Online Supplementary Materials

A Tower Property

We return to the calculation of (2.65) and (2.66). Recall that we consider the case when F_t^P has rank $m < n$ and F_t^E has rank $k < n$ for all $t = 0, 1, \ldots, T-1$.

Matching LHS with RHS. We can see the first term (quadratic in $\widehat{x}_{T-1}^P$) on the LHS and RHS are the same. Thus we are left to show that the constants terms in (2.65) and (2.66) are the same. We prove this in the following steps.

Step 1: Expanding L_1 and L_2 on the LHS. We first calculate the three terms in (2.65) that involves L_1 and L_2 using (2.63) and (2.64). The first term is given by

$$
\begin{aligned}
\mathrm{Tr}\left(L_1^\top Q_T^P L_1 \widehat{\Sigma}_{T-1}^E\right) \quad = \quad & \underbrace{\mathrm{Tr}\left((B_{T-1}^E F_{T-1}^E)^\top Q_T^P B_{T-1}^E F_{T-1}^E \widehat{\Sigma}_{T-1}^E\right)}_{(1a)} \\
& + \underbrace{\mathrm{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P \widehat{\Sigma}_{T-1}^E\right)}_{(1b)} \\
& + 2\,\mathrm{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P B_{T-1}^E F_{T-1}^E \widehat{\Sigma}_{T-1}^E\right)
\end{aligned}
$$

For the second term, we first note that by adding and subtracting A_{T-1}, we obtain

$$
\begin{aligned}
L_2 \quad &= \quad -(I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P - B_{T-1}^E F_{T-1}^E - K_T^P H_T^P A_{T-1} \\
&= \quad -(I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P - B_{T-1}^E F_{T-1}^E + (I - K_T^P H_T^P) A_{T-1} - A_{T-1}
\end{aligned}
$$

Then, the second term is given by

$$
\begin{aligned}
\mathrm{Tr}\left(L_2^\top Q_T^P L_2 \widehat{\Sigma}_{T-1}^P\right) \quad = \quad & \underbrace{\mathrm{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P \widehat{\Sigma}_{T-1}^P\right)}_{(2a)} \\
& + \underbrace{\mathrm{Tr}\left((A_{T-1} + B_{T-1}^E F_{T-1}^E)^\top Q_T^P (A_{T-1} + B_{T-1}^E F_{T-1}^E) \widehat{\Sigma}_{T-1}^P\right)}_{(2b)} \\
& + \underbrace{\mathrm{Tr}\left(((I - K_T^P H_T^P) A_{T-1})^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \widehat{\Sigma}_{T-1}^P\right)}_{(2c)} \\
& + 2\,\mathrm{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P (A_{T-1} + B_{T-1}^E F_{T-1}^E) \widehat{\Sigma}_{T-1}^P\right) \\
& \underbrace{- 2\,\mathrm{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \widehat{\Sigma}_{T-1}^P\right)}_{(2d)} \\
& - 2\,\mathrm{Tr}\left((A_{T-1} + B_{T-1}^E F_{T-1}^E)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \widehat{\Sigma}_{T-1}^P\right)
\end{aligned}
$$

Similarly the third term is given by

$$
\begin{aligned}
2\,\mathrm{Tr}\left(L_1^\top Q_T^P L_2 \widetilde{\Sigma}_{T-1}^{(P,E)}\right) \quad = \quad & \underbrace{-2\,\mathrm{Tr}\left((B_{T-1}^E F_{T-1}^E)^\top Q_T^P (A_{T-1} + B_{T-1}^E F_{T-1}^E) \widetilde{\Sigma}_{T-1}^{(P,E)}\right)}_{(3a)} \\
& - 2\,\mathrm{Tr}\left((B_{T-1}^E F_{T-1}^E)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P \widetilde{\Sigma}_{T-1}^{(P,E)}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+2\,\mathrm{Tr}\left((B^E_{T-1}F^E_{T-1})^\top Q^P_T(I-K^P_TH^P_T)A_{T-1}\widetilde{\Sigma}^{(P,E)}_{T-1}\right)\\
&-2\,\mathrm{Tr}\left(((I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1})^\top Q^P_T(A_{T-1}+B^E_{T-1}F^E_{T-1})\widetilde{\Sigma}^{(P,E)}_{T-1}\right)\\
&-\underbrace{2\,\mathrm{Tr}\left(((I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1})^\top Q^P_T(I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1}\widetilde{\Sigma}^{(P,E)}_{T-1}\right)}_{(3b)}\\
&+\underbrace{2\,\mathrm{Tr}\left(((I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1})^\top Q^P_T(I-K^P_TH^P_T)A_{T-1}\widetilde{\Sigma}^{(P,E)}_{T-1}\right)}_{(3c)}
\end{aligned}
$$

