
NONLINEAR MARKOVIAN STOCHASTIC APPROXIMATION

August 3, 2025

Mohammadhadi Hadavi
Prof. Hoi-To Wai - Chinese University of Hong Kong
Prof. Wenlong Mou - University of Toronto

1 Preliminaries

Notations The Euclidean norm is denoted by $\|\cdot\|$. The lowercase letter c and its derivatives c', c_0 , etc. denote universal numerical constants, whose value may change from line to line. As we are primarily interested in dependence of α and k , we adopt the following big- O notation: $\|f\| = \mathcal{O}(h(\alpha, k))$ if it holds that $\|f\| \leq s \cdot \|h(\alpha, k)\|$ for some constant $s > 0$.

We use of the following iteration scheme:

$$\theta_{t+1} = \theta_t + \alpha (g(\theta_t, X_{t+1}) + \xi_{t+1}(\theta_t))$$

where $g: \mathbb{R}^d \times \mathcal{X} \rightarrow \mathbb{R}^d$ is a deterministic function, $\{\xi_k\}_{k \geq 1}$ are *i.i.d* zero-mean random fields, and $\alpha > 0$ is a constant stepsize. We shall omit the superscript (α) in θ_k when the dependence on α is clear from the context.

In our settings, $\{X_n, n \in \mathbb{N}\}$ is a *state-dependent* (or controlled) Markov chain, *i.e.*, for any bounded measurable function $g: \mathbb{R}^d \times \mathcal{X} \rightarrow \mathbb{R}^d$,

$$\mathbb{E}[g(\theta_n, X_{n+1}) | \mathcal{F}_n] = P_{\theta_n} g(\theta_n, X_n) = \int g(\theta_n, x) P_{\theta}(X_n, dx),$$

where $P_{\theta}: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}_+$ is a Markov kernel such that, for each $\theta \in \Theta$, P_{θ} has a unique stationary distribution π_{θ} . Also, \mathcal{F}_n denotes the filtration generated by the random variables $(\theta_0, \{\xi_{m+1}\}_{m \leq n-1}, \{X_m\}_{m \leq n})$.

1.1 Assumptions

Assumption 1 For each $X \in \mathcal{X}$, the function $g(\theta, X)$ is three times continuously differentiable in θ with uniformly bounded first to third derivatives, *i.e.*, $\sup_{\theta \in \mathbb{R}^d} \|g^{(i)}(\theta, X)\| < \infty$ for $i = 1, 2, 3, X \in \mathcal{X}$. Moreover, there exists a constant $L_1 > 0$ such that (1) $\|g^{(i)}(\theta, X) - g^{(i)}(\theta', X)\| \leq L_1$, for all $\theta, \theta' \in \mathbb{R}^d, i = 0, 1, 2$ and $X \in \mathcal{X}$, and (2) $\|g^{(i)}(0, X)\| \leq L_1$ for all $X \in \mathcal{X}$ and $i = 0, 1, 2$.

Assumption 1 implies that $g(\theta, X)$ is L_1 -Lipschitz w.r.t θ uniformly in X . The above assumption immediately implies that the growth of $\|g\|$ and $\|\bar{g}\|$ will be at most linear in θ , *i.e.*, $\|g(\theta, X)\| \leq L_1(\|\theta - \theta^*\| + 1)$ and $\|\bar{g}(\theta)\| \leq L_1(\|\theta - \theta^*\| + 1)$. Similar implications also hold for the first and second derivatives in the same manner.

Assumption 2 There exists $\mu > 0$ such that $\langle \theta - \theta', \bar{g}(\theta) - \bar{g}(\theta') \rangle \leq -\mu \|\theta - \theta'\|^2, \forall \theta, \theta' \in \mathbb{R}^d$.

Consequently, the target equation $\bar{g}(\theta) = 0$ has a unique solution θ^* .

Denote by \mathcal{F}_k the filtration generated by $\{X_{t+1}, \theta_t, \xi_{t+1}\}_{t=0}^{k-1} \cup \{X_{k+1}, \theta_k\}$.

Assumption 3 Let $p \in \mathbb{Z}_+$ be given. The noise sequence $(\xi_k)_{k \geq 1}$ is a collection of *i.i.d* random fields satisfying the following conditions with $L_{2,p} > 0$:

$$\mathbb{E}[\xi_{k+1}(\theta) | \mathcal{F}_k] = 0 \quad \text{and} \quad \mathbb{E}^{1/(2p)}[\|\xi_1(\theta)\|^{2p}] \leq L_{2,p}(\|\theta - \theta^*\| + 1), \quad \forall \theta \in \mathbb{R}^d.$$

Define $C(\theta) = \mathbb{E}[\xi_1(\theta)^{\otimes 2}]$ and assume that $C(\theta)$ is at least twice differentiable. There also exists $M_c, k_c \geq 0$ such that for $\theta \in \mathbb{R}^d$, we have $\max_{i=1,2} \|C^{(i)}(\theta)\| \leq M_c\{1 + \|\theta - \theta^*\|^{k_c}\}$. In the sequel, we set $L := L_1 + L_2$, and without loss of generality, we assume $L \geq 2\mu$ for some technical reasons.

Assumption 4 There exists a Borel measurable function $\hat{g} : \mathbb{R}^d \times \mathcal{X} \rightarrow \mathbb{R}^d$ where for each $\theta \in \mathbb{R}^d, X \in \mathcal{X}$,

$$\hat{g}(\theta, X) - P_\theta \hat{g}(\theta, X) = g(\theta, X) - \bar{g}(\theta).$$

Assumption 5 There exists $L_{PH}^{(0)} < \infty$ and $L_{PH}^{(1)} < \infty$ such that, for all $\theta \in \mathbb{R}^d$ and $X \in \mathcal{X}$, one has $\|\hat{g}(\theta, X)\| \leq L_{PH}^{(0)}$, $\|P_\theta \hat{g}(\theta, X)\| \leq L_{PH}^{(0)}$. Moreover, for $(\theta, \theta') \in \mathcal{H}^2$,

$$\sup_{X \in \mathcal{X}} \|P_\theta \hat{g}(\theta, X) - P_{\theta'} \hat{g}(\theta', X)\| \leq L_{PH}^{(1)} \|\theta - \theta'\|.$$

Assumption 6 For any $\theta, \theta' \in \mathbb{R}^d$, we have $\sup_{X \in \mathcal{X}} \|P_\theta(X, \cdot) - P_{\theta'}(X, \cdot)\|_{TV} \leq L_P \|\theta - \theta'\|$.

Assumption 7 For any $\theta, \theta' \in \mathbb{R}^d$, we have $\sup_{X \in \mathcal{X}} \|g(\theta, X) - g(\theta', X)\| \leq L_H \|\theta - \theta'\|$.

Assumption 8 There exists $\rho < 1$, $K_P < \infty$ such that

$$\sup_{\theta \in \mathbb{R}^d, X \in \mathcal{X}} \|P_\theta^n(X, \cdot) - \pi_\theta(\cdot)\|_{TV} \leq \rho^n K_P,$$

Assumption 9 For any $\theta, \theta' \in \mathbb{R}^d$, we have $\sup_{X \in \mathcal{X}} \|\pi_\theta(X) - \pi_{\theta'}(X)\| \leq L_S \|\theta - \theta'\|$.

Lemma 1 Assume that assumptions 6-8 hold. Then, for any $\theta \in \mathbb{R}^d$ and $X \in \mathcal{X}$,

$$\|\hat{g}(\theta, X)\| \leq \frac{\sigma K_P}{1 - \rho},$$

$$\|P_\theta \hat{g}(\theta, X)\| \leq \frac{\sigma \rho K_P}{1 - \rho}.$$

Moreover, for any $\theta, \theta' \in \mathbb{R}^d$ and $X \in \mathcal{X}$,

$$\|P_\theta \hat{g}(\theta, X) - P_{\theta'} \hat{g}(\theta', X)\| \leq \|\theta - \theta'\|,$$

where

$$L_{PH}^{(1)} = \frac{K_P^2 \sigma L_P}{(1 - \rho)^2} (2 + K_P) + \frac{K_P}{1 - \rho} L_H.$$

Proof of this lemma can be found in [3], Lemma 7.

2 Error Bound

2.1 Base Case

We first prove the following lemma because we are going to use that calculation in many different parts of the proof:

Lemma2. Using Assumptions 1, 3, 5, and for sufficiently small α and $t \geq 1$, we have

$$\mathbb{E}[\|\theta_t - \theta_{t-1}\|] \leq \alpha L (\mathbb{E}[\|\theta_{t-1} - \theta^*\|] + 1).$$

Proof. We have

$$\begin{aligned}
\mathbb{E}[\|\theta_t - \theta_{t-1}\|] &\leq \alpha \mathbb{E}[\|g(\theta_{t-1}, X_t) + \xi_t(\theta_{t-1})\|] \\
&\leq \alpha L_1 (\mathbb{E}[\|\theta_{t-1} - \theta^*\|] + 1) + \alpha L_2 (\mathbb{E}[\|\theta_{t-1} - \theta^*\|] + 1) \\
&\leq \alpha L (\mathbb{E}[\|\theta_{t-1} - \theta^*\|] + 1)
\end{aligned}$$

where the first line follows from ??, second line from the Lipschitzness condition and the assumption of

$$\mathbb{E}^{1/2}[\|\xi_t(\theta_{t-1})\|^2 | \mathcal{F}_{t-1}] \leq L_2 (\mathbb{E}[\|\theta_{t-1}\|] + 1),$$

and the third line from ??.

For the base case analysis, we can write:

$$\begin{aligned}
&\mathbb{E}[\|\theta_{k+1} - \theta^*\|^2] - \mathbb{E}[\|\theta_k - \theta^*\|^2] = \\
&2\alpha \mathbb{E}[\langle \theta_k - \theta^*, g(\theta_k, X_{k+1}) \rangle] + \alpha^2 \mathbb{E}[\|g(\theta_k, X_{k+1})\|^2] + \alpha^2 \mathbb{E}[\|\xi_{k+1}(\theta_k)\|^2] = \\
&2\alpha \mathbb{E}[\langle \theta_k - \theta^*, g(\theta_k, X_{k+1}) - \bar{g}(\theta_k) \rangle] + 2\alpha \mathbb{E}[\langle \theta_k - \theta^*, \bar{g}(\theta_k) \rangle] + \alpha^2 \mathbb{E}[\|g(\theta_k, X_{k+1})\|^2] + \alpha^2 \mathbb{E}[\|\xi_{k+1}(\theta_k)\|^2].
\end{aligned}$$

It is easy to see that under Strong Monotonicity assumption, we have

$$\langle \theta_k - \theta^*, \bar{g}(\theta_k) \rangle = \langle \theta_k - \theta^*, \bar{g}(\theta_k) - \bar{g}(\theta^*) \rangle \leq -\mu \|\theta_k - \theta^*\|^2.$$

Additionally, under Assumption 1 and 3, we have the following upper bound

$$\begin{aligned}
&\alpha^2 (\mathbb{E}[\|g(\theta_k, X_{k+1})\|^2] + \mathbb{E}[\|\xi_{k+1}(\theta_k)\|^2]) \\
&\leq \alpha^2 \left(L_1^2 \mathbb{E}[(\|\theta_k - \theta^*\| + 1)^2] + L_2^2 \mathbb{E}[(\|\theta_k - \theta^*\| + 1)^2] \right) \\
&\leq 2\alpha^2 L^2 (\mathbb{E}[\|\theta_k - \theta^*\|^2] + 1).
\end{aligned}$$

Therefore, we have

$$\mathbb{E}[\|\theta_{k+1} - \theta^*\|^2] \leq (1 - 2\alpha(-\alpha L^2 + \mu)) \mathbb{E}[\|\theta_k - \theta^*\|^2] + 2\alpha^2 L^2 + 2\alpha \mathbb{E}[\langle \theta_k - \theta^*, g(\theta_k, X_{k+1}) - \bar{g}(\theta_k) \rangle]$$

Solving this recursion gives us the following inequality:

$$\begin{aligned}
\mathbb{E}[\|\theta_{k+1} - \theta^*\|^2] &\leq (1 - 2\alpha(-\alpha L^2 + \mu))^{k+1} \mathbb{E}[\|\theta_0 - \theta^*\|^2] \\
&\quad + \sum_{t=0}^k (1 - 2\alpha(-\alpha L^2 + \mu))^t 2\alpha^2 L^2 \\
&\quad + \sum_{t=0}^k 2\alpha (1 - 2\alpha(-\alpha L^2 + \mu))^{k-t} \mathbb{E}[\langle \theta_t - \theta^*, g(\theta_t, X_{t+1}) - \bar{g}(\theta_t) \rangle].
\end{aligned}$$

For notational simplicity we define $\gamma_t := 2\alpha(1 - 2\alpha(-\alpha L^2 + \mu))^{k-t}$ for $0 \leq t \leq k$.

