the base case, we can now invoke Lemmas 19 and 42 to conclude the strong stability of K_{τ_n} and finish the induction. Notice that this together with the above equation also show the algorithm does not abort.

The induction proves the first part of the lemma, i.e., all controller are strongly-stable. Now, we can apply Lemma 39 once more to conclude that $\|x_t\|^2 \leq x_b$ for all $\tau_0 \leq t \leq T$ and together with Lemma 18 this concludes the second claim of the lemma.

Finally, the third claim is now an immediate corollary of the Lemma 19. ■

A.3. Proof of Lemma 8

Recall that $E_i = \big\{\|\Delta_{A_{\tau_i}}\| \leq \varepsilon_0 2^{-i}\big\}$, and further denote $S_i = \Big\{\|x_{\tau_i}\|^2 \leq x_b\Big\}$. Trivially, we have that $\mathcal{E}_A \subseteq E_i \cap S_i$.

Now, define $\tilde{x}_{\tau_i} = x_{\tau_i}$ and for $\tau_i < t \leq \tau_{i+1} - 1$

$$\tilde{x}_t = \big(A_\star + B_\star K_{\tau_i}\big)\tilde{x}_{t-1} + w_t.$$

Since on $\mathcal{E}_A$ the algorithm does not abort, we have that

$$\mathbb{1}\{\mathcal{E}_A\}J_i = \mathbb{1}\{\mathcal{E}_A\}\sum_{t=\tau_i}^{\tau_{i+1}-1}\tilde{x}_t^T\big(Q + K_{\tau_i}^T R K_{\tau_i}\big)\tilde{x}_t \leq \mathbb{1}\{E_i \cap S_i\}\sum_{t=\tau_i}^{\tau_{i+1}-1}\tilde{x}_t^T\big(Q + K_{\tau_i}^T R K_{\tau_i}\big)\tilde{x}_t.$$

Noticing that E_i, S_i, and K_{τ_i} are completely determined by x_{τ_i}, A_{τ_i} we use total expectation to get that

$$\mathbb{E}\big[\mathbb{1}\{\mathcal{E}_A\}J_i\big] \le \mathbb{E}\left[\mathbb{1}\{E_i \cap S_i\}\mathbb{E}\left[\sum_{t=\tau_i}^{\tau_{i+1}-1} \tilde{x}_t^T\big(Q + K_{\tau_i}^T R K_{\tau_i}\big)\tilde{x}_t \;\middle|\; x_{\tau_i}, A_{\tau_i}\right]\right].$$

Now, by Lemma 42, E_i implies that K_{τ_i} is (κ, γ)–strongly stable and so we can use Lemma 40 to get that

$$\begin{aligned}
\mathbb{E}\big[\mathbb{1}\{\mathcal{E}_A\}J_i\big] &\le (\tau_{i+1} - \tau_i)\mathbb{E}\big[\mathbb{1}\{E_i\}J\big(K_{\tau_i}\big)\big] + \frac{2\alpha_1\kappa^4}{\gamma}\mathbb{E}\Big[\mathbb{1}\{S_i\}\|x_{\tau_i}\|^2\Big] \\
&\le (\tau_{i+1} - \tau_i)\mathbb{E}\big[\mathbb{1}\{E_i\}J\big(K_{\tau_i}\big)\big] + 4\alpha_1\kappa^6 x_b,
\end{aligned}$$

where the second transition also used that $\gamma^{-1} = 2\kappa^2$ and the third used our choice of $x_b \ge \sigma^2\kappa^4$.

A.4. Proof of Lemma 9 (R_2 upper bound)

We first need the following lemma.

Lemma 20 (Expected abort state)**.** *Suppose that $\mathbb{P}(\tau_{abort} \le T) \le T^{-2}$. Then we have that*

$$\mathbb{E}\Big[\|x_{\tau_{abort}}\|^2\mathbb{1}\{\tau_{abort} < T\}\Big] \le \big(1 + 8\vartheta^2\big)\big(\kappa^2 + \kappa_0^2\big)x_b T^{-2}.$$

Proof. First, by the lemmas assumption, we can apply Lemma 35 to get that

$$\mathbb{E}\left[\mathbb{1}\{\tau_{\text{abort}} \le T\}\max_{1\le t\le T}\|w_t\|^2\right] \le 5d\sigma^2 T^{-2}\log 3T.$$

Now, notice that $\|A_\star + B_\star K\| \le 2\vartheta\|K\|$ and split into two cases. First, if $\tau_{\text{abort}} > \tau_0$ then by definition of τ_{abort} we have that

$$\|x_{\tau_{\text{abort}}}\| = \big\|\big(A_\star + B_\star K_{\tau_{\text{abort}}-1}\big)x_{\tau_{\text{abort}}-1} + w_{\tau_{\text{abort}}-1}\big\| \le 2\vartheta\kappa\sqrt{x_b} + \max_{1\le s\le T}\|w_t\|,$$

and taking expectation we get that

$$\mathbb{E}\Big[\mathbb{1}\{\tau_0 < \tau_{\text{abort}} \le T\}\|x_{\tau_{\text{abort}}}\|^2\Big] \le 8\vartheta^2\kappa^2 x_b T^{-2} + 5d\sigma^2 T^{-2}\log 3T \le \big(1 + 8\vartheta^2\big)\kappa^2 x_b T^{-2}.$$

On the other hand if $\tau_{\text{abort}} = \tau_0$ then

$$\|x_{\tau_{\text{abort}}}\| = \|(A_\star + B_\star K_0)x_{\tau_0-1} + w_{\tau_0-1}\| \le 2\vartheta\kappa_0\|x_{\tau_0-1}\| + \max_{1\le t\le T}\|w_t\| \le (4\vartheta + 1)\kappa_0^4\max_{1\le t\le T}\|w_t\|,$$

where the last transition used Lemma 38 and $\gamma_0^{-1} = 2\kappa_0^2$. Taking expectation we get that

$$\mathbb{E}\Big[\mathbb{1}\{\tau_{\text{abort}} = \tau_0\}\|x_{\tau_{\text{abort}}}\|^2\Big] \le 20\big(1 + 8\vartheta^2\big)\kappa_0^8 d\sigma^2 T^{-2}\log 3T \le \big(1 + 8\vartheta^2\big)\kappa_0^2 x_b T^{-2},$$

and combining both cases yields the final bound. ■

Proof of Lemma 9. First, recall the decomposition of R_2.

$$R_2 \le \mathbb{E}\left[\mathbb{1}\{\mathcal{E}_A^c\}\sum_{t=\tau_0}^{\tau_{\text{abort}}-1} c_t\right] + \mathbb{E}\left[\sum_{t=\tau_{\text{abort}}}^{T} c_t\right].$$

For $\tau_0 \le t < \tau_{\text{abort}}$ we have that $\|x_t\|^2 \le x_b$ and $\|K_t\| \le \kappa$ and so we get that

$$c_t = x_t^T\big(Q + K_t^T R K_t\big)x_t \le \|x_t\|^2\Big(\|Q\| + \|R\|\|K_t\|^2\Big) \le 2\alpha_1\kappa^2 x_b.$$

By Lemma 7 we have that $\mathbb{P}\big(\mathcal{E}_A^c\big) \le T^{-2}$ and so we get that

$$\mathbb{E}\left[\mathbb{1}\{\mathcal{E}_A^c\}\sum_{t=\tau_0}^{\tau_{\text{abort}}-1} c_t\right] \le \mathbb{E}\big[\mathbb{1}\{\mathcal{E}_A^c\}2\alpha_1\kappa^2 x_b T\big] = 2\alpha_1\kappa^2 x_b T\,\mathbb{P}\big(\mathcal{E}_A^c\big) \le 2\alpha_1\kappa^2 x_b T^{-1}, \tag{12}$$

bounding the first term of R_2. Next, for $t \geq \tau_{\text{abort}}$ we have that $K_t = K_0$ and so we can apply Lemma 40 to relate the expected cost of this period to that of the steady state cost of K_0. we get that

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=\tau_{\text{abort}}}^{T} c_t\right] &= \mathbb{E}\left[\mathbb{E}\left[\sum_{t=\tau_{\text{abort}}}^{T} x_t^T\left(Q + K_0^T R K_0\right)x_t \;\middle|\; \tau_{\text{abort}}, x_{\tau_{\text{abort}}}\right]\right] \\
&\leq \mathbb{E}\left[\mathbb{1}\{\tau_{\text{abort}} \leq T\}\left(TJ(K_0) + \frac{2\alpha_1\kappa_0^4}{\gamma_0}\|x_{\tau_{\text{abort}}}\|^2\right)\right] \\
&= TJ(K_0)\mathbb{P}(\tau_{\text{abort}} \leq T) + 4\alpha_1\kappa_0^6\mathbb{E}\left[\|x_{\tau_{\text{abort}}}\|^2\mathbb{1}\{\tau_{\text{abort}} \leq T\}\right],
\end{aligned}$$

where the last transition used $\gamma_0^{-1} = 2\kappa_0^2$. Now, by Lemma 7 we know that on $\mathcal{E}_A$ the algorithm does not abort. We conclude that $\{\tau_{\text{abort}} \leq T\} \subseteq \mathcal{E}_A^c$ which in turn implies $\mathbb{P}(\tau_{\text{abort}} \leq T) \leq \mathbb{P}\left(\mathcal{E}_A^c\right) \leq T^{-2}$. We get that

$$\mathbb{E}\left[\sum_{t=\tau_{\text{abort}}}^{T} c_t\right] \leq J(K_0)T^{-1} + 4\alpha_1\kappa_0^6\mathbb{E}\left[\|x_{\tau_{\text{abort}}}\|^2\mathbb{1}\{\tau_{\text{abort}} \leq T\}\right],$$

Finally, we use Lemma 20 and simplify to get that

$$\begin{aligned}
R_2 &\leq 2\alpha_1\kappa^2 x_b T^{-1} + J(K_0)T^{-1} + 4\alpha_1\kappa_0^6\left(1 + 8\vartheta^2\right)\left(\kappa^2 + \kappa_0^2\right)x_b T^{-2} \\
&= \left(J(K_0) + 2\alpha_1\kappa^2 x_b\right)T^{-1} + 4\alpha_1\kappa_0^6\left(1 + 8\vartheta^2\right)\left(\kappa^2 + \kappa_0^2\right)x_b T^{-2},
\end{aligned}$$

as desired. ∎

A.5. Proof of Lemma 10

Notice that for $t < \tau_0$ we have that $K_t = K_0$. Moreover, we have that $x_1 = 0$. Applying Lemma 40 we get that

$$R_3 = \mathbb{E}\left[\sum_{t=1}^{\tau_0 - 1} x_t^T\left(Q + K_0^T R K_0\right)x_t\right] \leq \tau_0 J(K_0).$$

B. Algorithm 2 Proofs

B.1. The Good Event

We begin by stating the probabilistic events that guarantee the "good" operation of the algorithm. To that end, it will be convenient to specify how the randomized actions during the warm-up stage are generated. For $t = 1, \ldots, T$ let $\eta_t \sim \mathcal{N}(0, \sigma^2 I)$ be i.i.d. samples generated before the algorithm starts. Define $\tilde{u}_t = K_0 x_t + \eta_t$ and if at time t the algorithm chooses at random, i.e., during warm-up, then choose $u_t = \tilde{u}_t$. These virtual actions are a convenient technical tool as they do not directly depend on the action chosen by the algorithm.