Now we observe that all the constant terms on RHS except $\mathrm{Tr}(\Gamma^\top_{T-1}Q_T\Gamma_{T-1}W)$ can be cancelled with (2b), (1a), and (3a) on LHS, that is

$$
\text{Constant terms on RHS} = (2b)+(1a)+(3a)+\mathrm{Tr}(\Gamma^\top_{T-1}Q_T\Gamma_{T-1}W). \tag{A.1}
$$

Therefore, our goal now is to show that $\mathrm{Tr}(\Gamma^\top_{T-1}Q_T\Gamma_{T-1}W)$ equals the rest of terms on LHS. Before matching the terms, we first merge and simplify some of the terms on LHS.

Step 2: Merging quadratic terms of Π^P_{T-1} on LHS. Recall that Π^P_{T-1} is defined as $\Pi^P_{T-1} = \big(\widehat{\Sigma}^P_{t-1}-\widetilde{\Sigma}^{(P,E)}_{t-1}\big)\big(\widehat{\Sigma}^{(P,E)}_{t-1}\big)^{-1}$. Thus we have

$$
\Pi^P_{T-1}\widehat{\Sigma}^{(P,E)}_{t-1} = \widehat{\Sigma}^P_{t-1}-\widetilde{\Sigma}^{(P,E)}_{t-1}. \tag{A.2}
$$

This provides a way to reduce the order of Π^P_{T-1} on LHS. Collecting terms which are quadratic in Π^P_{T-1} we obtain

$$
\begin{aligned}
(1b)+(2a)+(3b) &= \mathrm{Tr}\Big(\big((I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1}\big)^\top Q^P_T(I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1}\cdot\\
&\qquad \big(\widehat{\Sigma}^P_{T-1}+\widehat{\Sigma}^E_{T-1}-\widetilde{\Sigma}^{(P,E)}_{T-1}-\widetilde{\Sigma}^{(E,P)}_{T-1}\big)\Big)\\
&= \mathrm{Tr}\Big(\big((I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1}\big)^\top Q^P_T(I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1}\widehat{\Sigma}^{(P,E)}_{T-1}\Big), \qquad (A.3)\\
&= \mathrm{Tr}\Big(\big((I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1}\big)^\top Q^P_T(I-K^P_TH^P_T)A_{T-1}\big(\widehat{\Sigma}^P_{t-1}-\widetilde{\Sigma}^{(P,E)}_{t-1}\big)\Big), \qquad (A.4)
\end{aligned}
$$

where (A.3) holds by definition of $\widehat{\Sigma}^{(P,E)}_{T-1}$ given in (2.38), and (A.4) holds by (A.2). Now we observe that the term in (A.4) will be cancelled with a half of the sum $(2d)+(3c)$, since

$$
(2d)+(3c) = 2\,\mathrm{Tr}\left(((I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1})^\top Q^P_T(I-K^P_TH^P_T)A_{T-1}\big(\widetilde{\Sigma}^{(P,E)}_{T-1}-\widehat{\Sigma}^P_{T-1}\big)\right).
$$

Therefore in summary, we have

$$
(1b)+(2a)+(3b)+(2d)+(3c) = \mathrm{Tr}\left(((I-K^P_TH^P_T)A_{T-1}\Pi^P_{T-1})^\top Q^P_T(I-K^P_TH^P_T)A_{T-1}\big(\widetilde{\Sigma}^{(P,E)}_{T-1}-\widehat{\Sigma}^P_{T-1}\big)\right).
$$

Step 3: Merging the last three terms on LHS. Going back to (2.65), we now consider the last three constant terms that do not involve L_1 and L_2. We now add the first two terms to (2c), which leads to

