F Proof of Lemma 5

Note that in our case the operators Q and R are the same both for the exact and Nyströmized dynamics, and are assumed to be positive definite. The structure of the proof is as follows. First, we write down a fixed-point operator equation, $X = \Phi X$, for a suitable operator Φ, that has a unique solution, given by the error on the Riccati operator due to the Nyström approximation, $\tilde{P} - P$. Then, we define a set of perturbations X with bounded norm, and we show that Φ maps this set to itself, and is a contraction. This implies that the fixed point of Φ, which is shown to be $\tilde{P} - P$, lies in the set we defined, and has bounded norm. We begin by defining

$$N : \mathcal{H}_1 \to \mathcal{H}_1,\ N = BR^{-1}B^*, \tag{F.1}$$

$$\tilde{N} : \mathcal{H}_1 \to \mathcal{H}_1,\ \tilde{N} = \tilde{B}R^{-1}\tilde{B}^*. \tag{F.2}$$

Proposition 10 *Given P as the self-adjoint positive semi-definite solution of (36), let us consider the following set of self-adjoint perturbations*

$$\mathcal{S} = \{X : \mathcal{H}_1 \to \mathcal{H}_1, X = X^*, P + X \succcurlyeq 0\}. \tag{F.3}$$

Let $X \in \mathcal{S}$. Let $L = A + BK$, where K is defined in (34). Then, it holds

$$F(P + X, A, B) = \underbrace{X - L^*XL}_{=:\mathcal{T}X} + \underbrace{L^*X[I + N(P + X)]^{-1}NXL}_{=:\mathcal{R}X}.$$

Proof This statement can be found, in the matrix setting, in [26,32]. We report here the derivation for completeness. By definition of P it holds $F(P, A, B) = 0$ thus:

$$\begin{aligned}
&X - L^*XL + L^*X[I + N(P + X)]^{-1}NXL \\
&= X - L^*XL + L^*X[I + N(P + X)]^{-1}NXL + F(P, A, B) && \text{(F.4)}\\
&= X - L^*XL + L^*X[I + N(P + X)]^{-1}NXL + P - A^*P(I + NP)^{-1}A - Q && \text{(F.5)}\\
&= (P + X) - Q + \Big[L^*X[I + N(P + X)]^{-1}NXL - L^*XL - A^*P(I + NP)^{-1}A\Big] && \text{(F.6)}\\
&= (P + X) - Q + \Big[L^*X\Big([I + N(P + X)]^{-1}NX - I\Big)L - A^*P(I + NP)^{-1}A\Big] && \text{(F.7)}\\
&= (P + X) - Q + \Big[L^*X\Big([I + N(P + X)]^{-1}NX - [I + N(P + X)]^{-1}[I + N(P + X)]\Big)L && \text{(F.8)}\\
&\quad - A^*P(I + NP)^{-1}A\Big] \\
&= (P + X) - Q - \Big[L^*X\Big([I + N(P + X)]^{-1}[I + NP]\Big)L + A^*P(I + NP)^{-1}A\Big] && \text{(F.9)}\\
&= (P + X) - Q - A^*\Big[((I + NP)^*)^{-1}X[I + N(P + X)]^{-1} + P(I + NP)^{-1}\Big]A && \text{(F.10)}\\
&= (P + X) - Q - A^*\Big[(I + PN)^{-1}X + P(I + NP)^{-1}[I + N(P + X)]\Big][I + N(P + X)]^{-1}A && \text{(F.11)}\\
&= (P + X) - Q - A^*\Big[(I + PN)^{-1}X + P(I + NP)^{-1}NX + P\Big][I + N(P + X)]^{-1}A && \text{(F.12)}\\
&= (P + X) - Q - A^*(P + X)(I + N(P + X))^{-1}A && \text{(F.13)}\\
&= F(P + X, A, B), && \text{(F.14)}
\end{aligned}$$

which concludes the proof. □

We can now study the operator $\mathcal{T}$ appearing in Proposition 10.

Proposition 11 *Let $\mathcal{L}(\mathcal{H}_1)$ denote the set of continuous linear operators from $\mathcal{H}_1$ to itself. The operator $\mathcal{T} : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_1)$ appearing in Proposition 10 is invertible. Under assumptions 2 and 3, the fixed point equation*

$$X = \mathcal{T}^{-1}[F(X + P, A, B) - F(X + P, \tilde{A}, \tilde{B}) - \mathcal{R}X]$$

has a unique solution in $\mathcal{S}$, namely $X = P - \tilde{P}$.

Proof The operator $\mathcal{T}$ can be rewritten as

$$\mathcal{T}X = (I - \mathcal{D})X \tag{F.15}$$

where $\mathcal{D}X = L^*XL : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_1)$ is linear ($\mathcal{D}(X+Y) = L^*XL + L^*YL, \mathcal{D}(\alpha X) = \alpha L^*XL$). Let $\mathcal{L}(\mathcal{L}(\mathcal{H}_1))$ denote the set of continuous linear operators from $\mathcal{L}(\mathcal{H}_1)$ to itself. Then $\mathcal{T} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1))$ (because A, B, Q, R, and K, defined in (34), appearing in the expression of L are bounded). We denote the spectrum of $\mathcal{D}$ as

$$\sigma(\mathcal{D}) = \{\lambda : \mathcal{D} - \lambda I \text{ is not invertible in } \mathcal{L}(\mathcal{L}(\mathcal{H}_1))\}, \tag{F.16}$$

and its spectral radius $\rho(\mathcal{D}) = \sup_{\lambda\in\sigma(\mathcal{D})} |\lambda|$. The operator $\mathcal{T}$ is invertible in $\mathcal{L}(\mathcal{L}(\mathcal{H}_1))$ iff 1 is not in the spectrum $\sigma(\mathcal{D})$, and we now show that this holds by proving $\rho(\mathcal{D}) < 1$.

We can observe that since K in (34) is a stabilizing gain by assumption, it holds [20, page 4] $\rho(L) < 1$. The spectral radius of $\mathcal{D}$ can be computed with the Gelfand's formula, using the fact that $L^k, (L^*)^k$ are bounded for any k:

$$\begin{aligned}
\rho(\mathcal{D}) &= \lim_{k\to\infty} \|\mathcal{D}^k\|^{1/k} && \text{(F.17)}\\
&= \lim_{k\to\infty} \sup_{\|X\|=1} \|\mathcal{D}^k X\|^{1/k} && \text{(F.18)}\\
&= \lim_{k\to\infty} \sup_{\|X\|=1} \|{L^*}^k X L^k\|^{1/k} && \text{(F.19)}\\
&\leq \lim_{k\to\infty} \|{L^*}^k\|^{1/k} \|L^k\|^{1/k} && \text{(F.20)}\\
&= \rho(L)^2 && \text{(F.21)}\\
&< 1. && \text{(F.22)}
\end{aligned}$$

Now, we can consider the following equation:

$$F(X+P, A, B) - F(X+P, \tilde{A}, \tilde{B}) = \mathcal{T}X + \mathcal{R}X. \tag{F.23}$$

By Proposition 10 we have

$$F(X+P, A, B) = \mathcal{T}X + \mathcal{R}X, \tag{F.24}$$

and thus it must hold

$$F(X+P, \tilde{A}, \tilde{B}) = 0. \tag{F.25}$$

By definition of $\tilde{P}$ in (37), $X = \tilde{P} - P$ satisfies this equation, but it is also known [20, Theorem 9] that under Assumptions 2 and 3, (F.25) admits a unique self-adjoint positive semi-definite solution. Moreover as both P and $\tilde{P}$ are self-ajdoint, $X = \tilde{P} - P$ is self-adjoint as well and thus the unique solution of (F.25) in $\mathcal{S}$.

Also, note that (F.23), for a suitable operator $\Phi : \mathcal{S} \to \mathcal{L}(\mathcal{H}_1)$, can be rewritten as

$$X = \underbrace{\mathcal{T}^{-1}[F(X+P, A, B) - F(X+P, \tilde{A}, \tilde{B}) - \mathcal{R}X]}_{=:\Phi X}, \tag{F.26}$$

so that $\tilde{P} - P$ is the unique fixed point of Φ. $\square$

We have shown that the operator Φ is well defined and has a unique fixed point in $\mathcal{S}$ equal to $\tilde{P} - P$, and we will now show that the latter is upper bounded by ϵ in operator norm, where ϵ is the error on the matrices of the dynamical system, similarly to [32]. As mentioned at the beginning of this section, in the following, we will define a set $\mathcal{S}_\nu$ of perturbations $X \in \mathcal{S}$ with bounded norm. Furthermore, we will show that the function Φ is a mapping from $\mathcal{S}_\nu$ to itself, and that it is a contraction (i.e. α-Lipschitz, $\alpha < 1$). Therefore, since $\mathcal{S}_\nu$ is a closed subset of $\mathcal{S}$, by the Banach fixed point theorem, Φ has a fixed point in $\mathcal{S}_\nu$. Since the fixed point of Φ in $\mathcal{S}$ has been proven to be $\tilde{P} - P$, this means that the error on the P operator due to the Nyström approximation is bounded, and the bound depends on the error rate of the system's matrices. More formally, let us define:

$$\mathcal{S}_\nu = \{X : \|X\| \leq \nu, X = X^*, P + X \succcurlyeq 0\}. \tag{F.27}$$

We can now prove some technical bounds that are used in [32]. We can define

$$\begin{aligned}
\Delta A &:= \tilde{A} - A, && \text{(F.28)}\\
\Delta B &:= \tilde{B} - B, && \text{(F.29)}\\
\Delta N &:= \tilde{N} - N, && \text{(F.30)}\\
\Delta P &:= \tilde{P} - P. && \text{(F.31)}
\end{aligned}$$

where we recall that $N, \tilde{N}$ are defined in F.1, F.2.