Now, recall that

$$B_t = \arg\min_B \sum_{s=1}^{t-1} \|(x_{s+1} - A_\star x_s) - B u_s\|^2 + \lambda\|B\|_F^2,$$

and denote $\Delta_{B_t} = B_t - B_\star$, $V_t^u = \lambda I + \sum_{s=1}^{t-1} u_t u_t^T$. Further recalling that $\tau_i = \tau_0 4^i$ for $0 \leq i \leq n_T$ and $\tau_{n_T+1} = T + 1 \leq \tau_0 4^{n_T+1}$, we define the following events

$$\mathcal{E}_{B_{ols}} = \left\{\mathrm{Tr}\left(\Delta_{B_t}^T V_t^u \Delta_{B_t}\right) \le 4\sigma^2 d \log\left(4T^3 \frac{\det\left(V_t^u\right)}{\det\left(V_1^u\right)}\right) + 2\lambda k \vartheta^2, \text{ for all } t \ge 1\right\}, \tag{13}$$

$$\mathcal{E}_{B_x} = \left\{\sum_{t=\tau_{i-1}}^{\tau_i - 1} x_t x_t^T \succeq \frac{(\tau_i - \tau_{i-1})\sigma^2}{40} I, \text{ for all } 1 \le i \le n_T\right\}, \tag{14}$$

$$\mathcal{E}_{B_w} = \left\{\max_{1 \le t \le T} \|w_t\| \le \sigma\sqrt{15 d \log 4T}\right\} \tag{15}$$

$$\mathcal{E}_{B_u} = \left\{\sum_{t=1}^{\tau_i - 1} \tilde{u}_t \tilde{u}_t^T \succeq \frac{(\tau_i - 1)\sigma^2}{40} I, \text{ for all } 0 \le i \le n_T\right\}, \tag{16}$$

$$\mathcal{E}_{B_\eta} = \left\{\max_{1 \le t \le T} \|\eta_t\| \le \sigma\sqrt{15 d \log 4T}\right\}. \tag{17}$$

Then we have the following lemma.

Lemma 21. *Let $\mathcal{E}_B = \mathcal{E}_{B_{ols}} \cap \mathcal{E}_{B_x} \cap \mathcal{E}_{B_w} \cap \mathcal{E}_{B_u} \cap \mathcal{E}_{B_\eta}$, and suppose that $T \ge 600 d \log 48T$. Then we have that $\mathbb{P}\left(\mathcal{E}_B\right) \ge 1 - T^{-2}$.*

Proof. First, we describe the parameter estimation error in terms of Lemma 6. To that end, let $z_t = u_t$, $y_{t+1} = x_{t+1} - A_\star x_t$, $V_t^u = \lambda I + \sum_{s=1}^{t-1} u_t u_t^T$, and $\Delta_{B_t} = B_t - B_\star$ Indeed, we have $y_{t+1} = B_\star x_t + w_t$, $w_t \sim \mathcal{N}(0, \sigma^2 I)$, and $\|B_\star\|_F^2 \le k\|B_\star\|^2 \le k\vartheta^2$ and so taking Lemma 6 with $\delta = \frac{1}{4}T^{-2}$, recalling that $T \ge d$, and simplifying, we get that $\mathbb{P}\left(\mathcal{E}_{B_{ols}}\right) \ge 1 - \frac{1}{4}T^{-2}$.

Next, for $\mathcal{E}_{B_x}$, we apply Lemma 36 to the sequence x_t with the filtration $\mathcal{F}_t = \sigma(x_1, u_1, \ldots, x_t, u_t)$. Notice that given x_{t-1}, u_{t-1} we have $x_t \sim \mathcal{N}(A_\star x_{t-1} + B_\star u_{t-1}, \sigma^2 I)$ and hence we also get

$$\mathbb{E}\left[x_t x_t^T \mid \mathcal{F}_{t-1}\right] \succeq (A_\star x_{t-1} + B_\star u_{t-1})(A_\star x_{t-1} + B_\star u_{t-1})^T + \sigma^2 I \succeq \sigma^2 I.$$

Notice that our choice of τ_0 ensures the minimal sum size assumption. We thus apply Lemma 36 for each $1 \le i \le n_T$ with $\delta = \frac{1}{4}T{-}3$ and apply a union bound to get that $\mathbb{P}\left(\mathcal{E}_{B_x}\right) \ge 1 - \frac{1}{4}n_T T^{-3}$. Repeating the same process for $\tilde{u}_t$ we also have that $\mathbb{P}\left(\mathcal{E}_{B_u}\right) \ge 1 - \frac{1}{4}n_T T^{-3}$.

Finally, for $\mathcal{E}_{B_w}, \mathcal{E}_{B_\eta}$ we apply Lemma 34 with $\delta = \frac{1}{4}T^{-2}$ to get that $\mathbb{P}\left(\mathcal{E}_{B_w}\right) \ge 1 - \frac{1}{4}T^{-2}$ and $\mathbb{P}\left(\mathcal{E}_{B_\eta}\right) \ge 1 - \frac{1}{4}T^{-2}$.

The final result is obtained by taking a union bound over the events and noticing that $2n_T \le T$. ■

B.2. Proof of Lemma 11

The proof is implied by the last part of the following lemma.

Lemma 22 (Algorithm 2 good warm-up)**.** *On $\mathcal{E}_B$ we have that*

1. $\|x_t\| \le \sigma\kappa_0^3(1+\vartheta)\sqrt{60 d \log 4T}$, *for all* $1 \le t \le \tau_{n_s}$;
2. $\|u_t\|^2 \le \lambda$, *for all* $1 \le t < \tau_{n_s}$;
3. $V_{\tau_i}^u \succeq \frac{\tau_i \sigma^2}{40} I$, *for all* $0 \le i \le n_s$;
4. $\|\Delta_{B_{\tau_i}}\| \le \varepsilon_0 2^{-i}$, *for all* $0 \le i \le n_s$.

Proof. Recall the definition of η_t from Appendix B.1 and define $\tilde{w}_t = w_t + B\eta_t$. then for $t \le \tau_{n_s}$ we have that

$$x_t = A_\star x_{t-1} + B_\star \tilde{u}_{t-1} + w_{t-1} = A_\star x_{t-1} + B_\star K_0 x_{t-1} + \underbrace{w_{t-1} + B_\star \eta_{t-1}}_{\tilde{w}_{t-1}},$$

i.e., we can consider x_t as a sequence generated from running the controller K_0 on a linear system with noise sequence $\tilde{w}_t$. We can then apply Lemma 38 to get that

$$\|x_t\| \le \frac{\kappa_0}{\gamma_0} \max_{1 \le s \le T} \|\tilde{w}_s\| \qquad , \text{ for all } 1 \le t \le \tau_{n_s}.$$

Now, on $\mathcal{E}_B$ we have the noise bounds in Eq. (17) and Eq. (15) and so we have that

$$\max_{1\le s\le T}\|\tilde{w}_s\| \le \max_{1\le s\le T}\|w_s\| + \|B_\star\| \max_{1\le s\le T}\|\eta_s\| \le \sigma(1+\vartheta)\sqrt{15d\log 4T}.$$

Combining the above and recalling that $\gamma_0^{-1} = 2\kappa_0^2$ we conclude that

$$\|x_t\| \le \sigma\kappa_0^3(1+\vartheta)\sqrt{60d\log 4T} \qquad \text{, for all } 1 \le t \le \tau_{n_s},$$

proving the first claim of the lemma. Next, for $1 \le t < \tau_{n_s}$ we have that $u_t = \tilde{u}_t = K_0 x_t + \eta_t$ and so

$$\|u_t\| \le \kappa_0\|x_t\| + \|\eta_t\| \le \sigma\kappa_0^4(2+\vartheta)\sqrt{60d\log 4T} \le \sqrt{\lambda},$$

proving the second claim. Next, notice that for $0 \le i \le n_s$ we have that $V_{\tau_i}^u = \lambda I + \sum_{s=1}^{\tau_i - 1}\tilde{u}_s\tilde{u}_s^T$. Since $\mathcal{E}_B$ holds, we can use the warm-up actions lower bound in Eq. (16) to get that

$$V_{\tau_i}^u \succeq \left(\lambda + \frac{(\tau_i - 1)\sigma^2}{40}\right) I \succeq \frac{\tau_i\sigma^2}{40} I \qquad \text{, for all } 0 \le i \le n_s,$$

proving the third claim. For the final claim, we first use the lower bound on $V_{\tau_i}^u$ to get that

$$\|\Delta_{B_{\tau_i}}\|^2 \le \mathrm{Tr}\left(\Delta_{B_{\tau_i}}^T \Delta_{B_{\tau_i}}\right) \le \frac{40}{\tau_i\sigma^2}\mathrm{Tr}\left(\Delta_{B_{\tau_i}}^T V_{\tau_i}^u \Delta_{B_{\tau_i}}\right).$$

Next, we apply Eq. (13) to get that

$$\|\Delta_{B_{\tau_i}}\|^2 \le \frac{1}{\tau_i}\left(160d\log\left(4T^3\frac{\det\left(V_{\tau_i}^u\right)}{\det\left(V_1^u\right)}\right) + \frac{80\lambda k\vartheta^2}{\sigma^2}\right).$$

Now, using the second claim of the lemma, we can use Lemma 37 to get that $\log\frac{\det\left(V_{\tau_i}^u\right)}{\det\left(V_1^u\right)} \le k\log T$, and applying it to the above and simplifying we get that

$$\begin{aligned}
\|\Delta_{B_{\tau_i}}\|^2 &\le \frac{1}{\tau_i}\left(160dk\log\left(4T^4\right) + \frac{80\lambda k\vartheta^2}{\sigma^2}\right) \\
&\le \frac{1}{\tau_i}\left(640dk\log(4T) + \frac{80\lambda k\vartheta^2}{\sigma^2}\right) \\
&\le \frac{1}{\tau_i}\frac{80\lambda k\left(1+\vartheta^2\right)}{\sigma^2} \le \frac{\varepsilon_0^2\tau_0}{\tau_i} = \varepsilon_0^2 4^{-i},
\end{aligned}$$

thus concluding the proof. ■

B.3. Proof of Lemma 12

The proof is broken into the following lemmas. The first two claims are concluded by putting together Lemmas 22 and 25 and the third is given by Lemma 26.

Before proceeding, we need the two following lemmas.

Lemma 23 (Algorithm 2 warm-up length)**.** *On $\mathcal{E}_B$ we have that* $\max\left\{0, \log_2\frac{\mu_0}{\mu_\star}\right\} \le n_s \le 2 + \max\left\{0, \log_2\frac{\mu_0}{\mu_\star}\right\}$.

