Supplement to: Efficient stochastic optimisation by unadjusted Langevin Monte Carlo. Application to maximum marginal likelihood and empirical Bayesian estimation.

Valentin De Bortoli *[1], Alain Durmus † [1], Marcelo Pereyra ‡ [2], and Ana Fernandez Vidal §[2]

[1]CMLA - École normale supérieure Paris-Saclay, CNRS, Université Paris-Saclay, 94235 Cachan, France.
[2]School of Mathematical and Computer Sciences, Heriot Watt University & Maxwell Institute for Mathematical Sciences.

September 10, 2020

S1 Postponed proofs

We first derive the following technical lemmas.

Lemma S1. *Let $t \in (0,1)$ and $\gamma \in (0,\bar{\gamma}]$ with $\bar{\gamma} > 0$ then $\sum_{n\in\mathbb{N}} t^{n\gamma} \leq t^{-\bar{\gamma}} \log^{-1}(1/t)\gamma^{-1}$ and $\sum_{n\in\mathbb{N}} n t^{n\gamma} \leq t^{-\bar{\gamma}} \log^{-2}(1/t)\gamma^{-2}$.*

Proof. Let $t \in (0,1)$ and $\gamma \in (0,\bar{\gamma}]$ with $\bar{\gamma} > 0$. Using that $e^u - 1 \leq u e^u$ for all $u \geq 0$, we have

$$\sum_{n\in\mathbb{N}} t^{n\gamma} = -(t^\gamma - 1)^{-1} \leq -\gamma^{-1}\log^{-1}(t)\exp(-\log(t)\gamma) \leq t^{-\bar{\gamma}} \log^{-1}(1/t)\gamma^{-1} ,$$

and

$$\sum_{n\in\mathbb{N}} n t^{n\gamma} = t^\gamma (t^\gamma - 1)^{-2} \leq t^\gamma \{\gamma^{-1}\log^{-1}(t)\exp(-\log(t)\gamma)\}^2 \leq t^{-\bar{\gamma}} \log^{-2}(1/t)\gamma^{-2} ,$$

which completes the proof. □

*Email: valentin.debortoli@cmla.ens-cachan.fr
†Email: alain.durmus@cmla.ens-cachan.fr
‡Email: m.pereyra@hw.ac.uk
§Email: af69@hw.ac.uk

Lemma S2. *For any probability measures μ, ν on $\mathcal{B}(\mathbb{R}^d)$, measurable function $V : \mathbb{R}^d \to [1, +\infty)$ such that $\mu(V) + \nu(V) < +\infty$ and $a \in (0, 1)$, we have*

$$\|\mu - \nu\|_{V^a} \leq 2\|\mu - \nu\|_V^a \ .$$

Proof. Let $a \in (0, 1]$. The statement is trivial if $\mu = \nu$. We just need to consider the case where $\mu \neq \nu$. Define $\xi = |\mu - \nu| / (|\mu - \nu| (\mathbb{R}^d))$. Using [5, Definition D.3.1] we get that

$$\begin{aligned}
\|\mu - \nu\|_{V^a} &= (1/2)\xi(V^a) \times |\mu - \nu| (\mathbb{R}^d) \\
&\leq (1/2)\xi(V)^a \times |\mu - \nu| (\mathbb{R}^d) \\
&\leq 2^{a-1}\|\mu - \nu\|_V^a \times [|\mu - \nu| (\mathbb{R}^d)]^{1-a} \ ,
\end{aligned}$$

which concludes the proof using that $a \leq 1$. □

Jensen's inequality implies that **H**1-(i) holds for $V \leftarrow V^a$ with $a \in (0, 1]$ since $A_1 \geq 1$. Lemma S2 implies that **H**1-(ii) holds replacing V by V^a, ρ by ρ^a and A_2 by $2A_2$. Similarly **H**1-(iii) holds replacing V by V^a and $\mathbf{\Psi}(\gamma)$ by $2\mathbf{\Psi}(\gamma)$.

S1.1 Proof of Theorem 3

Consider $(\eta_n)_{n\in\mathbb{N}}$ defined for any $n \in \mathbb{N}$ by

$$\eta_n = m_n^{-1} \sum_{k=1}^{m_n} \{H_{\theta_n}(X_k^n) - \pi_{\theta_n}(H_{\theta_n})\} \ . \tag{S1}$$

The proof of Theorem 3 relies on the two following lemmas. We consider the following decomposition for any $n \in \mathbb{N}$, $\eta_n = \eta_n^{(1)} + \eta_n^{(2)}$, where

$$\eta_n^{(1)} = \mathbb{E}\left[\eta_n | \mathcal{F}_{n-1}\right] \ , \qquad \eta_n^{(2)} = \eta_n - \mathbb{E}\left[\eta_n | \mathcal{F}_{n-1}\right] \ . \tag{S2}$$

We now give upper bounds on $\mathbb{E}[\|\eta_n^{(1)}\|]$, $\mathbb{E}[\|\eta_n^{(1)}\|^2]$ and $\mathbb{E}[\|\eta_n^{(2)}\|^2]$.

Lemma S3. *Assume* **A**1, **H**1 *and that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/2}(x)$. Then we have for any $n \in \mathbb{N}$*

$$\begin{aligned}
&\mathbb{E}\left[\|\eta_n^{(1)}\|\right] \leq B_1 \mathbb{E}\left[V^{1/2}(X_0^0)\right] / (m_n\gamma_n) + \mathbf{\Psi}(\gamma_n) \ ; \\
&\mathbb{E}\left[\|\eta_n^{(1)}\|^2\right] \leq 2B_1^2 \mathbb{E}\left[V(X_0^0)\right] / (m_n\gamma_n)^2 + 2\mathbf{\Psi}(\gamma_n)^2 \ ,
\end{aligned}$$

with

$$B_1 = 2A_1A_2\rho^{-\bar{\gamma}} / \log(1/\rho) \ .$$

Proof. Using the definition of $(\mathcal{F}_n)_{n\in\mathbb{N}}$, see (7), the Markov property, **H**1-(ii)-(iii), Lemma S2, Jensen's inequality and that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$,