The second term above is just a geometric series. According to Lemma 12 of [2], this equals to $\frac{\alpha L^2 [1 - (1 - 2\alpha(-\alpha L^2 + \mu))^{k+1}]}{-\alpha L^2 + \mu}$ which is of $\mathcal{O}(\alpha^2)$.

Now, we can upper bound the third summand using below decomposition:

$$\mathbb{E} \left[\sum_{t=0}^k \gamma_t \langle \theta_t - \theta^*, g(\theta_t, X_{t+1}) - \bar{g}(\theta_t) \rangle \right] = A_1 + A_2 + A_3 + A_4 + A_5$$

with

$$\begin{aligned} A_1 &:= \mathbb{E} \left[\sum_{t=1}^k \gamma_t \langle \theta_t - \theta^*, \hat{g}(\theta_t, X_{t+1}) - P_{\theta_t} \hat{g}(\theta_t, X_t) \rangle \right], \\ A_2 &:= \mathbb{E} \left[\sum_{t=1}^k \gamma_t \langle \theta_t - \theta^*, P_{\theta_t} \hat{g}(\theta_t, X_t) - P_{\theta_{t-1}} \hat{g}(\theta_{t-1}, X_t) \rangle \right], \\ A_3 &:= \mathbb{E} \left[\sum_{t=1}^k \gamma_t \langle \theta_t - \theta_{t-1}, P_{\theta_{t-1}} \hat{g}(\theta_{t-1}, X_t) \rangle \right], \\ A_4 &:= \mathbb{E} \left[\sum_{t=1}^k (\gamma_t - \gamma_{t-1}) \langle \theta_{t-1} - \theta^*, P_{\theta_{t-1}} \hat{g}(\theta_{t-1} - \theta^*, X_t) \rangle \right], \\ A_5 &:= \mathbb{E} [\gamma_0 \langle \theta_0 - \theta^*, \hat{g}(\theta_0, X_1) \rangle] - \mathbb{E} [\gamma_k \langle \theta_k - \theta^*, P_{\theta_k} \hat{g}(\theta_k, X_{k+1}) \rangle] \end{aligned}$$

For A_1 , we note that $\hat{g}(\theta_t, X_{t+1}) - P_{\theta_t} \hat{g}(\theta_t, X_t)$ is a martingale difference sequence [cf. ?] and therefore we have $A_1 = 0$ by taking the total expectation.

For A_2 , applying Cauchy-Schwarz inequality and ??, we have

$$\begin{aligned} A_2 &\leq \sum_{t=1}^k L_{PH}^{(1)} \gamma_t \mathbb{E} [\|\theta_t - \theta^*\| \|\theta_t - \theta_{t-1}\|] \\ &= \sum_{t=1}^k \alpha L_{PH}^{(1)} \gamma_t \mathbb{E} [\|\theta_t - \theta^*\| \|g(\theta_{t-1}, X_t) + \xi_t(\theta_{t-1})\|] \\ &\leq \sum_{t=1}^k \alpha L_{PH}^{(1)} \gamma_t \mathbb{E} [(\|\theta_t - \theta_{t-1}\| + \|\theta_{t-1} - \theta^*\|) (\|g(\theta_{t-1}, X_t)\| + \|\xi_t(\theta_{t-1})\|)] \\ &\leq \sum_{t=1}^k \alpha L_{PH}^{(1)} \gamma_t (L_1 (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^2] + \mathbb{E} [\|\theta_t - \theta_{t-1}\| \|\theta_{t-1} - \theta^*\|] + \mathbb{E} [\|\theta_{t-1} - \theta^*\|] + \mathbb{E} [\|\theta_t - \theta_{t-1}\|]) \\ &\quad + \mathbb{E} [\|\theta_t - \theta_{t-1}\| \|\xi_t(\theta_{t-1})\|] + \mathbb{E} [\|\theta_{t-1} - \theta^*\| \|\xi_t(\theta_{t-1})\|]) \end{aligned}$$

where the second line follows from ?? and the third line follows from the triangle inequality. Now we upper the compound terms in the last line's parentheses:

$$\begin{aligned} \mathbb{E} [\|\theta_t - \theta_{t-1}\| \|\theta_{t-1} - \theta^*\|] &= \mathbb{E} [\mathbb{E} [\|\theta_t - \theta_{t-1}\| \|\theta_{t-1} - \theta^*\| | \mathcal{F}_{t-1}]] \\ &\leq \mathbb{E} [\alpha L (\|\theta_{t-1} - \theta^*\| + 1) \|\theta_{t-1} - \theta^*\|] \\ &\leq \frac{\alpha L (3\mathbb{E} [\|\theta_{t-1} - \theta^*\|^2] + 1)}{2} \end{aligned}$$

where in the second line we used Lemma 2 and in the last line we used $u \leq \frac{u^2+1}{2}$.

$$\begin{aligned} \mathbb{E} [\|\theta_t - \theta_{t-1}\| \|\xi_t(\theta_{t-1})\|] &\leq \mathbb{E} [\alpha (\|g(\theta_{t-1}, X_t)\| + \|\xi_t(\theta_{t-1})\|) \|\xi_t(\theta_{t-1})\|] \\ &\leq \mathbb{E} [\alpha \|\xi_t(\theta_{t-1})\|^2 + \alpha L_1 (\|\theta_{t-1} - \theta^*\| + 1) \|\xi_t(\theta_{t-1})\|] \\ &\leq L (\mathbb{E} [\|\theta_{t-1} - \theta^*\|] + 1) \end{aligned}$$

where the first inequality follows from ??, second line from Lemma 2 and in the last line we used

boundedness property of the noise and sufficiently small α .

$$\begin{aligned}\mathbb{E}[\|\theta_{t-1} - \theta^*\| \|\xi_t(\theta_{t-1})\|] &= \mathbb{E}[\mathbb{E}[\|\theta_{t-1} - \theta^*\| \|\xi_t(\theta_{t-1})\| | \mathcal{F}_{t-1}]] \\ &\leq \mathbb{E}[L_2 \|\theta_{t-1} - \theta^*\| (\|\theta_{t-1} - \theta^*\| + 1)] \\ &\leq \frac{L_2(3\|\theta_{t-1} - \theta^*\|^2 + 1)}{2}\end{aligned}$$

where the second line follows from ?? and in the last line we used $u \leq \frac{u^2+1}{2}$.

Summing up all these bounds, we can write for A_2 :

$$\begin{aligned}A_2 &\leq \sum_{t=1}^k \alpha L_{PH}^{(1)} \gamma_t (2L(\mathbb{E}[\|\theta_{t-1} - \theta^*\|] + 1) + 2L(3\mathbb{E}[\|\theta_{t-1} - \theta^*\|^2] + 1) + \mathbb{E}[\|\theta_{t-1} - \theta^*\|^2]) \\ &\leq \sum_{t=1}^k \alpha L L_{PH}^{(1)} \gamma_t (8\mathbb{E}[\|\theta_{t-1} - \theta^*\|^2] + 5)\end{aligned}$$

which in the last line we again used $u \leq \frac{u^2+1}{2}$ property.

For A_3 , we obtain

$$\begin{aligned}A_3 &\leq \sum_{t=1}^k \gamma_t \mathbb{E}[\|\theta_t - \theta_{t-1}\| \|P_{\theta_{t-1}} \hat{g}(\theta_{t-1}, X_t)\|] \\ &\leq \sum_{t=1}^k \alpha L_{PH}^{(0)} \gamma_t \mathbb{E}[\|g(\theta_{t-1}, X_t) + \xi_t(\theta_{t-1})\|] \\ &\leq \sum_{t=1}^k \alpha L_{PH}^{(0)} \gamma_t (L_1(\mathbb{E}[\|\theta_{t-1} - \theta^*\|] + 1) + L_2(\mathbb{E}[\|\theta_{t-1} - \theta^*\|] + 1)) \\ &\leq \sum_{t=1}^k \alpha L L_{PH}^{(0)} \gamma_t (\mathbb{E}[\|\theta_{t-1} - \theta^*\|] + 1)\end{aligned}$$

where second line follows from ?? and third line follows from ?? .

For A_4 , we have

$$\begin{aligned}A_4 &\leq \sum_{t=1}^k |\gamma_t - \gamma_{t-1}| \mathbb{E}[\|\theta_{t-1} - \theta^*\| \|P_{\theta_{t-1}} \hat{g}(\theta_{t-1}, X_t)\|] \\ &\leq \sum_{t=1}^k L_{PH}^{(0)} |\gamma_t - \gamma_{t-1}| \mathbb{E}[\|\theta_{t-1} - \theta^*\|]\end{aligned}$$

Finally, for A_5 , we obtain

$$A_5 \leq L_{PH}^{(0)} (\gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|] - \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|])$$

which follows from Cacuhy-Schwarz inequality and ??.

Combining the above terms gives us:

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^k \gamma_t \langle \theta_t - \theta^*, g(\theta_t, X_{t+1}) - \bar{g}(\theta_t) \rangle\right] &\leq \sum_{t=0}^{k-1} \alpha L L_{PH}^{(1)} \gamma_{t+1} (5 + 8\mathbb{E}[\|\theta_t - \theta^*\|^2]) + \sum_{t=0}^{k-1} \alpha L L_{PH}^{(0)} \gamma_{t+1} (\mathbb{E}[\|\theta_{t-1} - \theta^*\|] + 1) + \\ &\quad \sum_{t=0}^{k-1} L_{PH}^{(0)} |\gamma_t - \gamma_{t+1}| \mathbb{E}[\|\theta_t - \theta^*\|] + L_{PH}^{(0)} (\gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|] - \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|])\end{aligned}$$

now it should be noticed that as long as the α satisfies $\alpha \leq \frac{\mu}{L^2}$, we have $\gamma_t \leq \gamma_{t+1}$. Thus, we can

simplify the above upper bound and write it this way:

$$\begin{aligned} \mathbb{E} \left[\sum_{t=0}^k \gamma_t \langle \theta_t - \theta^*, g(\theta_t, X_{t+1}) - \bar{g}(\theta_t) \rangle \right] &\leq \sum_{t=0}^{k-1} \alpha L L_{PH}^{(1)} \gamma_{t+1} (5 + 8\mathbb{E}[\|\theta_t - \theta^*\|^2]) + \\ &\quad \sum_{t=0}^{k-1} L_{PH}^{(0)} ((\alpha L + 1) \gamma_{t+1} - \gamma_t) \mathbb{E}[\|\theta_t - \theta^*\|] + \\ &\quad \sum_{t=0}^{k-1} \alpha L L_{PH}^{(0)} \gamma_{t+1} + L_{PH}^{(0)} (\gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|] - \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|]) \end{aligned}$$