Proof. First recall that by Lemma 22, we have that $\|\Delta_{B_{\tau_i}}\| \le \varepsilon_0 2^{-i}$ for all $0 \le i \le n_s$. Now, our choice of μ_0 implies that $\varepsilon_0 = \frac{\mu_0}{4\kappa C_0}$ and further recalling that $\mu_i = \mu_0 2^{-i}$, we apply Lemma 42 to get that

$$K_{\tau_i}K_{\tau_i}^T \succeq K_\star K_\star^T - \frac{\mu_i}{2}I \qquad \text{, for all } 0 \le i \le n_s \tag{18}$$

$$K_\star K_\star^T \succeq K_{\tau_i}K_{\tau_i}^T - \frac{\mu_i}{2}I \qquad \text{, for all } 0 \le i \le n_s. \tag{19}$$

Now, suppose in contradiction that $n_s > 0$ and $\mu_{n_s} < \frac{\mu_\star}{4}$. This means that $\mu_{n_s-1} < \frac{\mu_\star}{2}$ and so we can apply Eq. (18) to get that

$$K_{\tau_{n_s}-1}K_{\tau_{n_s}-1}^T \succeq \left(\mu_\star - \frac{\mu_{n_s-1}}{2}\right)I \succeq \left(2\mu_{n_s-1} - \frac{\mu_{n_s-1}}{2}\right)I = \frac{3}{2}\mu_{n_s-1}I,$$

which contradicts the fact that n_s is the first time the warm-up break condition is satisfied. We conclude that either $n_s = 0$ or $\mu_{n_s} \geq \frac{\mu_\star}{4}$. Plugging $\mu_{n_s} = \mu_0 2^{-n_s}$ the latter condition implies $n_s \leq 2 + \log_2 \frac{\mu_0}{\mu_\star}$ thus giving the lemma's upper bound.

Now for the lower bound, suppose in contradiction that $\mu_{n_s} > \mu_\star$ then by Eq. (19) we get that

$$K_\star K_\star^T \succeq \left(\frac{3}{2}\mu_{n_s} - \frac{\mu_{n_s}}{2}\right)I \succ \mu_\star I,$$

which contradicts the fact that $\mu_\star$ is the tight lower bound on the eigenvalues of $K_\star K_\star^T$. We conclude that $\mu_{n_s} \leq \mu_\star$ which in turn implies the desired lower bound. ∎

Lemma 24 (Algorithm 2 conditional control)**.** *Suppose $\mathcal{E}_B$ holds and fix some i such that $n_s \leq i \leq n_T$. If $\|u_t\|^2 \leq \lambda$ for all $1 \leq t \leq \tau_i - 1$, then K_{τ_i} is (κ, γ)–strongly stable and $K_{\tau_i}K_{\tau_i}^T \succeq \frac{\mu_\star}{2}I$.*

Proof. If $\|\Delta_{B_{\tau_i}}\| \leq \min\left\{\varepsilon_0, \frac{\mu_\star}{4\kappa C_0}\right\}$ then Lemma 42 immediately implies the desired result. We prove this estimation error bound thus concluding the proof.

To that end, notice that for $t \geq s$ we have $V_t^u \succeq V_s^u$. Using the lower bound on $V_{\tau_{n_s}}^u$ in Lemma 22 we get that

$$\mathrm{Tr}\left(\Delta_{B_{\tau_i}}^T V_{\tau_i}^u \Delta_{B_{\tau_i}}\right) \geq \mathrm{Tr}\left(\Delta_{B_{\tau_i}}^T V_{\tau_{n_s}}^u \Delta_{B_{\tau_i}}\right) \geq \|\Delta_{B_{\tau_i}}\|^2 \frac{\tau_{n_s}\sigma^2}{40},$$

and by changing sides and applying the parameter estimation bound in Eq. (13) we get that

$$\|\Delta_{B_{\tau_i}}\|^2 \leq \frac{1}{\tau_{n_s}}\left(160d\log\left(4T^3\frac{\det\left(V_{\tau_i}^u\right)}{\det\left(V_1^u\right)}\right) + \frac{80\lambda k\vartheta^2}{\sigma^2}\right) \tag{20}$$

Now, using the assumption on u_t, we can apply Lemma 37 to get that $\log\frac{\det V_{\tau_i}^u}{\det V_1^u} \leq k\log T$. Plugging this back into Eq. (20) and simplifying we get that

$$\|\Delta_{B_{\tau_i}}\| \leq \varepsilon_0 2^{-n_s},$$

and plugging in the lower bound on n_s from Lemma 23 gives the desired bound on the estimation error thus concluding the proof. ∎

Lemma 25 (Algorithm 2 bounded operation)**.** *On $\mathcal{E}_B$ we have that*

1. $\|x_t\|^2 \leq x_b$, *for all* $\tau_{n_s} \leq t \leq T$;
2. K_{τ_i} *is* (κ, γ)*–strongly stable, for all* $n_s \leq i \leq n_T$;
3. $K_{\tau_i}K_{\tau_i}^T \succeq \frac{1}{2}\mu_\star I$, *for all* $n_s \leq i \leq n_T$.

Proof. First, recall the bounds on x_t, u_t from Lemma 22, i.e.,

$$\|x_{\tau_{n_s}}\| \leq \sigma\kappa_0^3(1+\vartheta)\sqrt{60d\log 4T} \leq \sqrt{x_b},$$
$$\|u_t\|^2 \leq \lambda,\ \text{for all } 1 \leq t < \tau_{n_s}.$$

We prove by induction on n where $n_s \leq n \leq n_T$ that the claims of the lemma hold up to time τ_n and phase n respectively.

For the base case, $n = n_s$ the bounds above satisfy Lemma 24 and so we conclude that K_{τ_0} is strongly stable and that $K_{\tau_0}K_{\tau_0}^T \succeq \frac{1}{2}\mu_\star I$ thus satisfying the induction base.

Next, assume the induction hypothesis holds for $n-1$ and we show for n. By the induction hypothesis, the algorithm does not abort up to (including) time $\tau_{n-1}-1$. Moreover, it means that for all $n_s \le i \le n-1$ the controllers K_{τ_i} are (κ,γ)–strongly stable and so we can use Lemma 39 to get that

$$\|x_t\| \le 3\kappa \max\left\{\frac{\|x_{\tau_{n_s}}\|}{2}, \frac{\kappa}{\gamma}\max_{1\le s\le T}\|w_t\|\right\}, \qquad \text{for all } \tau_{n_s} \le t \le \tau_n,$$

and plugging in that $\gamma^{-1} = 2\kappa^2$, the bound for $\|x_{\tau_{n_s}}\|$ and the bound for the noise in Eq. (15) we get that

$$\|x_t\| \le \kappa\sigma \max\left\{\kappa_0^3(1+\vartheta), 2\kappa^3\right\}\sqrt{135d\log 4T} \le \sqrt{x_b}, \qquad \text{for all } \tau_{n_s} \le t \le \tau_n,$$

as desired for x_t. Notice that this ensures that the algorithm does not abort up to time $\tau_n - 1$. So, for $\tau_{n_s} \le t \le \tau_n - 1$ we have that $\|u_t\| = \|K_t x_t\| \le \|K_t\|\|x_t\| \le \kappa\sqrt{x_b} = \sqrt{\lambda}$, and thus Lemma 24 establishes the desired strong-stability and non-degeneracy of K_{τ_n}, finishing the induction.

Finally, using the strong stability of all controllers we apply Lemma 39 a final time to obtain the bound on x_t for all $\tau_{n_s} \le t \le T$. ■

Lemma 26 (Algorithm 2 parameter estimation)**.** *On $\mathcal{E}_B$ we have that $\|\Delta_{B_{\tau_i}}\| \le \varepsilon_0 \min\left\{2^{-n_s}, 2\mu_\star^{-1/2}2^{-i}\right\},\ \forall\, n_s < i \le n_T$.*

Proof. Recall that by Lemma 25, the algorithm does not abort on $\mathcal{E}_B$ and so for $\tau_i \le t \le \tau_{i+1}-1$ we have that $K_t = K_{\tau_i}$. This means we can decompose $V^u_{\tau_i}$ as

$$V^u_{\tau_i} = V^u_{\tau_{n_s}} + \sum_{t=\tau_{n_s}}^{\tau_i - 1} u_t u_t^T = V^u_{\tau_{n_s}} + \sum_{j=n_s}^{i-1}\sum_{t=\tau_j}^{\tau_{j+1}-1} u_t u_t^T = V^u_{\tau_{n_s}} + \sum_{j=n_s}^{i-1} K_{\tau_j}\left(\sum_{t=\tau_j}^{\tau_{j+1}-1} x_t x_t^T\right)K_{\tau_j}^T.$$

Next, we lower bound $V^u_{\tau_{n_s}}$ using Lemma 22 and the states using Eq. (14) and get that

$$V^u_{\tau_i} \succeq \frac{\tau_{n_s}\sigma^2}{40} I + \sum_{j=n_s}^{i-1}\left(\frac{\left(\tau_{j+1}-\tau_j\right)\sigma^2}{40}\right) K_{\tau_{n_s}2^{j-1}} K_{\tau_{n_s}2^{j-1}}^T,$$

and recalling that $K_{\tau_j}K_{\tau_j}^T \succeq \frac{\mu_\star}{2} I$ (see Lemma 25) we get that, assuming $i > n_s$,

$$V^u_{\tau_i} \succeq \frac{\sigma^2}{40}\left(\tau_{n_s} + \frac{\mu_\star}{2}\sum_{j=n_s}^{i-1}(\tau_{j+1}-\tau_j)\right) I = \frac{\sigma^2}{40}\left(\tau_{n_s} + \frac{\mu_\star}{2}\left(\tau_i - \tau_{n_s}\right)\right) I \succeq \frac{\sigma^2}{40}\max\left\{\tau_{n_s}, \frac{\mu_\star}{4}\tau_i\right\} I.$$

Now, apply this together with the parameter estimation bound in Eq. (13) to get that

$$\begin{aligned}
\|\Delta_{B_{\tau_i}}\|^2 &\le \mathrm{Tr}\left(\Delta_{B_{\tau_i}}^T \Delta_{B_{\tau_i}}\right) \\
&\le \frac{40}{\sigma^2 \max\left\{\tau_{n_s}, \frac{\mu_\star}{4}\tau_i\right\}} \mathrm{Tr}\left(\Delta_{B_{\tau_i}}^T V^u_{\tau_i} \Delta_{B_{\tau_i}}\right) \\
&\le \frac{1}{\max\left\{\tau_{n_s}, \frac{\mu_\star}{4}\tau_i\right\}}\left(160 d \log\left(4T^3 \frac{\det\left(V^u_{\tau_i}\right)}{\det\left(V^u_1\right)}\right) + \frac{80\lambda k \vartheta^2}{\sigma^2}\right).
\end{aligned}$$

Finally, from Lemma 22 we have that $\|u_t\|^2 \le \lambda$ for $1 \le t < \tau_{n_s}$ and from Lemma 25 we have that $\|x_t\|^2 \le x_b$ for $\tau_{n_s} \le t \le T$ and so $\|u_t\|^2 = \|K_t x_t\|^2 \le \kappa^2 x_b = \lambda$. Combining both claims, we apply Lemma 37 to get that $\log\frac{\det V^u_{\tau_i}}{\det V^u_1} \le k\log T$ and plugging this into the above equation we get

$$\|\Delta_{B_{\tau_i}}\|^2 \le \frac{1}{\max\left\{\tau_{n_s}, \frac{\mu_\star}{4}\tau_i\right\}}\left(640 dk\log(4T) + \frac{80\lambda\vartheta^2}{\sigma^2}\right) \le \frac{\tau_0\varepsilon_0^2}{\max\left\{\tau_{n_s}, \frac{\mu_\star}{4}\tau_i\right\}} = \varepsilon_0^2 \min\left\{4^{-n_s}, \frac{4}{\mu_\star}4^{-i}\right\},$$

where the second transition follows from our choice of τ_0. ■

B.4. Proof of Theorem 2

As in Algorithm 1, denote $J_i = \sum_{t=\tau_i}^{\tau_{i+1}-1} x_t^T \left(Q + K_{\tau_i}^T R K_{\tau_i}\right) x_t$. Recalling that warm-up lasts until phase n_s, we have the following decomposition of the regret:

$$\mathbb{E}[R_T] = R_1 + R_2 + R_3 - TJ_\star,$$

where

$$R_1 = \mathbb{E}\left[\sum_{i=n_s}^{n_T} \mathbb{1}\{\mathcal{E}_B\} J_i\right], \qquad R_2 = \mathbb{E}\left[\mathbb{1}\{\mathcal{E}_B^c\} \sum_{t=\tau_{n_s}}^{T} c_t\right], \qquad R_3 = \mathbb{E}\left[\sum_{t=1}^{\tau_{n_s}-1} c_t\right],$$

are the costs due to success, failure, and warm-up respectively. The following lemmas bound each of R_1, R_2, R_3 thus concluding the proof. The proofs for R_1, R_2 remain nearly the same but are provided for completeness. The proof of R_3 contains a few technical challenges, introduced by the randomness of the warm-up period duration.