$\|H_\theta(x)\| \leq V^{1/2}(x)$, we have for any $n \in \mathbb{N}^*$

$$
\begin{aligned}
\|\mathbb{E}\left[\eta_n | \mathcal{F}_{n-1}\right]\| &\leq m_n^{-1} \sum_{k=1}^{m_n} \left\| \mathrm{K}^k_{\gamma_n,\theta_n} H_{\theta_n}(X_0^n) - \pi_{\theta_n}\left(H_{\theta_n}\right)\right\| \\
&\leq m_n^{-1} \sum_{k=1}^{m_n} \left\| \left| \updelta_{X_0^n} \mathrm{K}^k_{\gamma_n,\theta_n} - \pi_{\theta_n} \right| (H_{\theta_n}) \right\| \\
&\leq m_n^{-1} \sum_{k=1}^{m_n} \left| \updelta_{X_0^n} \mathrm{K}^k_{\gamma_n,\theta_n} - \pi_{\theta_n} \right| (\|H_{\theta_n}\|) \\
&\leq m_n^{-1} \sum_{k=1}^{m_n} \left\{ \|\updelta_{X_0^n} \mathrm{K}^k_{\gamma_n,\theta_n} - \pi_{\gamma_n,\theta_n}\|_{V^{1/2}} \right\} + \|\pi_{\gamma_n,\theta_n} - \pi_{\theta_n}\|_{V^{1/2}} \\
&\leq m_n^{-1} \sum_{k=1}^{m_n} \left\{ 2A_2 \rho^{k\gamma_n} V^{1/2}(X^n_{m_n}) + \mathbf{\Psi}(\gamma_n) \right\} \\
&\leq \frac{2A_2 \rho^{-\bar{\gamma}} V^{1/2}(X^n_{m_n})}{\log(1/\rho)\gamma_n m_n} + \mathbf{\Psi}(\gamma_n) \, ,
\end{aligned}
$$

where for the last inequality we have used Lemma S1. In a similar manner, we have

$$
\left\|\mathbb{E}\left[\eta_0 | X_0^0\right]\right\| \leq \frac{2A_2 \rho^{-\bar{\gamma}} V^{1/2}(X_0^0)}{\log(1/\rho)\gamma_0 m_0} + \mathbf{\Psi}(\gamma_0) \, .
$$

We conclude using **H**1-(i) and that $(a+b)^2 \leq 2a^2 + 2b^2$ for $a, b \in \mathbb{R}$. $\square$

Lemma S4. *Assume* **A**1*,* **H**1 *and that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/2}(x)$. Then we have for any $n \in \mathbb{N}$*

$$
\mathbb{E}\left[\|\eta_n^{(2)}\|^2\right] \leq B_2 m_n^{-2} \gamma_n^{-1} \left(m_n + \gamma_n^{-1} \mathbb{E}\left[V(X_0^0)\right]\right) \, ,
$$

with $B_2 = 2(1+\bar{\gamma})^2 \max(B_{2,1}, B_{2,2})$ and

$$
\begin{aligned}
B_{2,1} &= 24 A_2^2 (1-\rho^{1/2})^{-2} A_3 \, , \\
B_{2,2} &= 4A_1 \left[1 + 6A_2^2(1-\rho^{1/2})^{-2}\left\{A_2(1-\rho)^{-1} + 2\right\} + A_2^2 \log^{-2}(1/\rho) + A_3^2\right] \, .
\end{aligned}
$$

Proof. Let $n \in \mathbb{N}^*$. We have using the Cauchy-Schwarz inequality

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\sum_{k=1}^{m_n} \left\{H_{\theta_n}(X_k^n) - \mathbb{E}\left[H_{\theta_n}(X_k^n) | \mathcal{F}_{n-1}\right]\right\}\right\|^2\right] \\
&\quad \leq 2\mathbb{E}\left[\left\|\sum_{k=1}^{m_n} \left\{H_{\theta_n}(X_k^n) - \pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\}\right\|^2\right] \\
&\quad + 2\mathbb{E}\left[\left\|\sum_{k=1}^{m_n} \left\{\mathbb{E}\left[H_{\theta_n}(X_k^n) | \mathcal{F}_{n-1}\right] - \pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\}\right\|^2\right] \qquad \text{(S3)}
\end{aligned}
$$

Using the Markov property, **H**1-(i)-(ii), Lemma S2, Lemma S1 and that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/2}(x)$ we obtain that

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\sum_{k=1}^{m_n}\left\{\mathbb{E}\left[H_{\theta_n}(X_k^n)|\mathcal{F}_{n-1}\right]-\pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\}\right\|^2\right] \\
&\qquad\leq \mathbb{E}\left[\left|\sum_{k=1}^{m_n}\mathbb{E}\left[\|\delta_{X_0^n}\mathrm{R}_{\gamma_n,\theta_n}-\pi_{\gamma_n,\theta_n}\|_{V^{1/2}}\middle|\mathcal{F}_{n-1}\right]\right|^2\right] \\
&\qquad\leq 4A_2^2\mathbb{E}\left[\left|\mathbb{E}\left[V^{1/2}(X_0^n)\middle|\mathcal{F}_{n-1}\right]\sum_{k=1}^{m_n}\rho^{k\gamma_n/2}\right|^2\right] \\
&\qquad\leq 4A_1A_2^2\gamma_n^{-2}\rho^{-2\bar{\gamma}}\log^{-2}(1/\rho)\mathbb{E}\left[V(X_0^0)\right] \; . \qquad\qquad (S4)
\end{aligned}
$$

We now give an upper-bound on the first term in the right-hand side of (S3). Consider for any $n \in \mathbb{N}$ the Euclidean division of m_n by $\lceil 1/\gamma_n \rceil$ there exist $q_n \in \mathbb{N}$ and $r_n \in \{0, \ldots, \lceil 1/\gamma_n \rceil - 1\}$ such that $m_n = q_n \lceil 1/\gamma_n \rceil + r_n$. Therefore using the Cauchy-Schwarz inequality we can derive the following decomposition

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\sum_{k=1}^{m_n}H_{\theta_n}(X_k^n)-\pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right] \\
&\leq 2\mathbb{E}\left[\left\|\sum_{j=1}^{r_n}H_{\theta_n}(X_{j+q_n\lceil 1/\gamma_n\rceil}^n)-\pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right] \\
&\quad+2\mathbb{E}\left[\left\|\sum_{j=1}^{\lceil 1/\gamma_n\rceil}\sum_{k=0}^{q_n-1}H_{\theta_n}(X_{j+k\lceil 1/\gamma_n\rceil}^n)-\pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right] \\
&\leq 2\mathbb{E}\left[\left\|\sum_{j=1}^{r_n}H_{\theta_n}(\bar{X}_{q_n}^{j,n})-\pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right] \\
&\quad+2\lceil 1/\gamma_n\rceil\sum_{j=1}^{\lceil 1/\gamma_n\rceil}\mathbb{E}\left[\left\|\sum_{k=0}^{q_n-1}H_{\theta_n}(\bar{X}_k^{j,n})-\pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right] \qquad (S5)
\end{aligned}
$$

