Convergence of stochastic approximation for Lyapunov stable dynamics: a proof from first principles.

Christophe Andrieu[1] - Éric Moulines[2] - Stanislav Volkov[1]
[1]School of Mathematics, University of Bristol
[2]ENST, Paris.

July 13, 2004

Abstract

In this short note we present a proof, aimed at beginners, of the convergence of the stochastic approximation recursion $\theta_{i+1} = \theta_i + \gamma_{i+1} h(\theta_i) + \gamma_{i+1}\xi_{i+1}$ under the classical 0-level Kushner-Clark noise condition when the underlying dynamic is Lyapunov stable. The technique of proof relies on simple calculus arguments and bypasses the need for the introduction of the associated continuous time ODE. Future work includes the extension of the result to the case where the 0-level Kushner-Clark condition is replaced with a r-level condition for $r > 0$ [2].

We study the convergence of the recursion

$$\theta_{i+1} = \theta_i + \gamma_{i+1} h(\theta_i) + \gamma_{i+1}\xi_{i+1},$$

for a given function $h : \Theta \to \mathbb{R}^{n_\theta}$, a sequence of stepsizes $\{\gamma_i\}$ and a noise sequence $\{\xi_i\}$ under simple and verifiable conditions. We define for any $T > 0$

$$m(n,T) := \max\left\{ j \geq n : \sum_{i=n+1}^{j} \gamma_i \leq T \right\}$$

with the convention that for any sequence $\{a_i\}$, $\sum_{i=n+1}^{n} a_i = 0$. The required conditions are as follows:

Condition 1 *We work here with the following assumptions:*

1. Θ *is an open subset of* $\mathbb{R}^{n_\theta}$ *for some integer* $n_\theta \geq 1$.
2. $h : \Theta \to \mathbb{R}^{n_\theta}$ *is continuous and there exists a continuously differentiable function* $w : \Theta \to [0, +\infty)$ *such that:*
   (a) $\mathcal{L} := \{\theta : \langle \nabla w(\theta), h(\theta)\rangle = 0\}$ *is non-empty.*
   (b) *For any* $\theta \in \Theta \setminus \mathcal{L}$, $\langle \nabla w(\theta), h(\theta)\rangle < 0$,
   (c) *The closure of* $w(\mathcal{L})$ *has an empty interior.*
3. $\{\gamma_i\} \subset \mathbb{R}^+$ *and* $\sum_{i=1}^{+\infty} \gamma_i = +\infty$ *and* $\lim_{i\to\infty} \gamma_i = 0$.
4. $\{\theta_i\} \subset \mathcal{K}$ *for some compact set* $\mathcal{K} \subset \Theta$ *such that* $\mathcal{K} \cap \mathcal{L} \neq \emptyset$.
5. *For any* $T > 0$

$$\limsup_{n\to\infty} \sup_{n \leq k \leq m(n,T)} \left| \sum_{i=n+1}^{k} \gamma_i \xi_i \right| = 0. \tag{1}$$

We define for any $n \geq 0$, $\Xi_n := \sup_{n\leq k\leq m(n,T)} \left|\sum_{i=n+1}^{k} \gamma_i\xi_i\right|$. Following [1] we prove that these conditions are sufficient to ensure the convergence of $\{\theta_i\}$ to $\mathcal{L}$. We will use the following notation: for any $\delta > 0$ and $\mathcal{K} \subset \Theta$ we define $\mathcal{K}_\delta = \{x \in \Theta : d(x,\mathcal{K}) \leq \delta\}$, where $d(x,\mathcal{K}) := \inf\{|x-y|, y \in \mathcal{K}\}$. For $\phi : \Theta \to \mathbb{R}^{n_\phi}$ will denote $\|\phi\|_{\mathcal{K}} = \sup_{x\in\mathcal{K}} |\phi(x)|$. For $a, b \geq 0$ we define $\iota(a,b) := [0,a] \times [0,b]$. Our main result is that $\lim_{i\to\infty} d(\theta_i, \mathcal{L}) = 0$. We preface the main result with several intermediate lemmas.

The following result is proved in Lemma 2.1 of [1].