Lemma 27. $R_1 - TJ_\star \le n_T\left(6C_0\varepsilon_0^2 \max\left\{1, 4\mu_\star^{-1}\right\}\tau_0 + 8\alpha_1\kappa^6 x_b\right)$.

Lemma 28. $R_2 \le \left(J(K_0) + 2\alpha_1\kappa^2 x_b\right)T^{-1} + 4\alpha_1\kappa_0^6\left(1 + 8\vartheta^2\right)\left(\kappa^2 + \kappa_0^2\right)x_b T^{-2}$.

Lemma 29. $R_3 \le (1 + \vartheta^2)\left(65J(K_0)\max\left\{1, \frac{\mu_0^2}{\mu_\star^2}\right\}\tau_0 + 80\alpha_1 d\sigma^2\kappa_0^{14}\log^2 3T\right)$.

B.4.1. Proof of Lemma 27

Proof. We begin by bounding $\mathbb{E}\left[\mathbb{1}\{\mathcal{E}_B\}J_i \mid n_s\right]$ for $n_s \le i \le n_T$. This follows exactly as in Lemma 8 but with some changes to the events E_i, and thus is repeated here. For $n_s \le i \le n_T$ define the events $S_i = \left\{\|x_{\tau_i}\|^2 \le x_b\right\}$ and

$$E_{n_s} = \left\{\|\Delta_{B_{\tau_{n_s}}}\| \le \varepsilon_0 2^{-n_s}\right\}, \qquad E_i = \left\{\|\Delta_{B_{\tau_i}}\| \le \varepsilon_0 \min\left\{2^{-n_s}, 2\mu_\star^{-1}2^{-i}\right\}\right\},\ \forall\, n_s < i \le n_T.$$

By Lemma 12, we have that $\mathcal{E}_B \subseteq E_i \cap S_i$. Now, define $\tilde{x}_{\tau_i} = x_{\tau_i}$ and for $\tau_i < t \le \tau_{i+1} - 1$ define

$$\tilde{x}_t = \left(A_\star + B_\star K_{\tau_i}\right)\tilde{x}_{t-1} + w_t.$$

Since on $\mathcal{E}_B$ the algorithm does not abort, we have that

$$\mathbb{1}\{\mathcal{E}_B\}J_i = \mathbb{1}\{\mathcal{E}_B\}\sum_{t=\tau_i}^{\tau_{i+1}-1}\tilde{x}_t^T\left(Q + K_{\tau_i}^T R K_{\tau_i}\right)\tilde{x}_t \le \mathbb{1}\{E_i \cap S_i\}\sum_{t=\tau_i}^{\tau_{i+1}-1}\tilde{x}_t^T\left(Q + K_{\tau_i}^T R K_{\tau_i}\right)\tilde{x}_t.$$

Noticing that E_i, S_i, and K_{τ_i} are completely determined by x_{τ_i}, B_{τ_i} we use total expectation to get that

$$\mathbb{E}\left[\mathbb{1}\{\mathcal{E}_B\}J_i \mid n_s\right] \le \mathbb{E}\left[\mathbb{1}\{E_i \cap S_i\}\mathbb{E}\left[\sum_{t=\tau_i}^{\tau_{i+1}-1}\tilde{x}_t^T\left(Q + K_{\tau_i}^T R K_{\tau_i}\right)\tilde{x}_t \,\middle|\, x_{\tau_i}, B_{\tau_i}\right] \,\middle|\, n_s\right],$$

where in the inner expectation we removed the conditioning on n_s since the $\tilde{x}_t$ are conditionally independent of n_s given x_{τ_i}. Now, by Lemma 42, E_i implies that K_{τ_i} is (κ, γ)–strongly stable and so we can use Lemma 40 to get that

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}\{\mathcal{E}_B\}J_i\right] &\le (\tau_{i+1} - \tau_i)\mathbb{E}\left[\mathbb{1}\{E_i\}J\left(K_{\tau_i}\right) \mid n_s\right] + \frac{2\alpha_1\kappa^4}{\gamma}\mathbb{E}\left[\mathbb{1}\{S_i\}\|x_{\tau_i}\|^2 \mid n_s\right] \\ &\le (\tau_{i+1} - \tau_i)\mathbb{E}\left[\mathbb{1}\{E_i\}J\left(K_{\tau_i}\right) \mid n_s\right] + 4\alpha_1\kappa^6 x_b, \end{aligned} \tag{21}$$

where the second transition also used that $\gamma^{-1} = 2\kappa^2$.

Now, by Lemma 4, on E_{n_s} we have that $J\left(K_{\tau_{n_s}}\right) \le J_\star + C_0\varepsilon_0^2 4^{-n_s}$ and on E_i where $n_s < i \le n_T$, we have that $J\left(K_{\tau_i}\right) \le J_\star + C_0\varepsilon_0^2\min\left\{4^{-n_s}, 4\mu_\star^{-1}4^{-i}\right\}$. Combining both cases we conclude that

$$\mathbb{1}\{E_i\}J\left(K_{\tau_i}\right) \le J_\star + C_0\varepsilon_0^2\max\left\{1, 4\mu_\star^{-1}\right\}4^{-i} \qquad ,\forall\, n_s \le i \le n_T,$$

and plugging this back into Eq. (21) and recalling that $\tau_{i+1} - \tau_i \leq 3\tau_i = 3\tau_0 4^i$ we have that

$$\mathbb{E}\left[\mathbb{1}\{\mathcal{E}_B\}J_i \mid n_s\right] \leq (\tau_{i+1} - \tau_i)J_\star + 3C_0\varepsilon_0^2 \max\left\{1, 4\mu_\star^{-1}\right\}\tau_0 + 4\alpha_1\kappa^6 x_b.$$

Finally, we sum over i to conclude that

$$\begin{aligned}
R_1 = \mathbb{E}\left[\sum_{i=n_s}^{n_T} \mathbb{E}\left[\mathbb{1}\{\mathcal{E}_B\}J_i \mid n_s\right]\right] &\leq \mathbb{E}\left[\sum_{i=n_s}^{n_T} (\tau_{i+1} - \tau_i)J_\star + 3C_0\varepsilon_0^2 \max\left\{1, 4\mu_\star^{-1}\right\}\tau_0 + 4\alpha_1\kappa^6 x_b\right] \\
&\leq \mathbb{E}\left[\left(\tau_{n_T+1} - \tau_{n_s}\right)J_\star + (n_T + 1 - n_s)\left(3C_0\varepsilon_0^2 \max\left\{1, 4\mu_\star^{-1}\right\}\tau_0 + 4\alpha_1\kappa^6 x_b\right)\right] \\
&\leq TJ_\star + n_T\left(6C_0\varepsilon_0^2 \max\left\{1, 4\mu_\star^{-1}\right\}\tau_0 + 8\alpha_1\kappa^6 x_b\right),
\end{aligned}$$

thus concluding the proof. ∎

B.4.2. PROOF OF LEMMA 28

The proof is identical to that of Lemma 9 where the initial warm-up duration τ_0 is replaced with τ_{n_s} and the uses of Lemmas 17 and 20 are replaced with Lemmas 21 and 30 respectively. We thus conclude by proving Lemma 30. To that end, recall that τ_{abort} is the time when the algorithm decides to abort, formally,

$$\tau_{\text{abort}} = \min\{t \geq \tau_{n_s} \mid \|x_t\|^2 > x_b \text{ or } \|K_t\| > \kappa\},$$

where we treat $\min \emptyset = T + 1$.

Lemma 30 (Expected abort state)**.** *Suppose that $\mathbb{P}(\tau_{abort} \leq T) \leq T^{-2}$. Then we have that*

$$\mathbb{E}\left[\|x_{\tau_{abort}}\|^2 \mathbb{1}\{\tau_{abort} < T\}\right] \leq \left(1 + 8\vartheta^2\right)\left(\kappa^2 + \kappa_0^2\right)x_b T^{-2}.$$

Proof. First, by the lemmas assumption, we can apply Lemma 35 to get that

$$\mathbb{E}\left[\mathbb{1}\{\tau_{\text{abort}} \leq T\} \max_{1\leq t\leq T}\|w_t\|^2\right] \leq 5d\sigma^2 T^{-2}\log 3T, \tag{22}$$

$$\mathbb{E}\left[\mathbb{1}\{\tau_{\text{abort}} \leq T\} \max_{1\leq t\leq T}\|B_\star\eta_t + w_t\|^2\right] \leq 5d\sigma^2\left(1 + \vartheta^2\right)T^{-2}\log 3T. \tag{23}$$

Now, notice that $\|A_\star + B_\star K\| \leq 2\vartheta\|K\|$ and split into two cases. First, if $\tau_{\text{abort}} > \tau_{n_s}$ then by definition of τ_{abort} we have that

$$\|x_{\tau_{\text{abort}}}\| = \left\|\left(A_\star + B_\star K_{\tau_{\text{abort}}-1}\right)x_{\tau_{\text{abort}}-1} + w_{\tau_{\text{abort}}-1}\right\| \leq 2\vartheta\kappa\sqrt{x_b} + \max_{1\leq s\leq T}\|w_t\|,$$

and taking expectation and applying Eq. (22) we get that

$$\mathbb{E}\left[\mathbb{1}\{\tau_{n_s} < \tau_{\text{abort}} \leq T\}\|x_{\tau_{\text{abort}}}\|^2\right] \leq 8\vartheta^2\kappa^2 x_b T^{-2} + 5d\sigma^2 T^{-2}\log 3T \leq \left(1 + 8\vartheta^2\right)\kappa^2 x_b T^{-2}.$$