Hence, using the derived upper bounds from the above terms, we have:

$$\begin{aligned} \mathbb{E}[\|\theta_{k+1} - \theta^*\|^2] &\leq \sum_{t=0}^{k-1} \alpha L L_{PH}^{(1)} \gamma_{t+1} (5 + 8\mathbb{E}[\|\theta_t - \theta^*\|^2]) + \sum_{t=0}^{k-1} L_{PH}^{(0)} ((\alpha L + 1) \gamma_{t+1} - \gamma_t) \mathbb{E}[\|\theta_t - \theta^*\|] + \\ &\quad (1 - 2\alpha(-\alpha L^2 + \mu)) \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|^2] + L_{PH}^{(0)} \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|] - L_{PH}^{(0)} \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|] + \\ &\quad 2\alpha^2 L L_{PH}^{(0)} \left[\frac{1 - (1 - 2\alpha(-\alpha L^2 + \mu))^k}{1 - (1 - 2\alpha(-\alpha L^2 + \mu))} \right] + \frac{\alpha L^2 [1 - (1 - 2\alpha(-\alpha L^2 + \mu))^{k+1}]}{-\alpha L^2 + \mu} \end{aligned}$$

for further notation simplicity we define $c_{1,t} := 2\alpha^2 L L_{PH}^{(0)} \left[\frac{1 - (1 - 2\alpha(-\alpha L^2 + \mu))^t}{1 - (1 - 2\alpha(-\alpha L^2 + \mu))} \right] + \frac{\alpha L^2 [1 - (1 - 2\alpha(-\alpha L^2 + \mu))^{t+1}]}{-\alpha L^2 + \mu}$ for $0 \leq t \leq k$. Now to write down this upper bound in a way in which it only depends on $\|\theta_0 - \theta^*\|$ related terms and constants, we can write:

$$\begin{aligned} \mathbb{E}[\|\theta_{k+1} - \theta^*\|^2] &\leq \sum_{t=0}^{k-1} \left[8\alpha L L_{PH}^{(1)} \gamma_{t+1} \mathbb{E}[\|\theta_t - \theta^*\|^2] + 5\alpha L L_{PH}^{(1)} \gamma_{t+1} \right] + \sum_{t=0}^{k-1} L_{PH}^{(0)} ((\alpha L + 1) \gamma_{t+1} - \gamma_t) \mathbb{E}[\|\theta_t - \theta^*\|] + \\ &\quad (1 - 2\alpha(-\alpha L^2 + \mu)) \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|^2] + L_{PH}^{(0)} \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|] - L_{PH}^{(0)} \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|] + c_{1,k} \\ &= \sum_{t=0}^{k-1} 8\alpha L L_{PH}^{(1)} \gamma_{t+1} \mathbb{E}[\|\theta_t - \theta^*\|^2] + 5\alpha L L_{PH}^{(1)} \sum_{t=0}^{k-1} \gamma_{t+1} + \sum_{t=0}^{k-1} L_{PH}^{(0)} ((\alpha L + 1) \gamma_{t+1} - \gamma_t) \mathbb{E}[\|\theta_t - \theta^*\|] + \\ &\quad (1 - 2\alpha(-\alpha L^2 + \mu)) \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|^2] + L_{PH}^{(0)} \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|] - L_{PH}^{(0)} \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|] + c_{1,k} \\ &= \sum_{t=0}^{k-1} 8\alpha L L_{PH}^{(1)} \gamma_{t+1} \mathbb{E}[\|\theta_t - \theta^*\|^2] + \sum_{t=0}^{k-1} L_{PH}^{(0)} ((\alpha L + 1) \gamma_{t+1} - \gamma_t) \mathbb{E}[\|\theta_t - \theta^*\|] + \\ &\quad (1 - 2\alpha(-\alpha L^2 + \mu)) \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|^2] + L_{PH}^{(0)} \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|] - L_{PH}^{(0)} \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|] + c_{1,k} + \\ &\quad \frac{10\alpha^2 L L_{PH}^{(1)} [1 - (1 - 2\alpha(-\alpha L^2 + \mu))^k]}{[1 - (1 - 2\alpha(-\alpha L^2 + \mu))]} \end{aligned}$$

where the last equality follows from the definition of γ_t s. Similarly we define $c_{2,t} := \frac{10\alpha^2 L L_{PH}^{(1)} [1 - (1 - 2\alpha(-\alpha L^2 + \mu))^t]}{[1 - (1 - 2\alpha(-\alpha L^2 + \mu))]}$ for $0 \leq t \leq k$. So we can write it as

$$\begin{aligned} \mathbb{E}[\|\theta_{k+1} - \theta^*\|^2] &\leq \sum_{t=0}^{k-1} 8\alpha L L_{PH}^{(1)} \gamma_{t+1} \mathbb{E}[\|\theta_t - \theta^*\|^2] + \sum_{t=0}^{k-1} L_{PH}^{(0)} ((\alpha L + 1) \gamma_{t+1} - \gamma_t) \mathbb{E}[\|\theta_t - \theta^*\|] + \\ &\quad (1 - 2\alpha(-\alpha L^2 + \mu)) \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|^2] + L_{PH}^{(0)} \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|] - L_{PH}^{(0)} \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|] + c_{1,k} + c_{2,k} \end{aligned}$$

Now for the second term on RHS, we note that since $L \geq 2\mu$,

$$(\alpha L + 1)\gamma_{t+1} - \gamma_t \leq 2\alpha L \gamma_{t+1}, \quad \mathbb{E}[\|\theta_t - \theta^*\|] \leq \sqrt{\mathbb{E}[\|\theta_t - \theta^*\|^2]},$$

and consequently

$$\begin{aligned} & \frac{1}{(1 - 2\alpha(-\alpha L^2 + \mu))^k} \sum_{t=0}^{k-1} L_{PH}^{(0)} ((\alpha L + 1)\gamma_{t+1} - \gamma_t) \mathbb{E}[\|\theta_t - \theta^*\|] \\ & \leq 4L_{PH}^{(0)} L \alpha^2 \sum_{t=0}^{k-1} \frac{1}{(1 - 2\alpha(-\alpha L^2 + \mu))^{t+1}} \sqrt{\mathbb{E}[\|\theta_t - \theta^*\|^2]} \\ & \leq 4L_{PH}^{(0)} L \alpha^2 \sum_{t=0}^{k-1} \frac{1}{(1 - 2\alpha(-\alpha L^2 + \mu))^{t+1}} (\mathbb{E}[\|\theta_t - \theta^*\|^2] + 1) \\ & \leq 4L_{PH}^{(0)} L \alpha^2 \cdot \sum_{t=0}^{k-1} \frac{1}{(1 - 2\alpha(-\alpha L^2 + \mu))^{t+1}} \mathbb{E}[\|\theta_t - \theta^*\|^2] + 2\alpha L L_{PH}^{(0)} \sum_{t=0}^{k-1} \gamma_{t+1}. \end{aligned}$$

We also note that

$$\frac{\gamma_k}{(1 - 2\alpha(-\alpha L^2 + \mu))^k} \mathbb{E}[\|\theta_k - \theta^*\|] \leq \alpha \frac{\mathbb{E}[\|\theta_k - \theta^*\|^2]}{(1 - 2\alpha(-\alpha L^2 + \mu))^k} + \frac{\alpha}{(1 - 2\alpha(-\alpha L^2 + \mu))^k}.$$

similarly

$$\frac{\gamma_0}{(1 - 2\alpha(-\alpha L^2 + \mu))^k} \mathbb{E}[\|\theta_0 - \theta^*\|] \leq \alpha \mathbb{E}[\|\theta_0 - \theta^*\|^2] + \alpha.$$

and we also define for $0 \leq t \leq k$

$$c_{3,t} := \frac{1}{-\alpha L^2 + \mu} \frac{2\alpha L L_{PH}^{(0)}}{(1 - 2\alpha(-\alpha L^2 + \mu))^t} + \frac{\alpha L_{PH}^{(0)}}{(1 - 2\alpha(-\alpha L^2 + \mu))^t} \alpha L_{PH}^{(0)}$$

to wrap up all the remainder terms.

Substituting back and rearranging with also defining $c'_{2,t} := \frac{c_{2,t}}{(1 - 2\alpha(-\alpha L^2 + \mu))^t}$ and $c'_{1,t} := \frac{c_{1,t}}{(1 - 2\alpha(-\alpha L^2 + \mu))^t}$, yields

$$\begin{aligned} \frac{\mathbb{E}[\|\theta_{k+1} - \theta^*\|^2]}{(1 - 2\alpha(-\alpha L^2 + \mu))^k} & \leq \frac{\alpha}{(1 - 2\alpha(-\alpha L^2 + \mu))^k} \mathbb{E}[\|\theta_k - \theta^*\|^2] + \sum_{t=0}^{k-1} \frac{\alpha \left(8L L_{PH}^{(1)} + 4\alpha (1 - 2\alpha(-\alpha L^2 + \mu))^{-1} L L_{PH}^{(0)} \right)}{(1 - 2\alpha(-\alpha L^2 + \mu))^t} \mathbb{E}[\|\theta_t - \theta^*\|^2] + \\ & \quad \alpha \mathbb{E}[\|\theta_0 - \theta^*\|^2] + c'_{1,k} + c'_{2,k} + c_{3,k}. \end{aligned}$$

for sufficiently small α s, we have

$$\alpha \left(\frac{3}{2} L L_{PH}^{(1)} + 2\alpha (1 - 2\alpha(-\alpha L^2 + \mu))^{-1} L L_{PH}^{(0)} \right) \leq 4\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2L L_{PH}^{(1)} \right)$$

using above simplification we can rewrite our upper bound as

$$\frac{\mathbb{E}[\|\theta_{k+1} - \theta^*\|^2]}{(1 - 2\alpha(-\alpha L^2 + \mu))^k} \leq 4\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2L L_{PH}^{(1)} \right) \sum_{t=0}^k \frac{\mathbb{E}[\|\theta_t - \theta^*\|^2]}{(1 - 2\alpha(-\alpha L^2 + \mu))^t} + \alpha \mathbb{E}[\|\theta_0 - \theta^*\|^2] + c'_{1,k} + c'_{2,k} + c_{3,k}.$$

For solving the above recursion, we first define $S_t := 4\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2L L_{PH}^{(1)} \right) \sum_{l=0}^t \frac{\mathbb{E}[\|\theta_l - \theta^*\|^2]}{(1 - 2\alpha(-\alpha L^2 + \mu))^l}$ for

$0 \leq t \leq k$. Also we use $C_t := c'_{1,t} + c'_{2,t} + c_{3,t}$ and for $0 \leq t \leq k$, defining constant terms. Now we can write

$$\frac{\mathbb{E}[\|\theta_{t+1} - \theta^*\|^2]}{(1 - 2\alpha(-\alpha L^2 + \mu))^t} \leq S_t + \alpha \mathbb{E}[\|\theta_0 - \theta^*\|^2] + C_t.$$

using this expansion, we should first notice that

$$\begin{aligned} \frac{S_t}{S_{t-1}} &= \frac{S_{t-1} + 4\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2LL_{PH}^{(1)} \right) \frac{\mathbb{E}[\|\theta_t - \theta^*\|^2]}{(1 - 2\alpha(-\alpha L^2 + \mu))^t}}{S_{t-1}} \\ &= 1 + 4\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2LL_{PH}^{(1)} \right) \frac{S_{t-1} + \alpha \mathbb{E}[\|\theta_0 - \theta^*\|^2] + C_{t-1}}{S_{t-1}} \\ &\leq 1 + 8\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2LL_{PH}^{(1)} \right). \end{aligned}$$

Now, since we have $S_0 = 4\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2LL_{PH}^{(1)} \right) \mathbb{E}[\|\theta_0 - \theta^*\|^2]$, thus

$$S_t \leq 4\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2LL_{PH}^{(1)} \right) \left[1 + 8\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2LL_{PH}^{(1)} \right) \right]^t \mathbb{E}[\|\theta_0 - \theta^*\|^2].$$

Substituting this upper bound into previous equations we get

$$\begin{aligned} \mathbb{E}[\|\theta_{k+1} - \theta^*\|^2] &\leq \\ (1 - 2\alpha(-\alpha L^2 + \mu))^k &\left[4\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2LL_{PH}^{(1)} \right) \left(1 + 8\alpha \left((\alpha L + 1) L_{PH}^{(0)} + 2LL_{PH}^{(1)} \right) \right)^k \mathbb{E}[\|\theta_0 - \theta^*\|^2] + \alpha \mathbb{E}[\|\theta_0 - \theta^*\|^2] + C_k \right] \end{aligned}$$

Choosing α sufficiently small, we can simplify the above inequality and write it as follows

$$\mathbb{E}[\|\theta_{k+1} - \theta^*\|^2] \leq \tilde{c}_1 \cdot (1 - 2\alpha\mu)^{k+1} \mathbb{E}[\|\theta_0 - \theta^*\|^2] + \tilde{c}_2 \cdot 2\alpha L \left(L_{PH}^{(0)} + \alpha L_{PH}^{(1)} \right)$$

where \tilde{c}_1 and \tilde{c}_2 are $\mathcal{O}(1)$ constants.