Proposition 12 *Assume $\|\Delta A\| \leq \epsilon$, and $\|\Delta B\| \leq \epsilon$. Then, for $\epsilon \leq \|B\|$, the following bound holds:*

$$\|\Delta N\| \leq 3\epsilon \|R^{-1}\| \|B\|.$$

Proof We can observe that:

$$\begin{aligned}
\|\Delta N\| &= \|BR^{-1}B^* - \tilde{B}R^{-1}\tilde{B}^*\| && \text{(F.32)}\\
&= \|BR^{-1}B^* - \tilde{B}R^{-1}B^* + \tilde{B}R^{-1}B^* - \tilde{B}R^{-1}\tilde{B}^*\| && \text{(F.33)}\\
&= \|(B - \tilde{B})R^{-1}B^* + \tilde{B}R^{-1}(B^* - \tilde{B}^*)\| && \text{(F.34)}\\
&\leq \|(B - \tilde{B})R^{-1}B^*\| + \|\tilde{B}R^{-1}(B^* - \tilde{B}^*)\| && \text{(F.35)}\\
&\leq \|B - \tilde{B}\|\|R^{-1}\|(\|B\| + \|\tilde{B}\|) && \text{(F.36)}\\
&\leq \epsilon\|R^{-1}\|(\|B\| + \|\tilde{B}\|). && \text{(F.37)}
\end{aligned}$$

We can observe that, by choosing $\epsilon \leq \|B\|$, by the reversed triangle inequality:

$$\|\tilde{B}\| - \|B\| \leq \|\tilde{B} - B\| \leq \epsilon \leq \|B\|, \tag{F.38}$$

meaning that

$$\|\Delta N\| \leq 3\epsilon\|R^{-1}\|\|B\|, \tag{F.39}$$

as also shown in [32]. □

Proposition 13 *Let L, $\mathcal{T}$ be as defined in Proposition 10. Moreover, let*

$$\tau(L, \rho) = \sup\{\|L^k\|\rho^{-k} : k \geq 0\}.$$

By assumption K stabilizes the system, i.e. $\rho(L) < 1$. For any $\rho(L) \leq \rho < 1$ we have that

$$\|\mathcal{T}^{-1}\| \leq \frac{\tau(L, \rho)^2}{1 - \rho^2}.$$

Proof By definition we have that operator $\mathcal{T}$ can be expressed as $\mathcal{T} = I - \mathcal{D}$, where $\mathcal{D}X = L^*XL$. Note that L is bounded because A, B and K are bounded, and thus $\mathcal{D} : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_1)$ is also bounded as

$$\|\mathcal{D}\| = \sup_{\|X\|=1} \|\mathcal{D}X\| = \sup_{\|X\|=1} \|L^*XL\| \leq \|L\|^2. \tag{F.40}$$

Now, observe that, since the spectral radius of L is strictly smaller than 1 (stability), we can choose $\rho(L) < \rho < 1$ so that the quantity $\tau(L, \rho)$ is finite [32]. Indeed according to Gelfand's formula, we are guaranteed that there exists k_0 such that for any $k \geq k_0$, $\|L^k\| \leq \rho^k$. Hence, we can choose

$$\tau'(L, \rho) = \max\{1\} \cup \{\|L^k\|/\rho^k : k \in [0, \dots, k_0]\} \tag{F.41}$$

which involves a finite sequence of values, so that for any $k \in \mathbb{N}$ it holds $\|L^k\| \leq \tau'(L, \rho)\rho^k$, and $\tau(L, \rho) \leq \tau'(L, \rho) < \infty$. The Neumann series $\sum_{k=0}^{\infty} \mathcal{D}^k$ is (absolutely) convergent in operator norm, as

$$\begin{aligned}
\sum_{k=0}^{\infty} \|\mathcal{D}^k\| &= \sum_{k=0}^{\infty} \sup_{\|X\|=1} \|L^{*k} X L^k\| && \text{(F.42)} \\
&\leq \sum_{k=0}^{\infty} \|L^k\|^2 && \text{(F.43)} \\
&\leq \sum_{k=0}^{\infty} \left[\tau(L, \rho)\rho^k\right]^2 && \text{(F.44)} \\
&= \frac{\tau(L, \rho)^2}{1 - \rho^2}, && \text{(F.45)}
\end{aligned}$$

since we have a geometric series with $\rho^2 < 1$. Therefore, we can express $\mathcal{T}^{-1}$ using the Neumann series as

$$\mathcal{T}^{-1} = (I - \mathcal{D})^{-1} = \sum_{k=0}^{\infty} \mathcal{D}^k. \tag{F.46}$$

This yields the claimed result by the triangle inequality. □

Proposition 14 . *Let $\mathcal{R}$, L, N, X be the operators defined in Proposition 10. Then, we have*

$$\|\mathcal{R}X\| \leq \|L\|^2 \|X\|^2 \|N\|.$$

Proof The first part of the proof is the same as the one of Lemma 7 in [32]. Note that N and $P + X$ are self-adjoint, positive semi-definite operators. Let us now consider $\alpha > 0$. We have that

$$\begin{aligned}
0 \preccurlyeq (N + \alpha I)[I + (P + X + \alpha I)(N + \alpha I)]^{-1} &= [(N + \alpha I)^{-1} + P + X + \alpha I]^{-1} && \text{(F.47)} \\
&\preccurlyeq (N + \alpha I). && \text{(F.48)}
\end{aligned}$$

This implies the following norm inequality:

$$\|(N + \alpha I)[I + (P + X + \alpha I)(N + \alpha I)]^{-1}\| \leq \|N + \alpha I\|. \tag{F.49}$$

By definition, since limits preserve inequalities and by continuity of the norm, we have that

$$\begin{aligned}
\|N[I + (P + X)N]^{-1}\| &= \|\lim_{\alpha \to 0} (N + \alpha I)[I + (P + X + \alpha I)(N + \alpha I)]^{-1}\| && \text{(F.50)} \\
&= \lim_{\alpha \to 0} \|(N + \alpha I)[I + (P + X + \alpha I)(N + \alpha I)]^{-1}\| && \text{(F.51)} \\
&\leq \lim_{\alpha \to 0} \|N + \alpha I\| && \text{(F.52)} \\
&= \|N\|. && \text{(F.53)}
\end{aligned}$$

Hence,

$$\begin{aligned}\|\mathcal{R}X\| &= \|L^*X[I+N(P+X)]^{-1}NXL\| && \text{(F.54)}\\ &= \|L^*XN[I+(P+X)N]^{-1}XL\| && \text{(F.55)}\\ &\leq \|L\|^2\|X\|^2\|N\|. && \text{(F.56)}\end{aligned}$$

□

Proposition 15 *With the same notations and assumptions of Proposition 10, for $X \in \mathcal{S}_\nu$, we have that*

$$\|F(P+X,\tilde{A},\tilde{B}) - F(P+X,A,B)\| \leq \|A\|^2\|P+X\|^2\|\Delta N\| + 2\|A\|\|P+X\|\epsilon + \|P+X\|\epsilon^2.$$

Proof Let P_X be a shorthand for $P+X$. We can consider the following decomposition [32]:

$$\begin{aligned}F(P_X,\tilde{A},\tilde{B}) - F(P_X,A,B) =& A^*P_X(I+NP_X)^{-1}A - \tilde{A}^*P_X(I+\tilde{N}P_X)^{-1}\tilde{A} && \text{(F.57)}\\ =& A^*P_X(I+NP_X)^{-1}A - A^*P_X(I+\tilde{N}P_X)^{-1}A\\ &+ A^*P_X(I+\tilde{N}P_X)^{-1}A - A^*P_X(I+\tilde{N}P_X)^{-1}\tilde{A}\\ &+ A^*P_X(I+\tilde{N}P_X)^{-1}A - \tilde{A}^*P_X(I+\tilde{N}P_X)^{-1}A\\ &+ A^*P_X(I+\tilde{N}P_X)^{-1}\tilde{A} + \tilde{A}^*P_X(I+\tilde{N}P_X)^{-1}A\\ &- A^*P_X(I+\tilde{N}P_X)^{-1}A - \tilde{A}^*P_X(I+\tilde{N}P_X)^{-1}\tilde{A} && \text{(F.58)}\\ =& A^*P_X\left[(I+NP_X)^{-1} - (I+\tilde{N}P_X)^{-1}\right]A\\ &- A^*P_X(I+\tilde{N}P_X)^{-1}\Delta A\\ &- \Delta A^*P_X(I+\tilde{N}P_X)^{-1}A\\ &- \Delta A^*P_X(I+\tilde{N}P_X)^{-1}\Delta A && \text{(F.59)}\\ =& A^*P_X(I+NP_X)^{-1}\Delta NP_X(I+\tilde{N}P_X)^{-1}A\\ &- A^*P_X(I+\tilde{N}P_X)^{-1}\Delta A\\ &- \Delta A^*P_X(I+\tilde{N}P_X)^{-1}A\\ &- \Delta A^*P_X(I+\tilde{N}P_X)^{-1}\Delta A. && \text{(F.60)}\end{aligned}$$

By [32, Lemma 7], we get

$$\|F(P_X,\tilde{A},\tilde{B}) - F(P_X,A,B)\| \leq \|A\|^2\|P_X\|^2\|\Delta N\| + 2\|A\|\|P_X\|\epsilon + \|P_X\|\epsilon^2. \tag{F.61}$$

□

With the technical results introduced above, we are now ready to show that Φ maps $\mathcal{S}_\nu$ to itself, and is a contraction. We can firstly compute suitable upper bounds, merging the results above.