On the other hand if $\tau_{\text{abort}} = \tau_{n_s}$ then $u_{\tau_{\text{abort}}-1} = K_0 x_{\tau_{n_s}-1} + \eta_{\tau_{n_s}-1}$ and so we have that

$$\begin{aligned}
\|x_{\tau_{\text{abort}}}\| &= \left\|(A_\star + B_\star K_0)x_{\tau_{n_s}-1} + \left(B_\star\eta_{\tau_{n_s}-1} + w_{\tau_{n_s}-1}\right)\right\| \\
&\leq 2\vartheta\kappa_0\|x_{\tau_{n_s}-1}\| + \max_{1\leq t\leq T}\|B_\star\eta_t + w_t\| \\
&\leq (4\vartheta + 1)\kappa_0^4 \max_{1\leq t\leq T}\|B_\star\eta_t + w_t\|,
\end{aligned}$$

where the last transition used Lemma 38 and $\gamma_0^{-1} = 2\kappa_0^2$. Taking expectation and applying Eq. (23) we get that

$$\mathbb{E}\left[\mathbb{1}\{\tau_{\text{abort}} = \tau_{n_s}\}\|x_{\tau_{\text{abort}}}\|^2\right] \leq 80(1 + \vartheta)^2(1 + \vartheta^2)\kappa_0^8 d\sigma^2 T^{-2}\log 3T \leq \left(1 + \vartheta^2\right)\kappa_0^2 x_b T^{-2},$$

and combining both cases yields the final bound. ∎

B.4.3. Proof of Lemma 29

Proof. We begin by decomposing R_3. Notice that $n_s \le n_T + 1$ and so we have that

$$R_3 = \mathbb{E}\left[\sum_{t=1}^{\tau_0-1} c_t\right] + \mathbb{E}\left[\sum_{i=0}^{n_s-1}\sum_{t=\tau_i}^{\tau_{i+1}-1} c_t\right] = \mathbb{E}\left[\sum_{t=1}^{\tau_0-1} c_t\right] + \sum_{i=0}^{n_T}\mathbb{E}\left[\mathbb{1}\{n_s > i\}\sum_{t=\tau_i}^{\tau_{i+1}-1} c_t\right].$$

Now, define $J(K, W)$ to be the infinite horizon cost of playing controller K on the LQ system $(A_\star, B_\star)$ whose system noise has covariance $W \in \mathbb{R}^{d\times d}$. In terms of our notation so far, this means that $J(K) = J\big(K, \sigma^2 I\big)$. It is well known that $J(K, W) = \mathrm{Tr}(PW)$ where P is a positive definite solution to

$$P = Q + K^T RK + (A_\star + B_\star K)^T P(A_\star + B_\star K),$$

and thus does not depend on W.

Now, for $1 \le t < \tau_{n_s}$ we have that $x_{t+1} = (A_\star + B_\star K_0)x_t + (B_\star\eta_t + w_t)$, i.e., this is equivalent to an LQ system $(A_\star, B_\star)$ with noise covariance $\sigma^2\big(I + B_\star B_\star^T\big) \preceq \big(1+\vartheta^2\big)\sigma^2 I$ and controller K_0 and so we have that

$$J\big(K_0, \sigma^2\big(I + B_\star B_\star^T\big)\big) = \mathrm{Tr}\big(\sigma^2\big(I + B_\star B_\star^T\big)P\big) \le \big(1+\vartheta^2\big)\mathrm{Tr}\big(\sigma^2 P\big) = \big(1+\vartheta^2\big)J\big(K_0, \sigma^2\big) = \big(1+\vartheta^2\big)J(K_0).$$

With the above in mind, we bound the first term in the decomposition of R_3 using Lemma 40. We get that

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^{\tau_0-1} c_t\right] &\le \tau_0 J\big(K_0, \sigma^2\big(I + B_\star B_\star^T\big)\big) + \frac{2\alpha_1\kappa_0^4}{\gamma_0}\|x_1\|^2 \\ &\le \big(1+\vartheta^2\big)J(K_0)\tau_0.\end{aligned} \tag{24}$$

Next, recall that $\gamma_0^{-1} = 2\kappa_0^2$, denote the filtration of the history, $\mathcal{F}_t = \sigma(x_1, u_1, w_1, \ldots, x_t, u_t, w_t)$ and similarly apply Lemma 40 to get that

$$\mathbb{E}\left[\sum_{t=\tau_i}^{\tau_{i+1}-1} c_t \,\middle|\, \mathcal{F}_{\tau_i-1}\right] \le (1+\vartheta^2)J(K_0)(\tau_{i+1}-\tau_i) + 4\alpha_1\kappa_0^6\|x_{\tau_i}\|^2.$$

Now, using Lemmas 35 and 38 we get that

$$\mathbb{E}\Big[\mathbb{1}\{n_s > i\}\|x_{\tau_i}\|^2\Big] \le \mathbb{E}\left[\frac{\kappa_0^2}{\gamma_0^2}\max_{1\le t\le T}\|w_t + B_\star\eta_t\|^2\right] \le 20d(1+\vartheta^2)\sigma^2\kappa_0^8\log 3T.$$

Combining the last two inequalities and noticing that $\mathbb{1}\{n_s > i\}$ is $\mathcal{F}_{\tau_i-1}$ measurable we further have that

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}\{n_s > i\}\sum_{t=\tau_i}^{\tau_{i+1}-1} c_t\right] &= \mathbb{E}\left[\mathbb{1}\{n_s > i\}\mathbb{E}\left[\sum_{t=\tau_i}^{\tau_{i+1}-1} c_t \mid \mathcal{F}_{\tau_i-1}\right]\right] \\ &\le \mathbb{E}\Big[\mathbb{1}\{n_s > i\}\Big((1+\vartheta^2)J(K_0)(\tau_{i+1}-\tau_i) + 4\alpha_1\kappa_0^6\|x_{\tau_i}\|^2\Big)\Big] \\ &\le (1+\vartheta^2)\big(\mathbb{P}(n_s > i)J(K_0)(\tau_{i+1}-\tau_i) + 80\alpha_1 d\sigma^2\kappa_0^{14}\log 3T\big).\end{aligned} \tag{25}$$

Now, from Lemma 23 we know that $\mathbb{P}\Big(n_s > 2 + \max\Big\{0, \log_2\frac{\mu_0}{\mu_\star}\Big\}\Big) \le \mathbb{P}\big(\mathcal{E}_B^c\big) \le T^{-2}$, and recalling that $\tau_i = \tau_0 4^i$ we get that

$$\begin{aligned}\tau_0 + \sum_{i=0}^{n_T}(\tau_{i+1}-\tau_i)\mathbb{P}(n_s > i) &\le \tau_0 + \sum_{i=0}^{\left\lfloor 2+\max\left\{0,\log_2\frac{\mu_0}{\mu_\star}\right\}\right\rfloor}(\tau_{i+1}-\tau_i) + \sum_{i=0}^{n_T}(\tau_{i+1}-\tau_i)T^{-2} \\ &= \tau_0 4^{\left\lfloor 3+\max\left\{0,\log_2\frac{\mu_0}{\mu_\star}\right\}\right\rfloor} + \big(\tau_{n_T+1}-\tau_0\big)T^{-2} \\ &\le 64\tau_0\max\left\{1, \frac{\mu_0^2}{\mu_\star^2}\right\} + 4T^{-1}.\end{aligned} \tag{26}$$

Finally, combining Eqs. (24) to (26) we get that

$$
\begin{aligned}
R_3 &\le (1+\vartheta^2)\left(J(K_0)\left(\tau_0 + \sum_{i=0}^{n_T}(\tau_{i+1}-\tau_i)\mathbb{P}(n_s > i)\right) + 80\alpha_1 d\sigma^2\kappa_0^{14}(n_T+1)\log 3T\right) \\
&\le (1+\vartheta^2)\left(64J(K_0)\max\left\{1, \frac{\mu_0^2}{\mu_\star^2}\right\}\tau_0 + 4J(K_0)T^{-1} + 80\alpha_1 d\sigma^2\kappa_0^{14}\log^2 3T\right) \\
&\le (1+\vartheta^2)\left(65J(K_0)\max\left\{1, \frac{\mu_0^2}{\mu_\star^2}\right\}\tau_0 + 80\alpha_1 d\sigma^2\kappa_0^{14}\log^2 3T\right),
\end{aligned}
$$

where the second transition also used $n_T + 1 \le \log 3T$. ■

C. Lower Bound Proofs

The next lemma requires the following well known results in LQRs (see, e.g., Bertsekas, 1995). Consider the Q-function of the system with respect to $k_\star$, that in the one-dimensional case takes the form $F(x,u) = x^2 + u^2 + (ax+bu)^2 p_\star$. Using the form of $k_\star$ given in Eq. (1), and by simple algebra we obtain

$$
F(x_t, u_t) - F(x_t, k_\star x_t) = (1 + b^2 p_\star)(u_t - k_\star x_t)^2. \tag{27}
$$

Further, we have $F(x_t, k_\star x_t) = x_t^2 p_\star$ as both sides are equal to the value of the optimal policy $k_\star$ starting from state x_t. Finally, also recall that $J(k_\star) = \sigma^2 p_\star$. The following explains Eq. (27):

$$
\begin{aligned}
F(x_t,u_t) &= x_t^2 + ((u_t - k_\star x_t) + k_\star x_t)^2 + ((a+bk_\star)x_t + b(u_t - k_\star x_t))^2 p_\star \\
&= F(x_t, k_\star x_t) + (u_t - k_\star x_t)^2 + 2(u_t - k_\star x_t)k_\star x_t + b^2 p_\star (u_t - k_\star x_t)^2 + 2bp_\star(u_t - k_\star x_t)(a+bk_\star)x_t \\
&= F(x_t, k_\star x_t) + (1+b^2 p_\star)(u_t - k_\star x_t)^2 + 2x_t(u_t - k_\star x_t)(k_\star + bp_\star(a+bk_\star)) \\
&= F(x_t, k_\star x_t) + (1+b^2 p_\star)(u_t - k_\star x_t)^2 + 2x_t(u_t - k_\star x_t)(k_\star(1+b^2 p_\star) + bp_\star a) \\
&= F(x_t, k_\star x_t) + (1+b^2 p_\star)(u_t - k_\star x_t)^2,
\end{aligned}
$$

where the last transition used $k_\star(1+b^2 p_\star) = -bp_\star a$ (see Eq. (1)).