2.2 General Case

Similar to the previous case, we first prove a useful lemma:

Lemma3. *Using Assumptions 1, 3, 5, and $m, n \geq 1$, we have*

$$\mathbb{E}[\|\theta_t - \theta_{t-1}\|^m \|\theta_{t-1} - \theta^*\|^n] \leq 2^{2m+n-2} \alpha^m L^m (\mathbb{E}[\|\theta_{t-1} - \theta^*\|^{m+n}] + 1).$$

Also when $m = 0$, this upper bound can be written as

$$\mathbb{E}[\|\theta_{t-1} - \theta^*\|^n] \leq 2^{n-1} (\mathbb{E}[\|\theta_{t-1} - \theta^*\|^n] + 1). \quad (1)$$

Proof. We have

$$\begin{aligned}
\mathbb{E} [\|\theta_t - \theta_{t-1}\|^m \|\theta_{t-1} - \theta^*\|^n] &\leq \alpha^m \mathbb{E} [\|g(\theta_{t-1}, X_t) + \xi_t(\theta_{t-1})\|^m \|\theta_{t-1} - \theta^*\|^n] \\
&\leq 2^{m-1} \alpha^m \mathbb{E} [(\|g(\theta_{t-1}, X_t)\|^m + \|\xi_t(\theta_{t-1})\|^m) \|\theta_{t-1} - \theta^*\|^n] \\
&= 2^{m-1} \alpha^m \mathbb{E} [\mathbb{E} [(\|g(\theta_{t-1}, X_t)\|^m + \|\xi_t(\theta_{t-1})\|^m) | \mathcal{F}_{t-1}] \|\theta_{t-1} - \theta^*\|^n] \\
&\leq 2^{m-1} \alpha^m \mathbb{E} [(L_1^m \mathbb{E} [\|\theta_{t-1} - \theta^*\| + 1]^m) + L_2^m \mathbb{E} [\|\theta_{t-1} - \theta^*\| + 1]^m] \|\theta_{t-1} - \theta^*\|^n \\
&\leq 2^{m-1} \alpha^m L^m \mathbb{E} [(\|\theta_{t-1} - \theta^*\| + 1)^{m+n}] \\
&\leq 2^{2m+n-2} \alpha^m L^m (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{m+n}] + 1)
\end{aligned}$$

where the second and the last line follows from the mclaurin's inequality. Also the fourth line follows from ???. Proof for the other case is a trivial consequence of these calculations.

Now to do the proof in this case, we assume that the moment bound in [??] has been proven for $k \leq n-1$, we now proceed to show that the desired moment convergence holds for n with $2 \leq n \leq p$.

We start with the following decomposition of $\|\theta_{k+1} - \theta^*\|^{2n}$

$$\begin{aligned}
\|\theta_{k+1} - \theta^*\|^{2n} &= (\|\theta_k - \theta^*\|^2 + 2\alpha \langle \theta_k - \theta^*, g(\theta_k, X_{k+1}) + \xi_{k+1}(\theta_k) \rangle + \alpha^2 \|g(\theta_k, X_{k+1}) + \xi_{k+1}(\theta_k)\|^2)^n \\
&= \sum_{\substack{i,j,l \\ i+j+l=n}} \binom{n}{i,j,l} \|\theta_k - \theta^*\|^{2i} (2\alpha \langle \theta_k - \theta^*, g(\theta_k, X_{k+1}) + \xi_{k+1}(\theta_k) \rangle)^j (\alpha \|g(\theta_k, X_{k+1}) + \xi_{k+1}(\theta_k)\|)^{2l}
\end{aligned}$$

We note the following cases.

1. $i = n, j = l = 0$. In this case, the summand is simply $\|\theta_k - \theta^*\|^{2i}$.
2. When $i = n-1, j = 1$ and $l = 0$. In this case, the summand is of order α , i.e.,

$$\alpha \cdot 2n \langle \theta_k - \theta^*, g(\theta_k, X_{k+1}) + \xi_{k+1}(\theta_k) \rangle^j \|\theta_k - \theta^*\|^{2(n-1)}.$$

We can further decompose it as

$$\begin{aligned}
&2n\alpha \langle \theta_k - \theta^*, g(\theta_k, X_{k+1}) + \xi_{k+1}(\theta_k) \rangle \|\theta_k - \theta^*\|^{2(n-1)} \\
&= \underbrace{2n\alpha \langle \theta_k - \theta^*, g(\theta_k, X_{k+1}) - \bar{g}(\theta_k) + \xi_{k+1}(\theta_k) \rangle \|\theta_k - \theta^*\|^{2(n-1)}}_{T_1} + \underbrace{2n\alpha \langle \theta_k - \theta^*, \bar{g}(\theta_k) \rangle \|\theta_k - \theta^*\|^{2(n-1)}}_{T_2}.
\end{aligned}$$

Note that, when (X_k) is i.i.d or from a martingale noise, we have

$$\mathbb{E}[T_1 | \theta_k] = 0$$

However, when (X_k) is Markovian, the above inequality does not hold and T_1 requires careful analysis.

Nonetheless, under the strong monotonicity assumption, we have

$$T_2 \leq -2n\alpha\mu \|\theta_k - \theta^*\|^{2n}.$$

3. For the remaining terms, we see that they are of higher orders of α . Therefore, when α is selected sufficiently small, these terms do not raise concern.

Therefore, to prove the desired moment bound, we spend the remaining section analyzing T_1 . Immediately, we note that

$$\begin{aligned}\mathbb{E}[T_1] &= \mathbb{E}\left[2n\alpha\langle\theta_k - \theta^*, g(\theta_k, X_{k+1}) - \bar{g}(\theta_k) + \mathbb{E}[\xi_{k+1}(\theta_k)|\theta_k]\rangle\|\theta_k - \theta^*\|^{2(n-1)}\right] \\ &= 2n\alpha\mathbb{E}\left[\underbrace{\langle\theta_k - \theta^*, g(\theta_k, X_{k+1}) - \bar{g}(\theta_k)\rangle\|\theta_k - \theta^*\|^{2(n-1)}}_{T'_1}\right].\end{aligned}$$

Subsequently, we focus on analyzing T'_1 ; but before that, we write the general recursion of the error bound. First, we define $T'_{1,t} := \langle\theta_t - \theta^*, g(\theta_t, X_{t+1}) - \bar{g}(\theta_t)\rangle\|\theta_t - \theta^*\|^{2(n-1)}$ to make T'_1 dependent on the iteration index. Now, following the above decomposition and taking the expectations, we have:

$$\mathbb{E}[\|\theta_{k+1} - \theta^*\|^{2n}] \leq \mathbb{E}[\|\theta_k - \theta^*\|^{2n}] + 2n\alpha\mathbb{E}[T'_{1,k}] - 2n\alpha\mu\mathbb{E}[\|\theta_k - \theta^*\|^{2n}] + o(\alpha) = (1 - 2n\alpha\mu)\mathbb{E}[\|\theta_k - \theta^*\|^{2n}] + 2n\alpha\mathbb{E}[T'_{1,k}] + o(\alpha)$$

similarly to the previous case we define $\gamma_t := 2n\alpha(1 - 2n\alpha\mu)^{k-t}$ for $0 \leq t \leq k$. Solving the above recursion will give us

$$\mathbb{E}[\|\theta_{k+1} - \theta^*\|^{2n}] \leq \sum_{t=0}^k \gamma_t \mathbb{E}[T'_{1,t}] + \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|^{2n}] + o(\alpha)$$

We have to upper bound the first term in the RHS above. For this purpose, we use a similar decomposition to our base case analysis:

$$\sum_{t=0}^k \gamma_t \mathbb{E}[T'_{1,t}] = \mathbb{E}\left[\sum_{t=0}^k \gamma_t \langle\theta_t - \theta^*, g(\theta_t, X_{t+1}) - \bar{g}(\theta_t)\rangle\|\theta_t - \theta^*\|^{2(n-1)}\right] = A_1 + A_2 + A_3 + A_4 + A_5$$

with

$$\begin{aligned}A_1 &:= \mathbb{E}\left[\sum_{t=1}^k \gamma_t \langle\theta_t - \theta^*, \hat{g}(\theta_t, X_{t+1}) - P_{\theta_t} \hat{g}(\theta_t, X_t)\rangle\|\theta_t - \theta^*\|^{2(n-1)}\right], \\ A_2 &:= \mathbb{E}\left[\sum_{t=1}^k \gamma_t \langle\theta_t - \theta^*, P_{\theta_t} \hat{g}(\theta_t, X_t) - P_{\theta_{t-1}} \hat{g}(\theta_{t-1}, X_t)\rangle\|\theta_t - \theta^*\|^{2(n-1)}\right], \\ A_3 &:= \mathbb{E}\left[\sum_{t=1}^k \gamma_t \langle\theta_t - \theta_{t-1}, P_{\theta_{t-1}} \hat{g}(\theta_{t-1}, X_t)\rangle\|\theta_t - \theta^*\|^{2(n-1)}\right], \\ A_4 &:= \mathbb{E}\left[\sum_{t=1}^k (\gamma_t - \gamma_{t-1}) \langle\theta_{t-1} - \theta^*, P_{\theta_{t-1}} \hat{g}(\theta_{t-1} - \theta^*, X_t)\rangle\|\theta_t - \theta^*\|^{2(n-1)}\right], \\ A_5 &:= \mathbb{E}\left[\sum_{t=1}^k \gamma_{t-1} \langle\theta_{t-1} - \theta^*, P_{\theta_{t-1}} \hat{g}(\theta_{t-1} - \theta^*, X_t)\rangle(\|\theta_t - \theta^*\|^{2(n-1)} - \|\theta_{t-1} - \theta^*\|^{2(n-1)})\right], \\ A_6 &:= \mathbb{E}[\gamma_0 \langle\theta_0 - \theta^*, \hat{g}(\theta_0, X_0)\rangle\|\theta_0 - \theta^*\|^{2(n-1)}] - \mathbb{E}[\gamma_k \langle\theta_k - \theta^*, P_{\theta_k} \hat{g}(\theta_k, X_{k+1})\rangle\|\theta_k - \theta^*\|^{2(n-1)}].\end{aligned}$$

For A_1 , we note that $\hat{g}(\theta_t, X_{t+1}) - P_{\theta_t} \hat{g}(\theta_t, X_t)$ is a martingale difference sequence [cf. ?] and therefore we have $A_1 = 0$ by taking the total expectation.

For A_2 , applying Cauchy-Schwarz inequality and ??, we have

$$\begin{aligned}
A_2 &\leq \sum_{t=1}^k L_{PH}^{(1)} \gamma_t \mathbb{E} [\|\theta_t - \theta_{t-1}\| \|\theta_t - \theta^*\|^{2n-1}] \\
&\leq \sum_{t=1}^k L_{PH}^{(1)} \gamma_t \mathbb{E} [\|\theta_t - \theta_{t-1}\| (\|\theta_t - \theta_{t-1}\| + \|\theta_{t-1} - \theta^*\|)^{2n-1}] \\
&\leq \sum_{t=1}^k 2^{2n-2} L_{PH}^{(1)} \gamma_t \mathbb{E} [\|\theta_t - \theta_{t-1}\|^{2n} + \|\theta_t - \theta_{t-1}\| \|\theta_{t-1} - \theta^*\|^{2n-1}] \\
&\leq \sum_{t=1}^k 2^{2n-2} L_{PH}^{(1)} \gamma_t (2^{4n} \alpha^{2n} L^{2n} (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n}] + 1) + 2^{2n-1} \alpha L (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n}] + 1)) \\
&\leq \sum_{t=1}^k 2^{4n-2} \alpha L L_{PH}^{(1)} \gamma_t (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n}] + 1)
\end{aligned}$$

where the second line follows from triangle inequality, third line from mclaurin's inequality, fourth line from Lemma 3, and the last line for $\alpha < \frac{1}{4L}$.