Lemma 2 *Assume Condition 1. Then for any compact set $\mathcal{K} \subset \Theta$ such that $0 < \inf_{\theta \in \mathcal{K}} |\langle \nabla w, h \rangle|$ and any $\varepsilon > -\sup_{\theta \in \mathcal{K}} \langle \nabla w, h \rangle$ there exists $\lambda > 0$ and $\beta > 0$ such that for any $\rho, |\xi| \in \iota(\lambda, \beta)$ and $\theta \in \mathcal{K}$ we have $w(\theta + \rho h(\theta) + \rho\xi) \leq w(\theta) - \rho\varepsilon$.*

We adapt Lemma 2.4 of [1] to the weak noise condition.

Lemma 3 *Assume Condition 1. Let $\mathcal{N} \subset \Theta$ be an open neighbourhood of $\mathcal{L} \cap \mathcal{K}$.Then there exist positive constants $\delta_0, \delta \leq \delta_0$, ε and λ (depending only on the set $\mathcal{N}$), such that for any $\delta' \in (0, \delta]$, any $\lambda' \in (0, \lambda]$ and any $T > 0$ one can find an integer N and define sequences $\left\{ \left\{ \bar{\theta}_{i,n} \right\}_{i \geq 0} \subset \mathcal{K}_{\delta'} \subset \Theta \right\}_{n \geq N}$ such that for any $n \geq N$*

$$\sup_{n \leq j \leq m(n,T)} \left| \theta_j - \bar{\theta}_{j,n} \right| \leq \delta', \quad \gamma_n \leq \lambda' \text{ and}$$

$$w(\bar{\theta}_{j,n}) \leq w(\bar{\theta}_{j-1,n}) - \gamma_j \varepsilon \text{ for all integers } j \in \{n+1, \ldots, m(n,T) \wedge \bar{\tau}_{\mathcal{N}}(n)\},$$

where $\bar{\tau}_{\mathcal{N}}(n) := \inf \left\{ k : n \leq k \leq m(n,T), \bar{\theta}_{k,n} \notin \mathcal{N} \right\}$.

Proof. Let $\delta_0 > 0$ be such that $\mathcal{K}_{\delta_0} \subset \Theta$. The set $\mathcal{K}_{\delta_0} \setminus \mathcal{N}$ is compact and $\sup_{\mathcal{K}_{\delta_0} \setminus \mathcal{N}} \langle \nabla w, h \rangle < 0$. By Lemma 2.1 in [1], choosing $\varepsilon > 0$ such that $\sup_{\mathcal{K}_{\delta_0} \setminus \mathcal{N}} \langle \nabla w, h \rangle < -\varepsilon$, one may find $\lambda > 0$ and $\beta > 0$ small enough so that for any $(\rho, |\zeta|) \in \iota(\lambda, \beta)$ and $\theta \in \mathcal{K}_{\delta_0} \setminus \mathcal{N}$

$$w(\theta + \rho h(\theta) + \rho\zeta) \leq w(\theta) - \rho\varepsilon. \tag{2}$$

Now one may choose $0 < \delta \leq \delta_0$ such that for all $(\theta, \bar{\theta}) \in \mathcal{K} \times \mathcal{K}_\delta$ satisfying $|\theta - \bar{\theta}| \leq \delta$,

$$\left| h(\theta) - h(\bar{\theta}) \right| \leq \beta \tag{3}$$

from the uniform continuity of h on $\mathcal{K}_\delta$. Under Eq. (1), for any $(\lambda', \delta') \in \iota(\lambda, \delta)$ and $T > 0$ there exists an integer $N = N(\lambda', \delta')$ such that for any $n \geq N$ we have $(\gamma_n, \Xi_n) \in \iota(\lambda', \delta')$. Now for any $n \geq N$ we define $\left\{ \bar{\theta}_{i,n} \right\}$ as $\bar{\theta}_{n,n} = \theta_n$,

$$\bar{\theta}_{i,n} = \bar{\theta}_{i-1,n} + \gamma_i h(\theta_{i-1}), \tag{4}$$