Setting for any $j \in \{1, \ldots, \lceil 1/\gamma_n \rceil\}$ and $k \in \{0, \ldots, q_n - 1\}$, $\bar{X}_k^{j,n} = X_{j+k\lceil 1/\gamma_n \rceil}^n$. We now bound the two terms in the right-hand side. First, using the Cauchy-Schwarz inequality and **H**1-(i)-(iii), the fact that $r_n \leq \lceil 1/\gamma_n \rceil$ and that for any

$\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/2}(x)$ we have

$$\mathbb{E}\left[\left\|\sum_{j=1}^{r_n} H_{\theta_n}(\bar{X}_{q_n}^{j,n}) - \pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right]$$

$$\leq r_n \sum_{j=1}^{r_n} \mathbb{E}\left[\left\|H_{\theta_n}(\bar{X}_{q_n}^{j,n}) - \pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right]$$

$$\leq \lceil 1/\gamma_n \rceil^2 \left(2A_1 \mathbb{E}\left[V(X_0^0)\right] + 2A_3^2\right) . \tag{S6}$$

Now consider the solution of the *Poisson equation* [14, Section 17.4.1] associated with $\mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n \rceil}$, $x \mapsto \hat{H}_{\gamma_n,\theta_n}(x)$ defined for any $x \in \mathbb{R}^d$ by

$$\hat{H}_{\gamma_n,\theta_n}(x) = \sum_{\ell \in \mathbb{N}} \left(\mathrm{K}_{\gamma_n,\theta_n}^{\ell \lceil 1/\gamma_n \rceil} H_{\theta_n}(x) - \pi_{\gamma_n,\theta_n}(H_{\theta_n})\right) .$$

Note that by **H**1-(ii), Lemma S2 and since for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/2}(x)$, we have that for any $x \in \mathbb{R}^d$

$$\left\|\hat{H}_{\gamma_n,\theta_n}(x)\right\| \leq 2A_2(1-\rho^{1/2})^{-1} V^{1/2}(x) , \tag{S7}$$

and in addition for any $x \in \mathbb{R}^d$

$$\hat{H}_{\gamma_n,\theta_n}(x) - \mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n \rceil} \hat{H}_{\gamma_n,\theta_n}(x) = H_{\theta_n}(x) - \pi_{\gamma_n,\theta_n}(H_{\theta_n}) .$$

Therefore, we have for any $j \in \{1, \ldots, \lceil 1/\gamma_n \rceil\}$

$$\begin{aligned}
\sum_{k=0}^{q_n-1} \left(H_{\theta_n}(\bar{X}_k^{j,n}) - \pi_{\gamma_n,\theta_n}(H_{\theta_n})\right) &= \sum_{k=0}^{q_n-1} \left(\hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n}) - \mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n \rceil} \hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n})\right) \\
&= \sum_{k=0}^{q_n-2} \left(\hat{H}_{\gamma_n,\theta_n}(\bar{X}_{k+1}^{j,n}) - \mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n \rceil} \hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n})\right) \\
&\quad + \hat{H}_{\gamma_n,\theta_n}(\bar{X}_0^{j,n}) - \mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n \rceil} \hat{H}_{\gamma_n,\theta_n}(\bar{X}_{q_n-1}^{j,n}) .
\end{aligned} \tag{S8}$$

Combining the Cauchy-Schwarz inequality and (S8) we obtain that

$$\mathbb{E}\left[\left\|\sum_{k=0}^{q_n-1} H_{\theta_n}(\bar{X}_k^{j,n}) - \pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right] \leq 3(\mathrm{C}_1 + \mathrm{C}_2) , \tag{S9}$$

with

$$\mathrm{C}_1 = \mathbb{E}\left[\left\|\hat{H}_{\gamma_n,\theta_n}(\bar{X}_0^{j,n})\right\|^2 + \mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n \rceil} \left\|\hat{H}_{\gamma_n,\theta_n}(\bar{X}_{q_n-1}^{j,n})\right\|^2\right] ;$$

$$\mathrm{C}_2 = \mathbb{E}\left[\left\|\sum_{k=0}^{q_n-2} \hat{H}_{\gamma_n,\theta_n}(\bar{X}_{k+1}^{j,n}) - \mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n \rceil} \hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n})\right\|^2\right] .$$

First, using (S7) and **H**1-(i) we get that

$$\begin{aligned}
\mathrm{C}_1 &\leq 4A_2^2(1-\rho^{1/2})^{-2}\left\{\mathbb{E}\left[V(X_j^n)\right]+\mathbb{E}\left[\mathrm{K}_{\gamma_n,\theta_n}V(X_{q_n+j-1}^n)\right]\right\} \\
&\leq 8A_1A_2^2(1-\rho^{1/2})^{-2}\mathbb{E}\left[V(X_0^0)\right] \ . \qquad \text{(S10)}
\end{aligned}$$

We now give an upper-bound on C_2. For any $j \in \{1,\ldots,r_n\}$ let $(\mathcal{G}_{j,k})_{k\in\{0,q_n-2\}}$ generated by $\mathcal{F}_{n-1}$ and the sequence of random variables $X_0^n,\ldots,X_{k\lceil 1/\gamma_n\rceil+j}^n$. Using the Markov property we have for any $k \in \{0,\ldots,q_n-2\}$ and $j \in \{1,\ldots,r_n\}$

$$\mathbb{E}\left[\hat{H}_{\gamma_n,\theta_n}(X_{k+1}^{j,n})\middle|\mathcal{G}_{j,k}\right] = \mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n\rceil}\hat{H}_{\gamma_n,\theta_n}(X_k^{j,n}) \ .$$

