A Appendices

A.1 Proofs for decomposition and scaling

Proof of Lemma 2.2. Recall the summation by parts formula: for scalar sequences $\{a_k, b_k\}$,

$$\sum_{k=0}^{N} a_{k+1}[b_{k+1} - b_k] = a_{k+1}b_{k+1} - a_1 b_0 - \sum_{k=1}^{N}[a_{k+1} - a_k]b_k \tag{44}$$

This is applied to (29b), beginning with

$$\widetilde{\theta}^{\mathcal{T}}_{N+1} = \sum_{n=0}^{N} \alpha_{n+1} A \widetilde{\theta}^{\mathcal{T}}_n + \sum_{n=0}^{N} \alpha_{n+1}[Z_{n+1} - Z_{n+2}]$$

Hence with $a_k = \alpha_k$ and $b_k = Z_{k+1}$, the identity (44) implies

$$\begin{aligned}\sum_{n=0}^{N} \alpha_{n+1}[Z_{n+1} - Z_{n+2}] &= Z_1 - \alpha_{N+1}Z_{N+2} + \sum_{n=1}^{N}[\alpha_{n+1} - \alpha_n]Z_{n+1} \\ &= Z_1 - \alpha_{N+1}Z_{N+2} - \sum_{n=1}^{N}\alpha_{n+1}\alpha_n Z_{n+1}\end{aligned}$$

By substitution, and using $\widetilde{\theta}^{\mathcal{T}}_0 = 0$,

$$\widetilde{\theta}^{\mathcal{T}}_{N+1} = Z_1 - \alpha_{N+1}Z_{N+2} + \sum_{n=1}^{N} \alpha_{n+1}\big[A\widetilde{\theta}^{\mathcal{T}}_n - \alpha_n Z_{n+1}\big]$$

With $\Xi_n := \widetilde{\theta}^{\mathcal{T}}_n + \alpha_n Z_{n+1}$ for $n \geq 1$ we finally obtain for $N \geq 1$,

$$\Xi_{N+1} = Z_1 + \sum_{n=1}^{N} \alpha_{n+1}\big[A\Xi_n - \alpha_n[I + A]Z_{n+1}\big]$$

which is equivalent to (30). □

Proof of Lemma 2.3. Consider the Taylor series expansion:

$$\begin{aligned}\frac{(n+1)^\varrho}{n^\varrho} &= (1 + n^{-1})^\varrho = 1 + \varrho n^{-1} - \tfrac{1}{2}\varrho(1-\varrho)n^{-2} + O(n^{-3}) \\ &= 1 + \varrho(n+1)^{-1} + \varrho n^{-1}(n+1)^{-1} - \tfrac{1}{2}\varrho(1-\varrho)n^{-2} + O(n^{-3})\end{aligned}$$

where the second equation uses $n^{-1} - (n+1)^{-1} = n^{-1}(n+1)^{-1}$. With $\alpha_n = 1/n$, the following bound follows:

$$(n+1)^\varrho = n^\varrho\big[1 + \alpha_{n+1}(\varrho + \varepsilon(n,\varrho))\big]$$

where $\varepsilon(n,\varrho) = O(n^{-1})$, and $\varepsilon(n,\varrho) > 0$ for all n.

Multiplying both sides of (3) by $(n+1)^\varrho$, we obtain

$$\widetilde{\theta}^{\varrho}_{n+1} = \widetilde{\theta}^{\varrho}_n + \alpha_{n+1}\big[\varrho_n\widetilde{\theta}^{\varrho}_n + A(n,\varrho)\widetilde{\theta}^{\varrho}_n + (n+1)^\varrho\Delta_{n+1}\big]$$

where $\varrho_n = \varrho + \varepsilon(n,\varrho)$ and $A(n,\varrho) = (1 + n^{-1})^\varrho A$. □

Lemma A.1. *Let $\varrho_0 > 0, L \geq 0$ be fixed real numbers. Then the following holds for each $n \geq 1$ and $1 \leq n_0 < n$:*

$$\prod_{k=n_0}^{n} [1 - \varrho_0 \alpha_k + L^2 \alpha_k^2] \leq K_{A.1} \frac{n_0^{\varrho_0}}{(n+1)^{\varrho_0}}$$

where $K_{A.1} = \exp(\varrho_0 + L^2 \sum_{k=1}^{\infty} \alpha_k^2)$.

Proof. By the inequality $1 - x \leq \exp(-x)$,

$$\prod_{k=n_0}^{n} [1 - \varrho_0 \alpha_k + L^2 \alpha_k^2] \leq \exp(-\varrho_0 \sum_{k=n_0}^{n} \alpha_k) \exp(L^2 \sum_{k=n_0}^{n} \alpha_k^2) \leq \exp(-\varrho_0) K \exp(-\varrho_0 \sum_{k=n_0}^{n} \alpha_k)$$

The remainder of the proof involves establishing the bound

$$\exp(-\varrho_0 \sum_{k=n_0}^{n} \alpha_k) \leq \exp(\varrho_0) \frac{n_0^{\varrho_0}}{(n+1)^{\varrho_0}} \tag{45}$$

For $n_0 = 1$ this follows from the bound $\sum_{k=1}^{n} \alpha_k \geq \ln(n+1)$, and for $n_0 \geq 2$ the bound (45) follows from $\sum_{k=n_0}^{n} \alpha_k > \ln(n+1) - \ln(n_0 - 1) - 1$. □

Lemma A.2. *Under Assumptions A1-A3, let $\lambda = -\varrho_0 + ui$ denote an eigenvalue of the matrix A with largest real part. Then*

$$\lim_{n\to\infty} n^{2\varrho} \mathsf{E}[\widetilde{\theta}_n^{\intercal} \widetilde{\theta}_n] = 0\,, \qquad \varrho < \varrho_0 \text{ and } \varrho \leq \tfrac{1}{2}$$

Proof. Recall the decomposition of $\widetilde{\theta}_n$ in (31): $\widetilde{\theta}_n = \widetilde{\theta}_n^{(1)} + \widetilde{\theta}_n^{(2)} + \widetilde{\theta}_n^{(3)}$, with $\widetilde{\theta}_n^{(1)}, \widetilde{\theta}_n^{(2)}$ evolving as

$$\widetilde{\theta}_{n+1}^{(1)} = \widetilde{\theta}_n^{(1)} + \alpha_{n+1}\big[A\widetilde{\theta}_n^{(1)} + \Delta_{n+2}^m\big]\,, \qquad \widetilde{\theta}_0^{(1)} = \widetilde{\theta}_0 \tag{46a}$$

$$\widetilde{\theta}_{n+1}^{(2)} = \widetilde{\theta}_n^{(2)} + \alpha_{n+1}\big[A\widetilde{\theta}_n^{(2)} - \alpha_n[I + A]Z_{n+1}\big]\,, \qquad \widetilde{\theta}_1^{(2)} = Z_1 \tag{46b}$$

For fixed $\varrho < \varrho_0$ and $\varrho \leq \frac{1}{2}$, Let $T > 0$ solve the Lyapunov equation $[A + \varrho I]T + T[A + \varrho I]^{\intercal} + I = 0$, which exists since $A + \varrho I$ is Hurwitz. Define the norm of $\widetilde{\theta}_n$ by $\|\widetilde{\theta}_n\|_T := \sqrt{\mathsf{E}[\widetilde{\theta}_n^{\intercal} T \widetilde{\theta}_n]}$.

First consider $\widetilde{\theta}_n^{(1)}$. Since the martingale difference Δ_{n+2}^m is uncorrelated with $\widetilde{\theta}_n^{(1)}$, denoting $e_n = \|\widetilde{\theta}_n^{(1)}\|_T^2, b_{n+2} = \|\Delta_{n+2}^m\|_T^2$, we obtain the following from (46a):

$$e_{n+1} = \|[I + \alpha_{n+1} A]\widetilde{\theta}_n^{(1)}\|_T^2 + b_{n+2} \tag{47}$$

Letting $\lambda_\circ > 0$ denote the largest eigenvalue of T, we arrive at the following simplification of the first term in (47)

$$\begin{aligned}\|[I + \alpha_{n+1} A]\widetilde{\theta}_n^{(1)}\|_T^2 &= \mathsf{E}\big[(\widetilde{\theta}_n^{(1)})^{\intercal}[T - 2\alpha_{n+1}\varrho T - \alpha_{n+1} I + \alpha_{n+1}^2 A T A^{\intercal}]\widetilde{\theta}_n^{(1)}\big] \\ &\leq \mathsf{E}\big[(\widetilde{\theta}_n^{(1)})^{\intercal}[T - 2\alpha_{n+1}\varrho T - \frac{1}{\lambda_\circ}\alpha_{n+1} T + \alpha_{n+1}^2 A T A^{\intercal}]\widetilde{\theta}_n^{(1)}\big] \\ &\leq [1 - 2\alpha_{n+1}\varrho - \alpha_{n+1}/\lambda_\circ + \alpha_{n+1}^2 L^2]\|\widetilde{\theta}_n\|_T^2\end{aligned} \tag{48}$$

where L denotes the induced operator norm of A with respect to the norm $\|\cdot\|_T$. We then obtain the following recursive bound from (47) and (48)

$$e_{n+1} \leq [1 - (2\varrho + 1/\lambda_\circ)\alpha_{n+1} + L^2 \alpha_{n+1}^2] e_n + \alpha_{n+1}^2 K$$

where $K = \sup_{n\ge1} b_n$. K is finite since b_n converges to $\mathsf{E}_\pi[(\Delta_n^m)^\intercal T \Delta_n^m]$ geometrically fast.

Consequently, for each $n \ge 1$,

$$e_{n+1} \le e_0 \prod_{k=1}^{n+1}[1-(2\varrho+1/\lambda_\circ)\alpha_k + L^2\alpha_k^2] + K\sum_{k=1}^{n+1}\alpha_k^2 \prod_{l=k+1}^{n+1}[1-(2\varrho+1/\lambda_\circ)\alpha_l + L^2\alpha_l^2]$$

By Lemma A.1,

$$e_{n+1} \le e_1 K_{A.1}\frac{1}{(n+2)^{2\varrho+1/\lambda_\circ}} + \frac{KK_{A.1}}{(n+2)^{2\varrho+1/\lambda_\circ}}\sum_{k=1}^{n+1}\alpha_k^{2-2\varrho-1/\lambda_\circ}$$

Therefore, $e_{n+1}\to 0$ at rate at least $n^{-2\varrho}$.

For $\widetilde{\theta}_n^{(2)}$, we use similar arguments. We obtain the following from (46b) by the triangle inequality.

$$\|\widetilde{\theta}_{n+1}^{(2)}\|_T \le \|[I+\alpha_{n+1}A]\widetilde{\theta}_n^{(2)}\|_T + \alpha_n\alpha_{n+1}\|[I+A]Z_{n+1}\|_T$$

Using the same argument as in (48), along with the inequality $\sqrt{1+x}\le 1+\frac{1}{2}x$,

$$\begin{aligned}\|[I+\alpha_{n+1}A]\widetilde{\theta}_n^{(2)}\|_T &\le \|\widetilde{\theta}_n^{(2)}\|_T\sqrt{1-2\alpha_{n+1}\varrho-\alpha_{n+1}/\lambda_\circ+\alpha_{n+1}^2L^2}\\ &\le \|\widetilde{\theta}_n^{(2)}\|_T(1-\alpha_{n+1}\varrho-\alpha_{n+1}/(2\lambda_\circ)+\tfrac{1}{2}\alpha_{n+1}^2L^2)\end{aligned}$$

Denote $K' = \sup_{n\ge1}\|[I+A]Z_{n+1}\|_T$.

$$\|\widetilde{\theta}_{n+1}^{(2)}\|_T \le [1-(\varrho+1/(2\lambda_\circ))\alpha_{n+1}+\tfrac{1}{2}\alpha_{n+1}^2L^2]\|\widetilde{\theta}_n^{(2)}\|_T + \alpha_n\alpha_{n+1}K'$$

Then by the same argument for the martingale difference term, we can show that $\|\widetilde{\theta}_n^{(2)}\|_T\to 0$ at rate at least $n^{-\varrho}$.

Given $\|\widetilde{\theta}_n^{(3)}\|_T = \alpha_n\|Z_{n+1}\|_T$ converges to zero at rate $1/n$, the proof is completed by the triangle inequality. □

A.2 Proof of Thm. 2.4

Denote $\mathrm{Cov}\,(\theta_n^{(i)}) = \mathsf{E}[\widetilde{\theta}_n^{(i)}(\widetilde{\theta}_n^{(i)})^\intercal]$ and $\Sigma_n^{\varrho,(i)} = \mathsf{E}[\widetilde{\theta}^{\varrho,(i)}(\widetilde{\theta}^{\varrho,(i)})^\intercal] = n^{2\varrho}\mathrm{Cov}\,(\theta_n^{(i)})$ for each i in (33). The proof proceeds by establishing the convergence rate for each $\mathrm{Cov}\,(\theta_n^{(i)})$. The main challenges are the first two: $\mathrm{Cov}\,(\theta_n^{(1)})$ and $\mathrm{Cov}\,(\theta_n^{(2)})$, for which explicit bounds are obtained by studying recursions of the scaled sequences. Bounding $\widetilde{\theta}_n^{(3)} = -\alpha_n Z_{n+1}$ is trivial.