$$
\begin{aligned}
&\mathrm{Tr}((K^P_T)^\top Q^P_TK^P_TG^P)+\mathrm{Tr}(\Gamma^\top_{T-1}(H^P_T)^\top(K^P_T)^\top Q^P_TK^P_TH^P_T\Gamma_{T-1}W)+(2c)\\
=\;&\mathrm{Tr}\Big((K^P_T)^\top Q^P_TK^P_T\big(H^P_T\Gamma_{T-1}W\Gamma^\top_{T-1}(H^P_T)^\top+G^P\big)\Big)\\
&+\mathrm{Tr}\left(((I-K^P_TH^P_T)A_{T-1})^\top Q^P_T(I-K^P_TH^P_T)A_{T-1}\widehat{\Sigma}^P_{T-1}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \operatorname{Tr}\Big((K_T^P)^\top Q_T^P K_T^P\big(H_T^P\Gamma_{T-1}W\Gamma_{T-1}^\top(H_T^P)^\top + G^P + H_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P A_{T-1}^\top (H_T^P)^\top\big)\Big)\\
&\quad -2\operatorname{Tr}(A_{T-1}^\top Q_T^P K_T^P H_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P) + \operatorname{Tr}(A_{T-1}^\top Q_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P)
\end{aligned} \tag{A.5}
$$

To simply the first term in the above sum, we use the following two facts.

Fact 1. We first note that

$$
\big(\widehat{\Sigma}_T^P\big)^- = A_{T-1}(\widehat{\Sigma}_{T-1}^P)^+ A_{T-1}^\top + \Gamma_{T-1}W\Gamma_{T-1}^\top \tag{A.6}
$$

$$
= A_{T-1}\Big(\widehat{\Sigma}_{T-1}^P - (\widehat{\Sigma}_{T-1}^P - \widetilde{\Sigma}_{T-1}^{(P,E)})(\widehat{\Sigma}_{t-1}^{(P,E)})^{-1}\big(\widehat{\Sigma}_{t-1}^P - \widetilde{\Sigma}_{T-1}^{(P,E)}\big)^\top\Big)A_{T-1}^\top + \Gamma_{T-1}W\Gamma_{T-1}^\top, \tag{A.7}
$$

where (A.6) holds by (2.21e) and (A.7) holds by (2.21c). Thus by rearranging terms in (A.6) and since $\Pi_{T-1}^P = \big(\widehat{\Sigma}_{t-1}^P - \widetilde{\Sigma}_{t-1}^{(P,E)}\big)\big(\widehat{\Sigma}_{t-1}^{(P,E)}\big)^{-1}$, we have

$$
A_{T-1}\widehat{\Sigma}_{t-1}^P A_{T-1}^\top + \Gamma_{T-1}W\Gamma_{T-1}^\top = \big(\widehat{\Sigma}_T^P\big)^- + A_{T-1}(\widehat{\Sigma}_{T-1}^P - \widetilde{\Sigma}_{T-1}^{(P,E)})(\Pi_{T-1}^P)^\top A_{T-1}^\top. \tag{A.8}
$$

Fact 2. By definition of K_t^i in (2.21f), we have

$$
K_T^P\left(H_T^P\big(\widehat{\Sigma}_T^P\big)^-(H_T^P)^\top + G^P\right) = \big(\widehat{\Sigma}_T^P\big)^-(H_T^P)^\top.
$$

Therefore,

$$
(K_T^P)^\top Q_T^P K_T^P\left(H_T^P\big(\widehat{\Sigma}_T^P\big)^-(H_T^P)^\top + G^P\right) = (K_T^P)^\top Q_T^P\big(\widehat{\Sigma}_T^P\big)^-(H_T^P)^\top. \tag{A.9}
$$

Having the above two facts, we can plug in (A.8) and (A.9) into (A.5) to get

$$
\begin{aligned}
&\operatorname{Tr}((K_T^P)^\top Q_T^P K_T^P G^P) + \operatorname{Tr}(\Gamma_{T-1}^\top (H_T^P)^\top (K_T^P)^\top Q_T^P K_T^P H_T^P \Gamma_{T-1}W) + (2c)\\
= &\operatorname{Tr}\Big((K_T^P)^\top Q_T^P K_T^P\big(H_T^P\big((\widehat{\Sigma}_T^i)^- + A_{T-1}(\widehat{\Sigma}_{T-1}^P - \widetilde{\Sigma}_{T-1}^{(P,E)})(\Pi_{T-1}^P)^\top A_{T-1}^\top\big)(H_T^P)^\top + G^P\big)\Big)\\
&-2\operatorname{Tr}(A_{T-1}^\top Q_T^P K_T^P H_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P) + \operatorname{Tr}(A_{T-1}^\top Q_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P)\\
= &\operatorname{Tr}\Big((K_T^P)^\top Q_T^P\big(\widehat{\Sigma}_T^P\big)^-(H_T^P)^\top\Big) + \operatorname{Tr}\Big((K_T^P)^\top Q_T^P K_T^P H_T^P A_{T-1}(\widehat{\Sigma}_{T-1}^P - \widetilde{\Sigma}_{T-1}^{(P,E)})(\Pi_{T-1}^P)^\top A_{T-1}^\top (H_T^P)^\top\Big)\\
&-2\operatorname{Tr}(A_{T-1}^\top Q_T^P K_T^P H_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P) + \operatorname{Tr}(A_{T-1}^\top Q_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P)\\
= &\operatorname{Tr}\Big(Q_T^P\big((\widehat{\Sigma}_T^P)^- - \widehat{\Sigma}_T^P\big)\Big) + \operatorname{Tr}\Big((K_T^P)^\top Q_T^P K_T^P H_T^P A_{T-1}(\widehat{\Sigma}_{T-1}^P - \widetilde{\Sigma}_{T-1}^{(P,E)})(\Pi_{T-1}^P)^\top A_{T-1}^\top (H_T^P)^\top\Big)\\
&-2\operatorname{Tr}(A_{T-1}^\top Q_T^P K_T^P H_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P) + \operatorname{Tr}(A_{T-1}^\top Q_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P)
\end{aligned} \tag{A.10}
$$