Proposition 16 *With the same notations and assumptions of Proposition 10, for $X \in \mathcal{S}_\nu$, $\nu \leq 1/2$, $\epsilon \leq \|B\|$, and $\rho(L) \leq \rho < 1$, it holds that*

$$\|\Phi X\| \leq \frac{\tau(L,\rho)^2}{1-\rho^2}\left[\|L\|^2\|N\|\nu^2 + 3\epsilon(\|P\|+1)^2(\|A\|+1)^2(\|R^{-1}\|+1)(\|B\|+1)\right].$$

Proof We already know from Proposition 12 that $\|\Delta N\| \leq 3\epsilon\|R^{-1}\|\|B\|$. Moreover, $\|P_X\| \leq \|P\| + \nu \leq \|P\| + 1$ by the triangle inequality and given that $X \in \mathcal{S}_\nu$ and $\nu \leq 1/2$. According to Proposition 11, 13, 14 and using $\epsilon \leq \|B\|$:

$$\begin{aligned}
\|\Phi X\| &= \|\mathcal{T}^{-1}[F(X+P,A,B) - F(X+P,\tilde{A},\tilde{B}) - \mathcal{R}X]\| && \text{(F.62)}\\
&\leq \|\mathcal{T}^{-1}\| \left[\|F(X+P,A,B) - F(X+P,\tilde{A},\tilde{B})\| + \|\mathcal{R}X\|\right] && \text{(F.63)}\\
&\leq \frac{\tau(L,\rho)^2}{1-\rho^2}\left[\|L\|^2\|N\|\nu^2 + \|A\|^2\|P_X\|^2 3\epsilon\|R^{-1}\|\|B\| + 2\|A\|\|P_X\|\epsilon + \|P_X\|\epsilon^2\right] && \text{(F.64)}\\
&\leq \frac{\tau(L,\rho)^2}{1-\rho^2}\left[\|L\|^2\|N\|\nu^2 + \epsilon\|P_X\|\left(3\|A\|^2\|P_X\|\|R^{-1}\|\|B\| + 2\|A\| + \epsilon\right)\right] && \text{(F.65)}\\
&\leq \frac{\tau(L,\rho)^2}{1-\rho^2}\left[\|L\|^2\|N\|\nu^2 + \epsilon\|P_X\|\left(3\|A\|^2\|P_X\|\|R^{-1}\|\|B\| + 2\|A\| + \|B\|\right)\right]. && \text{(F.66)}
\end{aligned}$$

We use a rough upper-bounded on the second term of this last expression to get claimed result, which is more convenient to work with:

$$\begin{aligned}
&3\epsilon(\|P\|+1)^2(\|A\|+1)^2(\|R^{-1}\|+1)(\|B\|+1) && \text{(F.67)}\\
&\quad= 3\epsilon(\|P\|+1)^2(\|A\|+1)^2(\|R^{-1}\|\|B\| + \|R^{-1}\| + \|B\| + 1) && \text{(F.68)}\\
&\quad= 3\epsilon(\|P\|+1)^2(\|A\|^2 + 2\|A\| + 1)(\|R^{-1}\|\|B\| + \|R^{-1}\| + \|B\| + 1) && \text{(F.69)}\\
&\quad= 3\epsilon(\|P\|^2 + 2\|P\| + 1)(\|A\|^2 + 2\|A\| + 1)(\|R^{-1}\|\|B\| + \|R^{-1}\| + \|B\| + 1) && \text{(F.70)}\\
&\quad\geq 3\epsilon\|A\|^2(\|P\|+1)^2\|R^{-1}\|\|B\| + 12\|A\|(\|P\|+1)\epsilon + 6(\|P\|+1)\|B\|\epsilon) && \text{(F.71)}\\
&\quad\geq \|A\|^2\|P_X\|^2 3\epsilon\|R^{-1}\|\|B\| + 2\|A\|\|P_X\|\epsilon + \|P_X\|\epsilon^2. && \text{(F.72)}
\end{aligned}$$

□

Proposition 17 *With the same assumptions and notations of Proposition 10, if $X_1, X_2 \in \mathcal{S}_\nu$, $\nu \leq \min\{1/2, \|N\|^{-1}\}$, $\|\Delta A\| \leq \epsilon$, $\|\Delta B\| \leq \epsilon$, $\epsilon \leq \|B\|$, and $\rho(L) \leq \rho < 1$, it holds that*

$$\begin{aligned}
\|\Phi X_1 - \Phi X_2\| \leq 32\frac{\tau(L,\rho)^2}{1-\rho^2}[&((\|A\|+1)^2(\|P\|+1)^3(\|B\|+1)^3(\|R\|^{-1}+1)^2\|X_1 - X_2\|\epsilon\\
&+ \|L\|^2\nu\|N\|\|X_1 - X_2\|].
\end{aligned}$$

Proof Let us choose $\nu \leq \min\{1/2, \|N\|^{-1}\}$, and observe that:

$$\begin{aligned}
\|\mathcal{R}X_1 - \mathcal{R}X_2\| \leq& \|L^* X_1 N[I + (P+X_1)N]^{-1}X_1 L - L^* X_2 N[I + (P+X_2)N]^{-1}X_2 L\| && \text{(F.73)}\\
\leq& \|L\|^2\|X_1 N[I + (P+X_1)N]^{-1}X_1 - X_1 N[I + (P+X_1)N]^{-1}X_2\\
&+ X_1 N[I + (P+X_1)N]^{-1}X_2 - X_2 N[I + (P+X_1)N]^{-1}X_2\\
&+ X_2 N[I + (P+X_1)N]^{-1}X_2 - X_2 N[I + (P+X_2)N]^{-1}X_2\| && \text{(F.74)}\\
\leq& \|L\|^2 2\nu\|N\|\|X_1 - X_2\|\\
&+ \|X_2 N[I + (P+X_1)N]^{-1}(X_2 N - X_1 N)[I + (P+X_2)N]^{-1}X_2\| && \text{(F.75)}\\
=& \|L\|^2 2\nu\|N\|\|X_1 - X_2\|\\
&+ \|X_2 N[I + (P+X_1)N]^{-1}(X_2 - X_1)N[I + (P+X_2)N]^{-1}X_2\| && \text{(F.76)}\\
\leq& \|L\|^2\left\{2\nu\|N\| + \nu^2\|N\|^2\right\}\|X_1 - X_2\| && \text{(F.77)}\\
\leq& 3\|L\|^2\nu\|N\|\|X_1 - X_2\|. && \text{(F.78)}
\end{aligned}$$

Again, let P_X be a shorthand for $P+X$. Having defined for convenience $\mathcal{G}X = F(P_X, \tilde{A}, \tilde{B}) - F(P_X, A, B)$, and using the definition of Φ in (F.26), it holds

$$\|\Phi X_1 - \Phi X_2\| = \|\mathcal{T}^{-1}\| \|\mathcal{G}X_1 - \mathcal{G}X_2 + \mathcal{R}X_2 - \mathcal{R}X_1\| \tag{F.79}$$
$$\leq \frac{\tau(L,\rho)^2}{1-\rho^2} [\|\mathcal{G}X_1 - \mathcal{G}X_2\| + 3\|L\|^2 \nu \|N\| \|X_1 - X_2\|]. \tag{F.80}$$

Now, we can study $\mathcal{G}$. Note that, as in [32] $\|(I+NP_X)^{-1}\| \leq 2\|P_X\|$, since $\nu \leq \frac{1}{2}$. Then,

$$\begin{aligned}
\|\mathcal{G}X_1 - \mathcal{G}X_2\| =& \|A^* P_{X1}(I+NP_{X1})^{-1}A - \tilde{A}^* P_{X1}(I+\tilde{N}P_{X1})^{-1}\tilde{A} \\
& - A^* P_{X2}(I+NP_{X2})^{-1}A + \tilde{A}^* P_{X2}(I+\tilde{N}P_{X2})^{-1}\tilde{A}\| && \text{(F.81)}\\
=& \|A^* P_{X1}(I+NP_{X1})^{-1}\Delta N P_{X1}(I+\tilde{N}P_{X1})^{-1}A - A^* P_{X2}(I+NP_{X2})^{-1}\Delta N P_{X2}(I+\tilde{N}P_{X2})^{-1}A && \text{(F.82)}\\
& - A^* P_{X1}(I+\tilde{N}P_{X1})^{-1}\Delta A + A^* P_{X2}(I+\tilde{N}P_{X2})^{-1}\Delta A && \text{(F.83)}\\
& - \Delta A^* P_{X1}(I+\tilde{N}P_{X1})^{-1}A + \Delta A^* P_{X2}(I+\tilde{N}P_{X2})^{-1}A && \text{(F.84)}\\
& - \Delta A^* P_{X1}(I+\tilde{N}P_{X1})^{-1}\Delta A + \Delta A^* P_{X2}(I+\tilde{N}P_{X2})^{-1}\Delta A\|. && \text{(F.85)}
\end{aligned}$$