Therefore, for any $j \in \{1,\ldots,r_n\}$, $\hat{H}_{\gamma_n,\theta_n}(X_{k+1}^{j,n}) - \mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n\rceil}\hat{H}_{\gamma_n,\theta_n}(X_k^{j,n})$ is a martingale increment with respect to $(\mathcal{G}_{j,k})_{k\in\{0,q_n-2\}}$, Combining this result with the Markov property implies that for any $k \in \{0,\ldots,q_n-2\}$ and $j \in \{1,\ldots,r_n\}$,

$$\begin{aligned}
\mathrm{C}_2 &= \sum_{k=0}^{q_n-2}\mathbb{E}\left[\mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n\rceil}\left\|\hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n}) - \mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n\rceil}\hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n})\right\|^2\right] \\
&= \sum_{k=0}^{q_n-2}\mathbb{E}\left[\mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n\rceil}\left\|\hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n})\right\|^2 - \left\|\mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n\rceil}\hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n})\right\|^2\right] \ . \qquad \text{(S11)}
\end{aligned}$$

Define for any $x \in \mathbb{R}^d$, $g_n(x) = \|\hat{H}_{\gamma_n,\theta_n}(x)\|^2$. Using (S11), **H** 1-(ii)-(iii) and (S7) we obtain that

$$\begin{aligned}
\mathrm{C}_2 &= \sum_{k=0}^{q_n-2}\mathbb{E}\left[\mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n\rceil}\left\|\hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n})\right\|^2 - \left\|\mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n\rceil}\hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n})\right\|^2\right] \\
&\leq \sum_{k=0}^{q_n-2}\mathbb{E}\left[\mathrm{K}_{\gamma_n,\theta_n}^{\lceil 1/\gamma_n\rceil}\left\|\hat{H}_{\gamma_n,\theta_n}(\bar{X}_k^{j,n})\right\|^2\right] \\
&\leq \mathbb{E}\left[\sum_{k=0}^{q_n-2}\mathbb{E}\left[\mathrm{K}_{\gamma_n,\theta_n}^{(k+1)\lceil 1/\gamma_n\rceil}g_n(\bar{X}_0^{j,n}) - \pi_{\gamma_n,\theta_n}(g_n)\middle|\mathcal{G}_{j,0}\right]\right] + \sum_{k=0}^{q_n-2}\pi_{\gamma_n,\theta_n}(g_n) \\
&\leq \frac{4A_2^2}{(1-\rho^{1/2})^2}\left\{\sum_{k=0}^{q_n-2}\mathbb{E}\left[\mathbb{E}\left[\|\delta_{X_j^n}\mathrm{K}_{\gamma_n,\theta_n}^{(k+1)\lceil 1/\gamma_n\rceil} - \pi_{\gamma_n,\theta_n}\|_V\middle|\mathcal{G}_{j,0}\right]\right]\right. \\
&\qquad\qquad \left. + \sum_{k=0}^{q_n-2}\pi_{\gamma_n,\theta_n}(V)\right\} \\
&\leq 4A_2^2(1-\rho^{1/2})^{-2}\left\{A_2(1-\rho)^{-1}\mathbb{E}\left[V(X_j^n)\right] + q_nA_3\right\} \\
&\leq 4A_2^2(1-\rho^{1/2})^{-2}\left\{A_1A_2(1-\rho)^{-1}\mathbb{E}\left[V(X_0^0)\right] + q_nA_3\right\} \ . \qquad \text{(S12)}
\end{aligned}$$

Therefore, using (S10) and (S12) in (S9) we obtain that

$$\mathbb{E}\left[\left\|\sum_{k=0}^{q_n-1} H_{\theta_n}(\bar{X}_k^{j,n}) - \pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right] \tag{S13}$$
$$\leq 12A_2^2(1-\rho^{1/2})^{-2}\left[\left\{A_1A_2(1-\rho)^{-1}\mathbb{E}\left[V(X_0^0)\right] + q_nA_3\right\} + 2\mathbb{E}\left[V(X_0^0)\right]\right] .$$

As a consequence, using (S6) and (S13) in (S5) we get that

$$\begin{aligned}
\mathbb{E}\left[\left\|\sum_{k=1}^{m_n} H_{\theta_n}(X_k^n) - \pi_{\gamma_n,\theta_n}(H_{\theta_n})\right\|^2\right] &\leq 4\left\lceil 1/\gamma_n\right\rceil^2\left(A_1\mathbb{E}\left[V(X_0^0)\right] + A_3^2\right) \\
&+ 24\left\lceil 1/\gamma_n\right\rceil^2 A_2^2(1-\rho^{1/2})^{-2}\left\{A_1\mathbb{E}\left[V(X_0^0)\right]\left(A_2(1-\rho)^{-1}+2\right) + q_nA_3\right\} \\
\leq \Big[\gamma_n^{-2}\Big(A_1\mathbb{E}\left[V(X_0^0)\right]&\left[24A_2^2(1-\rho^{1/2})^{-2}\left\{A_2(1-\rho)^{-1}+2\right\}+4\right] + 4A_3^2\Big) \\
+24A_2^2(1-\rho^{1/2})^{-2}A_3m_n/\gamma_n\Big]&(1+\bar{\gamma})^2
\end{aligned} \tag{S14}$$

Combining (S4) and (S14) in (S3) we obtain that

$$\begin{aligned}
&\mathbb{E}\left[\left\|\sum_{k=1}^{m_n} H_{\theta_n}(X_k^n) - \mathbb{E}\left[H_{\theta_n}(X_k^n)\right]\right\|^2\right] \leq 8\gamma_n^{-2}A_1A_2^2\rho^{-2\bar{\gamma}}\log^{-2}(1/\rho)\mathbb{E}\left[V(X_0^0)\right] \\
&+ 2\Big[\gamma_n^{-2}\Big(A_1\mathbb{E}\left[V(X_0^0)\right]\left[24A_2^2(1-\rho^{1/2})^{-2}\left\{A_2(1-\rho)^{-1}+2\right\}+4\right] + 4A_3^2\Big) \\
&+24A_2^2(1-\rho^{1/2})^{-2}A_3m_n/\gamma_n\Big](1+\bar{\gamma})^2 \\
&\leq 2(1+\bar{\gamma})^2\Big(A_1\mathbb{E}\left[V(X_0^0)\right]\Big[24A_2^2(1-\rho^{1/2})^{-2}\left\{A_2(1-\rho)^{-1}+2\right\} \\
&+4\left\{1+A_2^2\log^{-2}(1/\rho)\right\}\Big] + 4A_3^2\Big)\gamma_n^{-2} + 48A_2^2(1-\rho^{1/2})^{-2}A_3(1+\bar{\gamma})^2(m_n/\gamma_n) ,
\end{aligned}$$

which concludes the proof for $n \neq 0$. The same inequality holds in the case where $n = 0$. □

We now turn to the proof of Theorem 3.

of Theorem 3. The proof is an application of [1, Theorem 2, Theorem 3].