The martingale difference term

Proposition A.3. *Under (A1)-(A3),*

(i) *If Real*$(\lambda) < -\frac{1}{2}$ *for every eigenvalue* λ *of* A*, then*

$$\mathrm{Cov}\,(\theta_n^{(1)}) = n^{-1}\Sigma_\theta + O(n^{-1-\delta})$$

where $\delta = \delta(\frac{1}{2}I + A, \Sigma_\Delta) > 0$*, and* Σ_θ *is the solution to the Lyapunov equation* (4)*.*

(ii) *Suppose there is an eigenvalue* λ *of* A*, that satisfies* $-\varrho_0 = $ *Real*$(\lambda) > -\frac{1}{2}$*. Let* $v\neq 0$ *denote the corresponding left eigenvector, and suppose moreover that* $\Sigma_\Delta v \neq 0$*. Then,* $\mathsf{E}[|v^\intercal\widetilde{\theta}_n^{(1)}|^2]$ *converges to* 0 *at rate* $n^{-2\varrho_0}$*.* □

Proof of Prop. A.3 (i) Recall that $\{\Delta_n^m\}$ is a martingale difference sequence. It is thus an uncorrelated sequence for which $\widetilde{\theta}_n^{(1)}$ and Δ_{n+k}^m are uncorrelated for $k \geq 2$. The following recursion is obtained from these facts and (29a)

$$\mathrm{Cov}\,(\theta_{n+1}^{(1)}) = \mathrm{Cov}\,(\theta_n^{(1)}) + \alpha_{n+1}\Big[\mathrm{Cov}\,(\theta_n^{(1)})A^\intercal + A\mathrm{Cov}\,(\theta_n^{(1)}) + \alpha_{n+1}[A\mathrm{Cov}\,(\theta_n^{(1)})A^\intercal + \Sigma_{\Delta_{n+2}}]\Big]$$

Multiplying each side by $n+1$ gives

$$\begin{aligned}(n+1)\mathrm{Cov}\,(\theta_{n+1}^{(1)}) =& n\mathrm{Cov}\,(\theta_n^{(1)}) + \mathrm{Cov}\,(\theta_n^{(1)}) + \mathrm{Cov}\,(\theta_n^{(1)})A^\intercal + A\mathrm{Cov}\,(\theta_n^{(1)}) \\ &+ \alpha_{n+1}[A\mathrm{Cov}\,(\theta_n^{(1)})A^\intercal + \Sigma_{\Delta_{n+2}}] \\ =& n\mathrm{Cov}\,(\theta_n^{(1)}) + \alpha_{n+1}\Big[(1+\frac{1}{n})[n\mathrm{Cov}\,(\theta_n^{(1)}) + n\mathrm{Cov}\,(\theta_n^{(1)})A^\intercal + An\mathrm{Cov}\,(\theta_n^{(1)})] \\ &+ A\mathrm{Cov}\,(\theta_n^{(1)})A^\intercal + \Sigma_{\Delta_{n+2}}\Big]\end{aligned}$$

The following argument will be used repeatedly through this Appendix: the recursion for $n\mathrm{Cov}\,(\theta_n^{(1)})$ is a *deterministic* SA recursion for $n\mathrm{Cov}\,(\theta_n^{(1)})$, and is regarded as an Euler approximation to the stable linear system

$$\tfrac{d}{dt}\mathcal{X}(t) = (1+e^{-t})[\mathcal{X}(t) + A\mathcal{X}(t) + \mathcal{X}(t)A^\intercal] + \Sigma_\Delta + e^{-t}A\mathcal{X}(t)A^\intercal \tag{49}$$

Stability follows from the assumption that $\frac{1}{2}I + A$ is Hurwitz. The standard justification of the Euler approximation is through the choice of timescale: let $t_n = \sum_{k=1}^n \alpha_k$ and let $\mathcal{X}^n(t)$ denote the solution to this ODE on $[t_n,\infty)$ with $\mathcal{X}^n(t_n) = n\mathrm{Cov}\,(\theta_n^{(1)})$, $t \geq t_n$, for any $n \geq 1$. Using standard ODE arguments (Borkar, 2008),

$$\sup_{k\geq n} \|\mathcal{X}^n(t_k) - k\Sigma_k^{(1)}\| = O(1/n)$$

Exponential convergence of $\mathcal{X}$ to Σ_θ implies convergence of $\{n\mathrm{Cov}\,(\theta_n^{(1)})\}$ to zero at rate $1/n^\delta$ for some $\delta = \delta(\frac{1}{2}I + A, \Sigma_\Delta) > 0$. □

Proof of Prop. A.3 (ii) Denote $e_n^{\varrho_0} = \mathsf{E}[|v^\intercal \widetilde{\theta}_n^{\varrho_0}|^2]$ and $\lambda = -\varrho_0 + ui$. We begin with the proof that

$$\liminf_{n\to\infty} e_n^{\varrho_0} > 0 \tag{50}$$

With $v^\intercal[I\lambda - A] = 0$, we have $v^\intercal[I\varrho_n + A(n,\varrho)] = [\varepsilon_v(n,\varrho_0) + ui]v^\intercal$, with $\varepsilon_v(n,\varrho_0) = O(n^{-1})$. Applying (34a) gives

$$v^\intercal\widetilde{\theta}_{n+1}^{\varrho_0,(1)} = v^\intercal\widetilde{\theta}_n^{\varrho_0,(1)} + \alpha_{n+1}\big[[\varepsilon_v(n,\varrho_0) + ui]v^\intercal\widetilde{\theta}_n^{\varrho_0,(1)} + (n+1)^{\varrho_0}v^\intercal\Delta_{n+2}^m\big]$$

Let $\overline{v}$ denote the conjugate of v. Consequently, with $\sigma_n^2(v) = v^\intercal\Sigma_{\Delta_n}\overline{v}$,

$$e_{n+1}^{\varrho_0} = \big[[1+\varepsilon_v(n,\varrho_0)/(n+1)]^2 + u^2/(n+1)^2\big]e_n^{\varrho_0} + (n+1)^{2\varrho_0-2}\sigma_{n+2}^2(v)$$

V-uniform ergodicity implies that $\sigma_n^2(v) \to v^\intercal\Sigma_\Delta\overline{v} > 0$ as $n\to\infty$ at a geometric rate. Fix $n_0 > 0$ so that $\sigma_{n_0}^2(v) > 0$, and hence also $e_{n_0+1}^{\varrho_0} > 0$. We also assume that $1 + \varepsilon_v(n,\varrho_0)/(n+1) > 0$ for $n \geq n_0$, which is possible since $\varepsilon_v(n,\varrho_0) = O(n^{-1})$.

For $N > n_0$ we obtain the uniform bound

$$\log(e_N^{\varrho_0}) \geq \log(e_{n_0+1}^{\varrho_0}) + 2\sum_{n=n_0+2}^{\infty}\log[1 - |\varepsilon_v(n,\varrho_0)|/(n+1)] > -\infty$$

which proves that $\liminf_{n\to\infty} e_n^{\varrho_0} = \liminf_{n\to\infty} v^\intercal \Sigma_n^{\varrho_0,(1)} \overline{v} > 0$.

The proof of an upper bound for $\varrho_0 < 1/2$: by concavity of the logarithm,

$$\log(e_{n+1}^{\varrho_0}) \le \log\bigl(\bigl[[1+\varepsilon_v(n,\varrho_0)/(n+1)]^2 + u^2/(n+1)^2\bigr] e_n^{\varrho_0}\bigr) + K(n+1)^{2\varrho_0-2}$$

where $K = \sup_{n>n_0} \bigl[[1+\varepsilon_v(n,\varrho_0)/(n+1)]^2 + u^2/(n+1)^2\bigr]^{-1} [e_n^{\varrho_0}]^{-1} \sigma_{n+2}^2(v)$. Using concavity of the logarithm once more gives

$$\log(e_{n+1}^{\varrho_0}) \le \log(e_n^{\varrho_0}) + 2\varepsilon_v(n,\varrho_0)/(n+1) + \frac{\varepsilon_v(n,\varrho_0)^2}{(n+1)^2} + \frac{u^2}{(n+1)^2} + K(n+1)^{2\varrho_0-2}$$

which gives the uniform upper bound

$$\log(e_N^{\varrho_0}) \le \log(e_{n_0+1}^{\varrho_0}) + \sum_{n=n_0+2}^{\infty} \Bigl(2\frac{|\varepsilon_v(n,\varrho_0)|}{n+1)} + \frac{\varepsilon_v(n,\varrho_0)^2}{(n+1)^2} + \frac{u^2}{(n+1)^2} + K(n+1)^{2\varrho_0-2}\Bigr) < \infty$$

This proves that $\limsup_{n\to\infty} e_n^{\varrho_0} = \limsup_{n\to\infty} v^\intercal \Sigma_n^{\varrho_0,(1)} \overline{v} < \infty$. □

The telescoping sequence term

Proposition A.4. *Under (A1)-(A3),*

(i) *If Real$(\lambda) < -\frac{1}{2}$ for every eigenvalue λ of A, then,* $\mathrm{Cov}\,(\theta_n^{(2)}) = O(n^{-1-\delta})$ *for some* $\delta = \delta(\frac{1}{2}I + A, \Sigma_\Delta) > 0$.

(ii) *Suppose there is an eigenvalue λ of A that satisfies $-\varrho_0 =$ Real$(\lambda) > -\frac{1}{2}$. Let $v \neq 0$ denote the corresponding left eigenvector, and suppose moreover that $\Sigma_\Delta v \neq 0$. Then,*

$$\limsup_{n\to\infty} n^{2\varrho_0} \mathsf{E}[|v^\intercal \widetilde{\theta}_n^{(2)}|^2] < \infty$$

□

Proof for Prop. A.4 (i) Denote $\mathcal{D}_n = \varepsilon(n,\varrho)I + A(n,\varrho) - A$. We can rewrite (34b) as

$$\begin{aligned} \widetilde{\theta}_{n+1}^{\varrho,(2)} &= \widetilde{\theta}_n^{\varrho,(2)} + \alpha_{n+1}\bigl[[\tfrac{1}{2}I + A]\widetilde{\theta}_n^{\varrho,(2)} + \mathcal{D}_n \widetilde{\theta}_n^{\varrho,(2)} - \alpha_n (n+1)^\varrho [I+A] Z_{n+1}\bigr] \\ &= \bigl[I + \alpha_{n+1}[\tfrac{1}{2}I + A]\bigr]\widetilde{\theta}_n^{\varrho,(2)} + \alpha_{n+1}\mathcal{D}_n \widetilde{\theta}_n^{\varrho,(2)} - \alpha_{n+1}\alpha_n (n+1)^\varrho [I+A] Z_{n+1} \end{aligned} \tag{51}$$

Let $T > 0$ solve the Lyapunov equation

$$[\tfrac{1}{2}I + A]^\intercal T + T[\tfrac{1}{2}I + A] + I = 0$$

As in the proof of Lemma A.2, a solution exists because $\frac{1}{2}I + A$ is Hurwitz. Adopting the familiar notation $\|\widetilde{\theta}_n^{\varrho,(2)}\|_T := \sqrt{\mathsf{E}[(\widetilde{\theta}_n^{\varrho,(2)})^\intercal T \widetilde{\theta}_n^{\varrho,(2)}]}$, the triangle inequality applied to (51) gives

$$\|\widetilde{\theta}_{n+1}^{\varrho,(2)}\|_T \le \|\bigl[I + \alpha_{n+1}[\tfrac{1}{2}I + A]\bigr]\widetilde{\theta}_n^{\varrho,(2)}\|_T + \alpha_{n+1}\|\mathcal{D}_n\|_T \|\widetilde{\theta}_n^{\varrho,(2)}\|_T + \alpha_{n+1}\alpha_n (n+1)^\varrho \|[I+A]Z_{n+1}\|_T \tag{52}$$

The first term can be simplified by the Lyapunov equation.

$$\begin{aligned} \|\bigl[I + \alpha_{n+1}[\tfrac{1}{2}I + A]\bigr]\widetilde{\theta}_n^{\varrho,(2)}\|_T^2 =& \mathsf{E}\bigl[(\widetilde{\theta}_n^{\varrho,(2)})^\intercal \bigl[T - \alpha_{n+1} I + \alpha_{n+1}^2 [\tfrac{1}{2}I + A]^\intercal T [\tfrac{1}{2}I + A]\bigr]\widetilde{\theta}_n^{\varrho,(2)}\bigr] \\ \le& \mathsf{E}\bigl[(\widetilde{\theta}_n^{\varrho,(2)})^\intercal \bigl[T - \frac{\alpha_{n+1}}{\lambda_\circ} T + \alpha_{n+1}^2 [\tfrac{1}{2}I + A]^\intercal T [\tfrac{1}{2}I + A]\bigr]\widetilde{\theta}_n^{\varrho,(2)}\bigr] \\ \le& \|\widetilde{\theta}_n^{\varrho,(2)}\|_T^2 - \frac{\alpha_{n+1}}{\lambda_\circ}\|\widetilde{\theta}_n^{\varrho,(2)}\|_T^2 + \alpha_{n+1}^2 L^2 \|\widetilde{\theta}_n^{\varrho,(2)}\|_T^2 \end{aligned}$$

where L is the induced operator norm of $\frac{1}{2}I + A$, and $\lambda_\circ > 0$ denotes its largest eigenvalue.

Consequently, by the inequality $\sqrt{1+x} \le 1 + \frac{1}{2}x$,

$$\|\big[I + \alpha_{n+1}[\tfrac{1}{2}I + A]\big]\widetilde{\theta}_n^{\varrho,(2)}\|_T \le \|\widetilde{\theta}_n^{\varrho,(2)}\|_T \sqrt{1 - \frac{\alpha_{n+1}}{\lambda_\circ} + \alpha_{n+1}^2 L^2} \le \|\widetilde{\theta}_n^{\varrho,(2)}\|_T (1 - \frac{\alpha_{n+1}}{2\lambda_\circ} + \tfrac{1}{2}\alpha_{n+1}^2 L^2)$$

Fix $n_0 > 0$ such that for $n \ge n_0$,

$$1 - \frac{\alpha_{n+1}}{2\lambda_\circ} + \tfrac{1}{2}\alpha_{n+1}^2 L^2 + \alpha_{n+1}\|\mathcal{D}_n\|_T \le 1 - \frac{\alpha_{n+1}}{4\lambda_\circ}$$

This is possible since $\|\mathcal{D}_n\|_T = O(n^{-1})$.

Denote $\delta = \min(\frac{1}{4\lambda_\circ}, \frac{1}{4})$ and $K = \sup_{n \ge n_0} \|[I + A]Z_{n+1}\|_T$, which is finite because $\|Z_{n+1}\|_T$ converges. We obtain the following from (52)

$$\begin{aligned}\|\widetilde{\theta}_{n+1}^{\varrho,(2)}\|_T \le& \|\widetilde{\theta}_n^{\varrho,(2)}\|_T (1 - \delta\alpha_{n+1}) + \alpha_{n+1}^{1/2}\alpha_n K \\ \le& \|\widetilde{\theta}_n^{\varrho,(2)}\|_T (1 - \delta\alpha_{n+1}) + \alpha_n^{3/2} K\end{aligned} \tag{53}$$

Apply (53) repeatedly for $n \ge n_0$

$$\begin{aligned}\|\widetilde{\theta}_{n+1}^{\varrho,(2)}\|_T \le& \|\widetilde{\theta}_{n_0}^{\varrho,(2)}\|_T \prod_{k=n_0+1}^{n+1}(1 - \delta\alpha_k) + K\sum_{k=n_0}^{n}\alpha_k^{3/2}\prod_{l=k+1}^{n}(1 - \delta\alpha_l) \\ \le& \|\widetilde{\theta}_{n_0}^{\varrho,(2)}\|_T \exp(\delta)\frac{n_0^\delta}{(n+2)^\delta} + \frac{K\exp(\delta)}{(n+1)^\delta}\sum_{k=n_0}^{n} k^{-\frac{3}{2}+\delta}\end{aligned}$$

where $\sum_{k=1}^\infty k^{-\frac{3}{2}+\delta} < \infty$ for $\delta \le 1/4$. Therefore, $\|\widetilde{\theta}_n^{\varrho,(2)}\|_T \to 0$ at rate at least $n^{-\delta}$.