To control the difference in (F.82), we can use the following inequality. Note that, by definition, $\|N\| \leq \|B\|^2 \|R^{-1}\|$. Then, leveraging [32, Lemma 7]:

$$\begin{aligned}
&\|A^* P_{X1}(I+NP_{X1})^{-1}\Delta N P_{X1}(I+\tilde{N}P_{X1})^{-1}A - A^* P_{X2}(I+NP_{X2})^{-1}\Delta N P_{X2}(I+\tilde{N}P_{X2})^{-1}A\| \\
&= \|A^* P_{X1}(I+NP_{X1})^{-1}\Delta N P_{X1}(I+\tilde{N}P_{X1})^{-1}A - A^* P_{X1}(I+NP_{X2})^{-1}\Delta N P_{X1}(I+\tilde{N}P_{X1})^{-1}A \\
&\quad + A^* P_{X1}(I+NP_{X2})^{-1}\Delta N P_{X1}(I+\tilde{N}P_{X1})^{-1}A - A^* P_{X2}(I+NP_{X2})^{-1}\Delta N P_{X1}(I+\tilde{N}P_{X1})^{-1}A \\
&\quad + A^* P_{X2}(I+NP_{X2})^{-1}\Delta N P_{X1}(I+\tilde{N}P_{X1})^{-1}A - A^* P_{X2}(I+NP_{X2})^{-1}\Delta N P_{X2}(I+\tilde{N}P_{X1})^{-1}A \\
&\quad + A^* P_{X2}(I+NP_{X2})^{-1}\Delta N P_{X2}(I+\tilde{N}P_{X1})^{-1}A - A^* P_{X2}(I+NP_{X2})^{-1}\Delta N P_{X2}(I+\tilde{N}P_{X2})^{-1}A\| && \text{(F.86)}\\
&\leq \|A\|^2(\|P\|+1)\|\Delta N\| \|P_{X_1}(I+NP_{X1})^{-1} - P_{X_1}(I+NP_{X2})^{-1}\| \\
&\quad + \|A\|^2 \|\Delta N\| (\|P\|+1)^3 \|X_1 - X_2\| \\
&\quad + \|A\|^2 \|\Delta N\| (\|P\|+1)^3 \|X_1 - X_2\| \\
&\quad + \|A\|^2(\|P\|+1)\|\Delta N\| \|P_{X_2}(I+NP_{X1})^{-1} - P_{X_2}(I+NP_{X2})^{-1}\| && \text{(F.87)}\\
&\leq \|A\|^2(\|P\|+1)\|\Delta N\| \|P_{X_1}(I+NP_{X1})^{-1}(I+NP_{X_2} - I - NP_{X_1})(I+NP_{X2})^{-1}\| \\
&\quad + \|A\|^2 \|\Delta N\| (\|P\|+1)^3 \|X_1 - X_2\| \\
&\quad + \|A\|^2 \|\Delta N\| (\|P\|+1)^3 \|X_1 - X_2\| \\
&\quad + \|A\|^2(\|P\|+1)\|\Delta N\| \|P_{X_2}(I+NP_{X_2})^{-1}(I+NP_{X_2} - I - NP_{X_1})(I+NP_{X1})^{-1}\| && \text{(F.88)}\\
&\leq \|A\|^2 2(\|P\|+1)^3 \|\Delta N\| \|N\| \|X_2 - X_1\| \\
&\quad + \|A\|^2 \|\Delta N\| (\|P\|+1)^3 \|X_1 - X_2\| \\
&\quad + \|A\|^2 \|\Delta N\| (\|P\|+1)^3 \|X_1 - X_2\| \\
&\quad + \|A\|^2 2(\|P\|+1)^3 \|\Delta N\| \|N\| \|X_2 - X_1\| && \text{(F.89)}\\
&\leq 12\|A\|^2(\|P\|+1)^3 \|B\|^3 \|R^{-1}\|^2 \|X_2 - X_1\| \epsilon \\
&\quad + 6\|A\|^2 \|B\| \|R^{-1}\| (\|P\|+1)^3 \|X_1 - X_2\| \epsilon. && \text{(F.90)}
\end{aligned}$$

With a similar reasoning, we can compute an upper bound for the difference in (F.83), as follows:

$$\begin{aligned}
&\|A^* P_{X2}(I+\tilde{N}P_{X2})^{-1}\Delta A - A^* P_{X1}(I+\tilde{N}P_{X1})^{-1}\Delta A\| \\
&= \|A^* P_{X2}(I+\tilde{N}P_{X2})^{-1}\Delta A - A^* P_{X2}(I+\tilde{N}P_{X1})^{-1}\Delta A \\
&\quad + A^* P_{X2}(I+\tilde{N}P_{X1})^{-1}\Delta A - A^* P_{X1}(I+\tilde{N}P_{X1})^{-1}\Delta A\| && \text{(F.91)}\\
&\leq 4\|A\|(\|P\|+1)^3 \|X_2 - X_1\| \|B\|^2 \|R^{-1}\| \epsilon \\
&\quad + 2\|A\| \|X_2 - X_1\| (\|P\|+1)^2 \epsilon. && \text{(F.92)}
\end{aligned}$$

The difference in (F.84) is the adjoint of the difference in (F.83), and can be upper bounded by the same factor. An upper bound on the difference in (F.85) can be computed as follows:

$$
\begin{aligned}
\|\Delta A^* P_{X2}(I+&\tilde{N}P_{X2})^{-1}\Delta A - \Delta A^* P_{X1}(I+\tilde{N}P_{X1})^{-1}\Delta A\| \\
&= \|\Delta A^* P_{X2}(I+\tilde{N}P_{X2})^{-1}\Delta A - \Delta A^* P_{X2}(I+\tilde{N}P_{X1})^{-1}\Delta A \\
&\quad + \Delta A^* P_{X2}(I+\tilde{N}P_{X1})^{-1}\Delta A - \Delta A^* P_{X1}(I+\tilde{N}P_{X1})^{-1}\Delta A\| && \text{(F.93)} \\
&\le 4(\|P\|+1)^3\|X_2 - X_1\|\|B\|^2\|R^{-1}\|\epsilon^2 \\
&\quad + 2\|X_2 - X_1\|(\|P\|+1)^2\epsilon^2 && \text{(F.94)} \\
&\le 4(\|P\|+1)^3\|X_2 - X_1\|\|B\|^3\|R^{-1}\|\epsilon \\
&\quad + 2\|X_2 - X_1\|(\|P\|+1)^2\|B\|\epsilon, && \text{(F.95)}
\end{aligned}
$$

where we used on the last line the assumption $\epsilon \le \|B\|$. The overall upper bound has 6 terms in which $\|A\|$ appears at most with degree 2, $\|P\|+1$ with degree 3, $\|B\|$ with degree 3, $\|R^{-1}\|$ with degree 2. With the same reasoning as for the other upper bound, we can write

$$
\|\mathcal{G}X_1 - \mathcal{G}X_2\| \le 12(\|A\|+1)^2(\|P\|+1)^3(\|B\|+1)^3(\|R\|^{-1}+1)^2\|X_1 - X_2\|\epsilon. \tag{F.96}
$$

To conclude,

$$
\begin{aligned}
\|\Phi X_1 - \Phi X_2\| \le 12\frac{\tau(L,\rho)^2}{1-\rho^2}[&(\|A\|+1)^2(\|P\|+1)^3(\|B\|+1)^3(\|R\|^{-1}+1)^2\|X_1 - X_2\|\epsilon \\
&+ \|L\|^2\nu\|N\|\|X_1 - X_2\|]. && \text{(F.97)}
\end{aligned}
$$

□

We are now ready to show that Φ is a contraction from $\mathcal{S}_\nu$ to itself. First, we show that it maps $\mathcal{S}_\nu$ to $\mathcal{S}_\nu$ (Lemma 18). Then, we show that it is a contraction on $\mathcal{S}_\nu$ (Lemma 19).

Lemma 18 *Let us choose ν as a function of ϵ, namely,*

$$
\nu = \min\left\{6\epsilon\frac{\tau(L,\rho)^2}{1-\rho^2}(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)(\|R^{-1}\|+1), \|N\|^{-1}, \frac{1}{2}\right\}.
$$

Moreover, let ϵ be small enough, namely,

$$
\epsilon \le \min\left\{\frac{1}{12}(\|L\|+1)^{-2}\frac{(1-\rho^2)^2}{\tau(L,\rho)^4}(\|A\|+1)^{-2}(\|P\|+1)^{-2}(\|B\|+1)^{-3}(\|R^{-1}\|+1)^{-2}, \|B\|\right\}.
$$

Lastly, let $\sigma_{min}(P) \ge 1$, and $\rho(L) \le \rho < 1$. Note that the condition on the singular values of P can be achieved by rescaling R and Q accordingly, as discussed by [32]. Then, for $X \in \mathcal{S}_\nu$, we have that

$$
\Phi X \in \mathcal{S}_\nu.
$$

Proof In order to prove the lemma, we need to first show that ΦX has norm upper bounded by ν, for $X \in \mathcal{S}_\nu$ and a suitable choice of ν. Then, we need to show that $P + \Phi X$ is self-adjoint and positive semi-definite. Starting from