(a) To apply [1, Theorem 2], it is enough to show that the following series converge a.s

$$\sum_{n=0}^{+\infty}\delta_{n+1}\langle\Pi_\Theta(\theta_n - \delta_{n+1}\nabla f(\theta_n)), \eta_n^{(i)}\rangle , \quad \sum_{n=0}^{+\infty}\delta_{n+1}\eta_n^{(i)} , \quad \sum_{n=0}^{+\infty}\delta_{n+1}^2\|\eta_n^{(i)}\|^2 .$$

where $i \in \{1,2\}$ and the sequences $(\eta_n^{(1)})_{n\in\mathbb{N}}$ and $(\eta_n^{(2)})_{n\in\mathbb{N}}$ are given in (S3).

In the case where $i=1$, since $(\Pi_\Theta(\theta_n - \delta_{n+1}\nabla f(\theta_n)))_{n\in\mathbb{N}}$ is bounded, we are reduced to proving that a.s $\sum_{n=0}^{+\infty} \delta_{n+1}\|\eta_n^{(1)}\| < +\infty$. Using (11), Lemma S3 and Fubini-Tonelli's theorem we obtain that

$$\mathbb{E}\left[\sum_{n\in\mathbb{N}} \delta_{n+1}\|\eta_n^{(1)}\|\right] = \sum_{n\in\mathbb{N}} \delta_{n+1}\mathbb{E}\left[\|\eta_n^{(1)}\|\right] < +\infty \; . \tag{S15}$$

We consider the case where $i=2$. Let $(S_n)_{n\in\mathbb{N}}$ and $(T_n)_{n\in\mathbb{N}}$ be defined for any $n\in\mathbb{N}$ by $S_n = \sum_{k=0}^n \delta_{k+1}\langle \Pi_\Theta(\theta_k - \delta_{k+1}\nabla f(\theta_k)), \eta_k^{(2)}\rangle$ and $T_n = \sum_{k=0}^n \delta_{k+1}\eta_n^{(2)}$ are $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale by definition of $(\eta_n^{(2)})_{n\in\mathbb{N}}$ in (S3) and $(\mathcal{F}_n)_{n\in\mathbb{N}}$ in (7). Therefore, using [19, Section 12.5], the Cauchy-Schwarz inequality and that the sequence $(\Pi_\Theta(\theta_n - \delta_{n+1}\nabla f(\theta_n)))_{n\in\mathbb{N}}$ is bounded, it suffices to show that $\sum_{n=0}^{+\infty} \delta_{n+1}^2\mathbb{E}[\|\eta_n^{(2)}\|^2] < +\infty$. Using Lemma S4 we get that

$$\sum_{n=0}^{+\infty} \delta_{n+1}^2\mathbb{E}[\|\eta_n^{(2)}\|^2] \le B_2\left(\sum_{n=0}^{+\infty} \delta_{n+1}^2/(m_n\gamma_n) + \mathbb{E}\left[V(X_0^0)\right]\sum_{n=0}^{+\infty} \delta_{n+1}^2/(m_n\gamma_n)^2\right) \; .$$

Combining this result and (S15) implies the stated convergence applying [1, Theorem 2].

(b) Applying [1, Theorem 3], the Cauchy-Schwarz inequality and using **A**1 we obtain that a.s for any $n\in\mathbb{N}$

$$\begin{aligned} &\sum_{k=1}^n \delta_k\left\{f(\theta_k) - \min_\Theta f\right\} \qquad\qquad (\text{S16})\\ &\le \frac{\|\theta_0-\theta^\star\|^2}{2} - \sum_{k=0}^{n-1}\delta_{k+1}\langle \Pi_\Theta(\theta_k - \delta_{k+1}\nabla f(\theta_k)) - \theta^\star, \eta_k\rangle + \sum_{k=0}^{n-1}\delta_{k+1}^2\|\eta_k\|^2\\ &\le 2M_\Theta^2 - \sum_{i=1}^2\sum_{k=0}^{n-1}\delta_{k+1}\langle \Pi_\Theta(\theta_k - \delta_{k+1}\nabla f(\theta_k)) - \theta^\star, \eta_k^{(i)}\rangle + 2\sum_{\mathrm{i}=1}^2\sum_{k=0}^{n-1}\delta_{k+1}^2\|\eta_k^{(i)}\|^2 \; . \end{aligned}$$

which implies by the proof of (a) that $\sup_{n\in\mathbb{N}}[\sum_{k=1}^n \delta_k\left\{f(\theta_k) - \min_\Theta f\right\}] < +\infty$ a.s. The proof is then completed upon dividing (S16) by $\sum_{k=1}^n \delta_k$.

□

S1.2 Proof of Theorem 4

Proof. Taking the expectation in (S16) and using that $\eta_n^{(2)}$ is a martingale increment with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$, we get that for every $n\in\mathbb{N}$

$$\begin{aligned} \mathbb{E}\left[\sum_{k=1}^n \delta_k\left\{f(\theta_k) - \min_\Theta f\right\}\right] &\le 2M_\Theta^2 + 2M_\Theta\sum_{k=0}^{n-1}\delta_{k+1}\mathbb{E}\left[\left\|\eta_k^{(1)}\right\|\right]\\ &\quad + 2\sum_{k=0}^{n-1}\delta_{k+1}^2\mathbb{E}\left[\left\|\eta_k^{(1)}\right\|^2\right] + 2\sum_{k=0}^{n-1}\delta_{k+1}^2\mathbb{E}\left[\left\|\eta_k^{(2)}\right\|^2\right] \; . \end{aligned}$$

Combining this result, Lemma S3 and Lemma S4 completes the proof. □

S1.3 Proof of Theorem 5

We now introduce some tools needed for the proof. By **A**2 and **H**1-(i)-(ii), for any $\theta \in \Theta$ and $\gamma \in (0, \bar{\gamma}]$, there exists a function $\hat{H}_{\gamma,\theta} : \mathbb{R}^d \to \mathbb{R}^{d_\theta}$ solution of the *Poisson equation*,

$$(\mathrm{Id} - \mathrm{K}_{\gamma,\theta})\hat{H}_{\gamma,\theta} = H_\theta - \pi_{\gamma,\theta}(H_\theta) \ , \tag{S17}$$

defined for any $x \in \mathbb{R}^d$ by

$$\hat{H}_{\gamma,\theta}(x) = \sum_{j \in \mathbb{N}} \{\mathrm{K}^j_{\gamma,\theta} H_\theta(x) - \pi_{\gamma,\theta}(H_\theta)\} \ . \tag{S18}$$