where the last equation holds since $\big(\widehat{\Sigma}_T^P\big)^-(H_T^P)^\top(K_T^P)^\top = \big(\widehat{\Sigma}_T^P\big)^- - \widehat{\Sigma}_T^P$ by (2.21h). Now by (A.10), the sum of (2c) and the last three terms in (2.65) is given by

$$
\begin{aligned}
&\operatorname{Tr}((K_T^P)^\top Q_T^P K_T^P G^P) + \operatorname{Tr}(\Gamma_{T-1}^\top (H_T^P)^\top (K_T^P)^\top Q_T^P K_T^P H_T^P \Gamma_{T-1}W) + \operatorname{Tr}(Q_T^P\widehat{\Sigma}_T^P) + (2c)\\
= &\operatorname{Tr}\Big(Q_T^P\big(\widehat{\Sigma}_T^P\big)^-\Big) + \operatorname{Tr}\Big((K_T^P)^\top Q_T^P K_T^P H_T^P A_{T-1}(\widehat{\Sigma}_{T-1}^P - \widetilde{\Sigma}_{T-1}^{(P,E)})(\Pi_{T-1}^P)^\top A_{T-1}^\top (H_T^P)^\top\Big)\\
&-2\operatorname{Tr}(A_{T-1}^\top Q_T^P K_T^P H_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P) + \operatorname{Tr}(A_{T-1}^\top Q_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P)
\end{aligned} \tag{A.11}
$$

Step 4: Collecting all the terms on LHS. Recall that (2b) + (1a) + (3a) can be cancelled with RHS in Step 1, the sum of (1b) + (2a) + (3b) + (2d) + (3c) is given in Step 2, and the sum of the last three terms on LHS and (2c) is given in Step 3. We are now ready to sum up all the terms on LHS using the simplified forms obtained from Steps 2 and 3.

Constant terms on LHS

$$
\begin{aligned}
= \quad & (2b) + (1a) + (3a) \\
& + \operatorname{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \big(\widetilde{\Sigma}_{T-1}^{(P,E)} - \widehat{\Sigma}_{T-1}^P)\right) \\
& +2 \operatorname{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P B_{T-1}^E F_{T-1}^E \widehat{\Sigma}_{T-1}^E\right) && \text{(A.12)} \\
& +2 \operatorname{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P (A_{T-1} + B_{T-1}^E F_{T-1}^E) \widehat{\Sigma}_{T-1}^P\right) && \text{(A.13)} \\
& -2 \operatorname{Tr}\left((A_{T-1} + B_{T-1}^E F_{T-1}^E)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \widehat{\Sigma}_{T-1}^P\right) && \text{(A.14)} \\
& -2 \operatorname{Tr}\left((B_{T-1}^E F_{T-1}^E)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P \widetilde{\Sigma}_{T-1}^{(P,E)}\right) && \text{(A.15)} \\
& +2 \operatorname{Tr}\left((B_{T-1}^E F_{T-1}^E)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \widetilde{\Sigma}_{T-1}^{(P,E)}\right) && \text{(A.16)} \\
& -2 \operatorname{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P (A_{T-1} + B_{T-1}^E F_{T-1}^E) \widetilde{\Sigma}_{T-1}^{(P,E)}\right) && \text{(A.17)} \\
& + \operatorname{Tr}\left(Q_T^P (\widehat{\Sigma}_T^P)^-\right) + \operatorname{Tr}\left((K_T^P)^\top Q_T^P K_T^P H_T^P A_{T-1} (\widehat{\Sigma}_{T-1}^P - \widetilde{\Sigma}_{T-1}^{(P,E)}) (\Pi_{T-1}^P)^\top A_{T-1}^\top (H_T^P)^\top\right) \\
& -2 \operatorname{Tr}(A_{T-1}^\top Q_T^P K_T^P H_T^P A_{T-1} \widehat{\Sigma}_{T-1}^P) + \operatorname{Tr}(A_{T-1}^\top Q_T^P A_{T-1} \widehat{\Sigma}_{T-1}^P).
\end{aligned}
$$