Proposition 16, we see that:

$$\|\Phi X\| \leq \frac{\tau(L,\rho)^2}{1-\rho^2}\left[\|L\|^2\|N\|\nu^2 + 3\epsilon(\|P\|+1)^2(\|A\|+1)^2(\|R^{-1}\|+1)(\|B\|+1)\right] \tag{F.98}$$

$$\leq \frac{\tau(L,\rho)^2}{1-\rho^2}\left[\|L\|^2\|N\|\left(6\epsilon\frac{\tau(L,\rho)^2}{1-\rho^2}(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)(\|R^{-1}\|+1)\right)^2\right.$$

$$\left. +3\epsilon(\|P\|+1)^2(\|A\|+1)^2(\|R^{-1}\|+1)(\|B\|+1)\right] \tag{F.99}$$

$$= \frac{\tau(L,\rho)^2}{1-\rho^2}\left[\|L\|^2\|N\|12\epsilon\frac{\tau(L,\rho)^4}{(1-\rho^2)^2}(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)(\|R^{-1}\|+1)+1\right]$$

$$\cdot 3\epsilon(\|P\|+1)^2(\|A\|+1)^2(\|R^{-1}\|+1)(\|B\|+1) \tag{F.100}$$

$$\leq \frac{\tau(L,\rho)^2}{1-\rho^2}\left[\|L\|^2 12\epsilon\frac{\tau(L,\rho)^4}{(1-\rho^2)^2}(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)^3(\|R^{-1}\|+1)^2+1\right]$$

$$\cdot 3\epsilon(\|P\|+1)^2(\|A\|+1)^2(\|R^{-1}\|+1)(\|B\|+1). \tag{F.101}$$

For the choice of ϵ in the statement of this lemma, the bracket term is bounded by 2, meaning that:

$$\|\Phi X\| \leq \frac{\tau(L,\rho)^2}{1-\rho^2}6\epsilon(\|P\|+1)^2(\|A\|+1)^2(\|R^{-1}\|+1)(\|B\|+1). \tag{F.102}$$

Note that, for our choice of ϵ,

$$\|\Phi X\| \leq \frac{1}{2}\frac{1-\rho^2}{\tau(L,\rho)^2}(\|R^{-1}\|+1)^{-1}(\|B\|+1)^{-2}(\|L\|+1)^{-2} \tag{F.103}$$

$$\leq \min\left\{\frac{1}{2}, \|N\|^{-1}\right\}, \tag{F.104}$$

since $\|N\| = \|BR^{-1}B^*\| \leq \|B\|^2\|R^{-1}\|$ and $\frac{1-\rho^2}{\tau(L,\rho)^2} \leq 1$. This can be noted by observing that $\tau(L,\rho)$ is a decreasing function of ρ and is equal to 1 for $\rho \geq \|L\|$ (see [32, Page 3]), and $1-\rho^2 < 1$ by our choice of ρ. All in all, this means that $\|\Phi X\| \leq \nu$. To conclude the proof, we can observe that, since P is positive semi-definite and $\sigma_{min}(P) \geq 1$ by assumption, for $a \in \mathcal{H}_1$,

$$\left\langle (P+\Phi X)a, a\right\rangle_{\mathcal{H}_1} = \langle Pa, a\rangle_{\mathcal{H}_1} + \langle \Phi X a, a\rangle_{\mathcal{H}_1} \tag{F.105}$$

$$\geq \langle Pa, a\rangle_{\mathcal{H}_1} - \|\Phi X\|\|a\|^2 \tag{F.106}$$

$$\geq \langle Pa, a\rangle_{\mathcal{H}_1} - \nu\|a\|^2 \tag{F.107}$$

$$= \|P^{1/2}a\|^2 - \nu\|a\|^2 \tag{F.108}$$

$$\geq [\sigma_{min}(P) - \nu]\|a\|^2 \tag{F.109}$$

$$\geq 0. \tag{F.110}$$

Hence, $P+\Phi X \succcurlyeq 0$. Moreover, ΦX is self-adoint because $\mathcal{T}, \mathcal{R}$ and $X \mapsto F(P+X, A, B)$ preserve self-adjointness. □

Lemma 19 *Let ν be defined as in Lemma 18, and $\rho(L) \leq \rho < 1$. Let*

$$12\frac{\tau(L,\rho)^4}{(1-\rho^2)^2}\left[(\|P\|+1)+(\|L\|+1)^2\right](\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)^3(\|R^{-1}\|+1)^2\epsilon < 1.$$

Then, $\exists \eta < 1$ such that, $\forall X_1, X_2 \in \mathcal{S}_\nu$,

$$\|\Phi X_1 - \Phi X_2\| \leq \eta\|X_1 - X_2\|.$$

Proof Let us plug our choice of ν from Lemma 18 in the bound of Proposition 17, to get

$$\begin{aligned}
&12\frac{\tau(L,\rho)^2}{1-\rho^2}\left[(\|A\|+1)^2(\|P\|+1)^3(\|B\|+1)^3(\|R^{-1}\|+1)^2\|X_1-X_2\|\epsilon+\|L\|^2\nu\|N\|\|X_1-X_2\|\right]\\
&\quad=12\frac{\tau(L,\rho)^2}{1-\rho^2}\left[(\|P\|+1)(\|B\|+1)^2(\|R^{-1}\|+1)+\frac{\tau(L,\rho)^2}{1-\rho^2}\|L\|^2\|N\|\right]\\
&\quad\quad\cdot(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)(\|R^{-1}\|+1)\epsilon\|X_1-X_2\| && \text{(F.111)}\\
&\quad\leq 12\frac{\tau(L,\rho)^2}{1-\rho^2}\left[(\|P\|+1)(\|B\|+1)^2(\|R^{-1}\|+1)+\frac{\tau(L,\rho)^2}{1-\rho^2}\|L\|^2(\|B\|+1)^2(\|R^{-1}\|+1)\right]\\
&\quad\quad\cdot(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)(\|R^{-1}\|+1)\epsilon\|X_1-X_2\| && \text{(F.112)}\\
&\quad=12\frac{\tau(L,\rho)^2}{1-\rho^2}\left[(\|P\|+1)+\frac{\tau(L,\rho)^2}{1-\rho^2}\|L\|^2\right]\\
&\quad\quad\cdot(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)^3(\|R^{-1}\|+1)^2\epsilon\|X_1-X_2\| && \text{(F.113)}\\
&\quad\leq 12\frac{\tau(L,\rho)^2}{1-\rho^2}\left[\frac{\tau(L,\rho)^2}{1-\rho^2}(\|P\|+1)+\frac{\tau(L,\rho)^2}{1-\rho^2}\|L\|^2\right]\\
&\quad\quad\cdot(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)^3(\|R^{-1}\|+1)^2\epsilon\|X_1-X_2\| && \text{(F.114)}\\
&\quad\leq 12\frac{\tau(L,\rho)^2}{1-\rho^2}\left[\frac{\tau(L,\rho)^2}{1-\rho^2}(\|P\|+1)+\frac{\tau(L,\rho)^2}{1-\rho^2}(\|L\|+1)^2\right]\\
&\quad\quad\cdot(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)^3(\|R^{-1}\|+1)^2\epsilon\|X_1-X_2\|. && \text{(F.115)}
\end{aligned}$$

Again, in the last step we used the fact that $\tau(L,\rho)$ is a decreasing function of ρ and is equal to 1 for $\rho\geq\|L\|$ (see [32, Page 3]), and $1-\rho^2<1$ by our choice of ρ. The operator Φ is a contraction if ϵ is small enough, namely,

$$12\frac{\tau(L,\rho)^4}{(1-\rho^2)^2}\left[(\|P\|+1)+(\|L\|+1)^2\right](\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)^3(\|R^{-1}\|+1)^2\epsilon<1. \tag{F.116}$$

□

Lemmas 18, 19 imply that, for ϵ small enough, the unique fixed point of Φ is in $\mathcal{S}_\nu$. According to Proposition 11, that fixed point is exactly the error on the Riccati operator due to the Nyström approximation, whose norm scales linearly in ϵ. To conclude the proof of Lemma 5, we can observe that the conditions stated in Lemmas 18, 19 yield an upper bound on ϵ:

$$\epsilon<\frac{1}{12}\frac{1}{(\|L\|+1)^2+(\|P\|+1)}\frac{(1-\rho^2)^2}{\tau(L,\rho)^4}(\|A\|+1)^{-2}(\|P\|+1)^{-2}(\|B\|+1)^{-3}(\|R^{-1}\|+1)^{-2}, \tag{F.117}$$

guaranteeing

$$\|P-\tilde{P}\|\leq 6\epsilon\frac{\tau(L,\rho)^2}{1-\rho^2}(\|A\|+1)^2(\|P\|+1)^2(\|B\|+1)(\|R^{-1}\|+1). \tag{F.118}$$

G Proof of Theorem 7

The proof follows the steps in [32, Theorem 1], adapted to the operator setting considered in this work. We begin by computing an upper bound on the error for the Riccati gain $\|K-\tilde{K}\|$. We then show that when this error is small enough, $\tilde{K}$ stabilizes the system with exact kernel defined in (13). Lastly, we use [17, Lemma 10] to upper bound the error on the objective functions. We begin by re-stating two technical lemmas that appear in [32, Section 2.3] generalized to the case in which the LQR objective contains operators.