for $n+1 \leq i \leq m(n,T)$, and $\bar{\theta}_{i,n} = 0$ otherwise. By construction, for $n+1 \leq i \leq m(n,T)$ we have

$$\theta_i - \bar{\theta}_{i,n} = \sum_{j=n+1}^{i} \gamma_j \xi_j,$$

which implies that for any $n \geq N$, $\sup_{n \leq j \leq m(n,T)} \left| \theta_j - \bar{\theta}_{j,n} \right| \leq \delta'$. The two first statements of the lemma follow. On the other hand, for any $n \geq N$ and $n+1 \leq i \leq m(n,T)$ we have

$$\bar{\theta}_{i,n} = \bar{\theta}_{i-1,n} + \gamma_i h(\bar{\theta}_{i-1,n}) + \gamma_i \left(h(\theta_{i-1}) - h(\bar{\theta}_{i-1,n}) \right), \tag{5}$$

and since $|\theta_{i-1} - \bar{\theta}_{i-1,n}| \leq \delta' \leq \delta$, Eq. (3) implies that $\left| h(\theta_{i-1}) - h(\bar{\theta}_{i-1,n}) \right| \leq \beta$. Consequently Eq. (2) implies that whenever $\bar{\theta}_{i-1,n} \in \mathcal{K}_\delta \setminus \mathcal{N}$, $w(\bar{\theta}_{i,n}) \leq w(\bar{\theta}_{i-1,n}) - \gamma_i \varepsilon$, which concludes the proof. ■

Lemma 4 *Under Condition 1, the limit points of $\{w(\theta_i)\}$ belong to $w(\mathcal{L} \cap \mathcal{K})$.*

Proof. For any $\alpha > 0$ define the set $[w(\mathcal{L} \cap \mathcal{K})]_\alpha = \{x \in \mathbb{R} : d(x, w(\mathcal{L} \cap \mathcal{K})) \leq \alpha\}$ and let $\eta > 0$ be such that for $\theta, \theta' \in \mathcal{K}_\delta$ such that $|\theta - \theta'| \leq \eta$, $|w(\theta) - w(\theta')| \leq \alpha/8$ - such a η exists since w is continuous and $\mathcal{K}_\delta$ is a compact set. Since $\|w\|_{\mathcal{K}} < +\infty$, for any $\alpha > 0$, $[w(\mathcal{L} \cap \mathcal{K})]_\alpha$ is a finite union of disjoint closed intervals of length at least 2α. For any integer n, we define $\bar{\tau}_\alpha(n) := \inf \left\{ k : n \leq k \leq m(n,T) : w(\bar{\theta}_{k,n}) \notin [w(\mathcal{L} \cap \mathcal{K})]_\alpha \right\}$ and $\sigma_\alpha(n) := \inf \{k \geq n : \theta_k \notin [w(\mathcal{L} \cap \mathcal{K})]_\alpha\}$, with the convention that $\inf\{\emptyset\} = +\infty$.

For any $\alpha > 0$ defi ne $\mathcal{N} := \left\{ \theta \in \mathcal{K} : w(\theta) \in [w(\mathcal{L} \cap \mathcal{K})]_{\alpha/4} \right\}$ and apply Lemma 3 for $\lambda' \leq \{\eta / (2 \|h\|_{\mathcal{K}})\} \wedge \lambda$, $0 < \delta' \leq \delta \wedge \{\eta/2\}$ and $T^* = 2 \|w\|_{\mathcal{K}_\delta} / \varepsilon$, where $\delta, \delta_0, \varepsilon$ and λ are defined in Lemma 3. This together with Condition 1 shows the existence of an integer $N = N(\lambda', \delta')$ and sequences $\left\{ \left\{ \bar{\theta}_{i,n} \right\}_{i \geq 0} \right\}_{n \geq N}$ satisfying for any $n \geq N$

$$\sup_{n \leq j \leq m(n,T^*)} \left| \theta_j - \bar{\theta}_{j,n} \right| = \sup_{n \leq k \leq m(n,T^*)} \left| \sum_{i=n+1}^{k} \gamma_i \xi_i \right| \leq \delta', \qquad \gamma_n \leq \lambda',$$

and

$$w(\bar{\theta}_{j,n}) \leq w(\bar{\theta}_{j-1,n}) - \gamma_j \varepsilon \text{ for } j \in \{n+1, \ldots, m(n,T^*) \wedge \bar{\tau}_{\mathcal{N}}(n)\},$$

where we have used that $\mathbb{I}_{\mathcal{N}}(\theta) \leq \mathbb{I}_{[w(\mathcal{L}\cap\mathcal{K})]_{\alpha/4}}(w(\theta))$ for any $\theta \in \mathcal{K}_\delta \supset \mathcal{K}$.