Lemma 31. *The expected regret can be written as*

$$
\mathbb{E}[R_T] = \mathbb{E}\left[\sum_{t=1}^{T}(1+b^2 p_\star)(u_t - k_\star x_t)^2\right] - \mathbb{E}\left[x_{T+1}^2 p_\star\right].
$$

Proof. Using the expressions for the Q-function of the system with respect to $k_\star$, we have that

$$
\begin{aligned}
R_T &= \sum_{t=1}^{T}\mathbb{E}\left[x_t^2 + u_t^2 - J(k_\star)\right] \\
&= \sum_{t=1}^{T}\mathbb{E}\left[F(x_t,u_t) - \left((ax_t + bu_t)^2 + w_t^2\right)p_\star\right] && \text{(since } J(k_\star) = \mathbb{E}[w_t^2 p_\star]\text{)} \\
&= \sum_{t=1}^{T}\mathbb{E}\left[F(x_t,u_t) - x_{t+1}^2 p_\star\right] \\
&= \sum_{t=1}^{T}\mathbb{E}[F(x_t,u_t) - F(x_t,k_\star x_t)] + \sum_{t=1}^{T}\mathbb{E}\left[x_t^2 p_\star - x_{t+1}^2 p_\star\right] && \text{(since } F(x_t,k_\star x_t) = x_t^2 p_\star\text{)} \\
&= \mathbb{E}\left[\sum_{t=1}^{T}(1+b^2 p_\star)(u_t - k_\star x_t)^2\right] + \mathbb{E}\left[x_1^2 p_\star\right] - \mathbb{E}\left[x_{T+1}^2 p_\star\right]. && \text{(using Eq. (27))}
\end{aligned}
$$

The lemma now follows from our assumption that $x_1 = 0$. ■

Lemma 32. *We have* $\mathbb{E}[x_{T+1}^2] \le \frac{5}{2}\left(b^2 \sum_{t=1}^T \mathbb{E}[(u_t - k_\star x_t)^2] + \sigma^2\right)$.

Proof. Denote $m = a + bk_\star$ and $v_t = u_t - k_\star x_t$ for all $t \ge 1$. Then, $x_{t+1} = ax_t + b(u_t - k_\star x_t + k_\star x_t) + w_t = mx_t + bv_t + w_t$, and by unfolding the recursion and using $x_1 = 0$ we obtain

$$x_{T+1} = \sum_{t=1}^T m^{T-t} bv_t + \sum_{t=1}^T m^{T-t} w_t,$$

hence

$$\mathbb{E}[x_{T+1}^2] \le 2b^2 \mathbb{E}\left(\sum_{t=1}^T m^{T-t} v_t\right)^2 + 2\mathbb{E}\left(\sum_{t=1}^T m^{T-t} w_t\right)^2,$$

Now, observe that

$$|m| = |a + bk_\star| = \left|a - b \cdot \frac{abp_\star}{1 + b^2 p_\star}\right| = \left|\frac{a}{1 + b^2 p_\star}\right| \le |a| \le \frac{1}{\sqrt{5}}.$$

Using this bound and the Cauchy-Schwartz inequality, we have

$$\mathbb{E}\left(\sum_{t=1}^T m^{T-t} v_t\right)^2 \le \sum_{t=1}^T m^{2(T-t)} \cdot \mathbb{E}\left[\sum_{t=1}^T v_t^2\right] \le \frac{1}{1 - m^2}\mathbb{E}\left[\sum_{t=1}^T v_t^2\right] \le \frac{5}{4}\mathbb{E}\left[\sum_{t=1}^T v_t^2\right].$$

Further, as the noise terms $w_1, \ldots, w_T$ are i.i.d. and have variance σ^2,

$$\mathbb{E}\left(\sum_{t=1}^T m^{T-t} w_t\right)^2 = \sum_{t=1}^T m^{2(T-t)} \mathbb{E}[w_t^2] \le \frac{1}{1 - m^2}\sigma^2 \le \frac{5}{4}\sigma^2.$$

Combining inequalities, the lemma follows. ■

Proof of Lemma 14. Since $1 + b^2 p_\star \ge 1$ and $p_\star \le 5/4$ (see Eq. (7)), Lemma 31 lower bounds the regret as

$$\mathbb{E}[R_T] \ge \mathbb{E}\left[\sum_{t=1}^T (u_t - k_\star x_t)^2\right] - \frac{5}{4}\mathbb{E}[x_{T+1}^2].$$

Plugging in the bound of Lemma 32 and the assumption that $b^2 = \epsilon \le 1/400$, we obtain

$$\mathbb{E}[R_T] \ge \frac{99}{100}\mathbb{E}\left[\sum_{t=1}^T (u_t - k_\star x_t)^2\right] - 4\sigma^2. \tag{28}$$

On the other hand, note that $u_t^2 \le 2(u_t - k_\star x_t)^2 + 2k_\star^2 x_t^2$, and so

$$\mathbb{E}\left[\sum_{t=1}^T u_t^2\right] \le 2\mathbb{E}\left[\sum_{t=1}^T (u_t - k_\star x_t)^2\right] + 2k_\star^2 \mathbb{E}\left[\sum_{t=1}^T x_t^2\right].$$

Further, since $J(k_\star) = \sigma^2 p_\star \le \frac{5}{4}\sigma^2$ we have

$$\mathbb{E}\left[\sum_{t=1}^T x_t^2\right] \le \mathbb{E}\left[\sum_{t=1}^T (x_t^2 + u_t^2)\right] = \mathbb{E}[R_T] + T\mathbb{E}[J(k_\star)] \le \mathbb{E}[R_T] + \frac{5}{4}\sigma^2 T.$$

Therefore,

$$\mathbb{E}\left[\sum_{t=1}^T u_t^2\right] \le 2\mathbb{E}\left[\sum_{t=1}^T (u_t - k_\star x_t)^2\right] + 2k_\star^2 \mathbb{E}[R_T] + \frac{5}{2}\sigma^2 k_\star^2 T. \tag{29}$$

Combining Eqs. (28) and (29) and recalling that $2k_\star^2 \le \epsilon \le 1$ (see Eq. (7)), results with

$$\mathbb{E}\left[\sum_{t=1}^T u_t^2\right] \le 2\big(\frac{100}{99}\mathbb{E}[R_T] + 5\sigma^2\big) + 2k_\star^2 \mathbb{E}[R_T] + \frac{5}{2}\sigma^2 k_\star^2 T \le 3\mathbb{E}[R_T] + \frac{5}{2}\sigma^2 k_\star^2 T + 12\sigma^2,$$

and changing sides yields the second part of the lemma, thus concluding the proof. ■

Proof of Lemma 16. Let Z be a standard Gaussian random variable. Then, using a standard Gaussian tail lower bound,

$$\mathbb{P}\left[|w_{t-1}| \geq \frac{2\sigma}{5}\right] = \mathbb{P}\left[|Z| \geq \frac{2}{5}\right] \geq \frac{17}{25}.$$

Now, recall that $x_t = ax_{t-1} + bu_{t-1} + w_{t-1}$ and notice that, as the learning algorithm is deterministic, both x_{t-1} and u_{t-1} are determined conditioned on $x_1, \ldots, x_{t-1}$. We next aim to lower bound $\mathbb{P}[|x_t| > 2\sigma/5 \mid x_1, \ldots, x_{t-1}]$ which we claim that, as w_{t-1} is a zero-mean Gaussian random variable, is minimized when $ax_{t-1} + bu_{t-1} = 0$. Therefore,

$$\mathbb{P}\left[|x_t| > \frac{2\sigma}{5} \;\middle|\; x_1, \ldots, x_{t-1}\right] \geq \mathbb{P}\left[|w_{t-1}| > \frac{2\sigma}{5}\right] \geq \frac{17}{25}.$$

Denote by $I_t = \mathbb{1}\{|x_t| > 2\sigma/5\}$. Then, by Azuma's concentration inequality we have that with probability at least 7/8,

$$\sum_{t=1}^{T} I_t \geq \sum_{t=1}^{T} \mathbb{E}\left[I_t \mid x_1, \ldots, x_{t-1}\right] - \sqrt{\frac{T}{2}\log 8} \geq \frac{17}{25}T - \sqrt{2T} \geq \frac{2}{3}T,$$

where for the last inequality we used the assumption that $T \geq 12000$. ■

Proof of Lemma 15. First, using Pinsker's inequality yields

$$\mathrm{TV}(\mathbb{P}_+[x^{(T)}], \mathbb{P}_-[x^{(T)}]) \leq \sqrt{\frac{1}{2}\mathrm{KL}(\mathbb{P}_+[x^{(T)}] \parallel \mathbb{P}_-[x^{(T)}])}\,, \tag{30}$$

and by the chain rule of the KL divergence

$$\mathrm{KL}(\mathbb{P}_+[x^{(T)}] \parallel \mathbb{P}_-[x^{(T)}]) = \sum_{t=1}^{T} \mathbb{E}\left[\mathrm{KL}(\mathbb{P}_+[x_t \mid x^{(t-1)}] \parallel \mathbb{P}_-[x_t \mid x^{(t-1)}])\right]. \tag{31}$$

Next, let $\mathbb{E}_+$ and $\mathbb{E}_-$ denote the expectations conditioned on whether $\chi = 1$ or $\chi = -1$ respectively. Observe that as the learning algorithm is deterministic, the sequence of actions $u_1, \ldots, u_{t-1}$ is determined given $x^{(t-1)}$. As such, given $x^{(t-1)}$, the random variable x_t is Gaussian with variance σ^2 and expectation $ax_{t-1} + \sqrt{\epsilon}\chi u_{t-1}$. Therefore, by a standard formula for the KL divergence between Gaussian random variables, we have

$$\begin{aligned}
\mathrm{KL}(\mathbb{P}_+[x_t \mid x^{(t-1)}] \parallel \mathbb{P}_-[x_t \mid x^{(t-1)}]) &= \frac{1}{2\sigma^2}\mathbb{E}_+\left((ax_{t-1} + \sqrt{\epsilon}u_{t-1}) - (ax_{t-1} - \sqrt{\epsilon}u_{t-1})\right)^2 \\
&= \frac{1}{2\sigma^2}\mathbb{E}_+\left(2\sqrt{\epsilon}u_{t-1}\right)^2 \\
&= \frac{2\epsilon}{\sigma^2}\mathbb{E}_+[u_{t-1}^2],
\end{aligned}$$

unless $t = 1$ in which case $\mathrm{KL}(\mathbb{P}_+[x_1] \parallel \mathbb{P}_-[x_1]) = 0$ since x_1 is fixed. Using this bound in Eq. (31) and substituting into Eq. (30) yields

$$\mathrm{TV}(\mathbb{P}_+[x^{(T)}], \mathbb{P}_-[x^{(T)}]) \leq \sqrt{\frac{\epsilon}{\sigma^2}\mathbb{E}_+\left[\sum_{t=1}^{T} u_t^2\right]}\,.$$

Similarly, switching the roles of $\mathbb{P}_+$ and $\mathbb{P}_-$, we get the bound

$$\mathrm{TV}(\mathbb{P}_+[x^{(T)}], \mathbb{P}_-[x^{(T)}]) \leq \sqrt{\frac{\epsilon}{\sigma^2}\mathbb{E}_-\left[\sum_{t=1}^{T} u_t^2\right]}\,.$$

Averaging the two inequalities, using the concavity of the square root, and since $\mathbb{E}[\cdot] = \frac{1}{2}\mathbb{E}_+[\cdot] + \frac{1}{2}\mathbb{E}_-[\cdot]$, we obtain our claim. ■

D. Technical Lemmas

D.1. Noise Bounds

The following theorem is a variant of the Hanson-Wright inequality (Hanson and Wright, 1971; Wright, 1973) which can be found in Hsu et al. (2012).