Note that using **H**1-(ii) and Lemma S2 we have for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$

$$\left\| \hat{H}_\theta(x) \right\| \leq C_{\hat{H}} \gamma^{-1} V^{1/4}(x) \ , \qquad C_{\hat{H}} = 8A_2 \log^{-1}(1/\rho) \rho^{-\bar{\gamma}/4} \ . \tag{S19}$$

Define for any $n \in \mathbb{N}$

$$\tilde{\eta}_n = H_{\theta_n}(\tilde{X}_{n+1}) - \pi_{\tilde{\theta}_n}(H_{\tilde{\theta}_n}) \ . \tag{S20}$$

Using (S17) an alternative expression of $(\tilde{\eta}_n)_{n \in \mathbb{N}}$ is given for any $n \in \mathbb{N}$ by

$$\begin{aligned} \tilde{\eta}_n &= \hat{H}_{\gamma_n,\tilde{\theta}_n}(\tilde{X}_{n+1}) - \mathrm{K}_{\gamma_n,\tilde{\theta}_n} \hat{H}_{\gamma_n,\tilde{\theta}_n}(\tilde{X}_{n+1}) + \pi_{\gamma_n,\tilde{\theta}_n}(H_{\tilde{\theta}_n}) - \pi_{\tilde{\theta}_n}(H_{\tilde{\theta}_n}) \\ &= \tilde{\eta}_n^{(a)} + \tilde{\eta}_n^{(b)} + \tilde{\eta}_n^{(c)} + \tilde{\eta}_n^{(d)} \ , \end{aligned}$$

where

$$\begin{aligned} \tilde{\eta}_n^{(a)} &= \hat{H}_{\gamma_n,\tilde{\theta}_n}(\tilde{X}_{n+1}) - \mathrm{K}_{\gamma_n,\tilde{\theta}_n} \hat{H}_{\gamma_n,\tilde{\theta}_n}(\tilde{X}_n) \ , \\ \tilde{\eta}_n^{(b)} &= \mathrm{K}_{\gamma_n,\tilde{\theta}_n} \hat{H}_{\gamma_n,\tilde{\theta}_n}(\tilde{X}_n) - \mathrm{K}_{\gamma_{n+1},\tilde{\theta}_{n+1}} \hat{H}_{\gamma_{n+1},\tilde{\theta}_{n+1}}(\tilde{X}_{n+1}) \ , \\ \tilde{\eta}_n^{(c)} &= \mathrm{K}_{\gamma_{n+1},\tilde{\theta}_{n+1}} \hat{H}_{\gamma_{n+1},\tilde{\theta}_{n+1}}(\tilde{X}_{n+1}) - \mathrm{K}_{\gamma_n,\tilde{\theta}_n} \hat{H}_{\gamma_n,\tilde{\theta}_n}(\tilde{X}_{n+1}) \ , \\ \tilde{\eta}_n^{(d)} &= \pi_{\gamma_n,\tilde{\theta}_n}(H_{\tilde{\theta}_n}) - \pi_{\tilde{\theta}_n}(H_{\tilde{\theta}_n}) \ . \end{aligned} \tag{S21}$$

To establish Theorem 5 we need to get estimates on moments of $\left\| \tilde{\eta}_n^{(i)} \right\|$ for $i \in \{a, b, c, d\}$. It is the matter of the following technical results.

Lemma S5. *Assume* **A**1 *and that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/4}(x)$. Then we have for any $n \in \mathbb{N}$, $\mathbb{E}\left[\|\tilde{\eta}_n\|^2\right] \leq C_1$, with*

$$C_1 = 2A_1 \mathbb{E}\left[V^{1/2}(\tilde{X}_0)\right] + 2 \sup_{\theta \in \Theta} \|\nabla f(\theta)\|^2 \ .$$

Proof. Using (S20), that $\|x + y\|^2 \leq 2(\|x\|^2 + \|y\|^2)$ for any $x, y \in \mathbb{R}^d$, **A**1, $\|H_\theta(x)\| \leq V^{1/2}(x)$, we get for any $k \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E}\left[\|\tilde{\eta}_k\|^2\right] &\leq 2\mathbb{E}[\|H_{\tilde{\theta}_k}(\tilde{X}_{k+1})\|^2] + 2\left[\pi_{\tilde{\theta}_k}(\|H_{\tilde{\theta}_k}\|)\right]^2 \\ &\leq 2A_1 \mathbb{E}\left[V^{1/2}(\tilde{X}_0)\right] + 2 \sup_{\theta \in \Theta} \|\nabla f(\theta)\|^2 \ . \end{aligned}$$

□

Lemma S6. *Assume* **A**1*,* **A**2*,* **H**1*,* **H**2 *and that for any* $\theta \in \Theta$ *and* $x \in \mathbb{R}^d$*,* $\|H_\theta(x)\| \leq V^{1/4}(x)$*. Then we have for any* $n \in \mathbb{N}$*,* $\mathbb{E}\left[\left\|\tilde{\eta}_n^{(a)}\right\|^2\right] \leq \tilde{C}_1 \gamma_n^{-2}$*, with*

$$\tilde{C}_1 = A_1 C_{\hat{H}}^2 \mathbb{E}\left[V^{1/2}(\tilde{X}_0)\right] .$$

Proof. By (S21), using (S19) and **H**1-(i) we get that for any $n \in \mathbb{N}^*$

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[\left\|\tilde{\eta}_n^{(a)}\right\|^2 \middle| \mathcal{F}_n\right]\right] \\
&\leq \mathbb{E}\left[\mathbb{E}\left[\left\|\hat{H}_{\gamma_n,\tilde{\theta}_n}(\tilde{X}_{n+1})\right\|^2 \middle| \mathcal{F}_n\right]\right] - \mathbb{E}\left[\left\|\mathrm{K}_{\gamma_n,\tilde{\theta}_n}\hat{H}_{\gamma_n,\tilde{\theta}_n}(\tilde{X}_n)\right\|^2\right] \\
&\leq A_1 C_{\hat{H}}^2 \gamma_n^{-2} \mathbb{E}\left[V^{1/2}(\tilde{X}_0)\right] ,
\end{aligned}$$

which concludes the proof. □

Lemma S7. *Assume* **A**1*,* **H**1 *and that for any* $\theta \in \Theta$ *and* $x \in \mathbb{R}^d$*,* $\|H_\theta(x)\| \leq V^{1/4}(x)$*. Then the following statements hold.*

