Proof of Lemma 1:

Proof. The proof for this lemma follows the proof of theorem 6 in [1]. Here we modify the proof to account for the changing C_t matrix. The proof is by induction. First, we prove that it holds for $t = 1$. Then we assume that it holds for $t = k$, we prove it holds for $t = k + 1$.

When $t = 1$, then

$$\begin{aligned}
\widehat{x}_1 &= A\widehat{x}_0 + K_0(y_0 - C_0\widehat{x}_0) = K_0 y_0 \\
K_0 &= (G_0 - AP_0C_0^T)(\Lambda_0^0 - C_0P_0C_0^T)^{-1} = G_0(\Lambda_0^0)^{-1} = \Delta_1^0(L_1^0)^{-1} \\
&\Rightarrow \widehat{x}_1 = \Delta_1^0(L_1^0)^{-1}y_0
\end{aligned}$$

When $t = 1$, we can also calculate P_1:

$$\begin{aligned}
P_1 &= AP_0A^T + (G_0 - AP_0C_0^T)(\Lambda_0^0 - C_0P_0C_0^T)^{-1}(G_0 - AP_0C_0^T)^T \\
&= G_0(\Lambda_0^0)^{-1}G_0^T \\
&= \Delta_1^0(L_1^0)^{-1}(\Delta_1^0)^T
\end{aligned}$$

Hence, (13) and (14) hold for $t = 1$. Next we prove that if (13) and (14) hold for $t = k$, then they hold for $t = k + 1$. In this process, we will use the following formula for matrix inversion [2].

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} =$$
$$\begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\ -(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}$$

Let's calculate P_{k+1} first:

$$\begin{aligned}
P_{k+1} &= \Delta_{k+1}^0(L_{k+1}^0)^{-1}(\Delta_{k+1}^0)^T \\
&= (A\Delta_k^0 \quad G_k)\begin{pmatrix} L_k^0 & (\Delta_k^0)^TC_k^T \\ C_k\Delta_k^0 & \Lambda_0^k \end{pmatrix}^{-1}\begin{pmatrix} (\Delta_k^0)^TA^T \\ G_k^T \end{pmatrix} \\
&= (A\Delta_k^0 \quad G_k)\begin{pmatrix} (L_k^0)^{-1} + (L_k^0)^{-1}(\Delta_k^0)^TC_k^T\Delta^{-1}C_k\Delta_k^0(L_k^0)^{-1} & -(L_k^0)^{-1}(\Delta_k^0)^TC_k^T\Delta^{-1} \\ -\Delta^{-1}C_k\Delta_k^0(L_k^0)^{-1} & \Delta^{-1} \end{pmatrix} \\
&\begin{pmatrix} (\Delta_k^0)^TA^T \\ G_k^T \end{pmatrix} \\
&\text{where } \Delta = \Lambda_0^k - C_k\Delta_k^0(L_k^0)^{-1}(\Delta_k^0)^TC_k^T = \Lambda_0^k - C_kP_kC_k^T. \\
&= A\Delta_k^0(L_k^0)^{-1}(\Delta_k^0)^TA^T \\
&+ (G_k - A\Delta_k^0(L_k^0)^{-1}(\Delta_k^0)^TC_k^T)\Delta^{-1}(G_k - A\Delta_k^0(L_k^0)^{-1}(\Delta_k^0)^TC_k^T)^T \\
&= AP_kA^T + (G_k - AP_kC_k^T)(\Lambda_0^k - C_kP_kC_k^T)^{-1}(G_k - AP_kC_k^T)^T
\end{aligned}$$

This proves (14). Next we will prove (13).