Lemma 20 (Lemma 1 of [32]) *Let f_1, f_2 be two μ-strongly convex twice differentiable functions on a Hilbert space. Let $\mathbf{x}_1=\arg\min_{\mathbf{x}} f_1(\mathbf{x})$ and $\mathbf{x}_2=\arg\min_{\mathbf{x}}=f_2(\mathbf{x})$. If $\|\frac{\partial f_1(\mathbf{x})}{\partial\mathbf{x}}|_{\mathbf{x}=\mathbf{x}_2}\|\leq\epsilon$, then $\|\mathbf{x}_1-\mathbf{x}_2\|\leq\frac{\epsilon}{\mu}$.*

Lemma 21 *Let $A_i : \mathcal{H}_1 \to \mathcal{H}_1$, $B_i : \mathbb{R}^{n_u} \to \mathcal{H}_1$, $i = 1, 2$. Let us define $f_i : \mathbb{R}^{n_u} \times \mathcal{H}_1 \to \mathbb{R}$, $f_i(\mathbf{u}, z) = \frac{1}{2}\langle \mathbf{u}, R\mathbf{u}\rangle_{\mathbb{R}^{n_u}} + \frac{1}{2}\langle A_i z + B_i\mathbf{u}, P_i(A_i z + B_i\mathbf{u})\rangle_{\mathcal{H}_1}$ where R, P_i are positive definite operators, $i = 1, 2$, $z \in \mathcal{H}_1$, $\mathbf{u} \in \mathbb{R}^{n_u}$. Let $K_i : \mathcal{H}_1 \to \mathbb{R}^{n_u}$ be the unique operator s.t. $\mathbf{u}_i = \arg\min_{\mathbf{u}} f_i(\mathbf{u}, z) = K_i z$ for any z. Let us define the quantity $\Gamma = 1 + \max\{\|A_1\|, \|B_1\|, \|P_1\|, \|K_1\|\}$, and $\|A_1 - A_2\| \le \epsilon$, $\|B_1 - B_2\| \le \epsilon$, and $\|P_1 - P_2\| \le \epsilon$, for $\epsilon \in [0, 1)$. Then,*

$$\|K_1 - K_2\| \le \frac{3\epsilon\Gamma^3}{\sigma_{min}(R)}. \tag{G.1}$$

Remark 22 *Each f_i in the statement of this lemma is the so-called state-action value function in the language of [17]. It corresponds to an LQR objective when picking an arbitrary control input $\mathbf{u}$ at time 0, and then proceeding with the optimal control policy (static state feedback gain) for the next time steps. It is known that, if at timestep 1 we start using the LQR optimal control, the LQR objective from that timestep onwards is equal to $\frac{1}{2}\langle z_1, Pz_1\rangle_{\mathcal{H}_1}$.*

Remark 23 *The gains K_1 (resp. K_2) in the statement of this lemma are the Riccati gains, obtained with the dynamics given by A_1, B_1 (resp. A_2, B_2).*

Proof The structure of the proof is as follows. We begin by showing that our choice of f_1, f_2 fulfills the hypotheses of 20. Then we apply such a lemma to upper bound the error of interest. Computing the gradient of f_i w.r.t. $\mathbf{u}$ yields:

$$\begin{aligned}
\frac{\partial f_i(\mathbf{u}, z)}{\partial \mathbf{u}} &= R\mathbf{u} + \frac{\partial}{\partial \mathbf{u}}\left\{\frac{1}{2}\langle A_i z + B_i\mathbf{u}, P_i(A_i z + B_i\mathbf{u})\rangle_{\mathcal{H}_1}\right\} && \text{(G.2)}\\
&= R\mathbf{u} + \frac{\partial}{\partial \mathbf{u}}\left\{\frac{1}{2}\langle A_i z, P_i A_i z\rangle_{\mathcal{H}_1} + \langle A_i z, P_i B_i\mathbf{u}\rangle_{\mathcal{H}_1} + \frac{1}{2}\langle B_i\mathbf{u}, P_i B_i\mathbf{u}\rangle_{\mathcal{H}_1}\right\} && \text{(G.3)}\\
&= R\mathbf{u} + B_i^* P_i A_i z + B_i^* P_i^{1/2} P_i^{1/2} B_i\mathbf{u} && \text{(G.4)}\\
&= B_i^* P_i A_i z + (B_i^* P_i B_i + R)\mathbf{u}. && \text{(G.5)}
\end{aligned}$$

After having derived the gradient expression above, we can bound two differences that will appear in the remainder of the proof. In the same way as in [32], we can observe that, having defined $\Lambda = \max\{\|P_1\|, \|B_1\|, \|A_1\|\}$, and considering the assumption that $\epsilon \in [0, 1)$:

$$\begin{aligned}
\|B_1^* P_1 B_1 - B_2^* P_2 B_2\| =& \|(B_1^* - B_2^*)P_1 B_1 + B_2^*(P_1 - P_2)B_1 + B_2^* P_2(B_1 - B_2)\| && \text{(G.6)}\\
=& \|(B_1^* - B_2^*)P_1 B_1 + B_2^*(P_1 - P_2)B_1 + B_2^* P_2(B_1 - B_2)\\
&- B_1^*(P_1 - P_2)B_1 + B_1^*(P_1 - P_2)B_1\\
&- B_1^* P_2(B_1 - B_2) + B_1^* P_2(B_1 - B_2)\| && \text{(G.7)}\\
=& \|(B_1^* - B_2^*)P_1 B_1 + (B_2^* - B_1^*)(P_1 - P_2)B_1 + (B_2^* - B_1^*)P_2(B_1 - B_2)\\
&+ B_1^*(P_1 - P_2)B_1 + B_1^* P_2(B_1 - B_2)\\
&- (B_2^* - B_1^*)P_1(B_1 - B_2) + (B_2^* - B_1^*)P_1(B_1 - B_2)\\
&- B_1^* P_1(B_1 - B_2) + B_1^* P_1(B_1 - B_2)\| && \text{(G.8)}\\
=& \|(B_1^* - B_2^*)P_1 B_1 + (B_2^* - B_1^*)(P_1 - P_2)B_1\\
&+ (B_2^* - B_1^*)(P_2 - P_1)(B_1 - B_2)\\
&+ B_1^*(P_1 - P_2)B_1 + B_1^*(P_2 - P_1)(B_1 - B_2)\\
&+ (B_2^* - B_1^*)P_1(B_1 - B_2)\\
&+ B_1^* P_1(B_1 - B_2)\| && \text{(G.9)}\\
\le& \epsilon\|P_1\|\|B_1\| + \epsilon^2\|B_1\| + \epsilon^3 + \|B_1\|^2\epsilon + \|B_1\|\epsilon^2 + \epsilon^2\|P_1\| + \|B_1\|\|P_1\|\epsilon && \text{(G.10)}\\
\le& \epsilon\Lambda^2 + \epsilon^2\Lambda + \epsilon^3 + \epsilon\Lambda^2 + \epsilon\Lambda + \epsilon^2\Lambda + \Lambda^2\epsilon && \text{(G.11)}\\
\le& \epsilon(3\Lambda^2 + 3\Lambda + 1) && \text{(G.12)}\\
\le& 3\epsilon\Gamma^2. && \text{(G.13)}
\end{aligned}$$

Similarly, we have

$$
\begin{aligned}
\|B_1^*P_1A_1 - B_2^*P_2A_2\| =& \|(B_1^* - B_2^*)P_1A_1 + B_2^*(P_1 - P_2)A_1 + B_2^*P_2(A_1 - A_2)\| && \text{(G.14)}\\
=& \|(B_1^* - B_2^*)P_1A_1 + B_2^*(P_1 - P_2)A_1 + B_2^*P_2(A_1 - A_2) \\
& - B_1^*(P_1 - P_2)A_1 + B_1^*(P_1 - P_2)A_1 \\
& - B_1^*P_2(A_1 - A_2) + B_1^*P_2(A_1 - A_2)\| && \text{(G.15)}\\
=& \|(B_1^* - B_2^*)P_1A_1 + (B_2^* - B_1^*)(P_1 - P_2)A_1 + (B_2^* - B_1^*)P_2(A_1 - A_2) \\
& + B_1^*(P_1 - P_2)A_1 + B_1^*P_2(A_1 - A_2) \\
& - (B_2^* - B_1^*)P_1(A_1 - A_2) + (B_2^* - B_1^*)P_1(A_1 - A_2) \\
& - B_1^*P_1(A_1 - A_2) + B_1^*P_1(A_1 - A_2)\| && \text{(G.16)}\\
=& \|(B_1^* - B_2^*)P_1A_1 + (B_2^* - B_1^*)(P_1 - P_2)A_1 \\
& + (B_2^* - B_1^*)(P_2 - P_1)(A_1 - A_2) \\
& + B_1^*(P_1 - P_2)A_1 + B_1^*(P_2 - P_1)(A_1 - A_2) \\
& + (B_2^* - B_1^*)P_1(A_1 - A_2) \\
& + B_1^*P_1(A_1 - A_2)\| && \text{(G.17)}\\
\le& \epsilon\|P_1\|\|A_1\| + \epsilon^2\|A_1\| + \epsilon^3 \\
& + \|B_1\|\|A_1\|\epsilon + \|B_1\|\epsilon^2 + \epsilon^2\|P_1\| + \|B_1\|\|P_1\|\epsilon && \text{(G.18)}\\
\le& \epsilon\Lambda^2 + \epsilon^2\Lambda + \epsilon^3 + \epsilon\Lambda^2 + \epsilon\Lambda + \epsilon^2\Lambda + \Lambda^2\epsilon && \text{(G.19)}\\
\le& \epsilon(3\Lambda^2 + 3\Lambda + 1) && \text{(G.20)}\\
\le& 3\epsilon\Gamma^2. && \text{(G.21)}
\end{aligned}
$$

We can then consider the following difference. For arbitrary $\mathbf{u}$, z, by using the gradient expression in (G.5) and the triangle inequality,

$$\left\|\frac{\partial f_1(\mathbf{u}, z)}{\partial \mathbf{u}} - \frac{\partial f_2(\mathbf{u}, z)}{\partial \mathbf{u}}\right\| \le 3\epsilon\Gamma^2(\|\mathbf{u}\| + \|z\|). \tag{G.22}$$