For A_3 , we obtain

$$\begin{aligned}
A_3 &\leq \sum_{t=1}^k \gamma_t \mathbb{E} [\|\theta_t - \theta_{t-1}\| \|P_{\theta_{t-1}} \hat{g}(\theta_{t-1}, X_t)\| \|\theta_t - \theta^*\|^{2(n-1)}] \\
&\leq \sum_{t=1}^k L_{PH}^{(0)} \gamma_t \mathbb{E} [\|\theta_t - \theta_{t-1}\| (\|\theta_t - \theta_{t-1}\| + \|\theta_{t-1} - \theta^*\|)^{2(n-1)}] \\
&\leq \sum_{t=1}^k 2^{2n-3} L_{PH}^{(0)} \gamma_t \mathbb{E} [\|\theta_t - \theta_{t-1}\|^{2n-1} + \|\theta_t - \theta_{t-1}\| \|\theta_{t-1} - \theta^*\|^{2(n-1)}] \\
&\leq \sum_{t=1}^k 2^{2n-3} L_{PH}^{(0)} \gamma_t (2^{4n-1} \alpha^{2n-1} L^{2n-1} (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n-1}] + 1) + 2^{2n-2} \alpha L (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n-1}] + 1)) \\
&\leq \sum_{t=1}^k 2^{4n-4} \alpha L L_{PH}^{(0)} \gamma_t (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n-1}] + 1)
\end{aligned}$$

where second line follows from triangle inequality, third line from mclaurin's inequality, fourth line from Lemma 3, and the last line for $\alpha < \frac{1}{4L}$.

For A_4 , we have

$$\begin{aligned}
A_4 &\leq \sum_{t=1}^k |\gamma_t - \gamma_{t-1}| \mathbb{E} [\|\theta_{t-1} - \theta^*\| \|P_{\theta_{t-1}} \hat{g}(\theta_{t-1}, X_t)\| \|\theta_t - \theta^*\|^{2(n-1)}] \\
&\leq \sum_{t=1}^k L_{PH}^{(0)} |\gamma_t - \gamma_{t-1}| \mathbb{E} [\|\theta_{t-1} - \theta^*\| \|\theta_t - \theta^*\|^{2(n-1)}] \\
&\leq \sum_{t=1}^k L_{PH}^{(0)} |\gamma_t - \gamma_{t-1}| \mathbb{E} [\|\theta_{t-1} - \theta^*\| (\|\theta_t - \theta_{t-1}\| + \|\theta_{t-1} - \theta^*\|)^{2(n-1)}] \\
&\leq \sum_{t=1}^k 2^{2n-3} L_{PH}^{(0)} |\gamma_t - \gamma_{t-1}| \mathbb{E} [\|\theta_t - \theta_{t-1}\|^{2(n-1)} \|\theta_{t-1} - \theta^*\| + \|\theta_{t-1} - \theta^*\|^{2n-1}] \\
&\leq \sum_{t=1}^k 2^{2n-3} L_{PH}^{(0)} |\gamma_t - \gamma_{t-1}| (2^{4n-5} \alpha^{2n-2} L^{2n-2} (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n-1}] + 1) + 2^{2n-2} (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n-1}] + 1)) \\
&\leq \sum_{t=1}^k 2^{4n-4} L_{PH}^{(0)} |\gamma_t - \gamma_{t-1}| (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n-1}] + 1)
\end{aligned}$$

where the third line follows from triangle inequality, fourth line from mclaurin's inequality, fifth line from Lemma 3, and the last line for $\alpha < \frac{1}{4L}$.

Now for A_5 , we have to first note that, using mean-value theorem and with $a \in [0, 1]$, we'll get

$$\begin{aligned}\|\theta_t - \theta^*\|^{2(n-1)} - \|\theta_{t-1} - \theta^*\|^{2(n-1)} &= \|\theta_t - \theta_{t-1}\| \cdot 2(n-1) \|a(\theta_t - \theta^*) + (1-a)(\theta_{t-1} - \theta^*)\|^{2n-3} \\ &\leq \|\theta_t - \theta_{t-1}\| \cdot 2(n-1) \|a(\theta_t - \theta_{t-1}) + \theta_{t-1} - \theta^*\|^{2n-3} \\ &\leq 2^{2n-3}(n-1) \|\theta_t - \theta_{t-1}\| (\|\theta_t - \theta_{t-1}\|^{2n-3} + \|\theta_{t-1} - \theta^*\|^{2n-3})\end{aligned}$$

where the last line follows using the mclaurin's inequality. Plugging in the above upper bound to A_5 gives us

$$\begin{aligned}A_5 &\leq \sum_{t=1}^k 2^{2n-3}(n-1) \gamma_{t-1} \mathbb{E} [\langle \theta_{t-1} - \theta^*, P_{\theta_{t-1}} \hat{g}(\theta_{t-1} - \theta^*, X_t) \rangle \|\theta_t - \theta_{t-1}\| (\|\theta_t - \theta_{t-1}\|^{2n-3} + \|\theta_{t-1} - \theta^*\|^{2n-3})] \\ &\leq \sum_{t=1}^k 2^{2n-3}(n-1) \gamma_{t-1} \mathbb{E} [\|\theta_{t-1} - \theta^*\| \|P_{\theta_{t-1}} \hat{g}(\theta_{t-1}, X_t)\| \|\theta_t - \theta_{t-1}\| (\|\theta_t - \theta_{t-1}\|^{2n-3} + \|\theta_{t-1} - \theta^*\|^{2n-3})] \\ &\leq \sum_{t=1}^k 2^{2n-3}(n-1) L_{PH}^{(0)} \gamma_{t-1} \mathbb{E} [\|\theta_{t-1} - \theta^*\| \|\theta_t - \theta_{t-1}\| (\|\theta_t - \theta_{t-1}\|^{2n-3} + \|\theta_{t-1} - \theta^*\|^{2n-3})] \\ &\leq \sum_{t=1}^k 2^{2n-3}(n-1) L_{PH}^{(0)} \gamma_{t-1} \mathbb{E} [\|\theta_t - \theta_{t-1}\|^{2n-2} \|\theta_{t-1} - \theta^*\| + \|\theta_t - \theta_{t-1}\| \|\theta_{t-1} - \theta^*\|^{2n-2}] \\ &\leq \sum_{t=1}^k 2^{2n-3}(n-1) L_{PH}^{(0)} \gamma_{t-1} (2^{4n-5} \alpha^{2n-2} L^{2n-2} (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n-1}] + 1) + 2^{2n-2} \alpha L (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n-1}] + 1)) \\ &\leq \sum_{t=1}^k 2^{4n-4}(n-1) \alpha L L_{PH}^{(0)} \gamma_{t-1} (\mathbb{E} [\|\theta_{t-1} - \theta^*\|^{2n-1}] + 1)\end{aligned}$$

where the fifth line follows from Lemma 3 and the last line follows for $\alpha < \frac{1}{4L}$.

Finally, for A_6 , we obtain

$$A_6 \leq L_{PH}^{(0)} (\gamma_0 \|\theta_0 - \theta^*\|^{2n-1} + \gamma_k \|\theta_k - \theta^*\|^{2n-1})$$

which follows from Cauchy-Schwarz inequality and ??.

Combining the above terms gives us:

$$\begin{aligned}\sum_{t=0}^k \gamma_t \mathbb{E} [T'_{1,t}] &= \mathbb{E} \left[\sum_{t=0}^k \gamma_t \langle \theta_t - \theta^*, g(\theta_t, X_{t+1}) - \bar{g}(\theta_t) \rangle \|\theta_t - \theta^*\|^{2(n-1)} \right] \\ &\leq \sum_{t=0}^{k-1} 2^{4n-2} \alpha L L_{PH}^{(1)} \gamma_{t+1} (\mathbb{E} [\|\theta_t - \theta^*\|^{2n}] + 1) + \sum_{t=0}^{k-1} 2^{4n-4} \alpha L L_{PH}^{(0)} \gamma_{t+1} (\mathbb{E} [\|\theta_t - \theta^*\|^{2n-1}] + 1) \\ &\quad + \sum_{t=0}^{k-1} 2^{4n-4} L_{PH}^{(0)} |\gamma_{t+1} - \gamma_t| (\mathbb{E} [\|\theta_t - \theta^*\|^{2n-1}] + 1) + \sum_{t=0}^{k-1} 2^{4n-4} (n-1) \alpha L L_{PH}^{(0)} \gamma_{t+1} (\mathbb{E} [\|\theta_t - \theta^*\|^{2n-1}] + 1) \\ &\quad + L_{PH}^{(0)} (\gamma_0 \|\theta_0 - \theta^*\|^{2n-1} + \gamma_k \|\theta_k - \theta^*\|^{2n-1}) \\ &= \sum_{t=0}^{k-1} 2^{4n-2} \alpha L L_{PH}^{(1)} \gamma_{t+1} (\mathbb{E} [\|\theta_t - \theta^*\|^{2n}] + 1) + \sum_{t=0}^{k-1} 2^{4n-4} L_{PH}^{(0)} ((n\alpha L + 1) \gamma_{t+1} - \gamma_t) (\mathbb{E} [\|\theta_t - \theta^*\|^{2n-1}] + 1) \\ &\quad + L_{PH}^{(0)} (\gamma_0 \|\theta_0 - \theta^*\|^{2n-1} + \gamma_k \|\theta_k - \theta^*\|^{2n-1})\end{aligned}$$

where in the last equality we used the fact that for $\alpha \leq \frac{\mu}{L^2}$, $\gamma_t \leq \gamma_{t+1}$.

Now, if we make use of the inequality $2|x|^3 \leq x^2 + x^4$ and consolidating the terms, for sufficiently small

α , we have

$$\begin{aligned}
\mathbb{E}[\|\theta_{k+1} - \theta^*\|^{2n}] &\leq \sum_{t=0}^{k-1} 2^{4n-5} \left[8\alpha L L_{PH}^{(1)} \gamma_{t+1} + L_{PH}^{(0)} ((n\alpha L + 1) \gamma_{t+1} - \gamma_t) \right] (\mathbb{E}[\|\theta_t - \theta^*\|^{2n}] + 1) \\
&\quad + \sum_{t=0}^{k-1} 2^{4n-5} L_{PH}^{(0)} ((n\alpha L + 1) \gamma_{t+1} - \gamma_t) (\mathbb{E}[\|\theta_t - \theta^*\|^{2(n-1)}] + 1) + \frac{L_{PH}^{(0)}}{2} (\gamma_0 \|\theta_0 - \theta^*\|^{2(n-1)} + \gamma_k \|\theta_k - \theta^*\|^{2(n-1)}) \\
&\quad + \left(\frac{L_{PH}^{(0)}}{2} + 1 \right) \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|^{2n}] + \frac{L_{PH}^{(0)}}{2} \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|^{2n}] + o(\alpha) \\
&\leq \sum_{t=0}^{k-1} 2^{4n-4} \alpha L \gamma_{t+1} \left[4L_{PH}^{(1)} + nL_{PH}^{(0)} \right] \mathbb{E}[\|\theta_t - \theta^*\|^{2n}] + \sum_{t=0}^{k-1} 2^{4n-4} n\alpha L L_{PH}^{(0)} \mathbb{E}[\|\theta_t - \theta^*\|^{2(n-1)}] \\
&\quad + \frac{L_{PH}^{(0)}}{2} (\gamma_0 \|\theta_0 - \theta^*\|^{2(n-1)} + \gamma_k \|\theta_k - \theta^*\|^{2(n-1)}) + \left(\frac{L_{PH}^{(0)}}{2} + 1 \right) \gamma_0 \mathbb{E}[\|\theta_0 - \theta^*\|^{2n}] + \frac{L_{PH}^{(0)}}{2} \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|^{2n}] + \tilde{c}_{1,k} + o(\alpha)
\end{aligned}$$

where the last inequality follows for $(n\alpha L + 1) \gamma_{t+1} - \gamma_t \leq 2n\alpha L \gamma_{t+1}$ and also with

$$\tilde{c}_{1,t} := \sum_{l=0}^{t-1} 2^{4n-4} n\alpha L \gamma_{l+1} \left[4L_{PH}^{(1)} + nL_{PH}^{(0)} \right] + \sum_{l=0}^{t-1} 2^{4n-4} n\alpha L L_{PH}^{(0)}$$

being defined for $1 \leq t \leq k-1$. It can be seen that for sufficiently small α , $\tilde{c}_{1,t}$ would be of order $4^n n\alpha L (nL_{PH}^{(0)} + 4\alpha L_{PH}^{(1)})$. Now from induction hypothesis we know that

$$\mathbb{E}[\|\theta_t - \theta^*\|^{2(n-1)}] \leq \tilde{C}_{n-1,1} \cdot (1 - 2n\alpha\mu)^t \mathbb{E}[\|\theta_0 - \theta^*\|^{2n}] + \tilde{C}_{n-1,2} \cdot 4^{n-1} (n-1) \alpha L \left((n-1) L_{PH}^{(0)} + 2\alpha L_{PH}^{(1)} \right).$$