We first prove four intermediate claims, which will lead to the conclusion that if $n \geq \sigma_{\alpha/2}(\bar{\tau}_{\alpha/4}(N))$ then $w(\theta_n) \in [w(\mathcal{L}\cap\mathcal{K})]_\alpha$.

1. For all $n \geq N$, $\bar{\tau}_{\alpha/4}(n) < m(n,T^*)$.
2. Let $I \in \mathbb{N}$ and $a_1 < b_1 < \cdots < a_I < b_I$ be such that $[w(\mathcal{L}\cap\mathcal{K})]_{\alpha/4} = \cup_{l=1}^{I}[a_l, b_l]$. For any $n \geq N$, $w(\theta_n) \geq a_1$.
3. Using the notation above, define $b(\theta_n) := \max\{b \in \{a_1, b_1, b_2, \ldots, b_I\} : b \leq w(\theta_n)\}$ for any $n \geq 1$. For $i \geq N$, if $w(\theta_i) \notin [w(\mathcal{L}\cap\mathcal{K})]_{\alpha/4}$ then for $j = i+1, \ldots, \bar{\tau}_{\alpha/4}(i) - 1$, $w(\theta_j) \in [b(\theta_i) - \alpha/4, w(\theta_i) + \alpha/8]$.
4. For any $i \geq N$, if $w(\theta_i) \in [w(\mathcal{L}\cap\mathcal{K})]_{\alpha/2}$ (and even if $w(\theta_{i+1}) \notin [w(\mathcal{L}\cap\mathcal{K})]_{\alpha/2}$) then $w(\theta_{i+1}) \in [w(\mathcal{L}\cap\mathcal{K})]_{3\alpha/4}$.

We prove these four points:

1. For any $n \geq N$, assuming $\bar{\tau}_{\alpha/4}(n) \geq m(n,T^*)$, we have

   $$0 \leq w(\bar{\theta}_{m(n,T^*),n}) \leq w(\bar{\theta}_{n,n}) - \varepsilon \sum_{j=n+1}^{m(n,T^*)} \gamma_j \leq \|w\|_{\mathcal{K}_\delta} - \varepsilon \sum_{j=n+1}^{m(n,T^*)} \gamma_j,$$

   which leads to a contradiction

   $$\|w\|_{\mathcal{K}_\delta} - \varepsilon \sum_{j=n+1}^{m(n,T^*)} \gamma_j \leq -\|w\|_{\mathcal{K}_\delta} + \gamma_{m(n,T^*)+1}\varepsilon \leq -\|w\|_{\mathcal{K}_\delta}/2$$

   for our choice of T^* and λ'. Therefore for $n \geq N$, $\bar{\tau}_{\alpha/4}(n) < m(n,T^*) < +\infty$.

2. If $w(\theta_n) < a_1$, then $w(\bar{\theta}_{j,n}) \leq w(\bar{\theta}_{j-1,n}) - \gamma_j\varepsilon$ for $n \leq j \leq m(n,T^*)$ yielding that $\bar{\tau}_{\alpha/4}(n) \geq m(n,T^*)$, which contradicts Step 1.