$$
\begin{aligned}
\widehat{x}_{k+1} &= \Delta^0_{k+1}(L^0_{k+1})^{-1}\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix} \\
&= \begin{pmatrix} A\Delta^0_k & G_k \end{pmatrix}\begin{pmatrix} L^0_k & (\Delta^0_k)^T C_k^T \\ C_k\Delta^0_k & \Lambda^k_0 \end{pmatrix}^{-1}\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix} \\
&= \begin{pmatrix} A\Delta^0_k & G_k \end{pmatrix} \\
&\begin{pmatrix} (L^0_k)^{-1} + (L^0_k)^{-1}(\Delta^0_k)^T C_k^T \Delta^{-1} C_k \Delta^0_k (L^0_k)^{-1} & -(L^0_k)^{-1}(\Delta^0_k)^T C_k^T \Delta^{-1} \\ -\Delta^{-1} C_k \Delta^0_k (L^0_k)^{-1} & \Delta^{-1} \end{pmatrix}\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix} \\
&= [A - (G_k - A\Delta^0_k(L^0_k)^{-1}(\Delta^0_k)^T C_k^T)\Delta^{-1}C_k^T]\Delta^0_k(L^0_k)^{-1}\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{k-1} \end{pmatrix} \\
&+ (G_k - A\Delta^0_k(L^0_k)^{-1}(\Delta^0_k)^T C_k^T)\Delta^{-1} y_k \\
&= A\widehat{x}_k + (G_k - AP_kC_k^T)(\Lambda^k_0 - C_kP_kC_k^T)^{-1}(y_k - C_k\widehat{x}_k) \\
&= A\widehat{x}_k + K_k(y_k - C_k\widehat{x}_k)
\end{aligned}
$$

This proves equation (13). □

Proof of Theorem 1:

Proof. From the definition of H^d_t and the ergodicity of y_t, as $N \to \infty$,

$$H^d_t = E[Y_{t+d,r|t+2d-1}Y^T_{t,r|t+d-1}]E[Y_{t,r|t+d-1}Y^T_{t,r|t+d-1}]^\dagger Y_{t,r|t+d-1} = N^t_d(L^t_d)^{-1}Y_{t,r|t+d-1}$$

It is easy to verify that

$$N^t_d = O_{t,r|d}\Delta^t_d$$

Hence, we have

$$H^d_t = O_{t,r|d}\Delta^t_d(L^t_d)^{-1}Y_{t,r|t+d-1}$$

By equation (15), we can write

$$H^d_t = O_{t,r|d}\widehat{X}_{t,r}$$

Since $O_{t,r|d}$ is full column rank, H^d_t has the same row space with $\widehat{X}_{t,r}$. The pair $(A, Q^{1/2})$ being controllable can guarantee that $\widehat{X}_{t,r}$ is full row rank, so H^d_t has the same column space with $O_{t,r|d}$. So we can do singular value decomposition for H^d_t and find its row space and column space, respectively. Then we can get $\widehat{O}_{t,r|d}$ and $\widehat{X}^{\text{svd}}_{t,r}$ up to a similarity transformation. □

Proof of Theorem 2:

Proof. By (21), we have

$$\widehat{O}^{\text{ref}}_{t,1|d}\Gamma_t = O^{\text{ref}}_{t,1|d}, \quad \widehat{O}^{\text{ref}}_{t+1,1|d}\Gamma_{t+1} = O^{\text{ref}}_{t+1,1|d}$$

So we have the following equation:

$$\widehat{O}^{\text{ref}}_t\Gamma_t = O^{\text{ref}}_t, \quad \widehat{O}^{\text{ref}}_{t+1}\Gamma_{t+1} = O^{\text{ref}}_{t+1} = O^{\text{ref}}_t$$

By the above equation, we have

$$\widehat{O}^{\text{ref}}_t\Gamma_t = \widehat{O}^{\text{ref}}_{t+1}\Gamma_{t+1} \Rightarrow \Gamma_{t+1}\Gamma_t^{-1} = (\widehat{O}^{\text{ref}}_{t+1})^\dagger\widehat{O}^{\text{ref}}_t$$

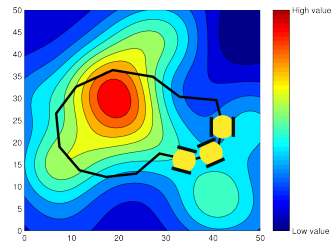

Figure 1: Multiple synchronized mobile sensors with a fixed network structure.