We observe that by adding (part of the terms in) (A.12), (A.13), (A.15), and (A.17) together, we can obtain the sum

$$
\begin{aligned}
& 2 \operatorname{Tr}\left(\left((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P\right)^\top Q_T^P B_{T-1}^E F_{T-1}^E \big(\widehat{\Sigma}_{T-1}^P + \widehat{\Sigma}_{T-1}^E - \widetilde{\Sigma}_{T-1}^{(P,E)} - \widetilde{\Sigma}_{T-1}^{(E,P)})\right) \\
= \quad & 2 \operatorname{Tr}\left(\left((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P\right)^\top Q_T^P B_{T-1}^E F_{T-1}^E \widehat{\Sigma}_{T-1}^{(P,E)}\right) \\
= \quad & 2 \operatorname{Tr}\left(\left((I - K_T^P H_T^P) A_{T-1}\right)^\top Q_T^P B_{T-1}^E F_{T-1}^E \big(\widehat{\Sigma}_{t-1}^P - \widetilde{\Sigma}_{t-1}^{(E,P)})\right)
\end{aligned}
$$

where in the last equation we use the trick (A.2) again. This sum will be further cancelled with part of (A.14) and (A.16). After these manipulations we have

$$
\begin{aligned}
& \text{Constant terms on LHS} \\
= \quad & (2b) + (1a) + (3a) \\
& + \operatorname{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \big(\widetilde{\Sigma}_{T-1}^{(P,E)} - \widehat{\Sigma}_{T-1}^P)\right) && \text{(A.18)} \\
& +2 \operatorname{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P A_{T-1} \widehat{\Sigma}_{T-1}^P\right) && \text{(A.19)} \\
& -2 \operatorname{Tr}\left(A_{T-1}^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \widehat{\Sigma}_{T-1}^P\right) && \text{(A.20)} \\
& -2 \operatorname{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P A_{T-1} \widetilde{\Sigma}_{T-1}^{(P,E)}\right) && \text{(A.21)} \\
& + \operatorname{Tr}\left(Q_T^P (\widehat{\Sigma}_T^P)^-\right) + \operatorname{Tr}\left((K_T^P)^\top Q_T^P K_T^P H_T^P A_{T-1} (\widehat{\Sigma}_{T-1}^P - \widetilde{\Sigma}_{T-1}^{(P,E)}) (\Pi_{T-1}^P)^\top A_{T-1}^\top (H_T^P)^\top\right) \\
& -2 \operatorname{Tr}(A_{T-1}^\top Q_T^P K_T^P H_T^P A_{T-1} \widehat{\Sigma}_{T-1}^P) + \operatorname{Tr}(A_{T-1}^\top Q_T^P A_{T-1} \widehat{\Sigma}_{T-1}^P). && \text{(A.22)}
\end{aligned}
$$

Now adding terms in (A.20) and (A.22) together we have

$$
\begin{aligned}
& -2 \operatorname{Tr}\left(A_{T-1}^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \widehat{\Sigma}_{T-1}^P\right) - 2 \operatorname{Tr}(A_{T-1}^\top Q_T^P K_T^P H_T^P A_{T-1} \widehat{\Sigma}_{T-1}^P) \\
& + \operatorname{Tr}(A_{T-1}^\top Q_T^P A_{T-1} \widehat{\Sigma}_{T-1}^P) \\
= \quad & - \operatorname{Tr}(A_{T-1}^\top Q_T^P A_{T-1} \widehat{\Sigma}_{T-1}^P)
\end{aligned}
$$