(a) There exists $C_3 \geq 0$ *such that for any* $n \in \mathbb{N}$ *and* $\theta \in \Theta$

$$\begin{aligned}
&\mathbb{E}\left[\left\|\sum_{k=0}^n \delta_{k+1}\langle a_{k+1}, \tilde{\eta}_k^b\rangle\right\|\right] \\
&\qquad \leq C_3\left[\sum_{k=0}^n |\delta_{k+1} - \delta_k|\,\gamma_k^{-1} + \sum_{k=0}^n \delta_{k+1}^2\gamma_k^{-1} + (\delta_{n+1}/\gamma_{n+1} - \delta_1/\gamma_1)\right] .
\end{aligned}$$

with $a_{k+1} = \Pi_\Theta\left[\tilde{\theta}_k - \delta_{k+1}\nabla f(\tilde{\theta}_k)\right] - \theta^\star$*,* $\theta^\star \in \arg\min_\Theta f$ *and*

$$C_3 = A_1 C_{\hat{H}}(4M_\Theta + \sup_{\theta\in\Theta}\|\nabla f(\theta)\| + 1 + \delta_1 L_f)\mathbb{E}\left[V^{1/4}(\tilde{X}_0)\right] .$$

(b) If (17) *holds then* $\sum_{k=0}^n \delta_{k+1}\langle a_{k+1}, \tilde{\eta}_k^{(b)}\rangle$ *converges a.s.*

Proof. By (S21) we have for any $n \in \mathbb{N}$ and $\theta \in \Theta$

$$\begin{aligned}
&\sum_{k=0}^n \delta_{k+1}\langle a_{k+1}, \tilde{\eta}_k^{(b)}\rangle \\
&= \sum_{k=0}^n \langle \delta_{k+1}a_{k+1}, \mathrm{K}_{\gamma_k,\tilde{\theta}_k}\hat{H}_{\gamma_k,\tilde{\theta}_k}(\tilde{X}_k) - \mathrm{K}_{\gamma_{k+1},\tilde{\theta}_{k+1}}\hat{H}_{\gamma_{k+1},\tilde{\theta}_{k+1}}(\tilde{X}_{k+1})\rangle \\
&= \sum_{k=1}^n \langle \delta_{k+1}a_{k+1} - \delta_k a_k, \mathrm{K}_{\gamma_k,\tilde{\theta}_k}\hat{H}_{\gamma_k,\tilde{\theta}_k}(\tilde{X}_k)\rangle \\
&- \langle \delta_{n+1}a_{n+1}, \mathrm{K}_{\gamma_{n+1},\tilde{\theta}_{n+1}}\hat{H}_{\gamma_{n+1},\tilde{\theta}_{n+1}}(\tilde{X}_{n+1})\rangle \\
&+ \langle \delta_1 a_1, \mathrm{K}_{\gamma_0,\tilde{\theta}_0}\hat{H}_{\gamma_0,\tilde{\theta}_0}(\tilde{X}_0)\rangle ,
\end{aligned} \tag{S22}$$

In addition, we have for any $n \in \mathbb{N}$, $\theta \in \Theta$ using **A**1, that Π_Θ is non-expansive, (15), **H**1-(i) and that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/4}(x)$

$$\begin{aligned}\|\delta_{n+1}a_{n+1} - \delta_n a_n\| &\leq 2M_\Theta \left|\delta_{n+1} - \delta_n\right| + \delta_{n+1} \left\|a_{n+1} - a_n\right\| \\ &\leq 2M_\Theta \left|\delta_{n+1} - \delta_n\right| + (1 + \delta_n L_f) \left\|\theta_{n+1} - \theta_n\right\| + \left|\delta_{n+1} - \delta_n\right| \left\|\nabla f(\theta_{n+1})\right\| \\ &\leq (2M_\Theta + \sup_{\theta \in \Theta} \|\nabla f(\theta)\|) \left|\delta_{n+1} - \delta_n\right| + \delta_{n+1}^2 (1 + \delta_{n+1} L_f) V^{1/4}(\tilde{X}_{n+1}) \ .\end{aligned} \tag{S23}$$

(a) Combining (S22), (S23), (S19), the Cauchy-Schwarz inequality and **H**1-(i) we get that

$$\begin{aligned}&\mathbb{E}\left[\left\|\sum_{k=0}^{n} \delta_{k+1} \langle a_k, \tilde{\eta}_k^{[b)} \rangle\right\|\right] \\ &\quad\leq (2M_\Theta + \sup_{\theta \in \Theta} \|\nabla f(\theta)\|) A_1 C_{\hat{H}} \mathbb{E}\left[V^{1/4}(\tilde{X}_0)\right] \sum_{k=0}^{n} \left|\delta_{k+1} - \delta_k\right| \gamma_k^{-1} \\ &\quad+ A_1 C_{\hat{H}} (1 + \delta_1 L_f) \mathbb{E}\left[V^{1/4}(\tilde{X}_0)\right] \sum_{k=0}^{n} \delta_{k+1}^2 \gamma_k^{-1} \\ &\quad+ 2A_1 M_\Theta C_{\hat{H}} \mathbb{E}\left[V^{1/4}(\tilde{X}_0)\right] \left\{\delta_{n+1}/\gamma_{n+1} + \delta_1/\gamma_1\right\} \ ,\end{aligned}$$

which concludes the proof of Lemma S7-(a).

(b) Assume now (17). We show that a.s the first term in (S22) is absolutely convergence and the second term converges to 0.