That is, the transformation matrix between $\widehat{O}^{\text{ref}}_{t+1}$ and $\widehat{O}^{\text{ref}}_{t}$ is given by $\Gamma = \Gamma_{t+1}\Gamma_t^{-1}$. Hence, $\widehat{O}_{t+1,1|d}\Gamma$ and $\widehat{O}_{t,1|d}$ will stay in the same basis. So we have

$$\widehat{A} = \left(\overline{\widehat{O}_{t+1,1|d}\Gamma}\right)^{\dagger} \underline{\widehat{O}}_{t,1|d}$$

□

Proof of Theorem 3:

Proof. The proof is straight forward. To prove this theorem, we use a concrete example. Let's consider the case of using three sensors to collect measurements in Figure 1. In this case, $N_c = 2$, $N_s = 3$. Let $r = 3$. So all the three sensors will execute the motion that they move forward for one step and stay for $r = 3$ steps and move again. Without loss of generality, let's assume that they start at time step $t = 1$ and sensor 1 starts at waypoint 1, sensor 2 starts at waypoint 2, and sensor 3 starts at waypoint 3. When $K_d = 2$, the corresponding extended observability matrix is given by

$$O^{\text{mult}}_{1,3|6} = \begin{pmatrix} C^{\text{mult}}_1 \\ C^{\text{mult}}_2 A \\ C^{\text{mult}}_3 A^2 \\ \hline C^{\text{mult}}_4 A^3 \\ C^{\text{mult}}_5 A^4 \\ C^{\text{mult}}_6 A^5 \end{pmatrix} \quad \bar{O}^{\text{mult}}_{1,3|6} = \begin{pmatrix} C^{\text{mult}}_1 \\ C^{\text{mult}}_2 A \\ C^2_3 A^2 \\ C^3_3 A^2 \\ C^{\text{mult}}_4 A^3 \\ C^{\text{mult}}_5 A^4 \end{pmatrix} \quad \underline{O}^{\text{mult}}_{1,3|6} = \begin{pmatrix} C^{\text{mult}}_2 A \\ C^{\text{mult}}_3 A^2 \\ C^1_4 A^3 \\ C^2_4 A^3 \\ C^{\text{mult}}_5 A^4 \\ C^{\text{mult}}_6 A^5 \end{pmatrix} \tag{1}$$

where $C^{\text{mult}}_t = \left[C^1_t; C^2_t; C^3_t\right]$ for all $t = 1, 2, \ldots, 6$. And $C^{\text{mult}}_1 = C^{\text{mult}}_2 = C^{\text{mult}}_3$, $C^{\text{mult}}_4 = C^{\text{mult}}_5 = C^{\text{mult}}_6$. We also have $C^2_3 = C^1_4$ and $C^3_3 = C^2_4$ since sensor 1 and sensor 2 are consecutive pairs $1 \twoheadrightarrow 2$ and sensor 2 and sensor 3 are also consecutive pairs $2 \twoheadrightarrow 3$. Hence, the number of rows in $\bar{O}^{\text{mult}}_{1,3|6}$ is 14, which is equal to $N_s(r-1)K_d + N_c(K_d - 1)$ when $N_s = 3, r = 3, K_d = 2, N_c = 2$. The $N_s(r-1)K_d$ rows come from $\left[C^{\text{mult}}_1; C^{\text{mult}}_2 A; C^{\text{mult}}_4 A^3; C^{\text{mult}}_5 A^4\right]$ and the $N_c(K_d - 1)$ rows come from $\left[C^2_3 A^2; C^3_3 A^2\right]$. This proof can be generalized to any N_s, N_c, r and K_d except that $N_s = 1$ and $r = 1$ at the same time. In this case, we only have one sensor and it keeps moving along the periodic trajectory and $N_c = 0$. We need to use a static reference sensor to learn the A matrix and $K_d \geq n + 1$. □

References

[1] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: theory, implementation, applications*. Kluwer academic publishers, 1996.

[2] T. Kailath, *Linear systems*, vol. 1. Prentice-Hall Englewood Cliffs, NJ, 1980.