Also, adding (A.18), (A.19), and (A.21) together leads to

$$
\begin{aligned}
& \operatorname{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P (I - K_T^P H_T^P) A_{T-1} \big(\widetilde{\Sigma}_{T-1}^{(P,E)} - \widehat{\Sigma}_{T-1}^P)\right) \\
& +2 \operatorname{Tr}\left(((I - K_T^P H_T^P) A_{T-1} \Pi_{T-1}^P)^\top Q_T^P A_{T-1} \widehat{\Sigma}_{T-1}^P\right)
\end{aligned}
$$

$$
\begin{aligned}
&-2\,\mathrm{Tr}\left(((I-K_T^P H_T^P)A_{T-1}\Pi_{T-1}^P)^\top Q_T^P A_{T-1}\widetilde{\Sigma}_{T-1}^{(P,E)}\right)\\
= \quad & -\mathrm{Tr}\left(((I-K_T^P H_T^P)A_{T-1}\Pi_{T-1}^P)^\top Q_T^P K_T^P H_T^P A_{T-1}\big(\widetilde{\Sigma}_{T-1}^{(P,E)}-\widehat{\Sigma}_{T-1}^P\big)\right)\\
&+\mathrm{Tr}\left(((I-K_T^P H_T^P)A_{T-1}\Pi_{T-1}^P)^\top Q_T^P A_{T-1}\big(\widehat{\Sigma}_{T-1}^P-\widetilde{\Sigma}_{T-1}^{(P,E)}\big)\right)
\end{aligned}
$$

Therefore, the constant terms on LHS can be further simplified as

$$
\begin{aligned}
& \text{Constant terms on LHS}\\
= \quad & (2b)+(1a)+(3a)\\
& -\mathrm{Tr}\left(((I-K_T^P H_T^P)A_{T-1}\Pi_{T-1}^P)^\top Q_T^P K_T^P H_T^P A_{T-1}\big(\widetilde{\Sigma}_{T-1}^{(P,E)}-\widehat{\Sigma}_{T-1}^P\big)\right)\\
& +\mathrm{Tr}\left(((I-K_T^P H_T^P)A_{T-1}\Pi_{T-1}^P)^\top Q_T^P A_{T-1}\big(\widehat{\Sigma}_{T-1}^P-\widetilde{\Sigma}_{T-1}^{(P,E)}\big)\right)\\
& -\mathrm{Tr}(A_{T-1}^\top Q_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P)+\mathrm{Tr}\left(Q_T^P\big(\widehat{\Sigma}_T^P\big)^-\right)\\
& +\mathrm{Tr}\left((K_T^P H_T^P A_{T-1}\Pi_{T-1}^P)^\top Q_T^P K_T^P H_T^P A_{T-1}(\widehat{\Sigma}_{T-1}^P-\widetilde{\Sigma}_{T-1}^{(P,E)})\right)\\
= \quad & (2b)+(1a)+(3a)\\
& -\mathrm{Tr}\left((A_{T-1}\Pi_{T-1}^P)^\top Q_T^P K_T^P H_T^P A_{T-1}\big(\widetilde{\Sigma}_{T-1}^{(P,E)}-\widehat{\Sigma}_{T-1}^P\big)\right)\\
& +\mathrm{Tr}\left(((I-K_T^P H_T^P)A_{T-1}\Pi_{T-1}^P)^\top Q_T^P A_{T-1}\big(\widehat{\Sigma}_{T-1}^P-\widetilde{\Sigma}_{T-1}^{(P,E)}\big)\right)\\
& -\mathrm{Tr}(A_{T-1}^\top Q_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P)+\mathrm{Tr}\left(Q_T^P\big(\widehat{\Sigma}_T^P\big)^-\right)\\
= \quad & (2b)+(1a)+(3a)\\
& +\mathrm{Tr}\left((A_{T-1}\Pi_{T-1}^P)^\top Q_T^P A_{T-1}\big(\widehat{\Sigma}_{T-1}^P-\widetilde{\Sigma}_{T-1}^{(P,E)}\big)\right)\\
& -\mathrm{Tr}(A_{T-1}^\top Q_T^P A_{T-1}\widehat{\Sigma}_{T-1}^P)+\mathrm{Tr}\left(Q_T^P\big(\widehat{\Sigma}_T^P\big)^-\right)\\
= \quad & (2b)+(1a)+(3a)+\mathrm{Tr}(\Gamma_{T-1}^\top Q_T^P\Gamma_{T-1}W) \qquad \text{(A.23)}
\end{aligned}
$$

where the second last equality holds since $\big(\widehat{\Sigma}_{T-1}^P-\widetilde{\Sigma}_{T-1}^{(P,E)}\big)(\Pi_{T-1}^P)^\top=\Pi_{T-1}^P\big(\widehat{\Sigma}_{T-1}^P-\widetilde{\Sigma}_{T-1}^{(P,E)}\big)^\top$, and the last equality holds by (A.7).