Thus, if we plug in this upper bound into the previous inequality, for $\alpha < \frac{1}{4^n}$ we will have

$$\begin{aligned}
\mathbb{E}[\|\theta_{k+1} - \theta^*\|^{2n}] &\leq \sum_{t=0}^{k-1} 2^{4n-4} \alpha L \gamma_{t+1} \left[4L_{PH}^{(1)} + nL_{PH}^{(0)} \right] \mathbb{E}[\|\theta_t - \theta^*\|^{2n}] \\
&\quad + \frac{L_{PH}^{(0)}}{2} \gamma_k \mathbb{E}[\|\theta_k - \theta^*\|^{2n}] + \tilde{c}_{2,k} \mathbb{E}[\|\theta_0 - \theta^*\|^{2n}] + \tilde{c}_{1,k} + \tilde{c}_{3,k} + o(\alpha)
\end{aligned}$$

in which $\tilde{c}_{3,k}$ would be of order $4^n n\alpha L (nL_{PH}^{(0)} + 4\alpha L_{PH}^{(1)})$ and $\tilde{c}_{2,k}$ is an $\mathcal{O}(1)$ constant dependent to k .

To solve the above recursion, we define $\tilde{S}_t := \frac{L_{PH}^{(0)}}{2} \gamma_t \mathbb{E}[\|\theta_t - \theta^*\|^{2n}] + \sum_{l=0}^{t-1} 2^{4n-4} \alpha L \gamma_{l+1} \left[4L_{PH}^{(1)} + nL_{PH}^{(0)} \right] \mathbb{E}[\|\theta_l - \theta^*\|^{2n}]$ for $1 \leq t \leq k$. Using this we can write

$$\mathbb{E}[\|\theta_{t+1} - \theta^*\|^{2n}] \leq \tilde{S}_t + \tilde{c}_{2,t} \mathbb{E}[\|\theta_0 - \theta^*\|^{2n}] + \tilde{c}_{1,t} + \tilde{c}_{3,t} + o(\alpha).$$

Now notice that we have (recall that $\gamma_t = 2n\alpha(1 - 2n\alpha\mu)^{k-t}$.)

$$\begin{aligned}
\frac{\tilde{S}_t}{\tilde{S}_{t-1}} &\leq \frac{\tilde{S}_{t-1} + \left(2^{4n-4} \alpha L \gamma_t \left[4L_{PH}^{(1)} + nL_{PH}^{(0)} \right] - \frac{L_{PH}^{(0)}}{2} \gamma_{t-1} \right) \mathbb{E}[\|\theta_{t-1} - \theta^*\|^{2n}] + \frac{L_{PH}^{(0)}}{2} \gamma_t \mathbb{E}[\|\theta_t - \theta^*\|^{2n}]}{\tilde{S}_{t-1}} \\
&= \frac{\tilde{S}_{t-1} + \left(2^{4n-4} \alpha L \gamma_t \left[4L_{PH}^{(1)} + nL_{PH}^{(0)} \right] - \frac{L_{PH}^{(0)}}{2} \gamma_{t-1} \right) \mathbb{E}[\|\theta_{t-1} - \theta^*\|^{2n}] + \frac{L_{PH}^{(0)}}{2} \gamma_t [\tilde{S}_{t-1} + \tilde{c}_{2,t-1} \mathbb{E}[\|\theta_0 - \theta^*\|^{2n}] + \tilde{c}_{1,t} + \tilde{c}_{3,t} + o(\alpha)]}{\tilde{S}_{t-1}}
\end{aligned}$$

it is not very hard to see that the if we take α small enough above fraction could get as close as possible

to 1. This means that we can write our final error bound as the following

$$\mathbb{E}[\|\theta_{k+1} - \theta^*\|^{2n}] \leq \tilde{C}_{n,1} \cdot (1 - 2n\alpha\mu)^{k+1} \mathbb{E}[\|\theta_0 - \theta^*\|^{2n}] + \tilde{C}_{n,2} \cdot 4^n n\alpha L \left(nL_{PH}^{(0)} + 4\alpha L_{PH}^{(1)} \right)$$

in which $\tilde{C}_{n,1}$ and $\tilde{C}_{n,2}$ are $\mathcal{O}(1)$ constants and also α has been taken small enough.

3 Bias Characterization

First, we recall that as a consequence of Assumption 2, the target equation $\bar{g}(\theta) = \mathbb{E}_{X \sim \pi_\theta} [g(\theta, X)] = 0$, has a unique solution which we denote by θ^* . Moreover, we define X^* to be the samples being drawn from the stationary distribution of the Markov chain related to θ^* ; *i.e.*, π_{θ^*} . Note that, from stationarity we know this is equal to the case that $\{X_n^*, n \in \mathbb{N}\}$ is a Markov chain initialized same as the other $\{X_n\}$ chain and evolving w.r.t. the P_{θ^*} . In this case we also take $\mathcal{F}_n = \{\theta_0, \{\xi_{m+1}\}_{m \leq n-1}, \{X_m\}_{m \leq n}, \{X_m^*\}_{m \leq n}\}$. For notational simplicity, we omit indexes related to the *optimal Chain* (*i.e.*, $\{X_n^*\}$).

Now, following the differentiability assumption of g in Assumption 1, we can apply Taylor expansion to g and note the following notation on residuals.

$$\begin{aligned} g(\theta, x) &= g(\theta^*, x) + g'(\theta^*, x)(\theta - \theta^*) + \frac{1}{2}g''(\theta^*, x)(\theta - \theta^*)^{\otimes 2} + R_3(\theta, x) \\ &= g(\theta^*, x) + g'(\theta^*, x)(\theta - \theta^*) + R_2(\theta, x). \end{aligned}$$

By Assumptions 1 and 2 and our boundedness of the error results, we note that the residual $R_n(\theta, X)$ satisfies

$$\sup_{x \in \mathcal{X}, \theta \in \mathbb{R}^d} \{\|R_n(\theta, x)\| / \|\theta - \theta^*\|^n\} < +\infty.$$

Hence, we have

$$\|R_n(\theta_\infty, X_{\infty+1})\|_{L^2(\pi_{\theta_\infty})} \lesssim \mathbb{E}[\|\theta_\infty - \theta^*\|^n] = \mathcal{O}(\alpha^{n/2}), \quad n = 2, 3, 4.$$

Lastly we denote

$$\begin{aligned} \bar{g}(\theta) &:= \mathbb{E}_{X \sim \pi_\theta} [g(\theta, X)], \quad \bar{g}_2(\theta) := \mathbb{E}_{X \sim \pi_\theta} [(g(\theta, X))^{\otimes 2}], \\ \bar{g}^{(1)}(\theta) &:= \mathbb{E}_{X \sim \pi_\theta} [g'(\theta, X)], \quad \bar{g}_2^{(1)}(\theta) := \mathbb{E}_{X \sim \pi_\theta} [(g'(\theta, X))^{\otimes 2}], \quad \bar{g}^{(2)}(\theta) := \mathbb{E}_{X \sim \pi_\theta} [g''(\theta, X)]. \end{aligned}$$

Proceeding to computing the bias, it is good to notice in the beginning that

$$\mathbb{E}[\theta_{\infty+1} - \theta^*] = \mathbb{E}[\theta_\infty - \theta^*] + \alpha (\mathbb{E}[g(\theta_\infty, X_{\infty+1})] + \mathbb{E}[\xi_{\infty+1}(\theta_\infty)]),$$

which immediately implies that $0 = \mathbb{E}[g(\theta_\infty, X_{\infty+1})]$. Writing the Taylor approximation again for this term and expanding it would give us:

$$\begin{aligned}
\mathbb{E}[g(\theta_\infty, X_{\infty+1})] &= \mathbb{E}[g(\theta^*, X_{\infty+1})] + \mathbb{E}[g'(\theta^*, X_{\infty+1})(\theta_\infty - \theta^*)] + \frac{1}{2}\mathbb{E}\left[(g''(\theta^*, X_{\infty+1})(\theta_\infty - \theta^*)^{\otimes 2})\right] + \mathbb{E}[R_3(\theta_\infty, X_{\infty+1})] \\
&= \mathbb{E}[g(\theta^*, X_{\infty+1}) - g(\theta^*, X^*)] + \mathbb{E}[g(\theta^*, X^*)] \\
&\quad + \mathbb{E}[(g'(\theta^*, X_{\infty+1}) - g'(\theta^*, X^*))(\theta_\infty - \theta^*)] + \mathbb{E}[g'(\theta^*, X^*)(\theta_\infty - \theta^*)] \\
&\quad + \frac{1}{2}\mathbb{E}\left[(g''(\theta^*, X_{\infty+1}) - g''(\theta^*, X^*))(\theta_\infty - \theta^*)^{\otimes 2}\right] + \frac{1}{2}\mathbb{E}\left[g''(\theta^*, X^*)(\theta_\infty - \theta^*)^{\otimes 2}\right] \\
&\quad + \mathbb{E}[R_3(\theta_\infty, X_{\infty+1})]
\end{aligned}$$

Firstly, note that from the definition we have $\mathbb{E}[g(\theta^*, X^*)] = 0$. Secondly, the two terms on the r.h.s. of the above ($\mathbb{E}[g'(\theta^*, X^*)(\theta_\infty - \theta^*)]$ and $\frac{1}{2}\mathbb{E}[g''(\theta^*, X^*)(\theta_\infty - \theta^*)^{\otimes 2}]$), are equivalent to the two terms in Bias Characterization of [1] which means that we do not need to analyze them here. Hence, it remains to analyze the three other terms.

To analyze the first term, we introduce a new variable $\tilde{X}_{n,k}$. This is the random variable which has initialized from X_0 , evolved according to the kernel P_{θ_i} for $i = 0, \dots, k-1$, and then with the kernel P_{θ^*} for $n-k$ times:

$$X_0 \xrightarrow{P_{\theta_0}} X_1 \xrightarrow{P_{\theta_1}} \dots \xrightarrow{P_{\theta_{k-1}}} X_k = \tilde{X}_{k,k} \xrightarrow{P_{\theta^*}} \dots \xrightarrow{P_{\theta^*}} \tilde{X}_{n-1,k} \xrightarrow{P_{\theta^*}} \tilde{X}_{n,k}$$

In addition, since we are interested in the difference between $g(\theta^*, X_{\infty+1})$ and $g(\theta^*, X^*)$, and the decision parameter of the function is fixed here, for notational simplicity we use $f(\cdot) := g(\theta^*, \cdot)$. Also note that $f(X_{\infty+1})$ and $f(X_\infty)$ are the same kind of objects in distribution. Using these new notations, we can write

$$\mathbb{E}[g(\theta^*, X_{\infty+1}) - g(\theta^*, X^*)] = \lim_{n \rightarrow \infty} \mathbb{E}[f(\tilde{X}_{n,n}) - f(\tilde{X}_{n,0})] = \lim_{n \rightarrow \infty} \mathbb{E}\left[\sum_{k=0}^{n-1} f(\tilde{X}_{n,k+1}) - f(\tilde{X}_{n,k})\right].$$

From the above we see that we are interested in calculating $f(\tilde{X}_{n,k+1}) - f(\tilde{X}_{n,k})$. Notice that $\tilde{X}_{n,k+1}$ and $\tilde{X}_{n,k}$ are the same random variables in their first k rounds of evolution. This means that we can write $\tilde{X}_{n,k+1} = P_{\theta^*}^{n-k-1} P_{\theta_k} \tilde{X}_{k,k}$ and $\tilde{X}_{n,k} = P_{\theta^*}^{n-k} \tilde{X}_{k,k}$.