In particular, for any z s.t. $\|z\| \le 1$, this inequality applied to $\mathbf{u}_1 = K_1 z$ yields by the first-order optimality condition of $\mathbf{u}_1$:

$$\left\|\frac{\partial f_1(\mathbf{u}, z)}{\partial \mathbf{u}}|_{\mathbf{u}=\mathbf{u}_1} - \frac{\partial f_2(\mathbf{u}, z)}{\partial \mathbf{u}}|_{\mathbf{u}=\mathbf{u}_1}\right\| = \left\|\frac{\partial f_2(\mathbf{u}, z)}{\partial \mathbf{u}}|_{\mathbf{u}=\mathbf{u}_1}\right\| \le 3\epsilon\Gamma^2(\|\mathbf{u}_1\| + 1). \tag{G.23}$$

Hence, still for any z with $\|z\| \le 1$, applying Lemma 20 to the functions f_1, f_2 defined in the statement, which are μ-strongly convex with $\mu \ge \sigma_{\min}(R)$, we get

$$
\begin{aligned}
\|K_1 z - K_2 z\| &\le \sigma_{\min}(R)^{-1} 3\epsilon\Gamma^2(\|\mathbf{u}_1\| + 1) && \text{(G.24)}\\
&\le \sigma_{\min}(R)^{-1} 3\epsilon\Gamma^2(\|K_1\|\|z\| + 1) && \text{(G.25)}\\
&\le \sigma_{\min}(R)^{-1} 3\epsilon\Gamma^2(\|K_1\| + 1) && \text{(G.26)}\\
&\le \sigma_{\min}(R)^{-1} 3\epsilon\Gamma^3. && \text{(G.27)}
\end{aligned}
$$

This yields the claimed result given that

$$\|K_1 - K_2\| = \sup_{\|z\| \le 1} \|(K_1 - K_2)z\|. \tag{G.28}$$

□

Having proved the lemma above, we now specialize it to the case of exact vs. Nyström kernel-based gains.

Proposition 24 *Let $\epsilon > 0$ be s.t. $\|A - \tilde{A}\| \le \epsilon$, $\|B - \tilde{B}\| \le \epsilon$, and $\|P - \tilde{P}\| \le g(\epsilon)$ with $g(\epsilon) \ge \epsilon$. Then, assuming R and Q are positive definite and $\sigma_{min}(R) \ge 1$, we have*

$$\|\tilde{K} - K\| \le 3\Gamma^3 g(\epsilon). \tag{G.29}$$

Moreover, choose ρ such that $\rho(L) \leq \rho < 1$. If ϵ is small enough s.t. the r.h.s. of (G.29) *is smaller than $\frac{1-\rho}{2\tau(L,\rho)}$, then*

$$\tau\left(A + B\tilde{K}, \frac{1+\rho}{2}\right) \leq \tau(L, \rho). \tag{G.30}$$

Proof The first part of the proposition is a corollary of Lemma 21 based on the condition $g(\epsilon) > \epsilon$ and $\sigma_{\min}(R) \geq 1$. The second part of the proposition follows from applying the first inequality of [32, Lemma 5] with $M = A + BK$ and $\Delta = B(\tilde{K} - K)$, which guarantees that, if

$$\|\tilde{K} - K\| \leq \frac{1-\rho}{2\tau(L,\rho)\|B\|}, \tag{G.31}$$

then

$$\begin{aligned}
\|(A + B\tilde{K})^k\| &= \|(A + BK + B\tilde{K} - BK)^k\| && \text{(G.32)}\\
&= \|(A + BK + B(\tilde{K} - K))^k\| && \text{(G.33)}\\
&\leq \tau(L,\rho)\left(\frac{1-\rho}{2} + \rho\right)^k && \text{(G.34)}\\
&= \tau(L,\rho)\left(\frac{1+\rho}{2}\right)^k. && \text{(G.35)}
\end{aligned}$$

□

Now we are ready to conclude the proof. Let us consider the following decomposition of the error on the LQR objective, which corresponds to the well known *performance difference lemma* in the reinforcement learning literature [24]. The following lemma is a restatement of [17, Lemma 10], with our notation and setup (i.e., with operators).

Lemma 25 *Let $\mathcal{J}$ and $\hat{\mathcal{J}}$ be as of* (38) *and* (39)*. Then, the following decomposition holds:*

$$\hat{\mathcal{J}} - \mathcal{J} = \lim_{T\to\infty} \sum_{i=0}^{T} \left\langle (\tilde{K} - K)z_i, (R + B^*PB)(\tilde{K} - K)z_i \right\rangle_{\mathcal{H}_1} \tag{G.36}$$

Proof We begin by using the same telescoping argument applied by [17]. Let A and B be defined as in (14), (15). Let us define, for an arbitrary initial condition $\hat{z}_0 \in \mathcal{H}_1$:

$$\hat{z}_{i+1} = A\hat{z}_i + B\tilde{\mathbf{u}}_i^{\text{opt}}, \tilde{\mathbf{u}}_i^{\text{opt}} = \tilde{K}\hat{z}_i, i \geq 0. \tag{G.37}$$

Similarly to (39), let us define, for an initial condition, at timestep i, $\hat{z}_i \in \mathcal{H}_1$:

$$\begin{aligned}
\hat{z}'_{i,j+1} &= A\hat{z}'_{i,j} + B\mathbf{u}_j^{\text{opt}}, \mathbf{u}_j^{\text{opt}} = K\hat{z}'_j, j \geq i, && \text{(G.38)}\\
\hat{z}'_{i,i} &= \hat{z}_i. && \text{(G.39)}
\end{aligned}$$

We can decompose the error of interest as follows, noting that, by the definition in (38),

$$\mathcal{J} = \sum_{j=0}^{T}\left(\left\langle \hat{z}'_{0,j}, Q\hat{z}'_{0,j}\right\rangle_{\mathcal{H}_1} + \left\langle \mathbf{u}_j^{\text{opt}}, R\mathbf{u}_j^{\text{opt}}\right\rangle_{\mathbb{R}^{n_u}}\right): \tag{G.40}$$

$$
\begin{aligned}
\hat{\mathcal{J}} - \mathcal{J} &= \lim_{T\to\infty} \left\{ \sum_{i=0}^{T} \left[\langle \hat{z}_i, Q\hat{z}_i \rangle_{\mathcal{H}_1} + \left\langle \tilde{\mathbf{u}}_i^{\text{opt}}, R\tilde{\mathbf{u}}_i^{\text{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right] - \mathcal{J} \right\} && \text{(G.41)} \\
&= \lim_{T\to\infty} \left\{ \sum_{i=0}^{T} \left[\langle \hat{z}_i, Q\hat{z}_i \rangle_{\mathcal{H}_1} + \left\langle \tilde{\mathbf{u}}_i^{\text{opt}}, R\tilde{\mathbf{u}}_i^{\text{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right. \right. \\
&\quad + \sum_{j=i}^{T} \left(\left\langle \hat{z}'_j, Q\hat{z}'_j \right\rangle_{\mathcal{H}_1} + \left\langle \mathbf{u}_j^{\text{opt}}, R\mathbf{u}_j^{\text{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right) \\
&\quad \left. \left. - \sum_{j=i}^{T} \left(\left\langle \hat{z}'_j, Q\hat{z}'_j \right\rangle_{\mathcal{H}_1} + \left\langle \mathbf{u}_j^{\text{opt}}, R\mathbf{u}_j^{\text{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right) \right] - \mathcal{J} \right\} && \text{(G.42)} \\
&= \lim_{T\to\infty} \left\{ \sum_{i=0}^{T} \left[\langle \hat{z}_i, Q\hat{z}_i \rangle_{\mathcal{H}_1} + \left\langle \tilde{\mathbf{u}}_i^{\text{opt}}, R\tilde{\mathbf{u}}_i^{\text{opt}} \right\rangle_{\mathbb{R}^{n_u}} + \sum_{j=i+1}^{T} \left(\left\langle \hat{z}'_{i,j}, Q\hat{z}'_{i,j} \right\rangle_{\mathcal{H}_1} + \left\langle \mathbf{u}_j^{\text{opt}}, R\mathbf{u}_j^{\text{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right) \right. \right. \\
&\quad \left. \left. - \sum_{j=i}^{T} \left(\left\langle \hat{z}'_{i,j}, Q\hat{z}'_{i,j} \right\rangle_{\mathcal{H}_1} + \left\langle \mathbf{u}_j^{\text{opt}}, R\mathbf{u}_j^{\text{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right) \right] \right\}. && \text{(G.43)}
\end{aligned}
$$

Now we can consider a single addend in the outer-most sum. Consider that by definition

$$
\tilde{\mathbf{u}}_i^{\text{opt}} = \tilde{K}\hat{z}_i. \tag{G.44}
$$

Moreover, note that

$$
\begin{aligned}
\hat{z}'_{i,i+1} &= A\hat{z}'_{i,i} + B\tilde{K}\hat{z}'_{i,i} && \text{(G.45)} \\
&= A\hat{z}_i + B\tilde{K}\hat{z}_i. && \text{(G.46)}
\end{aligned}
$$

Lastly, note that, subject to the Riccati-optimal state feedback law $\mathbf{u}^{\text{opt}}$, the following simplification is allowed, for $T \to \infty$:

$$
\begin{aligned}
\sum_{j=i+1}^{T} \left(\left\langle \hat{z}'_{i,j}, Q\hat{z}'_{i,j} \right\rangle_{\mathcal{H}_1} + \left\langle \mathbf{u}_j^{\text{opt}}, R\mathbf{u}_j^{\text{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right) &= \left\langle \hat{z}'_{i,i+1}, P\hat{z}'_{i,i+1} \right\rangle_{\mathcal{H}_1} && \text{(G.47)} \\
&= \left\langle (A + B\tilde{K})\hat{z}_i, P(A + B\tilde{K})\hat{z}_i \right\rangle_{\mathcal{H}_1}. && \text{(G.48)}
\end{aligned}
$$

