Supplementary Material

Proof of Lemma 2. We subtract (1) from (4) to get

$$\begin{aligned}\hat{x}_{i,t+1} - x_{t+1} &= a\bigg(\sum_{j\in\mathcal{N}_i} p_{ij}\hat{x}_{j,t} - x_t + \alpha(y_{i,t} - \hat{x}_{i,t})\bigg) - r_t \\ &= a\bigg(\sum_{j\in\mathcal{N}_i} p_{ij}(\hat{x}_{j,t} - x_t) + \alpha(y_{i,t} - \hat{x}_{i,t})\bigg) - r_t,\end{aligned}$$

where we used Assumption 1 in the latter step. Replacing $y_{i,t}$ from (2) in above, and simplifying using definition of $\hat{\xi}_{i,t}$, yields

$$\begin{aligned}\hat{\xi}_{i,t+1} &= a\bigg(\sum_{j\in\mathcal{N}_i} p_{ij}\hat{\xi}_{j,t} + \alpha(y_{i,t} - \hat{x}_{i,t})\bigg) - r_t \\ &= a\bigg(\sum_{j\in\mathcal{N}_i} p_{ij}\hat{\xi}_{j,t} - \alpha\hat{\xi}_{i,t}\bigg) + (a\alpha)w_{i,t} - r_t.\end{aligned}$$

Using definition (6) to write the above in the matrix form completes the proof for $\hat{\xi}_t$. The proof for $\tilde{\xi}_t$ follows precisely in the same fashion. □

Proof of Proposition 3. We start by the fact that the innovation and observation noise are zero mean, so (7) implies $\mathbb{E}[\hat{\xi}_{t+1}] = Q\mathbb{E}[\hat{\xi}_t]$, and $\mathbb{E}[\tilde{\xi}_{t+1}] = Q\mathbb{E}[\tilde{\xi}_t]$. Therefore, for mean stability of the linear equations, the spectral radius of Q must be less than unity. Considering the expression for Q from (8), for a fixed α we must have

$$|a| < \frac{1}{\rho(P - \alpha I_N)} = \frac{1}{\max\{1-\alpha, |\alpha - \lambda_N(P)|\}}. \tag{26}$$

To maximize the right hand side over α, we need to solve the min-max problem

$$\min_{\alpha}\bigg\{\max\{1-\alpha, |\alpha - \lambda_N(P)|\}\bigg\}.$$

Noting that $1-\alpha$ and $\alpha - \lambda_N(P)$ are straight lines with negative and positive slopes, respectively, the minimum occurs at the intersection of the two lines. Evaluating the right hand side of (26) at the intersection point $\alpha^* = \frac{1+\lambda_N(P)}{2}$, completes the proof. □

Proof of Theorem 5. We present the proof for $\widetilde{\text{MSD}}(P,\alpha)$ by observing that (7)

$$\mathbb{E}[\tilde{\xi}_{t+1}\tilde{\xi}_{t+1}^{\intercal}] = Q\mathbb{E}[\tilde{\xi}_t\tilde{\xi}_t^{\intercal}]Q^{\intercal} + \mathbb{E}[\tilde{s}_t\tilde{s}_t^{\intercal}],$$

since the innovation and observation noise are zero mean and uncorrelated. Therefore, letting $\tilde{S} = \mathbb{E}[\tilde{s}_t\tilde{s}_t^{\intercal}]$, since $\rho(Q) < 1$ by hypothesis, the steady state satisfies a Lyapunov equation as below

$$\tilde{\Sigma} = Q\tilde{\Sigma}Q^{\intercal} + \tilde{S}.$$

Let $Q = Q^{\intercal} = U\Lambda U^{\intercal}$ represent the Eigen decomposition of Q. Let also u_i denote the i-th eigenvector of Q corresponding to eigenvalue λ_i. Under stability of Q the solution of the Lyapunov equation is as follows

$$\begin{aligned}\tilde{\Sigma} &= \sum_{\tau=0}^{\infty} Q^{\tau}\tilde{S}Q^{\tau} \\ &= \sum_{\tau=0}^{\infty}\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i^{\tau}u_iu_i^{\intercal}\tilde{S}\lambda_j^{\tau}u_ju_j^{\intercal} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} u_iu_i^{\intercal}\tilde{S}u_ju_j^{\intercal}\sum_{\tau=0}^{\infty}\lambda_i^{\tau}\lambda_j^{\tau} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{u_iu_i^{\intercal}\tilde{S}u_ju_j^{\intercal}}{1-\lambda_i\lambda_j}.\end{aligned}$$