Combining (A.23) and (A.1), we can see that the constant terms on LHS and RHS are the same. Recall that we have already matched the terms that are quadratic in $\widehat{x}_{T-1}^P$ on both sides. Therefore, the tower property given in (2.59) holds.

B Bargaining Game with $n_1 \geq 2$

In Section 5 we examined a bargaining model where the players only have noisy observations of the true value of the good ($n_1=1$), and they can fully recover the opponent's state estimate. In this section we introduce a more complex setting, where the true value of the good depends on a set of factors and the players can only observe noisy versions of the factors. We will first construct the bargaining model which involves high dimensional factors, and then focus on the case when $n_1=2$. In contrast to Section 5, here players cannot fully recover their opponent's state estimate. We will finally present some numerical results to show the effect of observation noise and information corrections.

Dynamics of the Value of the Good When $n_1 > 1$. Assume that that the value of the good takes the form

$$p_t=\langle\theta,\xi_t\rangle, \qquad \text{(B.1)}$$

with $\xi_t \in \mathbb{R}^{n_1}$ a set of factors that determines the value of the good with coefficient $\theta \in \mathbb{R}^{n_1}$. For the value of common commodities, the factors could include weather, government policies, international events, consumer preferences, shifting input costs, and supply and demand imbalance. We assume the factors follow a simple model:

$$\xi_{t+1} = \xi_t + w_t,$$

where $\{w_t\}_{t=0}^{T-1}$ is a sequence of IID Gaussian random variables with zero mean and covariance $\overline{W} \in \mathbb{R}^{n_1 \times n_1}$.

Both the buyer and the seller do not have access to the true value of the good nor the precise value of the factors. Instead, they observe a noisy version of the factors using their private information. At time $t = 0$, player i ($i = B, S$) believes that the initial factor signal:

$$\xi_0 \sim \mathcal{N}(\widehat{\xi}_0^i, W_0^i), \tag{B.2}$$

and thereafter player i observes the following noisy factor signal:

$$z_{t+1}^i = \xi_{t+1} + w_{t+1}^i, \quad w_{t+1}^i \sim \mathcal{N}(0, G^i), \quad t = 0, 1, \cdots, T-1, \tag{B.3}$$

where $\{w_t^i\}_{t=1}^{T-1}$ is a sequence of IID random variables, and $\{w_t^B\}_{t=1}^{T-1}$ and $\{w_t^S\}_{t=1}^{T-1}$ are independent.

We Note that this case does not degenerate to the $n_1 = 1$ case as players have different noisy factor signals leading to different value estimates

The Bargaining Model when $n_1 = 2$. Now we focus on the case where the value of the good depends on a set of two factors and players can only observe a noisy version of the factors. The dynamics of the value of the good and the noisy signals are defined in (B.1)-(B.3). In contrast to the previous section, here the players are not able to fully recover their opponent's state estimate.

We set the coefficient $\theta = (\theta_1, \theta_2)^\top$ and the model parameters to be, for $t = 0, 1, \ldots, T-1$

$$A_t = I, \ B_t^B = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad B_t^S = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad W = \begin{bmatrix} \overline{W}^1 & 0 & 0 & 0 \\ 0 & \overline{W}^2 & 0 & 0 \\ 0 & 0 & \overline{W}^B & 0 \\ 0 & 0 & 0 & \overline{W}^S \end{bmatrix}, \quad Q_t^B = Q_t^S = 0.$$

and

$$Q_T^B = \begin{bmatrix} \beta_B(1+\delta_B)^2\theta_1^2 & \beta_B(1+\delta_B)^2\theta_1\theta_2 & -\beta_B(1+\delta_B)\theta_1 & 0 \\ \beta_B(1+\delta_B)^2\theta_1\theta_2 & -\beta_B(1+\delta_B)^2\theta_2^2 & -\beta_B(1+\delta_B)\theta_2 & 0 \\ -\beta_B(1+\delta_B)\theta_1 & -\beta_B(1+\delta_B)\theta_2 & \alpha_B + \beta_B & -\alpha_B \\ 0 & 0 & -\alpha_B & \alpha_B \end{bmatrix},$$

$$Q_T^S = \begin{bmatrix} \beta_S(1+\delta_S)^2\theta_1^2 & \beta_S(1+\delta_S)^2\theta_1\theta_2 & 0 & -\beta_S(1+\delta_S)\theta_1 \\ \beta_S(1+\delta_S)^2\theta_1\theta_2 & \beta_S(1+\delta_S)^2\theta_2^2 & 0 & -\beta_S(1+\delta_S)^2\theta_2 \\ 0 & 0 & \alpha_S & -\alpha_S \\ -\beta_S(1+\delta_S)\theta_1 & -\beta_S(1+\delta_S)\theta_2 & -\alpha_S & \alpha_S + \beta_S \end{bmatrix}.$$

Also we have $H_t^i = I$ for $i = S, B$.