The desired conclusion follows: letting $\lambda_\bullet > 0$ denote the smallest eigenvalue of T,

$$\Sigma_n^{\varrho,(2)} \le \mathsf{E}[(\widetilde{\theta}_n^{\varrho,(2)})^\intercal \widetilde{\theta}_n^{\varrho,(2)}]I \le \frac{1}{\lambda_\bullet}\|\widetilde{\theta}_n^{\varrho,(2)}\|_T^2 I$$

□

Proof for Prop. A.4 (ii) Multiplying both sides of (34b) by $v^\intercal$ gives

$$\begin{aligned}v^\intercal\widetilde{\theta}_{n+1}^{\varrho_0,(2)} &= v^\intercal\widetilde{\theta}_n^{\varrho_0,(2)} + \alpha_{n+1}\big[[\varepsilon_v(n,\varrho_0) + ui]v^\intercal\widetilde{\theta}_n^{\varrho_0,(2)} - (1 - \varrho_0 + ui)\alpha_n(n+1)^{\varrho_0}v^\intercal Z_{n+1}\big] \\ &= \big[1 + \alpha_{n+1}[\varepsilon_v(n,\varrho_0) + ui]\big]v^\intercal\widetilde{\theta}_n^{\varrho_0,(2)} - (1 - \varrho_0 + ui)\alpha_n\alpha_{n+1}(n+1)^{\varrho_0}v^\intercal Z_{n+1}\end{aligned} \tag{54}$$

With $\|v^\intercal\widetilde{\theta}_n^{\varrho_0,(2)}\|_2 := \sqrt{\mathsf{E}[|v^\intercal\widetilde{\theta}_n^{\varrho_0,(2)}|^2]}$, we obtain the following from (54) by the triangle inequality

$$\|v^\intercal\widetilde{\theta}_{n+1}^{\varrho_0,(2)}\|_2 \le \big|1 + \alpha_{n+1}[\varepsilon_v(n,\varrho_0) + ui]\big|\|v^\intercal\widetilde{\theta}_n^{\varrho_0,(2)}\|_2 + \big|1 - \varrho_0 + ui\big|\alpha_n\alpha_{n+1}(n+1)^{\varrho_0}\|v^\intercal Z_{n+1}\|_2 \tag{55}$$

By the inequality $\sqrt{1+x} \le 1 + \frac{1}{2}x$, we have

$$\big|1 + \alpha_{n+1}\varepsilon_v(n,\varrho_0) + \alpha_{n+1}ui\big| \le 1 + \alpha_{n+1}\varepsilon_v(n,\varrho_0) + \tfrac{1}{2}\alpha_{n+1}^2\varepsilon_v(n,\varrho_0)^2 + \tfrac{1}{2}\alpha_{n+1}^2 u^2$$

Fix $n_0 > 0$ such that for $n \ge n_0$,

$$1 + \alpha_{n+1}\varepsilon_v(n,\varrho_0) + \tfrac{1}{2}\alpha_{n+1}^2\varepsilon_v(n,\varrho_0)^2 + \tfrac{1}{2}\alpha_{n+1}^2 u^2 \le 1 + \alpha_{n+1}^{3/2}$$

which is possible since $\varepsilon_v(n,\varrho_0) = O(n^{-1})$. With $K = \sup_{n\geq n_0} |1-\varrho_0+ui| \|v^\intercal Z_{n+1}\|_2$, we obtain the following bound from (55):

$$\|v^\intercal \widetilde{\theta}^{\varrho_0,(2)}_{n+1}\|_2 \leq (1+\alpha_{n+1}^{3/2})\|v^\intercal \widetilde{\theta}^{\varrho_0,(2)}_{n}\|_2 + \alpha_n^{2-\varrho_0} K \tag{56}$$

Iterating (56) gives,

$$\begin{aligned}\|v^\intercal \widetilde{\theta}^{\varrho_0,(2)}_{n+1}\|_2 &\leq \|v^\intercal \widetilde{\theta}^{\varrho_0,(2)}_{n_0}\|_2 \prod_{k=n_0+1}^{n+1} (1+\alpha_k^{3/2}) + K\sum_{k=n_0}^{n} \alpha_k^{2-\varrho_0} \prod_{l=k+1}^{n} (1+\alpha_l^{3/2}) \\ &\leq \|v^\intercal \widetilde{\theta}^{\varrho_0,(2)}_{n_0}\|_2 \exp(\sum_{k=n_0+1}^{n+1} k^{-2/3}) + K\sum_{k=n_0}^{n} k^{-2+\varrho_0} \exp(\sum_{l=k+1}^{n} l^{-3/2})\end{aligned}$$

$\limsup_{n\to\infty} \|v^\intercal \widetilde{\theta}^{\varrho_0,(2)}_{n})\|_2 < \infty$, since it is assumed that $\varrho_0 < \frac{1}{2}$. □

Proof of Thm. 2.4 We obtain the convergence rate of $\mathrm{Cov}\,(\theta_n)$ based on

$$\mathrm{Cov}\,(\theta_n) = \sum_{i=1}^{3} \mathrm{Cov}\,(\theta_n^{(i)}) + \sum_{i=1}^{3}\sum_{j=1,j\neq i}^{3} \mathsf{E}[\widetilde{\theta}_n^{(i)}(\widetilde{\theta}_n^{(j)})^\intercal]$$

For case (i), by Prop. A.3 (i) and Prop. A.4 (i), there exists $\delta = \delta(\frac{1}{2}I + A, \Sigma_\Delta) > 0$ such that

$$\begin{aligned}\mathrm{Cov}\,(\theta_n^{(1)}) &= n^{-1}\Sigma_\theta + O(n^{-1-\delta}) \\ \mathrm{Cov}\,(\theta_n^{(2)}) &= O(n^{-1-\delta}) \\ \mathrm{Cov}\,(\theta_n^{(3)}) &= n^{-2}\Sigma_{Z_{n+1}}\end{aligned}$$

The cross terms between $\widetilde{\theta}_n^{(i)}$ and $\widetilde{\theta}_n^{(j)}$ for $i \neq j$ are of smaller orders than $O(1/n)$ by the Cauchy-Schwarz inequality. Therefore, for a possibly smaller $\delta > 0$,

$$\mathrm{Cov}\,(\theta_n) = n^{-1}\Sigma_\theta + O(n^{-1-\delta})$$

For case (ii), $\lim_{n\to 0} n^{2\varrho}\mathsf{E}[|v^\intercal \widetilde{\theta}_n|^2] = 0$ for each $\varrho < \varrho_0$ can be obtained from Prop. A.3 (ii) and Prop. A.4 (ii) directly by the triangle inequality. For $\varrho > \varrho_0$, the result $\lim_{n\to 0} n^{2\varrho}\mathsf{E}[|v^\intercal \widetilde{\theta}_n|^2] = \infty$ is established independently in Lemma A.13. □

A.3 Proof of Thm. 2.8

Denote the correlation between $\widetilde{\theta}_n^{(a)}$ and $\widetilde{\theta}_n^{(b)}$ as $R_n^{(a),(b)} = \mathsf{E}[\widetilde{\theta}_n^{(a)}(\widetilde{\theta}_n^{(b)})^\intercal]$, where $\widetilde{\theta}_n^{(a)}, \widetilde{\theta}_n^{(b)}$ are different terms in (42). The key results that help establish Thm. 2.8 are summarized in the following proposition.

Proposition A.5. *Under Assumptions (A1)-(A3), if Real$(\lambda) < -1$ for every eigenvalue of A, then there is $\delta > 0$ such that*

(i) $\mathrm{Cov}\,(\theta_n^{(1)}) = n^{-1}\Sigma_\theta + n^{-2}\Sigma_\sharp^{(1)} + O(n^{-2-\delta})$, *where $\delta = \delta(I + A, \Sigma_\Delta) > 0$, $\Sigma_\theta \geq 0$ is the unique solution to the Lyapunov equation (4), and $\Sigma_\sharp^{(1)} \geq 0$ solves the Lyapunov equation,*

$$[I + A]\Sigma + \Sigma[I + A]^\intercal + A\Sigma_\theta A^\intercal - \Sigma_\Delta = 0 \tag{57}$$

(ii) $R_n^{(2,1),(1)} + R_n^{(1),(2,1)} = n^{-2}\Sigma_\sharp^{(2)} + O(n^{-2-\delta})$, *where $\Sigma_\sharp^{(2)}$ solves the Lyapunov equation:*

$$[I + A]\Sigma + \Sigma[I + A]^\intercal - [I + A]\mathrm{Cov}_\pi(\widehat{\Delta}_n^m, \Delta_n^m) - \mathrm{Cov}_\pi(\Delta_n^m, \widehat{\Delta}_n^m)[I + A]^\intercal = 0 \tag{58}$$

(iii) $R_n^{(1),(3)} = -n^{-2}\mathsf{E}_\pi[\Delta_n^m \widehat{Z}_n^\intercal] + O(n^{-3})$.

□

Proof of Prop. A.5 (i) Since Δ^m_{n+2} is uncorrelated with $\widetilde{\theta}^{(1)}_n$, the following recursion follows from (29a):

$$\mathrm{Cov}\,(\theta^{(1)}_{n+1}) = \mathrm{Cov}\,(\theta^{(1)}_n) + \alpha_{n+1}\Big[\mathrm{Cov}\,(\theta^{(1)}_n)A^\intercal + A\mathrm{Cov}\,(\theta^{(1)}_n) + \alpha_{n+1}[A\mathrm{Cov}\,(\theta^{(1)}_n)A^\intercal + \Sigma_{\Delta_{n+2}}]\Big]$$

Take $\varrho = 1/2$ in the definition of $\widetilde{\theta}^{\varrho,(1)}$ and $\Sigma^{\varrho,(1)}_n = \mathsf{E}[\widetilde{\theta}^{\varrho,(1)}(\widetilde{\theta}^{\varrho,(1)})^\intercal] = n\mathrm{Cov}\,(\theta^{(1)}_n)$. Multiplying each side of the equation by $n+1$ gives

$$\Sigma^{\varrho,(1)}_{n+1} = \Sigma^{\varrho,(1)}_n + \alpha_{n+1}\Big[(1+\frac{1}{n})\big[\Sigma^{\varrho,(1)}_n + \Sigma^{\varrho,(1)}_n A^\intercal + A\Sigma^{\varrho,(1)}_n\big] + \frac{1}{n}A\Sigma^{\varrho,(1)}_n A^\intercal + \Sigma_{\Delta_{n+2}}\Big] \tag{59}$$

Recall that Σ_θ solves the Laypunov equation $\Sigma + \Sigma A^\intercal + A\Sigma + \Sigma_\Delta = 0$. Denoting $E_n = \Sigma^{\varrho,(1)}_n - \Sigma_\theta$, the following identity holds

$$\Sigma^{\varrho,(1)}_n + \Sigma^{\varrho,(1)}_n A^\intercal + A\Sigma^{\varrho,(1)}_n = E_n + E_n A^\intercal + AE_n - \Sigma_\Delta$$

Subtracting Σ_θ from both sides of (59) gives the recursion

$$\begin{aligned} E_{n+1} = E_n + \alpha_{n+1}\Big[(1+\frac{1}{n})\big[E_n + E_n A^\intercal + AE_n\big] + \frac{1}{n}AE_n A^\intercal \\ + \frac{1}{n}A\Sigma_\theta A^\intercal - \frac{1}{n}\Sigma_\Delta - \Sigma_\Delta + \Sigma_{\Delta_{n+2}}\Big] \end{aligned} \tag{60}$$

Similar to the decomposition in (29), we have $E_n = E^{(1)}_n + E^{(2)}_n$, each evolving as

$$E^{(1)}_{n+1} = E^{(1)}_n + \alpha_{n+1}\Big[(1+\frac{1}{n})\big[E^{(1)}_n + E^{(1)}_n A^\intercal + AE^{(1)}_n\big] + \frac{1}{n}AE^{(1)}_n A^\intercal + \frac{1}{n}\big[A\Sigma_\theta A^\intercal - \Sigma_\Delta\big]\Big] \tag{61a}$$

$$E^{(2)}_{n+1} = E^{(2)}_n + \alpha_{n+1}\Big[(1+\frac{1}{n})\big[E^{(2)}_n + E^{(2)}_n A^\intercal + AE^{(2)}_n\big] + \frac{1}{n}AE^{(2)}_n A^\intercal + \Sigma_{\Delta_{n+2}} - \Sigma_\Delta\Big] \tag{61b}$$

Since $\Sigma_{\Delta_{n+2}} - \Sigma_\Delta$ converges to zero geometrically fast, $\{E^{(1)}_n\}$ converges to zero faster than $\{E^{(2)}_n\}$.

Multiplying each side of (61a) by $n+1$ gives

$$\begin{aligned} (n+1)E^{(1)}_{n+1} &= (n+1)E^{(1)}_n + (1+\frac{1}{n})\big[E^{(1)}_n + E^{(1)}_n A^\intercal + AE^{(1)}_n\big] + \frac{1}{n}\big[AE^{(1)}_n A^\intercal + A\Sigma_\theta A^\intercal - \Sigma_\Delta\big] \\ &= nE^{(1)}_n + \frac{1}{n}\Big[(1+\frac{1}{n})\big[2nE^{(1)}_n + nE^{(1)}_n A^\intercal + AnE^{(1)}_n\big] + A\Sigma_\theta A^\intercal - \Sigma_\Delta + \mathcal{E}^{\bullet,(1)}_n\Big] \end{aligned}$$

with the error term $\mathcal{E}^{\bullet,(1)}_n = AE^{(1)}_n A^\intercal - E_n$. Note that $A\Sigma_\theta A^\intercal - \Sigma_\Delta = [A+I]\Sigma_\theta[A+I]^\intercal$ is positive definite.

The recursion for $\{nE^{(1)}_n\}$ is treated as in the proof of Prop. A.3 (i). Consider the matrix ODE,

$$\tfrac{d}{dt}\mathcal{X}(t) = (1+e^{-t})[2\mathcal{X}(t) + \mathcal{X}(t)A^\intercal + A\mathcal{X}(t)] + A\Sigma_\theta A^\intercal - \Sigma_\Delta + e^{-t}[A\mathcal{X}(t)A^\intercal - \mathcal{X}(t)] \tag{62}$$

Let $t_n = \sum_{k=1}^n 1/k$ and let $\mathcal{X}^n(t)$ denote the solution to this ODE on $[t_n,\infty)$ with $\mathcal{X}^n(t_n) = nE^{(1)}_n$, $t \ge t_n$, for any $n \ge 1$. We then obtain as previously,

$$\sup_{k\ge n}\|\mathcal{X}^n(t_k) - kE^{(1)}_k\| = O(1/n)$$

Recall that $\Sigma^{(1)}_\sharp \ge 0$ is the solution to the Lyapunov equation (57). Exponential convergence of $\mathcal{X}$ to $\Sigma^{(1)}_\sharp$ implies convergence of $\{nE^{(1)}_n\}$ at rate $1/n^\delta$ for $\delta = \delta(A+I,\Sigma_\Delta) > 0$. Therefore, $nE_n = \Sigma^{(1)}_\sharp + O(n^{-\delta})$.