Therefore, the $\tilde{\text{MSD}}$ defined in 4, can be computed as

$$\tilde{\text{MSD}} = \frac{1}{N}\text{Tr}(\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{u_i u_i^\intercal \tilde{S} u_j u_j^\intercal}{1-\lambda_i\lambda_j}) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{(u_j^\intercal u_i)(u_i^\intercal \tilde{S} u_j)}{1-\lambda_i\lambda_j} = \frac{1}{N}\sum_{i=1}^{N}\frac{u_i^\intercal \tilde{S} u_i}{1-\lambda_i^2},$$

where we used the fact that $u_j^\intercal u_i = 0$, for $i \neq j$, and $u_i^\intercal u_i = 1$, for any $i \in \mathcal{V}$. Taking into account that $Q = a(P - \alpha I_N)$ and $\tilde{S} = \sigma_r^2(\mathbf{1}_N\mathbf{1}_N^\intercal) + (a^2\alpha^2\sigma_w^2)P^2$, we derive

$$\begin{aligned}
\tilde{\text{MSD}} &= \frac{1}{N}\sum_{i=1}^{N}\frac{u_i^\intercal(\sigma_r^2(\mathbf{1}_N\mathbf{1}_N^\intercal) + (a^2\alpha^2\sigma_w^2)P^2)u_i}{1-\lambda_i^2} \\
&= \frac{1}{N}\sum_{i=1}^{N}\frac{(u_i^\intercal\mathbf{1}_N)^2\sigma_r^2}{1-\lambda_i^2} + \frac{1}{N}\sum_{i=1}^{N}\frac{a^2\alpha^2\sigma_w^2\lambda_i^2(P)}{1-\lambda_i^2} \\
&= \frac{\sigma_r^2}{1-a^2(1-\alpha)^2} + \frac{1}{N}\sum_{i=1}^{N}\frac{a^2\alpha^2\sigma_w^2\lambda_i^2(P)}{1-a^2(\lambda_i(P)-\alpha)^2},
\end{aligned}$$

where the last step is due to the facts that $\lambda_i = a(\lambda_i(P) - \alpha)$ and $\mathbf{1}_N/\sqrt{N}$ is one of the eigenvectors of Q with corresponding eigenvalue $a(1-\alpha)$, so it is orthogonal to other eigenvectors, i.e., $u_i^\intercal\mathbf{1}_N = 0$, for $u_i \neq \mathbf{1}_N/\sqrt{N}$. The proof for $\hat{\text{MSD}}$ follows in the same fashion. □

Proof of Theorem 8. The closed form solution of the error process (7) is,

$$\xi_{t+1} = Q^{t+1}\xi_0 + \sum_{\tau=0}^{t} Q^{t-\tau}s_\tau,$$

which implies

$$\begin{aligned}
\xi_{t+1}\xi_{t+1}^\intercal = Q^{t+1}\xi_0\xi_0^\intercal Q^{t+1} + Q^{t+1}\xi_0\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)^\intercal + \Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)\xi_0^\intercal Q^{t+1} \\
+ \Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)^\intercal,
\end{aligned} \tag{27}$$

since Q is symmetric. One can see that

$$\|\frac{1}{T}\sum_{t=1}^{T} Q^t\xi_0\xi_0^\intercal Q^t\| \leq \frac{1}{T}\Big(\frac{\|\xi_0\|^2}{1-\rho^2(Q)}\Big),$$

and

$$\|\frac{1}{T}\sum_{t=0}^{T-1} Q^{t+1}\xi_0\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)^\intercal\| \leq \frac{\|\xi_0\|s}{T}\sum_{t=0}^{T-1}\rho(Q)^{t+1}\sum_{\tau=0}^{t}\rho(Q)^{t-\tau} \leq \frac{1}{T}\Big(\frac{s\|\xi_0\|}{\big(1-\rho(Q)\big)^2}\Big).$$