For comparison we let α_B, α_S, β_B, β_S, δ_B, δ_S, T, and the penalty function to be the same as in Section. We also set $\theta_1 = \theta_2 = 1$. For the initial state we set $\xi_0 = (30, 20)^\top$, $x_0^B = 10$, $x_0^S = 90$. We also set $\overline{W}^1 = \overline{W}^2 = 4.5$ for the noise in the dynamics of the factors, and $\overline{W}^B = \overline{W}^S = 10^{-12}$.

To see the effect of the observation noise, we let the buyer have a much noisier observation of the true price by setting

$$G^B = \begin{bmatrix} 50 & 0 \\ 0 & 50 \end{bmatrix}, \quad G^S = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}.$$

For the players' belief about the initial state, we set $\widehat{x}_0^B = (25, 15)^\top$ and $\widehat{x}_0^S = (31, 20)^\top$ with

$$W_0^B = \begin{bmatrix} 50 & 0 \\ 0 & 50 \end{bmatrix}, W_0^S = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix},$$

thus the buyer has a far more inaccurate guess of the initial state.

Effect of Observation Noise. As for the case when $n_1 = 1$, the seller has a more accurate state estimate since he receives less noisy signals. The behaviour of the players is similar to that in the full information case.

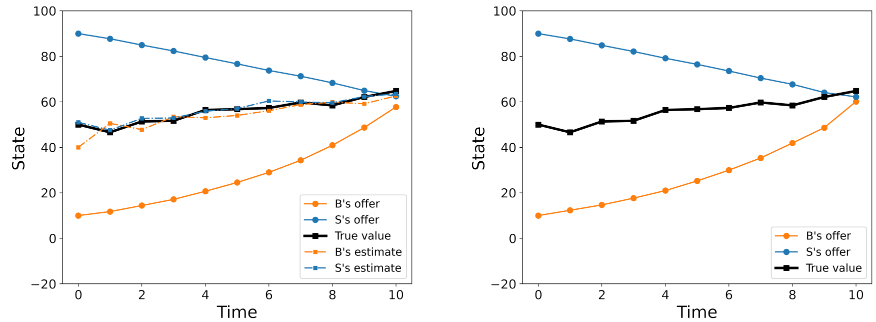

Figure 3: Comparison between the full observation (right) and partial observation (left) cases.

Effect of Information Corrections. With information corrections the buyer's estimate of the value is less affected by the noisy observations (see Figure 4) and players are more likely to achieve an agreement in each of the asymmetric and symmetric case (see Table 5). Compared with the case when $n_1 = 1$, the gap between the number of agreements achieved in the asymmetric and symmetric (inaccurate) case is larger, since players receive noisy signals from the different factors.

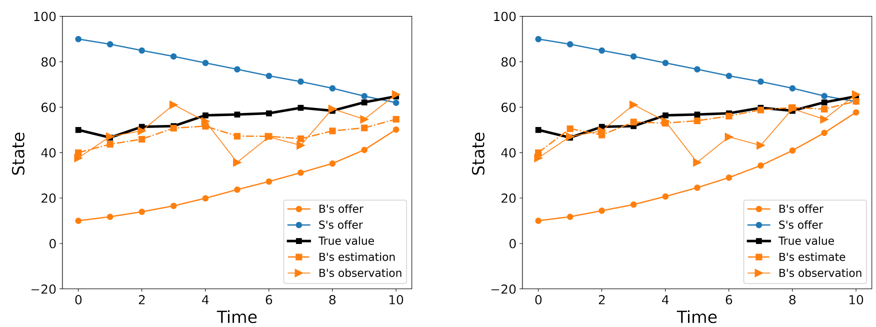

Figure 4: The buyer's price estimate with information corrections (right) and without information corrections (left).

	Asymmetric	Symmetric ("accurate")	Symmetric (inaccurate)
With IC	329	439	315
Without IC	237	440	207

Table 5: Number of agreements achieved in 500 experiments with and without information corrections (IC).