Given $\mathrm{Cov}\,(\theta^{(1)}_n) = n^{-1}\Sigma_\theta + n^{-1}E_n$, we have

$$\mathrm{Cov}\,(\theta^{(1)}_n) = n^{-1}\Sigma_\theta + n^{-2}\Sigma^{(1)}_\sharp + O(n^{-2-\delta})$$

□

Proof of Prop. A.5 (ii) We focus on $R_n^{(2,1),(1)}$ since $R_n^{(1),(2,1)} = [R_n^{(2,1),(1)}]^\intercal$. Recall the update forms of $\widetilde{\theta}_n^{(1)}$ and $\widetilde{\theta}_n^{(2,1)}$ in (29a) and (41a) respectively, where $\widetilde{\theta}_n^{(1)}$ is uncorrelated with the martingale difference sequence $\{\widehat{\Delta}_{n+k}^m\}$ for $k \geq 2$ and $\widetilde{\theta}_n^{(2,1)}$ is uncorrelated with $\{\Delta_{n+k}^m\}$ for $k \geq 2$. With $R_n^{(2,1),(1)} = \mathsf{E}[\widetilde{\theta}_n^{(2,1)}(\widetilde{\theta}_n^{(1)})^\intercal]$, the following is obtained from these facts:

$$
\begin{aligned}
R_{n+1}^{(2,1),(1)} = R_n^{(2,1),(1)} + \alpha_{n+1}\big[&R_n^{(2,1),(1)}A^\intercal + AR_n^{(2,1),(1)} + \alpha_{n+1}AR_n^{(2,1),(1)}A^\intercal \\
&- \alpha_n\alpha_{n+1}[I + A]\mathrm{Cov}\,(\widehat{\Delta}_{n+2}^m, \Delta_{n+2}^m)\big]
\end{aligned}
$$

Denote $C_n = nR_n^{(2,1),(1)}$. Multiplying both sides of the previous equation by $n+1$ gives

$$
C_{n+1} = C_n + \alpha_{n+1}\big[(1+n^{-1})[C_n + C_nA^\intercal + AC_n] + \alpha_nAC_nA^\intercal - \alpha_n[I + A]\mathrm{Cov}\,(\widehat{\Delta}_{n+2}^m, \Delta_{n+2}^m)\big]
$$

Multiplying each side of this equation by $n+1$ once more results in

$$
\begin{aligned}
(n+1)C_{n+1} &= (n+1)C_n + (1+n^{-1})[C_n + C_nA^\intercal + AC_n] + \alpha_nAC_nA^\intercal - \alpha_n[I + A]\mathrm{Cov}\,(\widehat{\Delta}_{n+2}^m, \Delta_{n+2}^m) \\
&= nC_n + \alpha_n\big[(1+n^{-1})[2nC_n + nC_nA^\intercal + AnC_n] - [I + A]\mathrm{Cov}_{\,\pi}(\widehat{\Delta}_{n+2}^m, \Delta_{n+2}^m) + \mathcal{D}_{n+1}^{(2)}\big]
\end{aligned}
$$

where the error term $\mathcal{D}_{n+1}^{(2)}$ consists of two components: $[I+A][\mathrm{Cov}_{\,\pi}(\widehat{\Delta}_{n+2}^m, \Delta_{n+2}^m) - \mathrm{Cov}\,(\widehat{\Delta}_{n+2}^m, \Delta_{n+2}^m)]$ that converges to zero at a geometric rate and $AC_nA^\intercal - C_n$.

As previously, this is approximated by the linear system

$$
\begin{aligned}
\tfrac{d}{dt}\mathcal{X}(t) =& (1+e^{-t})[2\mathcal{X}(t) + \mathcal{X}(t)A^\intercal + A\mathcal{X}(t)] + e^{-t}[A\mathcal{X}(t)A^\intercal - \mathcal{X}(t)] \\
&- [I + A]\mathrm{Cov}_{\,\pi}(\widehat{\Delta}_{n+2}^m, \Delta_{n+2}^m))
\end{aligned}
\tag{63}
$$

With the same argument used in (i), $\{nC_n + nC_n^\intercal\}$ converges to $\Sigma_\sharp^{(2)}$ in (58) at rate $1/n^\delta$ for $\delta = \delta(A+I) > 0$. Therefore, $nC_n + nC_n^\intercal = \Sigma_\sharp^{(2)} + O(n^{-\delta})$ and $R_n^{(2,1),(1)} = n^{-2}C_n = n^{-2}\Sigma_{\infty,C} + O(n^{-2-\delta})$. ☐

Proof of Prop. A.5 (iii) The third claim in Prop. A.5 is established through a sequence of lemmas. Start with the representation of $\widetilde{\theta}_{n+1}^{(3)}$ based on (39):

$$
\widetilde{\theta}_{n+1}^{(3)} = -\frac{1}{n+1}Z_{n+2} = -\frac{1}{n+1}\widehat{\Delta}_{n+3}^m + \frac{1}{n+1}(\widehat{Z}_{n+3} - \widehat{Z}_{n+2})
$$

Since $\widehat{\Delta}_{n+3}^m$ is uncorrelated with the sequence $\{\widetilde{\theta}_k^{(1)}\}$ for $k \leq n+1$, we have

$$
\mathsf{E}[\widetilde{\theta}_{n+1}^{(1)}(\widehat{\Delta}_{n+3}^m)^\intercal] = 0 \tag{64}
$$

Hence it suffices to consider the correlation between $\widetilde{\theta}_{n+1}^{(1)}$ and $\widehat{Z}_{n+3} - \widehat{Z}_{n+2}$. The formula for $\widetilde{\theta}_{n+1}^{(1)}$ for $n \geq 1$ is

$$
\widetilde{\theta}_{n+1}^{(1)} = \prod_{k=1}^{n+1}[I + \alpha_kA]\widetilde{\theta}_0 + \sum_{k=1}^{n+1}\alpha_k\prod_{l=k+1}^{n+1}[I + \alpha_lA]\Delta_{k+1}^m \tag{65}
$$

$\widetilde{\theta}_0\mathsf{E}[\widehat{Z}_{n+3}^\intercal - \widehat{Z}_{n+2}^\intercal]$ converges to zero geometrically fast under V-uniform ergodicity of $\boldsymbol{\Phi}$. Then we consider the expectation of the following:

$$
\begin{aligned}
&\sum_{k=1}^{n+1}\alpha_k\prod_{l=k+1}^{n+1}[I + \alpha_lA]\Delta_{k+1}^m[\widehat{Z}_{n+3}^\intercal - \widehat{Z}_{n+2}^\intercal] \\
&= \sum_{k=1}^{n+1}\alpha_k\prod_{l=k+1}^{n+1}[I + \alpha_lA]\big[\Delta_{k+2}^m\widehat{Z}_{n+3}^\intercal - \Delta_{k+1}^m\widehat{Z}_{n+2}^\intercal\big] + \sum_{k=1}^{n+1}\alpha_k\prod_{l=k+1}^{n+1}[I + \alpha_lA]\big[\Delta_{k+1}^m - \Delta_{k+2}^m\big]\widehat{Z}_{n+3}^\intercal
\end{aligned}
\tag{66}
$$

The definition of T is now based on the assumption that $I+A$ is Hurwitz: $T>0$ is the unique solution to the Lyapunov equation:

$$[A+I]T + T[A+I]^{\intercal} + I = 0$$

As previously, we denote $\|W\|_T^2 = \mathsf{E}[W^{\intercal} T W]$ for a random vector W, and denote by $\|M\|_T$ the induced operator norm of a matrix $M \in \mathbb{R}^{d\times d}$. In the following result the vector W is taken to be deterministic.

Lemma A.6. *Suppose the matrix $I+A$ is Hurwitz. Then there exists constant K such that the following holds for any $k\ge 1$ and all $n \ge k$*

$$\Big\| \prod_{l=k+1}^{n+1} [I+\alpha_l A] \Big\|_T \le K \frac{k}{n+2}$$

Proof. For any vector $W \in \mathbb{R}^d$ and $l \ge 1$, we have

$$\begin{aligned} \|[I+\alpha_l A]W\|_T^2 &= W^{\intercal}[T - 2\alpha_l T - \alpha_l I + \alpha_l^2 A^{\intercal} T A]W \\ &\le W^{\intercal}[TT - 2\alpha_l T + \alpha_l^2 A^{\intercal} T A]W \\ &\le (1 - 2\alpha_l + \alpha_l^2 L^2)\|W\|_T^2 \end{aligned}$$

where $L = \|A\|_T$. Hence

$$\|I+\alpha_l A\|_T \le \sqrt{1 - 2\alpha_l + \alpha_l^2 L^2} \le 1 - \alpha_l + \tfrac{1}{2}\alpha_l^2 L^2$$

Lemma A.1 completes the proof:

$$\Big\| \prod_{l=k+1}^{n+1} [I+\alpha_l A] \Big\|_T \le \prod_{l=k+1}^{n+1} \|[I+\alpha_l A]\|_T \le \prod_{l=k+1}^{n+1} [1 - \alpha_l + \tfrac{1}{2} L^2 \alpha_l^2] \le K_{A.1} \frac{k}{n+2}$$

□

To analyze $\mathsf{E}[\Delta_{k+2}^m \widehat{Z}_{n+3}^{\intercal}]$, consider the bivariate Markov chain $\Phi_n^* = (\Phi_n, \Phi_{n+1})$, $n \ge 0$, with state space $\mathsf{Z}^* = \mathsf{Z}\times\mathsf{Z}$. An associated *weighting function* $V^* : \mathsf{Z}\times\mathsf{Z} \to [1,\infty)$ is defined as $V^*(z,z') = V(z) + V(z')$.

Denote function $h^{k+1,n+2} : \mathsf{Z}^* \to \mathbb{R}^{d\times d}$ as $h^{k+1,n+2}(z',z'') = (\hat{f}(z'') - \mathsf{E}[\hat{f}(\Phi_{k+1}) \mid \Phi_k = z'])\mathsf{E}[\hat{\hat{f}}(\Phi_{n+2})^{\intercal} \mid \Phi_{k+1} = z'']$ and $h_{i,j}^{k+1,n+2} : \mathsf{Z}^* \to \mathbb{R}$ as the (i,j)-th entry of $h^{k+1,n+2}$ for $1 \le i,j \le d$. Note that $h^{k+1,n+2}(\Phi_k, \Phi_{k+1}) = \mathsf{E}[\Delta_{k+1}^m \widehat{Z}_{n+2} \mid \mathcal{F}_{k+1}]$

Lemma A.7. *Suppose Assumptions (A1) and (A3) hold. For each $1 \le i,j \le d$,*

(i) $h_{i,j}^{k+1,n+2} \in L_\infty^{V^*}$*, moreover there exists constant B such that*

$$\|h_{i,j}^{k+1,n+2}\|_{V^*} \le B \|\hat{f}_i\|_{\sqrt{V}} \|\hat{\hat{f}}_j\|_{\sqrt{V}} \rho^{n-k+1}$$

(ii) *Consequently, there exists constant B' such that*

$$\big|\mathsf{E}[h_{i,j}^{k+1,n+2}(\Phi_k,\Phi_{k+1}) \mid \Phi_0 = z] - \pi\big(h_{i,j}^{k+1,n+2}\big)\big| \le B' \|\hat{f}_i\|_{\sqrt{V}} \|\hat{\hat{f}}_j\|_{\sqrt{V}} V(z) \rho^{n+1}$$

Proof. By the definition of V^*-norm,

$$
\begin{aligned}
\|h_{i,j}^{k+1,n+2}\|_{V^*} &= \sup_{z',z''\in \mathsf{Z}} \frac{\big|\big[\hat{f}_i(z'') + \mathsf{E}[\hat{f}_i(\Phi_{k+1}) \mid \Phi_k = z']\big]\mathsf{E}[\hat{\hat{f}}_j(\Phi_{n+2}) \mid \Phi_{k+1} = z'']\big|}{V(z') + V(z'')} \\
&\le \sup_{z''\in \mathsf{Z}} \frac{\big|\hat{f}_i(z'')\mathsf{E}[\hat{\hat{f}}_j(\Phi_{n+2}) \mid \Phi_{k+1} = z'']\big|}{V(z'')} \\
&\quad + \sup_{z',z''\in \mathsf{Z}} \frac{\big|\mathsf{E}[\hat{f}_i(\Phi_{k+1}) \mid \Phi_k = z']\mathsf{E}[\hat{\hat{f}}_j(\Phi_{n+2}) \mid \Phi_{k+1} = z'']\big|}{V(z') + V(z'')}
\end{aligned}
$$

Given $\hat{\hat{f}}_j^2 \in L_\infty^V$ and the $\sqrt{V}$-uniform ergodicity of $\mathbf{\Phi}$ (Meyn and Tweedie, 2009, Lemma 15.2.9), there exists constant $B_{\sqrt{V}} < \infty$ such that

$$
\big|\mathsf{E}[\hat{\hat{f}}_j(\Phi_{n+2}) \mid \Phi_{k+1} = z'']\big| \le B_{\sqrt{V}} \|\hat{\hat{f}}_j\|_{\sqrt{V}} \sqrt{V(z'')}\rho^{n+1-k}
$$

Consequently,

$$
\sup_{z''\in \mathsf{Z}} \frac{|\hat{f}_i(z'')\big[\mathsf{E}[\hat{\hat{f}}_j(\Phi_{n+2}) \mid \Phi_{k+1} = z'']|}{V(z'')} \le \|\hat{f}_i\|_{\sqrt{V}} B_{\sqrt{V}} \|\hat{\hat{f}}_j\|_{\sqrt{V}} \rho^{n+1-k} \tag{67}
$$

By the inequality $V(z') + V(z'') \ge \sqrt{V(z')V(z'')}$ and the $\sqrt{V}$-uniform ergodicity of $\mathbf{\Phi}$ once more, we have

$$
\begin{aligned}
&\sup_{z',z''\in \mathsf{Z}} \frac{\big|\mathsf{E}[\hat{f}_i(\Phi_{k+1}) \mid \Phi_k = z']\mathsf{E}[\hat{\hat{f}}_j(\Phi_{n+2}) \mid \Phi_{k+1} = z'']\big|}{V(z') + V(z'')} \\
&\le \sup_{z'\in \mathsf{Z}} \frac{\big|\mathsf{E}[\hat{f}_i(\Phi_{k+1}) \mid \Phi_k = z']\big|}{\sqrt{V(z')}} \sup_{z''\in \mathsf{Z}} \frac{\big|\mathsf{E}[\hat{\hat{f}}_j(\Phi_{n+2}) \mid \Phi_{k+1} = z'']\big|}{\sqrt{V(z'')}} \le B_{\sqrt{V}}^2 \|\hat{f}_i\|_{\sqrt{V}} \|\hat{\hat{f}}_j\|_{\sqrt{V}} \rho^{n+2-k}
\end{aligned} \tag{68}
$$

Combining (67) and (68) gives

$$
\|h_{i,j}^{k+1,n+2}\|_{V^*} \le B \|\hat{f}_i\|_{\sqrt{V}} \|\hat{\hat{f}}_j\|_{\sqrt{V}} \rho^{n+1-k} \tag{69}
$$

with $B = B_{\sqrt{V}} + B_{\sqrt{V}}^2$.