3. From Lemma 3 and our choice of δ' for $j = i+1, \ldots, \bar{\tau}_{\alpha/4}(i)$

   $$w(\bar{\theta}_{j,i}) \leq w(\theta_i) - \varepsilon \sum_{k=i+1}^{j} \gamma_k \leq w(\theta_i), \qquad \left|w(\bar{\theta}_{j,i}) - w(\theta_j)\right| \leq \alpha/8,$$

   hence

   $$w(\theta_j) \leq w(\theta_i) + \alpha/8.$$

   Now since $\gamma_{j+1} \leq \lambda' \leq \eta/(2\|h\|_{\mathcal{K}})$ and from Eq. (4) we also have that for $j = i+1, \ldots, \bar{\tau}_{\alpha/4}(i)$

   $$\left|w(\bar{\theta}_{j,i}) - w(\bar{\theta}_{j+1,i})\right| \leq \alpha/8,$$

   so that $\{w(\bar{\theta}_{j,i}); j = i+1, \ldots, \bar{\tau}_{\alpha/4}(i)\}$ cannot "jump over" $[w(\mathcal{L}\cap\mathcal{K})]_{\alpha/2}$ whose intervals are at least of length α. Therefore, when $w(\bar{\theta}_{j,i})$ hits $[w(\mathcal{L}\cap\mathcal{K})]_{\alpha/4}$, it must lie in the segment $[a_k, b_k]$ where $b_k \equiv b(\theta_i)$ and moreover $w\left(\theta_{\bar{\tau}_{\alpha/4}(i)}\right)$ cannot lie further than $\alpha/8 + \alpha/8 = \alpha/4$ from the right endpoint of this segment. The fact that $\{w(\bar{\theta}_{j,i}); j = i+1, \ldots, \bar{\tau}_{\alpha/4}(i)\}$ is decreasing concludes the proof of Step 3.

4. With our choice of λ' and δ' and N,

   $$\begin{aligned}\left|\gamma_{i+1}h(\theta_i) + \gamma_{i+1}\xi_{i+1}\right| &\leq \gamma_{i+1}\|h\|_{\mathcal{K}} + \sup_{n+1\leq k\leq m(n,T)} \left|\sum_{i=n+1}^{k} \gamma_i \xi_i\right| \\ &\leq \eta/2 + \eta/2 = \eta,\end{aligned}$$

and consequently from our choice of η,

$$|w(\theta_{i+1}) - w(\theta_i)| \leq \frac{\alpha}{8} < \frac{\alpha}{4},$$

and the result follows.

Using these four statements we have that for any $n \geq N$, $\bar{\tau}_{\alpha/4}(n) < m(n, T^*)$, for $n \leq i \leq \bar{\tau}_{\alpha/4}(n)$,

$$w(\theta_i) \in [b(\theta_n) - \alpha/4, w(\theta_n) + \alpha/8] \text{ and for } \bar{\tau}_{\alpha/4}(n) \leq i \leq \sigma_{\alpha/2}(\bar{\tau}_{\alpha/4}(n)) - 1, w(\theta_i) \in [w(\mathcal{L} \cap \mathcal{K})]_{\alpha/2}. \quad (6)$$

We also have from Steps 3 and 4 that $w\left(\theta_{\sigma_{\alpha/2}\left(\bar{\tau}_{\alpha/4}(n)\right)}\right) \in [w(\mathcal{L} \cap \mathcal{K})]_{3\alpha/4}$. Let $n_k = \sigma_{\alpha/2}\left(\bar{\tau}_{\alpha/4}(n_{k-1})\right)$ for $k \geq 1$ and $n_0 = N$. We proceed by induction to prove that $w(\theta_n) \in [w(\mathcal{L} \cap \mathcal{K})]_\alpha$ for $n \geq \sigma_{\alpha/2}(\bar{\tau}_{\alpha/4}(N))$. First, $w(\theta_{n_1}) = w\left(\theta_{\sigma_{\alpha/2}\left(\bar{\tau}_{\alpha/4}(N)\right)}\right) \in [w(\mathcal{L} \cap \mathcal{K})]_{3\alpha/4}$. Assume that for $k \geq 1$, $w(\theta_{n_k}) \in [w(\mathcal{L} \cap \mathcal{K})]_{3\alpha/4}$, then from Eq. (6) for $i = n_k, \ldots, n_{k+1} - 1$, $w(\theta_i) \in [w(\mathcal{L} \cap \mathcal{K})]_\alpha$ and $w\left(\theta_{n_{k+1}}\right) \in [w(\mathcal{L} \cap \mathcal{K})]_{3\alpha/4}$, which proves that for all $n \geq \sigma_{\alpha/2}\left(\bar{\tau}_{\alpha/4}(N)\right)$, $w(\theta_n) \in [w(\mathcal{L} \cap \mathcal{K})]_\alpha$.