On the other hand, as we see in the proof of Theorem 5, letting $S = \mathbb{E}[s_\tau s_\tau^\intercal]$, we have $\Sigma = \sum_{\tau=0}^{\infty} Q^\tau S Q^\tau$. Based on definition (18), equation (27), and the bounds above, we derive

$$\begin{aligned}
R(T) \leq \frac{1}{T}\Big(\frac{\|\xi_0\|^2}{1-\rho^2(Q)}\Big) + \frac{1}{T}\Big(\frac{2s\|\xi_0\|}{\big(1-\rho(Q)\big)^2}\Big) \\
+ \frac{1}{T}\Big\|\sum_{t=0}^{T-1}\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)^\intercal - \sum_{\tau=0}^{t} Q^\tau S Q^\tau\Big\| + \|\frac{1}{T}\sum_{t=0}^{T-1}\sum_{\tau=t+1}^{\infty} Q^\tau S Q^\tau\|.
\end{aligned} \tag{28}$$

Let

$$H(s_0, ..., s_{T-1}) = \sum_{t=0}^{T-1}\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)^\intercal,$$

and observe that $\mathbb{E}[H(s_0,...,s_{T-1})] = \sum_{t=0}^{T-1}\sum_{\tau=0}^{t} Q^\tau S Q^\tau$. It can be verified that for any $0 \leq t < T$,

$$\|H(s_0,...,s_t...,s_{T-1}) - H(s_0,...,s_t',...,s_{T-1})\| \leq \frac{4s^2}{\big(1-\rho(Q)\big)^2}.$$

Thus, letting $Var = \frac{16Ts^4}{\big(1-\rho(Q)\big)^4}$, and appealing to Lemma 7, we get

$$\mathbb{P}\Big\{\Big\|\sum_{t=0}^{T-1}\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)^\intercal - \sum_{\tau=0}^{t} Q^\tau S Q^\tau\Big\| \geq c\Big\} \leq Ne^{-c^2/8\mathrm{Var}}.$$

Setting the probability above equal to δ, this implies that with probability at least $1-\delta$, we have

$$\frac{1}{T}\Big\|\sum_{t=0}^{T-1}\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)\Big(\sum_{\tau=0}^{t} Q^{t-\tau}s_\tau\Big)^\intercal - \sum_{\tau=0}^{t} Q^\tau S Q^\tau\Big\| \leq \frac{1}{\sqrt{T}}\frac{8s^2\sqrt{2\log\frac{N}{\delta}}}{(1-\rho(Q))^2}.$$

Moreover, we evidently have

$$\|\frac{1}{T}\sum_{t=0}^{T-1}\sum_{\tau=t+1}^{\infty} Q^\tau S Q^\tau\| \leq \frac{1}{T}\bigg(\frac{s^2}{\big(1-\rho^2(Q)\big)^2}\bigg).$$

Plugging the two bounds above in (28) completes the proof. □

Proof of Proposition 9. Considering the expression for MSD in Theorem 5, we have

$$\begin{aligned}\tilde{\mathrm{MSD}}(P,\alpha) - \tilde{\mathrm{MSD}}(P_{-\epsilon},\alpha) &= \tilde{W}_{MSD}(P,\alpha) - \tilde{W}_{MSD}(P_{-\epsilon},\alpha)\\ &\propto \sum_{i=1}^{N}\frac{\big(\lambda_i(P)-\lambda_i(P_{-\epsilon})\big)\big((1-\alpha^2a^2)(\lambda_i(P_{-\epsilon})+\lambda_i(P)) + 2a^2\alpha\lambda_i(P)\lambda_i(P_{-\epsilon})\big)}{\big(1-a^2(\lambda_i(P)-\alpha)^2\big)\big(1-a^2(\lambda_i(P_{-\epsilon})-\alpha)^2\big)}.\end{aligned}$$

Based on definitions (20) and (21), it follows from Weyl's eigenvalue inequality that $\lambda_k(P) - \lambda_k(P_{-\epsilon}) \leq \lambda_1\big(\epsilon\Delta P(i,j)\big) = 0$, for any $k \in \mathcal{V}$. Combined with the assumptions $P \geq 0$ and $|a\alpha| < 1$, this implies that the numerator of the expression above is always non-positive. The denominator is always positive due to stability of the error process $\tilde{\xi}_t$ in (7), and hence, $\tilde{\mathrm{MSD}}(P,\alpha) \leq \tilde{\mathrm{MSD}}(P_{-\epsilon},\alpha)$. □