For (ii), denote $g_{i,j}^{k,n+2} : \mathsf{Z} \to \mathbb{R}$ by the conditional expectation:

$$
g_{i,j}^{k,n+2}(z) = \mathsf{E}[h_{i,j}^{k+1,n+2}(\Phi_k, \Phi_{k+1}) \mid \Phi_k = z]
$$

This is bounded by a constant times V^*:

$$
\begin{aligned}
|g_{i,j}^{k,n+2}(z)| = \Big|\int h_{i,j}^{k+1,n+2}(z,z')P(z,dz')\Big| &\le \Big|\int \frac{h_{i,j}^{k+1,n+2}(z,z')}{V^*(z,z')} V^*(z,z')P(z,dz')\Big| \\
&\le \|h_{i,j}^{k+1,n+2}\|_{V^*}[V(z) + PV(z)]
\end{aligned}
$$

V-uniform ergodicity of $\mathbf{\Phi}$ is equivalent to the following drift condition (Meyn and Tweedie, 2009, Theorem 16.0.2): for some $\beta > 0, b < \infty$, and some “petite set” C,

$$
PV(z) - V(z) \le -\beta V(z) + b\mathbb{I}_C(z)\,, \qquad z \in \mathsf{Z}
$$

Consequently,

$$
[V(z) + PV(z)] \le [2V(z) + b] \le [2 + |b|]V(z)
$$

Therefore,

$$\|g_{i,j}^{k,n+2}\|_V \le [2+|b|]\|h_{i,j}^{k+1,n+2}\|_{V^*} \le [2+|b|]B\|\hat{f}_i\|_{\sqrt{V}}\|\hat{\hat{f}}_j\|_{\sqrt{V}}\rho^{n+1-k} \tag{70}$$

Thus $g_{i,j}^{k,n+2} \in L_\infty^V$. By V-uniform ergodicity of $\mathbf{\Phi}$ again,

$$\begin{aligned}\big|\mathsf{E}[g_{i,j}^{k,n+2}(\Phi_k) \mid \Phi_0 = z] - \pi\big(g_{i,j}^{k,n+2}\big)\big| &\le B_V\|g_{i,j}^{k,n+2}\|_V V(z)\rho^k \\ &\le B'\|\hat{f}_i\|_{\sqrt{V}}\|\hat{\hat{f}}_j\|_{\sqrt{V}}V(z)\rho^{n+1}\end{aligned}$$

with $B' = [2+|b|]B_V B$. The proof is then completed by applying the smoothing property of conditional expectation. □

Lemma A.8. *Under Assumptions (A1) and (A3), there exists $K < \infty$ such that the following hold*

$$\big\|\mathsf{E}[\Delta_{k+1}^m \widehat{Z}_{n+3}^\intercal]\big\|_T \le K\rho^{n+1-k} \tag{71a}$$

$$\big\|\mathsf{E}[\Delta_{k+1}^m \widehat{Z}_{n+2}^\intercal] - \mathsf{E}[\Delta_{k+2}^m \widehat{Z}_{n+3}^\intercal]\big\|_T \le K(1+\rho)\rho^{n+1} \tag{71b}$$

Proof. By the triangle inequality,

$$\big\|\mathsf{E}[\Delta_{k+1}^m \widehat{Z}_{n+2}^\intercal]\big\|_T \le \big\|\mathsf{E}[Z_{k+1}\widehat{Z}_{n+2}^\intercal]\big\|_T + \big\|\mathsf{E}\big[\mathsf{E}[Z_{k+1}|\mathcal{F}_k]\widehat{Z}_{n+2}^\intercal\big]\big\|_T$$

where both terms admit the geometric bound in (71a) following directly from the V-geometric mixing of $\mathbf{\Phi}$ (Meyn and Tweedie, 2009, Theorem 16.1.5).

For (71b), first notice that

$$\mathsf{E}[\Delta_{k+1}^m \widehat{Z}_{n+2}^\intercal] = \mathsf{E}\big[\mathsf{E}[\Delta_{k+1}^m \widehat{Z}_{n+2}^\intercal \mid \mathcal{F}_{k+1}]\big] = \mathsf{E}[h^{k+1,n+2}(\Phi_k, \Phi_{k+1})]$$

With Lemma A.7, we have for each (i,j)-th entry,

$$\Big|\mathsf{E}[h_{i,j}^{k+1,n+2}(\Phi_k, \Phi_{k+1}) \mid \Phi_0 = z] - \pi\big(h_{i,j}^{k+1,n+2}\big)\Big| \le B'\|\hat{f}_i\|_{\sqrt{V}}\|\hat{\hat{f}}_j\|_{\sqrt{V}}V(z)\rho^{n+1}$$

With fixed initial condition $\Phi_0 = z$, by equivalence of matrix norms, there exists a constant K such that

$$\Big\|\mathsf{E}[h^{k+1,n+2}(\Phi_k, \Phi_{k+1})] - \pi\big(h_{i,j}^{k+1,n+2}\big)\Big\|_T \le K\rho^{n+1}$$

(71b) then follows from the triangle inequality:

$$\big\|\mathsf{E}[\Delta_{k+1}^m \widehat{Z}_{n+2}^\intercal] - \mathsf{E}[\Delta_{k+2}^m \widehat{Z}_{n+3}^\intercal]\big\|_T \le K\rho^{n+1} + K\rho^{n+2} = K(1+\rho)\rho^{n+1}$$

□

Lemma A.9. *For fixed $\rho \in (0,1)$, there exists $K < \infty$ such that for all $n \ge 2$,*

$$\sum_{k=1}^{n-1}\frac{1}{k}\rho^{-k} \le K\frac{\rho^{-n}}{n}$$

Proof. Denote $\gamma = -\log\rho > 0$ and observe that the function $t^{-1}\exp(\gamma t)$ is increasing over $[1,\infty)$. The following holds for $n \ge 2$

$$\sum_{k=1}^{n-1}\frac{1}{k}\rho^{-k} = \sum_{k=1}^{n-1}\frac{1}{k}\exp(\gamma k) \le \int_1^n t^{-1}\exp(\gamma t)dt$$

Now consider the integral: for any $t_0 \in (1, n)$,

$$\begin{aligned}\int_1^n t^{-1}\exp(\gamma t)dt &\le \int_1^{t_0}\exp(\gamma t)dt + \int_{t_0}^n t_0^{-1}\exp(\gamma t)dt \\ &\le \gamma^{-1}\big[\exp(\gamma t_0) - \exp(\gamma) + \frac{\exp(\gamma n) - \exp(\gamma t_0)}{t_0}\big]\end{aligned}$$

Take $t_0 = n - \sqrt{n}$.

$$\begin{aligned}\exp(\gamma t_0) - \exp(\gamma) + \frac{\exp(\gamma n) - \exp(\gamma t_0)}{t_0} &= \exp(\gamma(n-\sqrt{n})) - \exp(\gamma) + \frac{\exp(\gamma n) - \exp(\gamma(n-\sqrt{n}))}{n-\sqrt{n}} \\ &\le K' n^{-1}\exp(\gamma n)\end{aligned}$$

where $K' = \sup_{t\ge 2} t\exp(-\gamma\sqrt{t}) - t\exp(\gamma - \gamma t) + [1 - \exp(-\gamma\sqrt{t})]/[1 - 1/\sqrt{t}]$. The proof is completed by setting $K = \gamma^{-1}K'$. □

Proof of Prop. A.5 (iii). Following (64), we have

$$R^{(1),(3)}_{n+1} = \mathsf{E}[\widetilde{\theta}^{(1)}_{n+1}(\widetilde{\theta}^{(3)}_{n+1})^{\intercal}] = \frac{1}{n+1}\mathsf{E}[\widetilde{\theta}^{(1)}_{n+1}[\widehat{Z}_{n+3} - \widehat{Z}_{n+2}]^{\intercal}] \tag{72}$$

This is bounded based on (66): Lemma A.6 and (71b) indicate that there exists some constant K such that

$$\sum_{k=1}^{n+1}\alpha_k \Big\|\prod_{l=k+1}^{n+1}[I + \alpha_l A]\Big\|_T \Big\|\mathsf{E}\big[\Delta^m_{k+2}\widehat{Z}^{\intercal}_{n+3} - \Delta^m_{k+1}\widehat{Z}^{\intercal}_{n+2}\big]\Big\|_T \le K\rho^{n+1} \tag{73}$$

For the second term in (66), it admits a simpler form

$$\begin{aligned}\sum_{k=1}^{n+1}\alpha_k\prod_{l=k+1}^{n+1}[I+\alpha_l A]\big[\Delta^m_{k+1} - \Delta^m_{k+2}\big]\widehat{Z}^{\intercal}_{n+3} &= \prod_{l=2}^{n+1}[I+\alpha_l A]\Delta^m_2\widehat{Z}^{\intercal}_{n+3} - \frac{1}{n+1}\Delta^m_{n+3}\widehat{Z}^{\intercal}_{n+3} \\ &\quad - \sum_{k=2}^{n+1}\alpha_{k-1}\alpha_k\prod_{l=k+1}^{n+1}[I+\alpha_l A][I+A]\Delta^m_{k+1}\widehat{Z}^{\intercal}_{n+3}\end{aligned}$$

where $\prod_{l=2}^{n+1}[I+\alpha_l A]\mathsf{E}[\Delta_2\widehat{Z}^{\intercal}_{n+3}] = O(\rho^n)$ and $\mathsf{E}[\Delta^m_{n+3}\widehat{Z}^{\intercal}_{n+3}]$ converges to its steady-state mean. For the remaining part, Lemma A.6 and (71a) together imply that

$$\begin{aligned}&\Big\|\sum_{k=2}^{n+1}\alpha_{k-1}\alpha_k\prod_{l=k+1}^{n+1}[I+\alpha_l A][I+A]\mathsf{E}[\Delta^m_{k+1}\widehat{Z}^{\intercal}_{n+3}]\Big\|_T \\ &\le \sum_{k=2}^{n+1}\alpha_{k-1}\alpha_k\prod_{l=k+1}^{n+1}\|I+\alpha_l A\|_T\|I+A\|_T\|\mathsf{E}[\Delta^m_{k+1}\widehat{Z}^{\intercal}_{n+3}]\|_T \\ &\le \frac{K'}{n+2}\sum_{k=2}^{n+1}\alpha_{k-1}\rho^{n+1-k}\end{aligned}$$

for some constant K'. By Lemma A.9, there exists another constant K'' such that

$$\frac{K'}{n+2}\sum_{k=2}^{n+1}\alpha_{k-1}\rho^{n-k} = \frac{K'\rho^n}{n+2}\sum_{k=1}^{n}\alpha_k\rho^{-k} \le \frac{K'K''\rho}{(n+1)(n+2)}$$

This combined with (73) shows that

$$\mathsf{E}[\widetilde{\theta}^{(1)}_{n+1}[\widehat{Z}_{n+3}-\widehat{Z}_{n+2}]^{\intercal}] = -(n+1)^{-1}\mathsf{E}_\pi[\Delta_n^m\widehat{Z}_n^{\intercal}] + O(\rho^{n+1})$$

Following (72), we obtain the desired result:

$$\mathsf{E}[\widetilde{\theta}^{(1)}_{n+1}(\widetilde{\theta}^{(3)}_{n+1})^{\intercal}] = -\frac{1}{(n+1)^2}\mathsf{E}_\pi[\Delta_n^m\widehat{Z}_n^{\intercal}] + O((n+1)^{-3})$$

□

Proof of Thm. 2.8 With the decomposition in (42), we have

$$\begin{aligned}\mathrm{Cov}\,(\theta_n) = \mathrm{Cov}\,(\theta^{(1)}_n) + \sum_{j=1}^{3}\mathrm{Cov}\,(\theta^{(2,j)}_n) + \mathrm{Cov}\,(\theta^{(3)}_n) + R^{(1),(3)}_n + R^{(3),(1)}_n \\ + \sum_{i\in\{1,3\}}\sum_{j=1}^{3}[R^{(2,j),(i)}_n + R^{(i),(2,j)}_n] + \sum_{j=1}^{3}\sum_{k=1,k\neq j}^{3}[R^{(2,j),(2,k)}_n + R^{(2,k),(2,j)}_n]\end{aligned}$$

$\mathrm{Cov}\,(\theta^{(2,1)}_n) = O(n^{-3})$, $\mathrm{Cov}\,(\theta^{(2,2)}_n) = O(n^{-5})$ and $\mathrm{Cov}\,(\theta^{(2,3)}_n) = O(n^{-4})$ by Thm. 2.4 (i). By the Cauchy-Schwarz inequality, the correlation terms involving $\widetilde{\theta}^{(2,2)}_n$ and $\widetilde{\theta}^{(2,3)}_n$ are $O(n^{-2.5})$, and $R^{(2,1),(3)}_n = O(n^{-2.5})$ is also $O(n^{-2.5})$. Prop. A.5 (ii) shows that $R^{(2,1),(3)}_n = O(n^{-3})$. Hence the covariance can be approximated as follows:

$$\mathrm{Cov}\,(\theta_n) = \mathrm{Cov}\,(\theta^{(1)}_n) + \mathrm{Cov}\,(\theta^{(3)}_n) + R^{(1),(3)}_n + R^{(3),(1)}_n + R^{(2,1),(1)}_n + R^{(1),(2,1)}_n + O(n^{-2.5})$$

By Prop. A.5, there exist $\delta(I+A,\Sigma_\Delta) > 0$ and $\delta(I+A) > 0$ such that

$$\begin{aligned}\mathrm{Cov}\,(\theta^{(1)}_n) &= n^{-1}\Sigma_\theta + n^{-2}\Sigma^{(1)}_\sharp + O(n^{-2-\delta})\\ \mathrm{Cov}\,(\theta^{(3)}_n) &= n^{-2}\Sigma_Z + O(\rho^n)\\ R^{(1),(3)}_n &= -n^{-2}\mathsf{E}_\pi[\Delta_n^m\widehat{Z}_n^{\intercal}] + O(n^{-3})\\ R^{(2,1),(1)}_n + R^{(1),(2,1)}_n &= n^{-2}\Sigma^{(2)}_\sharp + O(n^{-2-\delta})\end{aligned}$$

Putting those results together gives

$$\mathrm{Cov}\,(\theta_n) = n^{-1}\Sigma_\theta + n^{-2}\big(\Sigma^{(1)}_\sharp + \Sigma^{(2)}_\sharp + \Sigma_Z - \mathsf{E}_\pi[\Delta_n^m\widehat{Z}_n^{\intercal}] - \mathsf{E}_\pi[\widehat{Z}_n(\Delta_n^m)^{\intercal}]\big) + O(n^{-2-\delta})$$

for some $\delta > 0$, where $\Sigma_\sharp := \Sigma^{(1)}_\sharp + \Sigma^{(2)}_\sharp$ solves the Lyapunov equation (43). □

A.4 Unbounded moments

This section is devoted to the proof that $\lim_{n\to\infty}\mathsf{E}[|v^{\intercal}\widetilde{\theta}^{\varrho}_n|^2] = \infty$ for $\varrho > \varrho_0$ (see Thm. 2.4 (ii)). Since it suffices to show the result holds for $\varrho_0 < \varrho < \frac{1}{2}$, we assume $\varrho < \frac{1}{2}$ throughout. Recall that $\lambda = -\varrho_0 + ui$.

Consider the update of $\widetilde{\theta}^{\varrho}_n$ in (32). With $v^{\intercal}[\lambda I - A] = 0$, we have $v^{\intercal}[\varrho_n I + A_n] = v^{\intercal}[\varrho - \varrho_0 + \varepsilon_v(n,\varrho) + ui]$. Multiplying each side of (32) by $v^{\intercal}$ gives

$$\begin{aligned}v^{\intercal}\widetilde{\theta}^{\varrho}_{n+1} &= v^{\intercal}\widetilde{\theta}^{\varrho}_n + \alpha_{n+1}\big[[\varrho - \varrho_0 + \varepsilon_v(n,\varrho) + ui]v^{\intercal}\widetilde{\theta}^{\varrho}_n + (n+1)^{\varrho}v^{\intercal}\Delta_{n+1}\big]\\ &= [1 + \alpha_{n+1}\tilde{\varrho}_{n+1} + \alpha_{n+1}ui]v^{\intercal}\widetilde{\theta}^{\varrho}_n + (n+1)^{\varrho-1}v^{\intercal}\Delta_{n+1}\end{aligned}$$

with $\tilde{\varrho}_{n+1} = \varrho - \varrho_0 + \varepsilon_v(n, \varrho)$. Note that $\tilde{\varrho}_{n+1}$ is strictly positive for sufficiently large n.