Since $\{w(\theta_i)\}$ is bounded, it has limit points from the Bolzano-Weierstrass theorem. However we have just proved that these limit points must belong to $\cap_{\alpha>0} [w(\mathcal{L} \cap \mathcal{K})]_\alpha$, since $\alpha > 0$ was arbitrary. Because $\mathcal{L} \cap \mathcal{K}$ is a compact subset of $\mathbb{R}^{n_\theta}$ and w is continuous, $w(\mathcal{L} \cap \mathcal{K})$ is a compact subset of $\mathbb{R}$. This implies that $\cap_{\alpha>0} [w(\mathcal{L} \cap \mathcal{K})]_\alpha = w(\mathcal{L} \cap \mathcal{K})$, *i.e.* all limit points of $\{w(\theta_i)\}$ belong to the set $w(\mathcal{L} \cap \mathcal{K})$, yielding that $\lim_{i\to\infty} w(\theta_i)$ exists. Indeed, if it did not, there would be at least two separate limit points in $w(\mathcal{L} \cap \mathcal{K})$, say $w_1 < w_2$. However $\lim_{i\to\infty} |w(\theta_i) - w(\theta_{i-1})| = 0$ implies that all points of the interval $[w_1, w_2]$ are limit points for $\{w(\theta_i)\}$, which contradicts 2c of Condition 1. ■

Theorem 5 *Assume Condition 1. Then* $\lim_{i\to\infty} d(\theta_i, \mathcal{L}) = 0$.

Proof. Let $\mathcal{N} \subset \mathcal{K}$ be an arbitrary neighbourhood of $\mathcal{L} \cap \mathcal{K}$. Define for any integer n, $\bar{\tau}(n) := \inf\left\{j \geq n, \bar{\theta}_{j,n} \in \mathcal{N}\right\}$. Fix any $\alpha > 0$ and apply Lemma 3 and choose $\lambda' \in \left(0, \alpha/(9\|h\|_{\mathcal{K}_\delta}) \wedge \lambda\right]$ and $\delta' \in (0, \delta \wedge \alpha/3]$ such that $|w(\theta) - w(\theta')| \leq \varepsilon\alpha/(9\|h\|_{\mathcal{K}_\delta})$ as long as $\left|\theta - \theta'\right| \leq \delta'$ and $\theta, \theta' \in \mathcal{K}_\delta$; which is possible since w is uniformly continuous on the compact $\mathcal{K}_\delta$. Finally, let $T = \alpha/\left(3\|h\|_{\mathcal{K}_\delta}\right)$. By Lemmas 3 and 4 and Eq. (1) there exists an integer N and sequences $\left\{\left\{\bar{\theta}_{i,n}\right\}_{i\geq 0} \subset \mathcal{K}_{\delta'} \subset \Theta\right\}_{n\geq N}$ such that for any $n \geq N$ and any $j \in \{n, \ldots, m(n, T)\}$ the following five inequalities hold:

$$\begin{aligned} \left|\theta_j - \bar{\theta}_{j,n}\right| &\leq \delta', \\ \gamma_n &< \lambda', \end{aligned} \quad (7)$$

$$w(\bar{\theta}_{j,n}) \leq w(\bar{\theta}_{j-1,n}) - \gamma_j\varepsilon \text{ if } j \leq \bar{\tau}(n), \quad (8)$$

$$\left|w(\theta_n) - \lim_{i\to+\infty} w(\theta_i)\right| < \frac{\varepsilon\alpha}{18\|h\|_{\mathcal{K}_\delta}}, \quad (9)$$

and

$$\left|\sum_{i=n+1}^{j} \gamma_i \xi_i\right| \leq \frac{\alpha}{3}. \quad (10)$$