Similarly, observe that, for $T \to \infty$,

$$
\begin{aligned}
\sum_{j=i}^{T} \left(\left\langle \hat{z}'_{i,j}, Q\hat{z}'_{i,j} \right\rangle_{\mathcal{H}_1} + \left\langle \mathbf{u}_j^{\text{opt}}, R\mathbf{u}_j^{\text{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right) &= \left\langle \hat{z}'_{i,i}, P\hat{z}'_{i,i} \right\rangle_{\mathcal{H}_1} && \text{(G.49)} \\
&= \langle \hat{z}_i, P\hat{z}_i \rangle_{\mathcal{H}_1}. && \text{(G.50)}
\end{aligned}
$$

Hence, we can rewrite the addend as the following, for $T \to \infty$:

$$
\begin{aligned}
&\langle \hat{z}_i, Q\hat{z}_i \rangle_{\mathcal{H}_1} + \left\langle \tilde{\mathbf{u}}_i^{\mathrm{opt}}, R\tilde{\mathbf{u}}_i^{\mathrm{opt}} \right\rangle_{\mathbb{R}^{n_u}} + \sum_{j=i+1}^{T} \left(\left\langle \hat{z}'_{i,j}, Q\hat{z}'_{i,j} \right\rangle_{\mathcal{H}_1} + \left\langle \mathbf{u}_j^{\mathrm{opt}}, R\mathbf{u}_j^{\mathrm{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right) \\
&\qquad - \sum_{j=i}^{T} \left(\left\langle \hat{z}'_{i,j}, Q\hat{z}'_{i,j} \right\rangle_{\mathcal{H}_1} + \left\langle \mathbf{u}_j^{\mathrm{opt}}, R\mathbf{u}_j^{\mathrm{opt}} \right\rangle_{\mathbb{R}^{n_u}} \right) \\
&\quad = \left\langle \hat{z}_i, (Q + \tilde{K}^* R\tilde{K})\hat{z}_i \right\rangle_{\mathcal{H}_1} + \left\langle (A + B\tilde{K})\hat{z}_i, P(A + B\tilde{K})\hat{z}_i \right\rangle_{\mathcal{H}_1} - \langle \hat{z}_i, P\hat{z}_i \rangle_{\mathcal{H}_1} &\text{(G.51)} \\
&\quad = \left\langle \hat{z}_i, (Q + (\tilde{K} - K + K)^* R(\tilde{K} - K + K)\hat{z}_i \right\rangle_{\mathcal{H}_1} \\
&\qquad + \left\langle (A + B\tilde{K} - BK + BK)\hat{z}_i, P(A + B\tilde{K} - BK + BK)\hat{z}_i \right\rangle_{\mathcal{H}_1} - \langle \hat{z}_i, P\hat{z}_i \rangle_{\mathcal{H}_1} &\text{(G.52)} \\
&\quad = \langle \hat{z}_i, Q\hat{z}_i \rangle_{\mathcal{H}_1} + \left\langle \hat{z}_i, (\tilde{K} - K)^* R(\tilde{K} - K)\hat{z}_i \right\rangle_{\mathcal{H}_1} + \left\langle \hat{z}_i, K^* R(\tilde{K} - K)\hat{z}_i \right\rangle_{\mathcal{H}_1} \\
&\qquad + \left\langle \hat{z}_i, (\tilde{K} - K)^* RK\hat{z}_i \right\rangle_{\mathcal{H}_1} + \langle \hat{z}_i, K^* RK\hat{z}_i \rangle_{\mathcal{H}_1} \\
&\qquad + \left\langle B(\tilde{K} - K)\hat{z}_i, PB(\tilde{K} - K)\hat{z}_i \right\rangle_{\mathcal{H}_1} + \langle (A + BK)\hat{z}_i, P(A + BK)\hat{z}_i \rangle_{\mathcal{H}_1} \\
&\qquad + \left\langle B(\tilde{K} - K)\hat{z}_i, P(A + BK)\hat{z}_i \right\rangle_{\mathcal{H}_1} + \left\langle (A + BK)\hat{z}_i, PB(\tilde{K} - K)\hat{z}_i \right\rangle_{\mathcal{H}_1} \\
&\qquad - \langle \hat{z}_i, P\hat{z}_i \rangle_{\mathcal{H}_1} &\text{(G.53)} \\
&\quad = \left\langle \hat{z}_i, (\tilde{K} - K)^* (R + B^* PB)(\tilde{K} - K)\hat{z}_i \right\rangle_{\mathcal{H}_1} \\
&\qquad + 2 \left\langle \hat{z}_i, (\tilde{K} - K)^* [RK + B^* P(A + BK)]\hat{z}_i \right\rangle_{\mathcal{H}_1} &\text{(G.54)} \\
&\quad = \left\langle \hat{z}_i, (\tilde{K} - K)^* (R + B^* PB)(\tilde{K} - K)\hat{z}_i \right\rangle_{\mathcal{H}_1} \\
&\qquad + 2 \left\langle \hat{z}_i, (\tilde{K} - K)^* [(R + B^* PB)K + B^* PA]\hat{z}_i \right\rangle_{\mathcal{H}_1} &\text{(G.55)} \\
&\quad = \left\langle \hat{z}_i, (\tilde{K} - K)^* (R + B^* PB)(\tilde{K} - K)\hat{z}_i \right\rangle_{\mathcal{H}_1} \\
&\qquad + 2 \left\langle \hat{z}_i, (\tilde{K} - K)^* [(R + B^* PB)(-1)(R + B^* PB)^{-1} B^* PA + B^* PA]\hat{z}_i \right\rangle_{\mathcal{H}_1} &\text{(G.56)} \\
&\quad = \left\langle \hat{z}_i, (\tilde{K} - K)^* (R + B^* PB)(\tilde{K} - K)\hat{z}_i \right\rangle_{\mathcal{H}_1}. &\text{(G.57)}
\end{aligned}
$$

□

To conclude the proof of the theorem, we can apply Cauchy-Schwarz inequality, Lemma 21, and the fact that, since $\sigma_{\min}(R) \geq 1$, it holds

$$
\begin{aligned}
\|R + B^* PB\| &\leq \|R\| + \|B^* PB\| &\text{(G.58)} \\
&\leq \|R\| + \|R\| \|B^* PB\| &\text{(G.59)} \\
&\leq \sigma_{\max}(R)\Gamma^3, &\text{(G.60)}
\end{aligned}
$$

where we recall that $\Gamma := 1 + \max(\|A\|, \|P\|, \|K\|, \|B\|)$ in the statement of Theorem 7. We have that

$$\hat{\mathcal{J}} - \mathcal{J} = \lim_{T\to\infty} \sum_{i=0}^{T} \left\langle \hat{z}_i, (\tilde{K} - K)^*(R + B^*PB)(\tilde{K} - K)\hat{z}_i \right\rangle_{\mathcal{H}_1} \tag{G.61}$$

$$\leq \lim_{T\to\infty} \sum_{i=0}^{T} \|\hat{z}_i\|^2_{\mathcal{H}_1} \|\tilde{K} - K\|^2 \|R + B^*PB\| \tag{G.62}$$

$$\leq 9\sigma_{\max}(R)\Gamma^9 g(\epsilon)^2 \lim_{T\to\infty} \sum_{i=0}^{T} \|\hat{z}_i\|^2_{\mathcal{H}_1} \tag{G.63}$$

$$= 9\sigma_{\max}(R)\Gamma^9 g(\epsilon)^2 \lim_{T\to\infty} \sum_{i=0}^{T} \|(A + B\tilde{K})^i z_0\|^2_{\mathcal{H}_1} \tag{G.64}$$

$$\leq 9\sigma_{\max}(R)\Gamma^9 g(\epsilon)^2 \|z_0\|^2_{\mathcal{H}_1} \lim_{T\to\infty} \sum_{i=0}^{T} \|(A + B\tilde{K})^i\|^2 \tag{G.65}$$

$$\leq 9\sigma_{\max}(R)\Gamma^9 g(\epsilon)^2 \|z_0\|^2_{\mathcal{H}_1} \lim_{T\to\infty} \sum_{i=0}^{T} \tau(A + BK, \rho)^2 \left(\frac{1+\rho}{2}\right)^{2i} \tag{G.66}$$

$$\leq 9\sigma_{\max}(R)\Gamma^9 g(\epsilon)^2 \|z_0\|^2_{\mathcal{H}_1} \tau(A + BK, \rho)^2 \frac{1}{1 - \left(\frac{1+\rho}{2}\right)^2} \tag{G.67}$$

$$\leq 36\sigma_{\max}(R)\Gamma^9 g(\epsilon)^2 \|z_0\|^2_{\mathcal{H}_1} \frac{\tau(A + BK, \rho)^2}{1 - \rho^2}. \tag{G.68}$$

To conclude, note that, according to Assumption 1,

$$\|z_0\|^2_{\mathcal{H}_1} = \|k(x_0, \cdot)\|^2_{\mathcal{H}_1} \tag{G.69}$$

$$= \langle k(x_0, \cdot), k(x_0, \cdot)\rangle_{\mathcal{H}_1} \tag{G.70}$$

$$= k(x_0, x_0) \tag{G.71}$$

$$\leq \kappa^2. \tag{G.72}$$