Theorem 33. *Let $x \sim \mathcal{N}(0, I)$ be a Gaussian random vector,, let $A \in \mathbb{R}^{m\times n}$ and define $\Sigma = A^T A$. Then we have that*

$$\mathbb{P}\left(\|Ax\|^2 > \mathrm{Tr}(\Sigma) + 2\sqrt{\mathrm{Tr}\left(\Sigma^2\right)z} + 2\|\Sigma\|z\right) \le \exp(-z), \qquad \textit{for all } z \ge 0.$$

The following lemma is a direct corollary of Theorem 33.

Lemma 34. *Let $w_t \in \mathbb{R}^d$ for $t = 1, \ldots, T$ be i.i.d. random variables with distribution $\mathcal{N}(0, \sigma^2 I)$. Suppose that $T > 2$, then with probability at least $1-\delta$ we have that*

$$\max_{1\le t\le T}\|w_t\| \le \sigma\sqrt{5d\log\frac{T}{\delta}}.$$

Proof. Consider Theorem 33 with $A = \sigma I$ and thus $\Sigma = \sigma^2 I$. We then have that $\mathrm{Tr}(\Sigma) = d\sigma^2$, $\|\Sigma\| \le \sigma^2$ and $\mathrm{Tr}\left(\Sigma^2\right) \le \|\Sigma\|\mathrm{Tr}(\Sigma) \le d\sigma^4$. We conclude that for $z \ge 1$ we have that

$$\mathrm{Tr}(\Sigma) + 2\sqrt{\mathrm{Tr}\left(\Sigma^2\right)z} + 2\|\Sigma\|z \le \sigma^2 d + 2\sigma^2\sqrt{dz} + 2\sigma^2 z \le 5\sigma^2 dz.$$

Now, for $x \sim \mathcal{N}(0, I)$ we have that $w_t \overset{d}{=} Ax$ (equals in distribution). We thus have that for $z \ge 1$

$$\mathbb{P}\left(\|w_t\| > \sigma\sqrt{5dz}\right) \le \mathbb{P}\left(\|Ax\| > \sqrt{\mathrm{Tr}(\Sigma) + 2\sqrt{\mathrm{Tr}\left(\Sigma^2\right)z} + 2\|\Sigma\|z}\right) \le \exp(-z).$$

Denoting $z = \log\frac{T}{\delta}$, the assumption $T > 2$ ensures that $z \ge 1$ and thus $\mathbb{P}\left(\|w_t\| > \sigma\sqrt{5d\log\frac{T}{\delta}}\right) \le \frac{\delta}{T}$. Performing a union bound over $1 \le t \le T$ we conclude that

$$\mathbb{P}\left(\max_{1\le t\le T}\|w_t\| > \sigma\sqrt{5d\log\frac{T}{\delta}}\right) \le \delta,$$

and taking the complement we obtain the desired. ■

Lemma 35 (Expected maximum noise)**.** *Let E be an event such that $\mathbb{P}(E) \le \delta$ for some $\delta \in [0, 1]$ and let $w_t \in \mathbb{R}^d$ for $t = 1, \ldots, T$ be i.i.d. random variables with distribution $\mathcal{N}(0, \sigma^2 I)$. Suppose $T > 2$, then we have that*

1. $\mathbb{E}\left[\max_{1\le t\le T}\|w_t\|^2\right] \le 5\sigma^2 d\log 3T$;
2. $\mathbb{E}\left[\mathbb{1}\{E\}\max_{1\le t\le T}\|w_t\|^2\right] \le 5\sigma^2 d\delta\log\frac{3T}{\delta}$.

Proof. Recall that from Lemma 34 we have that for all $x \ge 5\sigma^2 d\log T$

$$\mathbb{P}\left(\max_{1\le t\le T}\|w_t\|^2 > x\right) \le T\exp\left(-\frac{x}{5\sigma^2 d}\right).$$

Applying the tail sum formula we get that

$$\begin{aligned}\mathbb{E}\left[\max_{1\le t\le T}\|w_t\|^2\right] &= \int_0^\infty \mathbb{P}\left(\max_{1\le t\le T}\|w_t\|^2 > x\right)dx \\ &\le 5\sigma^2 d\log T + \int_{5\sigma^2 d\log T}^\infty T\exp\left(-\frac{x}{5\sigma^2 d}\right)dx \\ &\le 5\sigma^2 d\log 3T,\end{aligned}$$

proving the first part of the lemma. For the second part notice that $\mathbb{P}\Big(\mathbb{1}\{E\} \max_{1\le t\le T}\|w_t\|^2 > x\Big) \le \min\Big\{\mathbb{P}(E), \mathbb{P}\Big(\max_{1\le t\le T}\|w_t\|^2 > x\Big)\Big\}$. So, applying the tail sum formula we get that

$$\begin{aligned}
\mathbb{E}\left[\mathbb{1}\{E\} \max_{1\le t\le T}\|w_t\|^2\right] &= \int_0^\infty \mathbb{P}\left(\mathbb{1}\{E\} \max_{1\le t\le T}\|w_t\|^2 > x\right)dx \\
&\le \int_0^{5\sigma^2 d \log\frac{T}{\delta}} \mathbb{P}(E)dx + \int_{5\sigma^2 d\log\frac{T}{\delta}}^\infty \mathbb{P}\left(\max_{1\le t\le T}\|w_t\|^2 > x\right)dx \\
&\le 5\sigma^2 d\delta \log\frac{T}{\delta} + \int_{5\sigma^2 d\log\frac{T}{\delta}}^\infty T\exp\left(-\frac{x}{5\sigma^2 d}\right)dx \\
&= 5\sigma^2 d\delta\left(1 + \log\frac{T}{\delta}\right) \\
&\le 5\sigma^2 d\delta \log\frac{3T}{\delta},
\end{aligned}$$

proving the second part and concluding the proof. ∎

D.2. Estimation auxiliary lemmas

The following is due to Cohen et al. (2019). Here we state the result for a general sequence of conditionally Gaussian vectors but the proof follows without change.

Lemma 36 (Theorem 20 of Cohen et al., 2019)**.** *Let z_t for $t = 1, 2, \ldots$ be a sequence random variables that is adapted to a filtration $\{\mathcal{F}_t\}_{t=1}^\infty$. Suppose that z_t are conditionally Gaussian on $\mathcal{F}_{t-1}$ and that $\mathbb{E}\left[z_t z_t^T \mid \mathcal{F}_{t-1}\right] \succeq \sigma_z^2 I$ for some fixed $\sigma_z^2 > 0$. Then for $t \ge 200d\log\frac{12}{\delta}$ we have that with probability at least $1-\delta$*

$$\sum_{s=1}^t z_s z_s^T \succeq \frac{t\sigma_z^2}{40} I.$$

Lemma 37. *Let $z_s \in \mathbb{R}^m$ for $s = 1, \ldots, t-1$ be such that $\|z_s\|^2 \le \lambda$. Define $V_t = \lambda I + \sum_{s=1}^{t-1} z_s z_s^T$ then we have that*

$$\log\frac{\det(V_t)}{\det(V_1)} \le m\log t.$$

Proof. First we have that

$$\|V_t\| \le \lambda + \sum_{s=1}^{t-1}\|z_s z_s^T\| = \lambda + \sum_{s=1}^{t-1}\|z_s\|^2 \le \lambda t.$$

Now, recall that $\det(V_t) \le \det\left(\|V_t\|^m\right)$ and so we have that

$$\log\frac{\det(V_t)}{\det(V_1)} \le \log\frac{\det\left(\|V_t\|^m\right)}{\lambda^m} \le \log\frac{\lambda^m t^m}{\lambda^m} = m\log t,$$

as desired. ∎

D.3. Strong Stability Lemmas

The following lemma bounds the norm of the state when playing a strongly stable controller. Its proof adapts techniques from (Cohen et al., 2019).

Lemma 38. *Suppose K is a (κ, γ)–strongly stable controller and s_0, s_1 are integers such that $1 \le s_0 < s_1 \le T$. Let x_s for $s = s_0, \ldots s_1$ be the sequence of states generated under the control K starting from x_{s_0}, i.e., $x_{s+1} = (A_\star + B_\star K)x_s + w_s$ for all $s_0 \le s < s_1$. Then we have that*

$$\|x_t\| \le \kappa(1-\gamma)^{t-s_0}\|x_{s_0}\| + \frac{\kappa}{\gamma}\max_{1\le t\le T}\|w_t\|, \qquad \text{for all } s_0 \le t \le s_1.$$

Proof. Denote $M = A_\star + B_\star K$ then for $s_0 < t \le s_1$ we have that $x_t = Mx_{t-1} + w_{t-1}$ and by expanding this equation we have

$$x_t = M^{t-s_0}x_{s_0} + \sum_{s=s_0}^{t-1} M^{t-(s+1)}w_s.$$

Recall that by strong stability we have that

$$\|M^s\| = \|HL^sH^{-1}\| \le \kappa(1-\gamma)^s.$$

To ease notation denote $W = \max_{1\le t\le T}\|w_t\|$. Then for $s_0 < t \le s_1$ we have that

$$\begin{aligned}\|x_t\| &\le \|M^{t-s_0}\|\|x_{s_0}\| + \sum_{s=s_0}^{t-1}\|M^{t-(s+1)}\|\|w_s\| \\ &\le \kappa(1-\gamma)^{t-s_0}\|x_{s_0}\| + \sum_{s=s_0}^{t-1}\kappa(1-\gamma)^{t-(s+1)}W \\ &\le \kappa(1-\gamma)^{t-s_0}\|x_{s_0}\| + \frac{\kappa}{\gamma}W.\end{aligned}$$

■

The following lemma bounds the norm of the state when playing a sequence of strongly stable controllers.

Lemma 39. *Suppose $K_1, \ldots, K_l$ are (κ, γ)-strongly stable controllers and $\{t_i\}_{i=1}^{l+1}$ are integers such that $1 \le t_1 < \ldots < t_{l+1} \le T$. Let x_t for $t = t_1, \ldots t_{l+1}$ be the sequence of states generated by starting from x_{t_1} and playing controller K_i at times $t_i \le t < t_{i+1}$, i.e., $x_{t+1} = (A_\star + B_\star K_i)x_t + w_t$ for all $t_i \le t < t_{i+1}$. Denote $\tau = \min_i\{t_{i+1} - t_i\}$ and suppose that $\tau \ge \gamma^{-1}\log(2\kappa)$, then we have that*

$$\|x_t\| \le 3\kappa \max\left\{\frac{1}{2}\|x_{t_1}\|, \frac{\kappa}{\gamma}\max_{1\le t\le T}\|w_t\|\right\}, \quad \forall\, t_1 \le t \le t_{l+1}.$$

Proof. For $0 < \gamma \le 1$ it is a well known fact that $\gamma \le -\log 1-\gamma$. Plugging this into the lower bound on τ and rearranging we get that $\kappa(1-\gamma)^\tau \le \frac{1}{2}$. Now, applying Lemma 38 with $s_0 = t_i$ and $s_1 = t_{i+1}$, and taking $t = t_{i+1}$ we have that

$$\begin{aligned}\|x_{t_{i+1}}\| &\le \kappa(1-\gamma)^{t_{i+1}-t_i}\|x_{t_i}\| + \frac{\kappa}{\gamma}W \\ &\le \kappa(1-\gamma)^\tau\|x_{t_i}\| + \frac{\kappa}{\gamma}W \\ &\le \frac{1}{2}\|x_{t_i}\| + \frac{\kappa}{\gamma}W,\end{aligned}$$

and solving this difference equation we get that

$$\|x_{t_i}\| \le \frac{2\kappa}{\gamma}W + \left(\|x_{t_1}\| - \frac{2\kappa}{\gamma}W\right)2^{1-i} \le \max\left\{\|x_{t_1}\|, \frac{2\kappa}{\gamma}W\right\}.$$

Plugging this result back into Lemma 38 we have that for $t_i < t \le t_{i+1}$

$$\begin{aligned}\|x_t\| &\le \kappa(1-\gamma)^{t-t_i}\max\left\{\|x_{t_1}\|, \frac{2\kappa}{\gamma}W\right\} + \frac{\kappa}{\gamma}W \\ &\le \kappa\max\left\{\|x_{t_1}\|, \frac{2\kappa}{\gamma}W\right\} + \frac{\kappa}{\gamma}W \\ &\le \kappa\max\left\{\frac{3\|x_{t_1}\|}{2}, \frac{3\kappa}{\gamma}W\right\},\end{aligned}$$

where the last inequality used the fact that $\kappa \ge 1$. This is true for all i and thus for all $t_1 \le t \le t_{l+1}$. ■

The next two lemmas require the following well known result in linear control theory (see, e.g., Bertsekas, 1995). We have that $J(K) = \sigma^2\mathrm{Tr}(P)$ where P is a positive definite solution of

$$P = Q + K^TRK + (A_\star + B_\star K)^TP(A_\star + B_\star K). \tag{32}$$

The following lemma relates the expected cost of playing controller K for t rounds to the infinite horizon cost of K.