We prove that $\bar{\tau}(n) < m(n, T)$ for $n \geq N$. Assume that for some $n \geq N$ we have $\bar{\tau}(n) \geq m(n, T)$. Then from Eq. (8),

$$w(\bar{\theta}_{m(n,T),n}) - w(\bar{\theta}_{n,n}) = \sum_{i=n+1}^{m(n,T)} \left[w(\bar{\theta}_{i,n}) - w(\bar{\theta}_{i-1,n})\right] \leq -\varepsilon \sum_{i=n+1}^{m(n,T)} \gamma_i.$$

From Eq. (7) we have

$$\begin{aligned} w(\theta_{m(n,T)}) - w(\theta_n) &= w(\theta_{m(n,T)}) - w(\bar{\theta}_{m(n,T),n}) + w(\bar{\theta}_{m(n,T),n}) - w(\bar{\theta}_{n,n}) + \underbrace{w(\bar{\theta}_{n,n}) - w(\theta_n)}_{=0} \\ &\leq \frac{\varepsilon\alpha}{9\|h\|_{\mathcal{K}_\delta}} - \varepsilon \sum_{i=n+1}^{m(n,T)} \gamma_i. \end{aligned} \quad (11)$$

In turn, from Eq. (9) we have that the LHS of Eq. (11) is bounded in absolute value by $2 \times \alpha\varepsilon/(18\,\|h\|_{\mathcal{K}_\delta})$, hence

$$-\frac{\varepsilon\alpha}{9\,\|h\|_{\mathcal{K}_\delta}} \leq \frac{\varepsilon\alpha}{9\,\|h\|_{\mathcal{K}_\delta}} - \varepsilon \sum_{i=n+1}^{m(n,T)} \gamma_i. \tag{12}$$

On the other hand, from our choice of λ' and the definition of $m(n,T)$ we have

$$T = \frac{\alpha}{3\,\|h\|_{\mathcal{K}_\delta}} \leq \sum_{i=n+1}^{m(n,T)+1} \gamma_i$$
$$\frac{\alpha}{3\,\|h\|_{\mathcal{K}_\delta}} - \gamma_{m(n,T)+1} \leq \sum_{i=n+1}^{m(n,T)} \gamma_i,$$

contradicting Eq. (12) since $\left|\gamma_{m(n,T)+1}\right| < \alpha/(9\,\|h\|_{\mathcal{K}_\delta})$. Consequently $\bar{\tau}(n) < m(n,T)$ for all $n \geq N$. Now, for any integers $0 \leq n \leq k$ we have

$$\theta_k - \theta_n = \sum_{i=n+1}^{k} \gamma_i h(\theta_{i-1}) + \sum_{i=n+1}^{k} \gamma_i \xi_i$$

so that for any integers n, k such that $n \leq k \leq m(n,T)$,

$$\begin{aligned} |\theta_k - \theta_n| &\leq \|h\|_{\mathcal{K}_\delta} \sum_{i=n+1}^{k} \gamma_i + \left| \sum_{i=n+1}^{k} \gamma_i \xi_i \right| \\ &\leq \frac{\alpha}{3} + \sup_{n \leq l \leq m(n,T)} \left| \sum_{i=n+1}^{l} \gamma_i \xi_i \right|. \end{aligned} \tag{13}$$

In turn this implies that for $n \geq N$,

$$\begin{aligned} d(\theta_n, \mathcal{N}) &\leq |\bar{\theta}_{\bar{\tau}(n),n} - \theta_{\bar{\tau}(n)}| + |\theta_{\bar{\tau}(n)} - \theta_n| \leq \delta' + \|h\|_{\mathcal{K}_\delta} T + \sup_{n \leq k \leq m(n,T)} \left| \sum_{i=n+1}^{k} \gamma_i \xi_i \right| \\ &\leq \frac{\alpha}{3} + \frac{\alpha}{3} + \frac{\alpha}{3} = \alpha \end{aligned}$$

which can be made arbitrarily small. ■

References

[1] C. Andrieu, É. Moulines and P. Priouret, "Stability of stochastic approximation under verifiable conditions," *to appear SIAM Journal on Control and Optimisation, 2004.*

[2] V. Tadić, "Asymptotic analysis of stochastic approximation algorithms under violated Kushner-Clark conditions with applications", *CDC 2000.*