For a fixed but arbitrary n_0 and each $n \ge n_0$, we have

$$\begin{aligned} v^\intercal \widetilde{\theta}^{\varrho}_{n+1} &= v^\intercal \widetilde{\theta}^{\varrho}_{n_0} \prod_{k=n_0+1}^{n+1} [1 + \alpha_k \tilde{\varrho}_k + \alpha_k u i] + \sum_{k=n_0+1}^{n+1} k^{\varrho-1} v^\intercal \Delta_k \prod_{l=k+1}^{n+1} [1 + \alpha_l \tilde{\varrho}_l + \alpha_l u i] \\ &= \Big[\prod_{k=n_0+1}^{n+1} [1 + \alpha_k \tilde{\varrho}_k + \alpha_k u i] \Big] \cdot \Big[v^\intercal \widetilde{\theta}^{\varrho}_{n_0} + \sum_{k=n_0+1}^{n+1} \frac{k^{\varrho-1}}{\prod_{l=n_0+1}^{k} [1 + \alpha_l \tilde{\varrho}_l + \alpha_l u i]} v^\intercal \Delta_k \Big] \\ &= \Big[\prod_{k=n_0+1}^{n+1} [1 + \alpha_k \tilde{\varrho}_k + \alpha_k u i] \Big] \cdot \Big[v^\intercal \widetilde{\theta}^{\varrho}_{n_0} + \sum_{k=n_0+1}^{n+1} \beta_k v^\intercal \Delta_k \Big] \end{aligned} \tag{74}$$

with $\beta_n = n^{\varrho-1} / \prod_{l=n_0+1}^{n} [1 + \alpha_l \tilde{\varrho}_l + \alpha_l u i]$.

The analysis of $\{v^\intercal \widetilde{\theta}^{\varrho}_n\}$ is mainly based on the random series appearing in (74), which requires the following three preliminary results:

Lemma A.10. *There exists some n_0 such that for each $n > n_0$,*

$$|\beta_n - \beta_{n+1}|^2 \le 4 |\beta_{n+1}|^2 \alpha_n^2 (1 + u^2)$$

Proof. Note that $|\beta_n - \beta_{n+1}|^2 = |\beta_{n+1}|^2 |\beta_n / \beta_{n+1} - 1|^2$, so it is sufficient to bound the second factor:

$$\begin{aligned} |\beta_n / \beta_{n+1} - 1|^2 &= |(1 + n^{-1})^{1-\varrho} [1 + \alpha_{n+1} \tilde{\varrho}_{n+1} + \alpha_{n+1} u i] - 1|^2 \\ &= |(1 + n^{-1})^{1-\varrho} [1 + \alpha_{n+1} \tilde{\varrho}_{n+1}] - 1 + (1 + n^{-1})^{1-\varrho} \alpha_{n+1} u i|^2 \end{aligned} \tag{75}$$

Consider the real part in (75): since $\varepsilon_v(n, \varrho) = O(n^{-1})$, there exists n_0 such that $|\varepsilon_v(n, \varrho)| \le \varrho - \varrho_0$ and $\tilde{\varrho}_{n+1} = \varrho - \varrho_0 + \varepsilon_v(n, \varrho) > 0$ for $n \ge n_0$. Consequently,

$$\begin{aligned} 0 \le (1 + n^{-1})^{1-\varrho} [1 + \alpha_{n+1} \tilde{\varrho}_{n+1}] - 1 &< (1 + n^{-1}) [1 + \alpha_{n+1} \tilde{\varrho}_{n+1}] - 1 \\ &\le n^{-1} (1 + \tilde{\varrho}_{n+1} + \alpha_{n+1} \tilde{\varrho}_{n+1}) \end{aligned}$$

Given $0 < \varrho - \varrho_0 < \frac{1}{2}$, we can increase n_0 if necessary, such that $1 + \tilde{\varrho}_{n+1} + \alpha_{n+1} \tilde{\varrho}_{n+1} \le 2$ for $n \ge n_0$. Then we have

$$(1 + n^{-1})^{1-\varrho} [1 + \alpha_{n+1} \tilde{\varrho}_{n+1}] - 1 \le 2 \alpha_n$$

For the imaginary part, observe that

$$(1 + n^{-1})^{1-\varrho} \alpha_{n+1} u = \alpha_n \frac{n^{\varrho}}{(n+1)^{\varrho}} u \le 2 u \alpha_n$$

The proof is completed by summing the bounds for the real and imaginary parts. □

Lemma A.11. *Suppose Assumptions A1 and A3 hold. With each $n_0 \ge 1$, the random series $\sum_{k=n_0+1}^{\infty} \beta_k v^\intercal \Delta_k$ converges a.s..*

Proof. Decompose the series into the sum of a martingale difference and telescoping sequence. The martingale difference sequence converges *almost surely* given $\{\beta_n\} \in \ell_2$; the telescoping series is absolutely convergent by Lemma A.10. □

Lemma A.12. *Suppose Assumptions A1 and A3 hold. Denote $Z^v_n = v^\intercal Z_n = v^\intercal \hat{f}(\Phi_n)$. There exists a deterministic constant $K > 0$, such that for all n_0 and each sequence $\gamma \in \ell_1 \subseteq \ell_2$,*

$$\mathsf{E} \big[\mathsf{Var} \big(\sum_{k=n_0+2}^{\infty} \gamma_{k-n_0-1} Z^v_k \mid \mathcal{F}_{n_0+1} \big) \big] \le K \sum_{k=1}^{\infty} |\gamma_k|^2 \tag{76}$$

Proof. First recall that $\mathsf{Var}\big(\sum_{k=n_0+2}^{\infty}\gamma_{k-n_0-1}Z_k^v \mid \mathcal{F}_{n_0+1}\big) \le \mathsf{E}\big[|\sum_{k=n_0+2}^{\infty}\gamma_{k-n_0-1}Z_k^v|^2 \mid \mathcal{F}_{n_0+1}\big]$, and hence by the Markov property,

$$\mathsf{E}\big[|\sum_{k=n_0+2}^{\infty}\gamma_{k-n_0-1}Z_k^v|^2 \mid \mathcal{F}_{n_0+1}\big] = \mathsf{E}_{z'}\big[|\sum_{k=1}^{\infty}\gamma_k Z_k^v|^2\big] = \lim_{n\to\infty}\mathsf{E}_{z'}\big[|\sum_{k=1}^{n}\gamma_k Z_k^v|^2\big]$$

where $z' = \Phi_n$, and the last equality holds by the assumption $\gamma \in \ell_1$ and dominated convergence. For each n, letting $\lceil\gamma\rceil^n = (\gamma_1,\dots,\gamma_n)$ denote γ truncated at index n, we have

$$\mathsf{E}_{z'}\big[|\sum_{k=1}^{n}\gamma_k Z_k^v|^2\big] = \sum_{k=1}^{n}|\gamma_k|^2\mathsf{E}_{z'}\big[|Z_k^v|^2\big] + \sum_{i=1}^{n}\sum_{j\neq i}^{n}\gamma_i^\dagger\gamma_j\mathsf{E}_{z'}\big[(Z_i^v)^\dagger Z_j^v\big] = (\lceil\gamma\rceil^n)^\dagger[R]_n\lceil\gamma\rceil^n \tag{77}$$

where $[R]_n \in \mathbb{C}^{n\times n}$ is the covariance matrix with each entry defined as $R(i,j) = \mathsf{E}_{z'}\big[(Z_i^v)^\dagger Z_j^v\big], 1 \le i,j \le n$; $[R]_n$ is Hermitian and positive semi-definite. With $\lambda_n \ge 0$ denoting the largest eigenvalue of $[R]_n$, we have

$$(\lceil\gamma\rceil^n)^\dagger[R]_n\lceil\gamma\rceil^n \le \lambda_n\sum_{k=1}^{n}|\gamma_k|^2 \le \lambda_n\sum_{k=1}^{\infty}|\gamma_k|^2 \tag{78}$$

By the Gershgorin circle theorem (Golub and Van Loan, 1996), the maximal eigenvalue is upper bounded by the maximum row sum of absolute values of entries:

$$\lambda_n \le \max_{i\in\{1,\dots,n\}}\sum_{j=1}^{n}|R(i,j)| \le \sup_{i\in\mathbb{Z}_+}\sum_{j=1}^{\infty}|R(i,j)|$$

For any i, observe that

$$\sum_{j=1}^{\infty}|R(i,j)| = \mathsf{E}_{z'}\big[|Z_i^v|^2\big] + \sum_{i<j}|R(i,j)| + \sum_{i>j}|R(i,j)|$$

Since V-uniform ergodicity of the Markov chain $\mathbf{\Phi}$ implies V-geometric mixing (Meyn and Tweedie, 2009, Theorem 16.1.5) and $|v^\intercal\hat{f}|^2 \in L_\infty^V$, there exist $B < \infty$ and $r \in (0,1)$ such that for each $i,k \in \mathbb{Z}_+$,

$$\Big|R(i,i+k) - \mathsf{E}_{z'}\big[(Z_i^v)^\dagger\big]\mathsf{E}_{z'}\big[Z_{i+k}^v\big]\Big| \le Br^k[1 + r^iV(z')]$$

Consequently,

$$\begin{aligned}\sum_{j=1}^{\infty}|R(i,j)| \le{}& \mathsf{E}_{z'}\big[|Z_i^v|^2\big] + \Big|\mathsf{E}_{z'}\big[(Z_i^v)^\dagger\big]\Big|\sum_{j=1}^{\infty}\Big|\mathsf{E}_{z'}[Z_j^v]\Big| \\ &+ \sum_{i<j}Br^{j-i}[1 + r^iV(z')] + \sum_{i>j}Br^{i-j}[1 + r^jV(z')]\end{aligned} \tag{79}$$

Given $|v^\intercal\hat{f}|^2 \in L_\infty^V$, by (23),

$$\mathsf{E}_{z'}\big[|Z_n^v|^2\big] \le \mathsf{E}_\pi\big[|Z_n^v|^2\big] + B_V\big\||v^\intercal\hat{f}|^2\big\|_V V(z')$$

The Markov chain $\mathbf{\Phi}$ is also $\sqrt{V}$-uniformly ergodic. By (23) for $\sqrt{V}$ and $|v^\intercal\hat{f}|^2 \in L_\infty^V$ once more,

$$\big|\mathsf{E}_{z'}[(Z_i^v)^\dagger]\big| \le B_{\sqrt{V}}\|v^\intercal\hat{f}\|_{\sqrt{V}}\sqrt{V(z')}\rho^j$$

Hence

$$\big|\mathsf{E}_{z'}[(Z_i^v)^\dagger]\big|\sum_{j=1}^{\infty}\big|\mathsf{E}_{z'}[Z_j^v]\big| \le B^2_{\sqrt{V}}\|v^\intercal \hat f\|^2_{\sqrt{V}}V(z')\rho^i\sum_{j=1}^{\infty}\rho^j \le B^2_{\sqrt{V}}\|v^\intercal \hat f\|^2_{\sqrt{V}}\frac{\rho}{1-\rho}V(z')$$

The other two terms on the right hand side of (79) are bounded as follows:

$$\sum_{j>i} Br^{j-i}[1+r^iV(z')] = \sum_{j>i} B[r^{j-i}+r^jV(z')] \le \frac{Br}{1-r}(1+V(z'))$$
$$\sum_{j<i} Br^{i-j}[1+r^jV(z')] = \big[\sum_{j<i} B[r^{i-j}]\big] + BV(z')(i-1)r^i \le \frac{Br}{1-r} + BV(z')\sup_i ir^i$$

where $\sup_i ir^i$ exists since $\lim_{n\to\infty} nr^n = 0$.

Consequently, there exists some deterministic constant K' independent of z' such that, the largest eigenvalues $\{\lambda_n\}$ are uniformly bounded

$$\sup_n \lambda_n \le K'V(z')$$

Combining this with (77) and (78) gives

$$\mathsf{E}_{z'}\big[|\sum_{k=1}^{\infty} Z_k^v|^2\big] \le K'V(z')\sum_{k=1}^{\infty}|\gamma_k|^2$$

Therefore,

$$\mathsf{E}\Big[\mathsf{E}\big[|\sum_{k=n_0+2}^{\infty}\gamma_{k-n_0-1}Z_k^v|^2 \mid \mathcal{F}_{n_0+1}\big] \mid \Phi_0 = z\Big] \le K'\mathsf{E}\big[V(\Phi_{n_0+1}) \mid \Phi_0 = z\big]\sum_{k=1}^{\infty}|\gamma_k|^2$$

By $V \in L_\infty^V$ and (23) again, $\mathsf{E}[V(\Phi_{n_0+1}) \mid \Phi_0 = z] \le \pi(V) + B_V V(z)$. The desired conclusion then follows by setting $K = K'(B_V V(z) + \pi(V))$. □

Lemma A.13. *Suppose Assumptions A1-A3 hold and $\Sigma_\Delta v \neq 0$. With $\{\widetilde\theta_n^\varrho\}$ updated via (32),*

$$\liminf_{n\to\infty}\mathsf{E}[|v^\intercal\widetilde\theta_n^\varrho|^2] = \infty\,, \qquad \varrho > \varrho_0$$

Proof. With fixed n_0, equation (74) gives a representation for $v^\intercal\widetilde\theta_{n+1}^\varrho$ for each $n \ge n_0$. It is obvious that $\liminf_{n\to\infty}\prod_{k=n_0+1}^n|1+\tilde\varrho_k\alpha_k+\alpha_k ui|^2 = \infty$. Hence it suffices to show that $\liminf_{n\to\infty}\mathsf{E}[|v^\intercal\widetilde\theta_{n_0}^\varrho + \sum_{k=n_0+1}^{n+1}\beta_k v^\intercal\Delta_k|^2]$ is strictly greater than zero.