Lemma 40. *Suppose K is a (κ, γ)–strongly stable controller and let x_s for $s = 1, \dots t$ be the sequence of states generated under the control K starting from x_1, i.e., $x_{s+1} = (A_\star + B_\star K)x_s + w_s$ for all $1 \le s < t$. Then we have that*

$$\mathbb{E}\left[\sum_{s=1}^{t} x_s^T \left(Q + K^T R K\right) x_s \;\middle|\; x_1\right] \le tJ(K) + \frac{2\alpha_1 \kappa^4}{\gamma} \|x_1\|^2.$$

Proof. To ease notation, assume, without loss of generality, that x_1 is deterministic. We thus omit the conditioning on x_1 in all expectation arguments.

First, recall that $x_{s+1} = (A_\star + B_\star K)x_s + w_s$ and $J(K) = \sigma^2 \mathrm{Tr}(P)$ where P satisfies Eq. (32). Then we have that

$$\begin{aligned}
\mathbb{E}\left[x_{s+1}^T P x_{s+1}\right] &= \mathbb{E}\left[((A_\star + B_\star K)x_s + w_s)^T P((A_\star + B_\star K)x_s + w_s)\right] \\
&= \mathbb{E}\left[((A_\star + B_\star K)x_s)^T P((A_\star + B_\star K)x_s)\right] + \mathbb{E}\left[w_s^T P w_s\right] \\
&= \mathbb{E}\left[x_s^T (A_\star + B_\star K)^T P (A_\star + B_\star K)x_s\right] + J(K).
\end{aligned}$$

Now, multiplying Eq. (32) by x_s from both sides and taking expectation we get that

$$\begin{aligned}
\mathbb{E}\left[x_s^T P x_s\right] &= \mathbb{E}\left[x_s^T \left(Q + K^T R K\right) x_s\right] + \mathbb{E}\left[x_s^T (A_\star + B_\star K)^T P (A_\star + B_\star K)x_s\right] \\
&= \mathbb{E}\left[x_s^T \left(Q + K^T R K\right) x_s\right] + \mathbb{E}\left[x_{s+1}^T P x_{s+1}\right] - J(K),
\end{aligned}$$

and changing sides and summing over s we get that

$$\mathbb{E}\left[x_1^T P x_1 - x_{t+1}^T P x_{t+1}\right] = \sum_{s=1}^{t} \mathbb{E}\left[x_s^T P x_s - x_{s+1}^T P x_{s+1}\right] = \mathbb{E}\left[\sum_{s=1}^{t} x_s^T \left(Q + K^T R K\right) x_s\right] - tJ(K),$$

and changing sides again we conclude that

$$\mathbb{E}\left[\sum_{s=1}^{t} x_s^T \left(Q + K^T R K\right) x_s\right] \le tJ(K) + \mathbb{E}\left[x_1^T P x_1\right] \le tJ(K) + \|x_1\|^2 \|P\|.$$

We conclude the proof by bounding $\|P\|$. To that end, recall that the strong stability of K implies that $A_\star + B_\star K = HLH^{-1}$ where $\|L\| \le 1 - \gamma$ and $\|H\|\|H^{-1}\| \le \kappa$. Applying Eq. (32) recursively we then have that

$$\begin{aligned}
\|P\| &= \left\| \sum_{s=0}^{\infty} \left((A_\star + B_\star K)^s\right)^T \left(Q + K^T R K\right) (A_\star + B_\star K)^s \right\| \\
&= \left\| \sum_{s=0}^{\infty} \left(HL^s H^{-1}\right)^T \left(Q + K^T R K\right) HL^s H^{-1} \right\| \\
&\le \|H\|^2 \|H^{-1}\|^2 \|Q + K^T R K\| \sum_{s=0}^{\infty} \|L\|^{2s} \\
&\le 2\alpha_1 \kappa^4 \sum_{s=0}^{\infty} (1 - \gamma)^s = \frac{2\alpha_1 \kappa^4}{\gamma},
\end{aligned}$$

thus concluding the proof. ■

The following lemma relates the infinite horizon cost of a controller to its strong stability parameters. Its proof is an adaptation of Lemma 18 in (Cohen et al., 2019) that fits our assumptions.

Lemma 41. *Suppose $J(K) < J$ then K is (κ, γ)–strongly stable with $\kappa = \sqrt{\frac{J}{\alpha_0 \sigma^2}}$ and $\gamma = \frac{\alpha_0 \sigma^2}{2J}$.*

Proof. Recall that $J(K) = \sigma^2 \mathrm{Tr}(P)$ where P satisfies Eq. (32). Using the bound $J(K) \le J$ we have that $\mathrm{Tr}(P) \le J/\sigma^2$ and thus also that $P \preceq (J/\sigma^2)I$. Recalling that $Q \succeq \alpha_0 I$ we get that $Q \succeq \frac{\alpha_0 \sigma^2}{J} P = 2\gamma P$. Recalling that R is positive definite and plugging back into Eq. (32) we get that

$$P \succeq 2\gamma P + (A_\star + B_\star K)^T P (A_\star + B_\star K),$$

rearranging the equation we get that

$$P^{-1/2}(A_\star + B_\star K)^T P (A_\star + B_\star K) P^{-1/2} \preceq (1 - 2\gamma) I.$$

Now, denote $H = P^{-1/2}$ and $L = P^{1/2}(A_\star + B_\star K)P^{-1/2}$ and notice that indeed $HLH^{-1} = A_\star + B_\star K$. Plugging into the above we get that

$$P^{-1/2}(A_\star + B_\star K)^T P (A_\star + B_\star K) P^{-1/2} = H\left(HLH^{-1}\right)^T H^{-1} H^{-1}\left(HLH^{-1}\right) H = L^T L \preceq (1 - 2\gamma) I,$$

and thus $\|L\| \leq \sqrt{1-2\gamma} \leq 1-\gamma$. Now recall that $P \preceq (J/\sigma^2) I$ and thus $\|H^{-1}\| = \|P^{1/2}\| \leq \sqrt{J/\sigma^2}$. Going back to Eq. (32) we also have that $P \succeq Q \succeq \alpha_0 I$ and thus $\|H\| = \|P^{-1/2}\| \leq \sqrt{1/\alpha_0}$. All together, we get that $\|H\|\|H^{-1}\| \leq \sqrt{J/\alpha_0\sigma^2} = \kappa$. Finally, recall that $R \succeq \alpha_0 I$ and thus going back to Eq. (32) we have that $P \succeq K^T R K \succeq \alpha_0 K^T K$ and thus $\|K\| \leq \sqrt{\|P\|/\alpha_0} \leq \sqrt{J/\alpha_0\sigma^2} = \kappa$, as desired. ■

The following lemma relates system parameter estimation bounds to properties of the resulting greedy controller.

Lemma 42. *Let $A \in \mathbb{R}^{d\times d}, B \in \mathbb{R}^{d\times k}$ and take $K = \mathcal{K}(A, B, Q, R)$. Denote $\Delta = \max\{\|A - A_\star\|, \|B - B_\star\|\}$, $\kappa = \sqrt{\frac{\nu + C_0\varepsilon_0^2}{\alpha_0\sigma^2}}$, and $\gamma = \frac{1}{2\kappa^2}$. Then we have that*

1. *If $\Delta \leq \varepsilon_0$ then K is (κ, γ)–strongly stable;*
2. *If $\Delta \leq \min\left\{\varepsilon_0, \frac{\mu}{4\kappa C_0}\right\}$ then $KK^T \succeq K_\star K_\star^T - \frac{\mu}{2} I$ and $K_\star K_\star^T \succeq KK^T - \frac{\mu}{2} I$;*
3. *If $\Delta \leq \min\left\{\varepsilon_0, \frac{\mu_\star}{4\kappa C_0}\right\}$ then $KK^T \succeq \frac{\mu_\star}{2} I$.*

Proof. First, if $\Delta \leq \varepsilon_0$ we can invoke Lemma 4 to get that $J(K) \leq J_\star + C_0\varepsilon_0^2 \leq \nu + C_0\varepsilon_0^2$ and so by Lemma 41, K is (κ, γ)–strongly stable, proving the first part of the lemma.

Second, if $\Delta \leq \min\left\{\varepsilon_0, \frac{\mu}{4\kappa C_0}\right\}$ then we can invoke Lemma 4 to get that $\|K - K_\star\| \leq \frac{\mu}{4\kappa}$. Moreover, by the first claim of the lemma, $K, K_\star$ are (κ, γ)–strongly stable and thus upper bounded by κ. Combining the above we get that

$$\begin{aligned} KK^T &= K_\star K_\star^T - \frac{1}{2}\left((K_\star + K)(K_\star - K)^T + (K_\star - K)(K_\star + K)^T\right) \\ &\succeq K_\star K_\star^T - \left(\|K_\star\| + \|K\|\right)\|K_\star - K\| I \\ &\succeq K_\star K_\star^T - \frac{2\kappa\mu}{4\kappa} I = K_\star K_\star^T - \frac{\mu}{2} I, \end{aligned}$$

and reversing the roles of K and $K_\star$ in the above yields $K_\star K_\star^T \succeq KK^T - \frac{\mu}{2} I$, thus proving the second part of the lemma.

Finally, if $\Delta \leq \min\left\{\varepsilon_0, \frac{\mu_\star}{4\kappa C_0}\right\}$, then recalling that $K_\star K_\star \succeq \mu_\star I$ and continuing from the second part we get that

$$KK^T \succeq K_\star K_\star^T - \frac{\mu_\star}{2} I \succeq \mu_\star I - \frac{\mu_\star}{2} I = \frac{\mu_\star}{2} I,$$

thus concluding the third and final part of the lemma. ■