Starting with the first one, we have

$$\begin{aligned}
\mathbb{E}[g(\theta^*, X_{\infty+1}) - g(\theta^*, X^*)] &= \mathbb{E}[\mathbb{E}[g(\theta^*, X_{\infty+1}) | \theta_\infty, X_\infty] - g(\theta^*, X^*)] \\
&= \mathbb{E}[P_{\theta_\infty} g(\theta^*, X_\infty) - g(\theta^*, X^*)] \\
&= \mathbb{E}[P_{\theta_\infty} g(\theta^*, X_\infty) - P_{\theta^*} g(\theta^*, X_\infty) + P_{\theta^*} g(\theta^*, X_\infty) - P_{\theta^*} g(\theta^*, X^*)] \\
&= \mathbb{E}[(P_{\theta_\infty} - P_{\theta^*}) g(\theta^*, X_\infty) + P_{\theta^*} (g(\theta^*, X_\infty) - g(\theta^*, X^*))],
\end{aligned}$$

where the third equality follows from the fact that X^* comes from the stationary distribution of θ^* . Now

notice that we have from Assumption 1 that $\mathbb{E}[\|g(\theta^*, X_\infty)\|] \leq L_1$. Additionally,

$$\mathbb{E}[\|P_{\theta_\infty} - P_{\theta^*}\|] \leq L_P \mathbb{E}[\|\theta_\infty - \theta^*\|],$$

from Assumption 7. For the remaining term we also have

$$\begin{aligned} \mathbb{E}[\|P_{\theta^*}(g(\theta^*, X_\infty) - g(\theta^*, X^*))\|] &= \mathbb{E}\left[\left\|\int_{\mathbb{R}} P_{\theta^*}[g(\theta^*, x)\pi_{\theta_\infty}(x) - g(\theta^*, x)\pi_{\theta^*}(x)]dx\right\|\right] \\ &\leq \mathbb{E}\left[L_S \|\theta_\infty - \theta^*\| \left\|\int_{\mathbb{R}} P_{\theta^*} g(\theta^*, x)dx\right\|\right] \\ &\leq L_1 L_S \mathbb{E}[\|\theta_\infty - \theta^*\|] \end{aligned}$$

where the second line follows from Assumption 8. Summing up all these would give us

$$\mathbb{E}[\|g(\theta^*, X_{\infty+1}) - g(\theta^*, X^*)\|] \leq L_1 (L_P + L_S) \mathbb{E}[\|\theta_\infty - \theta^*\|],$$

which is of $\mathcal{O}(\alpha^{1/2})$, considering our error bound.

To analyze the first derivative related terms, we use a method similar to [4] and [5] by conditioning on the information of τ steps before (what is the exact definition for this such that the TV distance between distributions becomes equal to α).

From Assumption 1 we know that $\|g(\theta_{\infty-\tau+k}, X_{\infty-\tau+k+1})\| \leq L_1 (\|\theta_{\infty-\tau+k} - \theta^*\| + 1)$. Thus we can write

$$\begin{aligned} &\mathbb{E}[\|(g'(\theta^*, X_{\infty+1}) - g'(\theta^*, X^*))(\theta_\infty - \theta^*)\|] \\ &\leq \mathbb{E}\left[\mathbb{E}\left[\|g'(\theta^*, X_{\infty+1}) - g'(\theta^*, X^*)\| \left\|\theta_{\infty-\tau} + \alpha \sum_{k=0}^{\tau-1} [g(\theta_{\infty-\tau+k}, X_{\infty-\tau+k+1}) + \xi_{\infty-\tau+k+1}(\theta_{\infty-\tau+k+1})] - \theta^*\right\| \middle| \mathcal{F}_{\infty-\tau}\right]\right] \\ &\leq \mathbb{E}\left[\mathbb{E}\left[\|g'(\theta^*, X_{\infty+1}) - g'(\theta^*, X^*)\| \left(\|\theta_{\infty-\tau} - \theta^*\| + \alpha \sum_{k=0}^{\tau-1} [g(\theta_{\infty-\tau+k}, X_{\infty-\tau+k+1}) + \xi_{\infty-\tau+k+1}(\theta_{\infty-\tau+k+1})]\right) \middle| \mathcal{F}_{\infty-\tau}\right]\right] \\ &\leq \mathbb{E}\left[\mathbb{E}\left[\|g'(\theta^*, X_{\infty+1}) - g'(\theta^*, X^*)\| \left[\|\theta_{\infty-\tau} - \theta^*\| + \alpha \sum_{k=0}^{\tau-1} (\|g(\theta_{\infty-\tau+k}, X_{\infty-\tau+k+1})\| + \|\xi_{\infty-\tau+k+1}(\theta_{\infty-\tau+k+1})\|)\right] \middle| \mathcal{F}_{\infty-\tau}\right]\right]. \end{aligned}$$

To handle the above expression, we separate two multiplications arising from the sum in the r.h.s term; one involving an α and the other one which is constant when being conditioned on $\mathcal{F}_{\infty-\tau}$. The second part which has an α coefficient, is easier to take care of and for that part we only upper bound the l.h.s term using the trivial upper bound that Assumption 1 gives us. Also for the other term, since the r.h.s term is constant when being conditioned, we use a similar procedure to the previous step and use the fact that in τ steps the chain has mixed enough to do the final bounding. Writing all of these in mathematical terms would give us

$$\begin{aligned} &\mathbb{E}[\mathbb{E}[\|g'(\theta^*, X_{\infty+1}) - g'(\theta^*, X^*)\| \|\theta_{\infty-\tau} - \theta^*\| \middle| \mathcal{F}_{\infty-\tau}]] \\ &= \mathbb{E}[\mathbb{E}[\|(P_{\theta_\infty} - P_{\theta^*})g'(\theta^*, X_\infty) + P_{\theta^*}(g'(\theta^*, X_\infty) - g'(\theta^*, X^*))\| \|\theta_{\infty-\tau} - \theta^*\| \middle| \mathcal{F}_{\infty-\tau}]] \\ &\leq \mathbb{E}[\mathbb{E}[(\|(P_{\theta_\infty} - P_{\theta^*})g'(\theta^*, X_\infty)\| + \|P_{\theta^*}(g'(\theta^*, X_\infty) - g'(\theta^*, X^*))\|) \|\theta_{\infty-\tau} - \theta^*\| \middle| \mathcal{F}_{\infty-\tau}]] \\ &\leq L_1 L_P \mathbb{E}[\mathbb{E}[\|\theta_\infty - \theta^*\| \|\theta_{\infty-\tau} - \theta^*\| \middle| \mathcal{F}_{\infty-\tau}]] + \mathbb{E}\left[\mathbb{E}\left[\left\|\int_{\mathbb{R}} P_{\theta^*}[g(\theta^*, x)\pi_{\theta_\infty}(x) - g(\theta^*, x)\pi_{\theta^*}(x)]dx\right\| \middle| \mathcal{F}_{\infty-\tau}\right] \|\theta_{\infty-\tau} - \theta^*\|\right] \\ &\leq \mathcal{O}(\alpha) \mathbb{E}[\|\theta_{\infty-\tau} - \theta^*\|] + L_1 L_S \mathbb{E}[\|\theta_\infty - \theta^*\| + \mathcal{O}(\alpha)] = \mathcal{O}(\alpha^2) + 2\mathcal{O}(\alpha) = \mathcal{O}(\alpha), \end{aligned}$$

in which the second line in the above has been done similar to the previous steps by replacing g' with g

and using the fact that equality of two expressions would result in equality of their norms. And,

$$\begin{aligned}
& \mathbb{E} \left[\mathbb{E} \left[\alpha \|g'(\theta^*, X_{\infty+1}) - g'(\theta^*, X^*)\| \sum_{k=0}^{\tau-1} (\|g(\theta_{\infty-\tau+k}, X_{\infty-\tau+k+1})\| + \|\xi_{\infty-\tau+k+1}(\theta_{\infty-\tau+k+1})\|) | \mathcal{F}_{\infty-\tau} \right] \right] \\
& \leq \mathbb{E} \left[\mathbb{E} \left[2\alpha L_1 \sum_{k=0}^{\tau-1} (\|g(\theta_{\infty-\tau+k}, X_{\infty-\tau+k+1})\| + \|\xi_{\infty-\tau+k+1}(\theta_{\infty-\tau+k+1})\|) | \mathcal{F}_{\infty-\tau} \right] \right] \\
& \leq 2\alpha L_1 \mathbb{E} \left[\sum_{k=0}^{\tau-1} (\|g(\theta_{\infty-\tau+k}, X_{\infty-\tau+k+1})\| + \|\xi_{\infty-\tau+k+1}(\theta_{\infty-\tau+k+1})\|) \right] = \mathcal{O}(\alpha\tau).
\end{aligned}$$

Where the last equality easily follows from our boundedness assumptions of the function g and the noise field. Thus, we showed that the first order term is of $\mathcal{O}(\alpha\tau)$.

Lastly, for the remaining second order term, using a Cauchy-Swarz inequality, we have from Assumption 1 that

$$\frac{1}{2} \mathbb{E} \left[\|(g''(\theta^*, X_{\infty+1}) - g''(\theta^*, X^*))(\theta_{\infty} - \theta^*)^{\otimes 2}\| \right] \leq \alpha L_1 \mathbb{E} \left[\|\theta_{\infty} - \theta^*\|^{\otimes 2} \right].$$

It has been showed in [1] that $\mathbb{E}[\|\theta_{\infty} - \theta^*\|^{\otimes 2}]$ is of $\mathcal{O}(\alpha^2\tau)$, which results in our term to be of $\mathcal{O}(\alpha^2\tau)$ and completes our proof.

Family of Transition Kernels: Theoretical Properties and Experimental Analysis

Below we summarize the nine kernels used in our experiments. For each kernel, we detail its theoretical properties and connect them to its observed behavior in the convergence plots.

1. Langevin (Gaussian) Kernel

1.1 Transition Rule. A discretization of the Ornstein-Uhlenbeck (OU) process: $P_{\theta}(x, dy) = \mathcal{N}\left(x - \frac{\epsilon^2}{2}(x - \theta), \epsilon^2\right)(dy)$.

1.2 Stationary Law. The stationary distribution is Gaussian, $\pi_{\theta} = \mathcal{N}(\theta, 1)$. The mean is $m(\theta) = \theta$.

1.3 Derivatives. $m'(\theta) = 1$ and $m''(\theta) = 0$. Both are constant and globally Lipschitz.

1.4 Experimental Analysis. As the archetypal "smooth" kernel, the Langevin process exhibits the expected $O(\alpha)$ bias behavior. The convergence curves for different stepsizes are clearly separated, indicating that smaller α values lead to significantly more accurate estimates. A subtle non-linear effect may be visible, as the improvement for the smallest stepsizes appears less pronounced, though this could also suggest that these runs require more iterations to reach their asymptotic floor.

2. Oscillatory Kernel

2.1 Transition Rule. An OU-type process where the mean function $m(\theta)$ oscillates wildly near zero:

$$m(\theta) = \begin{cases} \theta + \theta^2 \sin(1/\theta), & \theta \neq 0 \\ 0, & \theta = 0 \end{cases}$$

The transition is $P_{\theta}(x, dy) = \mathcal{N}\left(x - \frac{\epsilon^2}{2}(x - m(\theta)), \epsilon^2\right)(dy)$.