By Fatou's lemma,

$$\begin{aligned}\liminf_{n\to\infty}\mathsf{E}\big[|v^\intercal\widetilde\theta_{n_0}^\varrho + \sum_{k=n_0+1}^{n+1}\beta_k v^\intercal\Delta_k|^2\big] &\ge \mathsf{E}\big[\liminf_{n\to\infty}|v^\intercal\widetilde\theta_{n_0}^\varrho + \sum_{k=n_0+1}^{n+1}\beta_k v^\intercal\Delta_k|^2\big]\\ &= \mathsf{E}\big[|v^\intercal\widetilde\theta_{n_0}^\varrho + \sum_{k=n_0+1}^{\infty}\beta_k v^\intercal\Delta_k|^2\big]\\ &\ge \mathsf{Var}\,\big(v^\intercal\widetilde\theta_{n_0}^\varrho + \sum_{k=n_0+1}^{\infty}\beta_k v^\intercal\Delta_k\big)\end{aligned}$$

where the equality holds by Lemma A.11. By the law of total variance,

$$\begin{aligned}\mathsf{Var}\big(v^\intercal\widetilde{\theta}^{\varrho}_{n_0}+\sum_{k=n_0+1}^{\infty}\beta_k v^\intercal\Delta_k\big) &\geq \mathsf{E}\big[\mathsf{Var}\,(v^\intercal\widetilde{\theta}^{\varrho}_{n_0}+\sum_{k=n_0+1}^{\infty}\beta_k v^\intercal\Delta_k\mid\mathcal{F}_{n_0+1})\big]\\ &=\mathsf{E}\big[\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k v^\intercal\Delta_k\mid\mathcal{F}_{n_0+1})\big]\end{aligned}$$

Apply once more the decomposition based on Poisson's equation:

$$v^\intercal\Delta_n=\Delta^{vm}_{n+1}+Z^v_n-Z^v_{n+1}\,,\qquad n\geq 1\,,$$

where $Z^v_n=v^\intercal\hat{f}(\Phi_n)$ and $\Delta^{vm}_{n+1}=Z^v_{n+1}-\mathsf{E}[Z^v_{n+1}\mid\mathcal{F}_n]$ is a martingale difference. By the variance inequality $\mathsf{Var}\,(X+Y\mid\mathcal{F}_{n_0+1})\leq 2\mathsf{Var}\,(X\mid\mathcal{F}_{n_0+1})+2\mathsf{Var}\,(Y\mid\mathcal{F}_{n_0+1})$, we have

$$\begin{aligned}&\mathsf{E}\big[\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k v^\intercal\Delta_k\mid\mathcal{F}_{n_0+1})\big]\\ &\quad\geq\tfrac{1}{2}\mathsf{E}\big[\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k\Delta^{vm}_{k+1}\mid\mathcal{F}_{n_0+1})\big]-\mathsf{E}\big[\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k(Z^v_k-Z^v_{k+1})\mid\mathcal{F}_{n_0+1})\big]\end{aligned}\tag{80}$$

By the law of total variance once more,

$$\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k\Delta^{vm}_{k+1})=\mathsf{E}\big[\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k\Delta^{vm}_{k+1}\mid\mathcal{F}_{n_0+1})\big]+\mathsf{Var}\,\big(\mathsf{E}[\sum_{k=n_0+1}^{\infty}\beta_k\Delta^{vm}_{k+1}\mid\mathcal{F}_{n_0+1}]\big)$$

Note that $\lim_{n\to\infty}\mathsf{E}[\sum_{k=n_0+1}^{n}\beta_k\Delta^{vm}_{k+1}\mid\mathcal{F}_{n_0+1}]$ converges to zero *almost surely*. With $\{\beta_n\}\in\ell_2$ and the Jensen's inequality, we have for all n,

$$\big|\mathsf{E}[\sum_{k=n_0+1}^{n}\beta_k\Delta^{vm}_{k+1}\mid\mathcal{F}_{n_0+1}]\big|^2\leq\sum_{k=n_0+1}^{\infty}|\beta_k|^2\mathsf{E}[|\Delta^{vm}_{k+1}|^2\mid\mathcal{F}_{n_0+1}]<\infty$$

Then by the dominated convergence theorem, $\mathsf{E}\big[\big|\mathsf{E}[\sum_{k=n_0+1}^{\infty}\beta_k\Delta^{vm}_{k+1}\mid\mathcal{F}_{n_0+1}]\big|^2\big]=0$. Therefore,

$$\mathsf{Var}\,\big(\mathsf{E}[\sum_{k=n_0+1}^{\infty}\beta_k\Delta^{vm}_{k+1}\mid\mathcal{F}_{n_0+1}]\big)\leq\mathsf{E}\big[\big|\mathsf{E}[\sum_{k=n_0+1}^{\infty}\beta_k\Delta^{vm}_{k+1}\mid\mathcal{F}_{n_0+1}]\big|^2\big]=0$$

Hence,

$$\mathsf{E}\big[\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k\Delta^{vm}_{k+1}\mid\mathcal{F}_{n_0+1})\big]=\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k\Delta^{vm}_{k+1})=\sum_{k=n_0+1}^{\infty}|\beta_k|^2\sigma^2_{k+1}\tag{81}$$

where $\sigma^2_n=\mathsf{Var}\,(\Delta^{vm}_n)$.

For the telescoping term on the right hand side of (80), we have

$$\begin{aligned}\mathsf{E}\big[\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k(Z^v_k-Z^v_{k+1})\mid\mathcal{F}_{n_0+1})\big]&=\mathsf{E}\big[\mathsf{Var}\,(\beta_{n_0+1}Z^v_{n_0+1}-\sum_{k=n_0+2}^{\infty}(\beta_k-\beta_{k+1})Z^v_k\mid\mathcal{F}_{n_0+1})\big]\\ &=\mathsf{E}\big[\mathsf{Var}\,(\sum_{k=n_0+2}^{\infty}(\beta_k-\beta_{k+1})Z^v_k\mid\mathcal{F}_{n_0+1})\big]\end{aligned}\tag{82}$$

Given $\{\beta_n - \beta_{n+1}\} \in \ell_1$ by Lemma A.10, Lemma A.12 indicates that there exists some constant K independent of n_0 such that,

$$\mathsf{E}\Big[\mathsf{Var}\,\big(\sum_{k=n_0+2}^{\infty}(\beta_k - \beta_{k+1})\hat{Z}_k \mid \mathcal{F}_{n_0+1}\big)\Big] \le K \sum_{k=n_0+2}^{\infty} |\beta_k - \beta_{k+1}|^2$$

Combining (81) and (82) gives

$$\mathsf{E}[\mathsf{Var}\,(\sum_{k=n_0+1}^{\infty}\beta_k v^{\intercal}\Delta_k \mid \mathcal{F}_{n_0+1})] \ge \tfrac{1}{2}\sum_{k=n_0+1}^{\infty}|\beta_k|^2\sigma_{k+1}^2 - K\sum_{k=n_0+2}^{\infty}|\beta_k - \beta_{k+1}|^2$$

Since $|v^{\intercal}\hat{f}|^2 \in L_\infty^V$ and $\sigma_n^2 \to \sigma^2 = v^{\intercal}\Sigma_\Delta \overline{v} > 0$ at a geometric rate, we set n_0 sufficiently large such that Lemma A.10 holds and moreover for all $n \ge n_0$,

$$\sigma_n^2 \ge \tfrac{1}{2}\sigma^2, \qquad \frac{1}{4}\sigma^2 - 4K\alpha_n^2(1+u^2) \ge \frac{1}{8}\sigma^2\,,$$

Then,

$$\mathsf{E}\Big[\mathsf{Var}\,\big(\sum_{k=n_0+1}^{\infty}\beta_k v^{\intercal}\Delta_k \mid \mathcal{F}_{n_0+1}\big)\Big] \ge \frac{1}{8}\sigma^2\sum_{k=n_0+1}^{\infty}|\beta_k|^2$$

Therefore,

$$\liminf_{n\to\infty}\mathsf{E}\big[|v^{\intercal}\widetilde{\theta}_{n_0}^{\varrho} + \sum_{k=n_0+1}^{n}\beta_k v^{\intercal}\Delta_k|^2\big] \ge \frac{1}{8}\sigma^2\sum_{k=n_0+1}^{\infty}|\beta_k|^2 > 0$$

The desired conclusion then follows from (74):

$$\liminf_{n\to\infty}\mathsf{E}\big[|v^{\intercal}\widetilde{\theta}_{n+1}^{\varrho}|^2\big] \ge \liminf_{n\to\infty}\prod_{k=n_0+1}^{n}|1+\tilde{\varrho}_k\alpha_k+\alpha_k ui|^2 \cdot \liminf_{n\to\infty}\mathsf{E}\big[|v^{\intercal}\widetilde{\theta}_{n_0}^{\varrho} + \sum_{k=n_0+1}^{n}\beta_k v^{\intercal}\Delta_k|^2\big] = \infty$$

□

A.5 Coupling of Deterministic and Random Linear SA

Let $\widehat{\mathcal{A}}:\mathsf{Z}\to\mathbb{R}^{d\times d}$ denote the zero-mean solution to the following Poisson equation:

$$\mathsf{E}[\widehat{\mathcal{A}}(\Phi_{n+1}) \mid \Phi_n = z] = \widehat{\mathcal{A}}(z) - \mathcal{A}(z) + A\,, \qquad z\in\mathsf{Z}$$

which is a matrix version of (25). Denote $\Delta_{n+1}^{\mathcal{A}} = \widehat{\mathcal{A}}(\Phi_{n+1}) - \mathsf{E}[\widehat{\mathcal{A}}(\Phi_{n+1}) \mid \mathcal{F}_n]$ (a martingale difference sequence), and $\mathcal{A}_n = \widehat{\mathcal{A}}(\Phi_n)$. Then, from (35),

$$\begin{aligned}(A_{n+1}-A)\widetilde{\theta}_n^{\circ} &= [\Delta_{n+2}^{\mathcal{A}} + \mathcal{A}_{n+1} - \mathcal{A}_{n+2}]\widetilde{\theta}_n^{\circ}\\ &= \Delta_{n+2}^{\mathcal{A}}\widetilde{\theta}_n^{\circ} + \mathcal{A}_{n+1}\widetilde{\theta}_n^{\circ} - \mathcal{A}_{n+2}\widetilde{\theta}_{n+1}^{\circ} + \mathcal{A}_{n+2}(\widetilde{\theta}_{n+1}^{\circ} - \widetilde{\theta}_n^{\circ})\\ &= \Delta_{n+2}^{\mathcal{A}}\widetilde{\theta}_n^{\circ} + [\mathcal{A}_{n+1}\widetilde{\theta}_n^{\circ} - \mathcal{A}_{n+2}\widetilde{\theta}_{n+1}^{\circ}] + \alpha_{n+1}\mathcal{A}_{n+2}(A_{n+1}\widetilde{\theta}_n^{\circ} + \Delta_{n+1})\end{aligned}$$

The sequence $\{\mathcal{E}_n\}$ from (37) can be expressed as the sum

$$\mathcal{E}_n = \mathcal{E}_n^{(1)} + \mathcal{E}_n^{(2)} + \mathcal{E}_n^{(3)} + \mathcal{E}_n^{(4)}$$

where $\mathcal{E}_n^{(4)} = -\alpha_n \mathcal{A}_{n+1}\widetilde{\theta}_n^\circ$, and the first three sequences are solutions to the following linear systems:

$$\mathcal{E}_{n+1}^{(1)} = \mathcal{E}_n^{(1)} + \alpha_{n+1}[A\mathcal{E}_n^{(1)} + \Delta_{n+2}^{\mathcal{A}}\widetilde{\theta}_n^\circ]\,, \qquad \mathcal{E}_0^{(1)} = 0 \tag{83a}$$

$$\mathcal{E}_{n+1}^{(2)} = \mathcal{E}_n^{(2)} + \alpha_{n+1}[A\mathcal{E}_n^{(2)} - \alpha_n[I+A]\mathcal{A}_{n+1}\widetilde{\theta}_n^\circ]\,, \qquad \mathcal{E}_1^{(2)} = \mathcal{A}_1\widetilde{\theta}_0^\circ \tag{83b}$$

$$\mathcal{E}_{n+1}^{(3)} = \mathcal{E}_n^{(3)} + \alpha_{n+1}[A\mathcal{E}_n^{(3)} + \alpha_{n+1}\mathcal{A}_{n+2}(A_{n+1}\widetilde{\theta}_n^\circ + \Delta_{n+1})]\,, \qquad \mathcal{E}_0^{(3)} = 0 \tag{83c}$$

The second recursion arises through the arguments used in the proof of Lemma 2.2.

Recall that $\lambda = -\varrho_0 + ui$ is an eigenvalue of the matrix A with largest real part. For fixed $0 < \varrho < \varrho_0$, let $T \geq 0$ denote the unique solution to the Lyapunov equation

$$[\varrho I + A]T + T[\varrho I + A]^\intercal + I = 0 \tag{84}$$

As previously, the norm of random vector $E \in \mathbb{R}^d$ is defined as: $\|E\|_T = \sqrt{\mathsf{E}[E^\intercal T E]}$.

Lemma A.14. *Under Assumptions (A1)-(A4), there exist constants $L_{A.14}$ and $K_{A.14}$ such that, for all $n \geq 1$,*

(i) *The following holds for each $1 \leq i \leq 3$,*

$$\|\mathcal{E}_{n+1}^{(i)}\|_T^2 \leq (1 - 2\varrho\alpha_{n+1} + L_{A.14}^2\alpha_{n+1}^2)\|\mathcal{E}_n^{(i)}\|_T^2 + K_{A.14}\alpha_{n+1}^2(\|\mathcal{E}_n\|_T^2 + \|\widetilde{\theta}_n^\bullet\|_T^2 + 1)$$

(ii) *The following holds for $\mathcal{E}_n^{(4)}$,*

$$\|\mathcal{E}_{n+1}^{(4)}\|_T^2 \leq K_{A.14}\alpha_{n+1}^2(\|\mathcal{E}_n\|_T^2 + \|\widetilde{\theta}_n^\bullet\|_T^2 + 1)$$

The inequality below will be useful in proving Lemma A.14.

Lemma A.15. *For any real numbers a, b and all $c > 0$,*

$$(a+b)^2 \leq (1 + c^{-1})a^2 + (1+c)b^2$$

Proof. With $(a+b)^2 = a^2 + b^2 + 2ab$, the result follows directly from the inequality

$$2ab = 2(a/\sqrt{c})(\sqrt{c}b) \leq a^2/c + cb^2$$

□

Proof of Lemma A.14. First consider $\{\mathcal{E}_n^{(1)}\}$ updated via (83a). Since the martingale difference sequence $\Delta_{n+2}^{\mathcal{A}}$ is uncorrelated with $\widetilde{\theta}_n^\circ$ or $\mathcal{E}_n^{(1)}$, we have

$$\|\mathcal{E}_{n+1}^{(1)}\|_T^2 = \|[I + \alpha_{n+1}A]\mathcal{E}_n^{(1)}\|_T^2 + \alpha_{n+1}^2\|\Delta_{n+2}^{\mathcal{A}}\widetilde{\theta}_n^\circ\|_T^2$$

Using the fact that $T \geq 0$ solves the Lyapunov equation (84) gives

$$\|\mathcal{E}_{n+1}^{(1)}\|_T^2 \leq (1 - 2\varrho\alpha_{n+1} + L_1^2\alpha_{n+1}^2)\|\mathcal{E}_n^{(1)}\|_T^2 + \alpha_{n+1}^2\|\Delta_{n+2}^{\mathcal{A}}\widetilde{\theta}_n^\circ\|_T^2$$

where $L_1 = \|A\|_T$ (the induced operator norm). With $\widetilde{\theta}_n^\circ = \mathcal{E}_n + \widetilde{\theta}_n^\bullet$,

$$\|\Delta_{n+2}^{\mathcal{A}}\widetilde{\theta}_n^\circ\|_T^2 \leq 2\|\Delta_{n+2}^{\mathcal{A}}\|_T^2(\|\mathcal{E}_n\|_T^2 + \|\widetilde{\theta}_n^\bullet\|_T^2)$$

Consequently,

$$\|\mathcal{E}_{n+1}^{(1)}\|_T^2 \leq (1 - 2\varrho\alpha_{n+1} + L_1^2\alpha_{n+1}^2)\|\mathcal{E}_n^{(1)}\|_T^2 + K_1\alpha_{n+1}^2(\|\mathcal{E}_n\|_T^2 + \|\widetilde{\theta}_n^\bullet\|_T^2) \tag{85}$$

where $K_1 = \sup_n 2\|\Delta^{\mathcal{A}}_{n+2}\|_T^2$ is finite by the V-uniform ergodicity of $\mathbf{\Phi}$ applied to $\widehat{\mathcal{A}}^2_{i,j}$ (recall Thm. 2.1).