2.2 Stationary Law. $\pi_\theta = \mathcal{N}(m(\theta), 1)$.

2.3 Derivatives of $m(\theta)$ (for $\theta \neq 0$).

$$m'(\theta) = 1 + 2\theta \sin(1/\theta) - \cos(1/\theta),$$

$$m''(\theta) = 2 \sin(1/\theta) - \frac{2}{\theta} \cos(1/\theta) - \frac{1}{\theta^2} \sin(1/\theta).$$

2.4 Lipschitz Properties. $m'(\theta)$ is bounded on compact sets but is not globally Lipschitz. $m''(\theta)$ is unbounded near $\theta = 0$.

2.5 Experimental Analysis. This kernel is a canonical example of pathological non-smoothness. The experiments clearly demonstrate this: the convergence curves for all stepsizes are tightly clustered, showing almost no improvement in accuracy as α decreases. This aligns with the theoretical prediction of a slower, $O(\sqrt{\alpha})$ -like bias, where the benefits of a smaller stepsize are severely diminished by the violent oscillations of the kernel's derivatives.

3. OU Uniform Smooth Kernel

3.1 Transition Rule. An OU process driven by uniform noise: $y_{t+1} = x_t - \frac{\epsilon^2}{2}(x_t - \theta) + \xi_t$, $\xi_t \sim \mathcal{U}(-\epsilon, \epsilon)$.

3.2 Stationary Law. The mean of the stationary distribution is $\mathbb{E}_{\pi_\theta}[X] = \theta$.

3.3 Derivatives. The mean function is $m(\theta) = \theta$, with globally Lipschitz derivatives $m'(\theta) = 1$ and $m''(\theta) = 0$.

3.4 Experimental Analysis. While theoretically as smooth as the Langevin kernel, its experimental behavior is less ideal. The curves for different α are not as clearly separated, appearing only slightly better than the Oscillatory kernel. This suggests that while smoothness is a necessary condition for $O(\alpha)$ bias, the magnitude of the derivatives also matters. Here, the constant and relatively small derivatives may lead to a smaller leading constant in the bias term, making the differences between stepsizes harder to distinguish within the given number of iterations.

4. OU Uniform Sin Kernel

4.1 Transition Rule. An OU process with uniform noise and a mean function $m(\theta) = \frac{1}{2}|\sin \theta|$.

4.2 Stationary Law. The mean of the stationary distribution is $\mathbb{E}_{\pi_\theta}[X] = m(\theta)$.

4.3 Derivatives of $m(\theta)$ (for $\theta \neq k\pi, k \in \mathbb{Z}$).

$$m'(\theta) = \frac{1}{2} \cos \theta \cdot \begin{cases} 1 & \text{if } \sin \theta > 0 \\ -1 & \text{if } \sin \theta < 0 \end{cases}, \quad m''(\theta) = -\frac{1}{2}|\sin \theta|.$$

4.4 Lipschitz Properties. $m'(\theta)$ has jump discontinuities but is bounded. $m''(\theta)$ is continuous, bounded, and globally Lipschitz.

4.5 Experimental Analysis. This is one of the most interesting cases. Despite having a non-Lipschitz first derivative, its convergence behavior is remarkably similar to the smooth kernels, exhibiting a clear separation between the curves for different α . This strongly suggests that the nature of the

non-smoothness is critical. Because the "kinks" in the derivative are isolated and predictable, and the second derivative is well-behaved, the system retains the faster $O(\alpha)$ bias scaling.

5. OU Uniform SqrtOsc Kernel

5.1 Transition Rule. An OU process with uniform noise and a scaled oscillatory mean function:

$$m(\theta) = \begin{cases} 0.1(\theta + \theta^2 \sin(1/\theta)), & \theta \neq 0 \\ 0, & \theta = 0 \end{cases}$$

5.2 Derivatives. The derivatives of $m(\theta)$ are scaled versions of those from the Oscillatory Kernel (2). The first derivative is bounded but not Lipschitz, and the second derivative is unbounded.

5.3 Experimental Analysis. This kernel exhibits clear non-smooth properties, consistent with its highly irregular derivatives. As with the Oscillatory kernel, the final error levels are tightly clustered, indicating that decreasing the stepsize provides minimal benefit. This reinforces the conclusion that unbounded or wildly oscillating derivatives severely degrade the convergence rate of the bias.

6. Discrete Smooth 500 Kernel

6.1 Transition Rule. A 500-state birth-death process on a scaled state space $\mathcal{X} \subset [-1, 1]$. It is designed to mimic the OU dynamic. The probability of moving right from state x is a smooth logistic function of the difference $(\theta - x)$, creating a mean-reverting drift towards θ .

6.2 Stationary Law. For each θ , there is a unique stationary distribution π_θ . At $\theta = 0$, the process is a symmetric random walk on a centered state space, ensuring $\mathbb{E}_{\pi_0}[X] = 0$.

6.3 Derivatives. The transition probabilities are infinitely differentiable with respect to θ . Consequently, the mean of the stationary distribution, $m(\theta) = \mathbb{E}_{\pi_\theta}[X]$, is a smooth function with globally Lipschitz derivatives.

6.4 Experimental Analysis. As intended, this kernel successfully discretizes the properties of its continuous counterpart, 'OU Smooth'. The plot shows a clear separation between the different α curves, consistent with the $O(\alpha)$ bias expected from a smooth kernel. This demonstrates that discretization, when done carefully to preserve smoothness properties, does not fundamentally alter the convergence behavior.

7. RW Hinge 500 Kernel

7.1 Transition Rule. A random walk on $\{0, \dots, 499\}$ where the right-move probability is given by the hinge function $p(\theta) = \min(1, \max(0, \theta))$.

7.2 Stationary Law. For $0 < p(\theta) < 1$, the stationary distribution is $\pi_\theta(i) \propto \left(\frac{p(\theta)}{1-p(\theta)}\right)^i$.

7.3 Derivatives of $p(\theta)$.

$$p'(\theta) = \begin{cases} 1, & \text{if } 0 < \theta < 1, \\ 0, & \text{if } \theta < 0 \text{ or } \theta > 1, \end{cases} \quad p''(\theta) = 0 \quad (\text{a.e.}).$$

7.4 Lipschitz Properties. $p'(\theta)$ is not Lipschitz due to jump discontinuities at $\theta = 0$ and $\theta = 1$.

7.5 Experimental Analysis. This is another fascinating case. Like ‘OU Sin’, this is a non-smooth discrete kernel that exhibits behavior remarkably similar to the smooth kernels, with clear $O(\alpha)$ separation. This further supports the hypothesis that isolated, well-behaved "kinks" in the derivative do not degrade the bias to $O(\sqrt{\alpha})$. It is also observable that the smallest stepsize has not yet fully converged, suggesting that the combination of discrete states and small derivative magnitudes can lead to slow convergence dynamics.

8. RW ClipOsc 500 Kernel

8.1 Transition Rule. A random walk on $\{0, \dots, 499\}$ where the right-move probability is a clipped, oscillatory function of θ , $p(\theta) = \min(1, \max(0, r(\theta)))$, where $r(\theta) = \theta \sqrt{|\theta|} \sin(1/\theta)$.

8.2 Derivatives. The underlying function’s derivative $r'(\theta)$ oscillates unboundedly near $\theta = 0$, making the derivatives of the transition probability non-Lipschitz.

8.3 Experimental Analysis. This kernel clearly demonstrates its non-smooth properties. The final error levels show significant clustering, characteristic of a slower bias convergence rate. The plot also highlights extremely slow convergence for the two smallest stepsizes, which have not reached their asymptotic floors. This indicates that the combination of discrete states and pathological derivative behavior can make convergence exceptionally slow.

9. Two-State Centered Kernel

9.1 Transition Rule. A two-state Markov chain on $\{-0.5, 0.5\}$ where the probability of switching states is $p(\theta) = \cos^2(\theta)$.

9.2 Stationary Law. The stationary distribution is uniform, $\pi_\theta(-0.5) = \pi_\theta(0.5) = 1/2$, so its mean is always zero.

9.3 Derivatives. The derivatives of the transition probability $p(\theta)$ are $p'(\theta) = -\sin(2\theta)$ and $p''(\theta) = -2\cos(2\theta)$. Both are globally Lipschitz.

9.4 Experimental Analysis. As a simple, smooth discrete kernel, it displays the expected $O(\alpha)$ behavior with clear separation between the curves. However, like other discrete kernels, the final convergence appears slow, suggesting that reaching the asymptotic limit requires a very large number of iterations.

Key Findings and Theoretical Justification

- **Universal Ergodicity and Foundational Assumptions:** All nine kernels are ergodic and satisfy the foundational assumption that the transition probability P_θ is Lipschitz in θ with respect to the total variation norm. This ensures a valid theoretical framework for the stochastic approximation analysis.
- **Correlation of Asymptotic Bias with Derivative Regularity:** The simulation results reveal a clear correlation between the asymptotic behavior of the iterates and the regularity of the kernel’s parameter function derivatives. However, the initial hypothesis of a simple binary split between "smooth" ($O(\alpha)$ bias) and "non-smooth" ($O(\sqrt{\alpha})$ bias) kernels is insufficient. The experimental data suggests a more nuanced categorization based on the nature and magnitude of the derivative’s behavior.

- **Class 1: Smooth Kernels with Clear $O(\alpha)$ Bias.** This class includes the **Langevin**, **OU Smooth**, **Discrete Smooth 500**, and **Two-State** kernels. These kernels all possess globally Lipschitz first and second derivatives. As predicted by theory, they exhibit a distinct separation between the convergence curves for different stepsizes, confirming the expected $O(\alpha)$ scaling of the asymptotic bias. For the Langevin kernel, this separation is particularly clear. For OU Smooth and its discrete counterpart, the effect is visible but less pronounced, suggesting that the smaller magnitude of their derivatives results in a smaller leading constant in the bias term, making the curves appear closer. The slow final convergence of the discrete kernels also highlights the impact of the Markov chain’s mixing time.
- **Class 2: Pathological Kernels with Clustered Bias.** This class includes the **Oscillatory**, **OU SqrtOsc**, and **RW ClipOsc** kernels. These were designed with severe irregularities, such as wildly oscillating or unbounded derivatives near the origin. The experimental results are stark: the final error levels for all stepsizes are tightly clustered together. This indicates that decreasing the stepsize yields almost no improvement in accuracy, a hallmark of a much slower bias convergence rate, likely on the order of $O(\sqrt{\alpha})$. The extremely slow convergence of the RW ClipOsc kernel for small α further demonstrates how pathological derivative behavior, combined with the memory of a discrete chain, can severely hamper performance.
- **Class 3: Piecewise-Smooth Kernels Exhibiting $O(\alpha)$ Bias.** This is the most insightful category, containing the **OU Sin** and **RW Hinge 500** kernels. Theoretically, both are "non-smooth" because their first derivatives are not Lipschitz, possessing jump discontinuities. However, their experimental behavior is nearly indistinguishable from the truly smooth kernels, showing clear $O(\alpha)$ separation. This strongly suggests that the nature of the non-smoothness is paramount. Because the derivatives of these kernels are well-behaved and constant between isolated "kinks," the system retains the faster bias scaling. This challenges the simple smooth/non-smooth dichotomy and indicates that as long as the derivatives are not pathologically oscillatory, the algorithm’s performance remains robust.

References

- [1] D. Huo, Y. Zhang, Y. Chen, and Q. Xie. The collusion of memory and nonlinearity in stochastic approximation with constant stepsize. *arXiv preprint arXiv:2405.16732*, 2024.
- [2] M. Kaledin, E. Moulines, A. Naumov, V. Tadic, and H.-T. Wai. Finite time analysis of linear two-timescale stochastic approximation with markovian noise. In *Conference on Learning Theory*, pages 2144–2203. PMLR, 2020.
- [3] B. Karimi, B. Miasojedow, E. Moulines, and H.-T. Wai. Non-asymptotic analysis of biased stochastic approximation scheme. In *Conference on Learning Theory*, pages 1944–1974. PMLR, 2019.
- [4] Y. F. Wu, W. Zhang, P. Xu, and Q. Gu. A finite-time analysis of two time-scale actor-critic methods. *Advances in Neural Information Processing Systems*, 33:17617–17628, 2020.
- [5] S. Zou, T. Xu, and Y. Liang. Finite-sample analysis for sarsa with linear function approximation. *Advances in neural information processing systems*, 32, 2019.