For $\{\mathcal{E}^{(2)}_n\}$ updated by (83b), using Lemma A.15 with $c = n(n+1)$ gives

$$\begin{aligned}\|\mathcal{E}^{(2)}_{n+1}\|_T^2 \leq &(1+\alpha_n\alpha_{n+1})(1-2\varrho\alpha_{n+1}+L_1^2\alpha_{n+1}^2)\|\mathcal{E}^{(2)}_n\|_T^2 \\ &+ 2(\alpha_n\alpha_{n+1}+\alpha_n^2\alpha_{n+1}^2)\|[I+A]\mathcal{A}_{n+1}\|_T^2(\|\mathcal{E}_n\|_T^2+\|\widetilde{\theta}^\bullet_n\|_T^2)\end{aligned}$$

We can find L_2 and K_2 such that for all $n \geq 1$,

$$\begin{aligned}\alpha_{n+1}^2L_1^2+\alpha_n\alpha_{n+1}(1-2\varrho\alpha_{n+1}+L_1^2\alpha_{n+1}^2) &\leq L_2^2\alpha_{n+1}^2 \\ 2(\alpha_n\alpha_{n+1}+\alpha_n^2\alpha_{n+1}^2)\|[I+A]\mathcal{A}_{n+1}\|_T^2 &\leq K_2\alpha_{n+1}^2\end{aligned}$$

We then obtain the desired form for the sequence $\{\mathcal{E}^{(2)}_n\}$

$$\|\mathcal{E}^{(2)}_{n+1}\|_T^2 \leq (1-2\varrho\alpha_{n+1}+L_2^2\alpha_{n+1}^2)\|\mathcal{E}^{(2)}_n\|_T^2 + K_2\alpha_{n+1}^2(\|\mathcal{E}_n\|_T^2+\|\widetilde{\theta}^\bullet_n\|_T^2) \tag{86}$$

The same argument applies to $\{\mathcal{E}^{(3)}_n\}$ in (83c). Therefore, for some constants L_3 and K_3,

$$\|\mathcal{E}^{(3)}_{n+1}\|_T^2 \leq (1-2\varrho\alpha_{n+1}+L_3^2\alpha_{n+1}^2)\|\mathcal{E}^{(3)}_n\|_T^2 + K_3\alpha_{n+1}^2(\|\mathcal{E}_n\|_T^2+\|\widetilde{\theta}^\bullet_n\|_T^2+1) \tag{87}$$

A bound on the final term $\mathcal{E}^{(4)}_{n+1} = -\alpha_{n+1}\mathcal{A}_{n+2}\widetilde{\theta}^\circ_{n+1}$ is relatively easy.

$$\begin{aligned}\|\mathcal{E}^{(4)}_{n+1}\|_T^2 &= \|\alpha_{n+1}\mathcal{A}_{n+2}[\widetilde{\theta}^\circ_n+\alpha_{n+1}(A_{n+1}\widetilde{\theta}^\circ_n+\Delta_{n+1})]\|_T^2 \\ &\leq 2\alpha_{n+1}^2\|\mathcal{A}_{n+2}\|_T^2(\|I+\alpha_{n+1}A_{n+1}\|_T^2\|\widetilde{\theta}^\circ_n\|_T^2+\alpha_{n+1}^2\|\Delta_{n+1}\|_T^2)\end{aligned}$$

Hence there exists some constant K_4 such that

$$\|\mathcal{E}^{(4)}_{n+1}\|_T^2 \leq K_4\alpha_{n+1}^2(\|\mathcal{E}_n\|_T^2+\|\widetilde{\theta}^\bullet_n\|_T^2+1)$$

□

The results in Lemma A.14 lead to a rough bound on $\|\widetilde{\theta}^\circ_n\|_T^2$ presented in the following. This intermediate result will be used later to establish the refined bound in Thm. 2.6.

Lemma A.16. *Under Assumptions (A1)-(A4),*

$$\limsup_{n\to\infty} n^\varrho\|\widetilde{\theta}^\circ_n\|_T^2 < \infty\,, \qquad \text{for } \varrho < \varrho_0 \text{ and } \varrho \leq 1$$

Proof. Denote $\mathcal{E}^{\text{tot}}_n = \sum_{i=1}^4\|\mathcal{E}^{(i)}_n\|_T^2$. By Lemma A.14, we can find $n_0 \geq 1$ such that $1-2\varrho\alpha_{n+1}+L^2_{A.14}\alpha_{n+1}^2 > 0$ for $n \geq n_0$ and

$$\begin{aligned}\mathcal{E}^{\text{tot}}_{n+1} &\leq (1-2\varrho\alpha_{n+1}+L^2_{A.14}\alpha_{n+1}^2)\mathcal{E}^{\text{tot}}_n + 4K_{A.14}\alpha_{n+1}^2(\|\mathcal{E}_n\|_T^2+\|\widetilde{\theta}^\bullet_n\|_T^2+1) \\ &\leq (1-2\varrho\alpha_{n+1}+L^2_{A.14}\alpha_{n+1}^2)\mathcal{E}^{\text{tot}}_n + 4K_{A.14}\alpha_{n+1}^2(4\mathcal{E}^{\text{tot}}_n+\|\widetilde{\theta}^\bullet_n\|_T^2+1) \\ &\leq (1-2\varrho\alpha_{n+1}+L^2_{\text{tot}}\alpha_{n+1}^2)\mathcal{E}^{\text{tot}}_n + K_{\text{tot}}\alpha_{n+1}^2\end{aligned}$$

with $L^2_{\text{tot}} = L^2_{A.14}+16K_{A.14}$ and $K_{\text{tot}} = \sup_n 4K_{A.14}(\|\widetilde{\theta}^\bullet_n\|_T^2+1)$, which are finite by Lemma A.2 combined with Lemma A.14. Iterating this inequality gives, for $n \geq n_0$,

$$\mathcal{E}^{\text{tot}}_{n+1} \leq \mathcal{E}^{\text{tot}}_{n_0}\prod_{k=n_0+1}^{n+1}(1-2\varrho\alpha_k+L^2_{\text{tot}}\alpha_k^2) + K_{\text{tot}}\sum_{k=n_0+1}^{n+1}\alpha_k^2\prod_{l=k+1}^{n+1}(1-2\varrho\alpha_l+L^2_{\text{tot}}\alpha_l^2)$$

By Lemma A.1,

$$\mathcal{E}^{\text{tot}}_{n+1} \le \mathcal{E}^{\text{tot}}_{n_0} \frac{K_{A.1} n_0^{2\varrho}}{(n+2)^{2\varrho}} + \frac{K_{A.1} K_{\text{tot}}}{(n+2)^{2\varrho}} \sum_{k=n_0+1}^{n+1} k^{2\varrho-2}$$

The partial sum can be estimated by an integral: with $2\varrho - 2 \le 0$,

$$\sum_{k=n_0}^{n+1} k^{2\varrho-2} \le 1 + \int_{n_0}^{n+1} r^{2\varrho-2} dr = \begin{cases} 1 + [(n+1)^{2\varrho-1} - n_0^{2\varrho-1}]/(2\varrho-1)\,, & \text{if } \varrho \ne \frac{1}{2} \\ 1 + \ln(n+1) - \ln(n_0)\,, & \text{if } \varrho = \frac{1}{2} \end{cases} \tag{88}$$

Given $\varrho \le 1$,

$$n^\varrho \mathcal{E}^{\text{tot}}_n \le \mathcal{E}^{\text{tot}}_{n_0} \frac{K_{A.1} n_0^{2\varrho}}{(n+2)^{\varrho}} + \frac{K_{A.1} K_{\text{tot}}}{(n+2)^{\varrho}} \sum_{k=n_0+1}^{n+1} k^{2\varrho-2} < \infty$$

Consequently, $\limsup_{n\to\infty} n^\varrho \|\mathcal{E}_n\|_T^2 < \infty$ by the inequality $n^\varrho \|\mathcal{E}_n\|_T^2 \le 4 n^\varrho \mathcal{E}^{\text{tot}}_n$. Then we have

$$n^\varrho \|\widetilde{\theta}^\circ_n\|_T^2 \le 2 n^\varrho \|\mathcal{E}_n\|_T^2 + 2 n^\varrho \|\widetilde{\theta}^\bullet_n\|_T^2$$

where $n^\varrho \|\widetilde{\theta}^\bullet_n\|_T^2 \to 0$ as n goes to infinity by Lemma A.2. Hence $\limsup_{n\to\infty} n^\varrho \|\widetilde{\theta}^\circ_n\|_T^2 < \infty$. □

Proof of Thm. 2.6. First consider $\{\mathcal{E}^{(2)}_n\}$ updated via (83b). By the triangle inequality and the inequality $\sqrt{1-x} \le \frac{1}{2}x$,

$$\begin{aligned} \|\mathcal{E}^{(2)}_{n+1}\|_T &\le \|[I + \alpha_{n+1} A] \mathcal{E}^{(2)}_n\|_T + \alpha_n \alpha_{n+1} \|[I + A] \mathcal{A}_{n+1} \widetilde{\theta}^\circ_n\|_T \\ &\le (1 - \varrho \alpha_{n+1} + \tfrac{1}{2} L^2 \alpha_{n+1}^2) \|\mathcal{E}^{(2)}_n\|_T + \alpha_{n+1}^{2+\varrho/2} K \end{aligned}$$

where $L = \|A\|_T$ and $K = \sup_n 2 \|[I + A]\mathcal{A}_{n+1}\|_T \|\widetilde{\theta}^\circ_n\| / (n+1)^{\varrho/2}$, which is finite thanks to Lemma A.16. Hence, by Lemma A.1 once more,

$$\begin{aligned} \|\mathcal{E}^{(2)}_{n+1}\|_T &\le \|\mathcal{E}^{(2)}_1\|_T \prod_{k=2}^{n+1} [1 - \varrho \alpha_k + \tfrac{1}{2} L^2 \alpha_k^2] + K \sum_{k=2}^{n+1} \alpha_k^{2+\varrho/2} \prod_{l=k+1}^{n+1} [1 - \varrho \alpha_k + \tfrac{1}{2} L^2 \alpha_k^2] \\ &\le \|\mathcal{E}^{(2)}_1\|_T \frac{K_{A.1}}{(n+2)^\varrho} + \frac{K K_{A.1}}{(n+2)^\varrho} \sum_{k=2}^{n+1} k^{\varrho/2-2} \end{aligned}$$

With $\varrho \le 1$, we have $\sum_{k=1}^\infty k^{\varrho/2-2} \le \sum_{k=1}^\infty k^{-3/2} < \infty$. Hence $\limsup_{n\to\infty} n^\varrho \|\mathcal{E}^{(2)}_n\|_T < \infty$. Replacing $A_{n+1}\widetilde{\theta}^\circ_n + \Delta_{n+1}$ with $\widetilde{\theta}^\circ_{n+1} - \widetilde{\theta}^\circ_n$ in (83c), the same argument applies to $\{\mathcal{E}^{(3)}_n\}$ and we get $\limsup_{n\to\infty} n^\varrho \|\mathcal{E}^{(3)}_n\|_T < \infty$. The fact that $\limsup_{n\to\infty} n \|\mathcal{E}^{(4)}_{n+1}\|_T < \infty$ follows directly from definition $\mathcal{E}^{(4)}_n = -\alpha_n \mathcal{A}_{n+1} \widetilde{\theta}^\circ_n$ and Lemma A.16. Then we have, for each $2 \le i \le 4$,

$$\limsup_{n\to\infty} n^\varrho \|\mathcal{E}^{(i)}_n\|_T < \infty\,, \qquad \text{for } \varrho < \varrho_0 \text{ and } \varrho \le 1 \tag{89}$$

Now consider the martingale difference part $\{\mathcal{E}^{(1)}_n\}$. The following is directly obtained from (83a):

$$\begin{aligned} \|\mathcal{E}^{(1)}_{n+1}\|_T^2 \le& (1 - 2\varrho \alpha_{n+1} + L^2 \alpha_{n+1}^2) \|\mathcal{E}^{(1)}_n\|_T^2 + \alpha_{n+1}^2 \|\Delta^{\mathcal{A}}_{n+2}\|_T^2 \|\widetilde{\theta}^\circ_n\|_T^2 \\ \le& (1 - 2\varrho \alpha_{n+1} + L^2 \alpha_{n+1}^2) \|\mathcal{E}^{(1)}_n\|_T^2 + \alpha_{n+1}^2 \|\Delta^{\mathcal{A}}_{n+2}\|_T^2 \big[8 \sum_{i=1}^4 \|\mathcal{E}^{(i)}_n\|_T^2 + 2 \|\widetilde{\theta}^\bullet_n\|_T^2\big] \end{aligned}$$

From Lemma A.2 we have $\sup_n n^\delta \|\widetilde{\theta}^\bullet_n\|_T^2 < \infty$ for $\delta = \min(1, 2\varrho)$. Combining this with (89) implies that there exists some constant $K_{\mathcal{M}}$ such that for $\delta = \min(1, 2\varrho)$,

$$\|\Delta^{\mathcal{A}}_{n+2}\|_T^2 \big[8 \sum_{i=2}^{4} \|\mathcal{E}^{(i)}_n\|_T^2 + 2\|\widetilde{\theta}^\bullet_n\|_T^2\big] \leq K_{\mathcal{M}} \frac{1}{(n+1)^\delta}$$

Consequently,

$$\|\mathcal{E}^{(1)}_{n+1}\|_T^2 \leq (1 - 2\varrho\alpha_{n+1} + L^2_{\mathcal{M}}\alpha^2_{n+1})\|\mathcal{E}^{(1)}_n\|_T^2 + K_{\mathcal{M}}\alpha^{2+\delta}_{n+1}$$

where $L^2_{\mathcal{M}} = \sup_n L^2 + 8\|\Delta^{\mathcal{A}}_{n+2}\|_T^2$. With initial condition $\mathcal{E}_0 = 0$, iterating this inequality gives

$$\|\mathcal{E}^{(1)}_{n+1}\|_T^2 \leq K_{\mathcal{M}} \sum_{k=1}^{n+1} \alpha_k^{2+\delta} \prod_{l=k+1}^{n+1} [1 - 2\varrho\alpha_l + L^2_{\mathcal{M}}\alpha_l^2] \leq \frac{K_{\mathcal{M}} K_{A.1}}{(n+2)^{2\varrho}} \sum_{k=1}^{n+1} k^{-(2+\delta-2\varrho)}$$

With $2 + \delta - 2\varrho > 0$, the partial sum is bounded by an integral similar as (88):

$$\frac{1}{(n+2)^{2\varrho}} \sum_{k=1}^{n+1} k^{-(2+\delta-2\varrho)} = \begin{cases} O((n+1)^{-2\varrho}), & \text{if } \varrho \leq \frac{1}{2} \text{ and } \delta = 2\varrho \\ O((n+1)^{-2\varrho}), & \text{if } \frac{1}{2} < \varrho < 1 \text{ and } \delta = 1 \\ O((n+1)^{-2}), & \text{if } \varrho > 1 \text{ and } \delta = 1 \end{cases}$$

Therefore,

(i) If $\varrho_0 \leq 1$, then $\limsup_{n\to\infty} (n+1)^{2\varrho} \|\mathcal{E}^{(1)}_{n+1}\|_T^2 < \infty$ for $\varrho < \varrho_0$.

(ii) If $\varrho_0 > 1$, then $\limsup_{n\to\infty} (n+1)^2 \|\mathcal{E}^{(1)}_{n+1}\|_T^2 < \infty$.

Given that the same convergence rates hold for the other components in (89), the conclusion then follows. □