Using (S23), (S19), the Cauchy-Schwarz inequality and (17) we get that

$$\begin{aligned}&\mathbb{E}\left[\sum_{k=1}^{+\infty} \left|\langle \delta_{k+1} a_{k+1} - \delta_k a_k, \mathrm{K}_{\gamma_k, \tilde{\theta}_k} \hat{H}_{\gamma_k, \tilde{\theta}_k}(\tilde{X}_k) \rangle\right|\right] \\ &\leq (2M_\Theta + \sup_{\theta \in \Theta} \|\nabla f(\theta)\|) A_1 C_{\hat{H}} \mathbb{E}\left[V^{1/4}(\tilde{X}_0)\right] \sum_{k=0}^{+\infty} \left|\delta_{k+1} - \delta_k\right| \gamma_k^{-1} \\ &\quad+ A_1 C_{\hat{H}} (1 + \delta_1 L_f) \mathbb{E}\left[V^{1/4}(\tilde{X}_0)\right] \sum_{k=0}^{+\infty} \delta_{k+1}^2 \\ &< +\infty \ ,\end{aligned}$$

which implies that $(\langle \delta_{k+1} a_{k+1} - \delta_k a_k, \mathrm{K}_{\gamma_k, \tilde{\theta}_k} \hat{H}_{\gamma_k, \tilde{\theta}_k}(\tilde{X}_k) \rangle)_{k \in \mathbb{N}}$ is absolutely convergent a.s. We have that $\mathrm{K}_{\gamma_{n+1}, \tilde{\theta}_{n+1}} \|\hat{H}_{\gamma_{n+1}, \tilde{\theta}_{n+1}}(\tilde{X}_{n+1})\|$ is upper-bounded using (S19) by $\gamma_{n+1}^{-1} C_{\hat{H}} \mathrm{K}_{\gamma_{n+1}, \tilde{\theta}_{n+1}} V^{1/4}(\tilde{X}_{n+1})$. It follows that we have for any $\theta \in \Theta$, $\varepsilon > 0$, using the Markov inequality, the Cauchy-Schwarz inequality, (S19) and (17)

$$
\begin{aligned}
&\sum_{n\in\mathbb{N}} \mathbb{P}\left(\|a_{n+1}\|\,\delta_{n+1}\mathrm{K}_{\gamma_{n+1},\tilde{\theta}_{n+1}}\|\hat{H}_{\gamma_{n+1},\tilde{\theta}_{n+1}}(\tilde{X}_{n+1})\| \geq \varepsilon\right) \\
&\qquad \leq \sum_{n\in\mathbb{N}} \mathbb{P}\left(2C_{\hat{H}}M_\Theta\,\delta_{n+1}\,\gamma_{n+1}^{-1}\,\mathrm{K}_{\gamma_{n+1},\tilde{\theta}_{n+1}}V^{1/4}(\tilde{X}_{n+1}) \geq \varepsilon\right) \\
&\qquad \leq 4\varepsilon^{-2}M_\Theta^2 C_{\hat{H}}^2 A_1 \mathbb{E}\left[V^{1/2}(\tilde{X}_0)\right]\sum_{n\in\mathbb{N}}\delta_n^2\gamma_n^{-2} < +\infty\ ,
\end{aligned}
$$

Using the Borel-Cantelli lemma, we get $\lim_{n\to+\infty}\langle\delta_n a_n \mathrm{K}_{\gamma_n,\tilde{\theta}_n}\hat{H}_{\gamma_n,\tilde{\theta}_n}(\tilde{X}_n)\rangle = 0$ a.s. This completes the proof of convergence for any $\theta\in\Theta$ of the series $\sum_{k\in\mathbb{N}}\delta_{k+1}\langle a_{k+1},\tilde{\eta}_k^{(b)}\rangle$.

□

Lemma S8. *Assume* **A**1*,* **A**2*,* **H**1*,* **H**2 *and that for any $\theta\in\Theta$ and $x\in\mathbb{R}^d$, $\|H_\theta(x)\|\leq V^{1/4}(x)$. Then we have for any $n\in\mathbb{N}$*

$$
\mathbb{E}\left[\left\|\tilde{\eta}_n^{(c)}\right\|\right] \leq C_2\gamma_{n+1}^{-1}\left[\gamma_{n+1}^{-1}\left\{\boldsymbol{\Lambda}_1(\gamma_n,\gamma_{n+1})+\boldsymbol{\Lambda}_2(\gamma_n,\gamma_{n+1})\delta_{n+1}\right\}+\delta_{n+1}\right]\ ,
$$

with

$$
C_2 = 4A_1A_2\log^{-1}(1/\rho)\rho^{-\bar{\gamma}/2}\ \max\left[L_H, C_{c,1}+2A_2\log^{-1}(1/\rho)\rho^{-\bar{\gamma}/2}\right]\ , \tag{S24}
$$

where $C_{c,1}$ is given by

$$
C_{c,1} = 4A_1A_2\log^{-1}(1/\rho)\rho^{-\bar{\gamma}/2}\mathbb{E}\left[V(\tilde{X}_0)\right]\ . \tag{S25}
$$

Proof. We start by giving an upper-bound on $\|\pi_{\gamma_1,\theta_1}-\pi_{\gamma_2,\theta_2}\|_{V^{1/2}}$ for $\gamma_1,\gamma_2\in(0,\bar{\gamma}]$ with $\gamma_1>\gamma_2$ and, $\theta_1,\theta_2\in\Theta$. Let $g:\mathbb{R}^d\to\mathbb{R}^{d_\theta}$ be a measurable function satisfying $\sup_{x\in\mathbb{R}^d}\{\|g(x)\|/V^{1/2}(x)\}\leq 1$. Using **H**1-(i)-(ii), **H**2, Lemma S1 and Lemma S2, we get that for any $\gamma_1,\gamma_2\in(0,\bar{\gamma}]$ with $\gamma_1>\gamma_2$, $\theta_1,\theta_2\in\Theta$ and $\ell\in\mathbb{N}^*$

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\mathrm{K}_{\gamma_1,\theta_1}^\ell g(\tilde{X}_0)-\mathrm{K}_{\gamma_2,\theta_2}^\ell g(\tilde{X}_0)\right\|\right] \\
&= \left\|\sum_{j=0}^{\ell-1}\mathrm{K}_{\gamma_1,\theta_1}^j(\mathrm{K}_{\gamma_1,\theta_1}-\mathrm{K}_{\gamma_2,\theta_2})\left\{\mathrm{K}_{\gamma_2,\theta_2}^{(\ell-1-j)}g(x)-\pi_{\gamma_2,\theta_2}(f)\right\}\right\| \\
&\leq 2A_2\sum_{j=0}^{\ell-1}\rho^{(\ell-1-j)\gamma_2/2}\left\|\mathrm{K}_{\gamma_1,\theta_1}^j(\mathrm{K}_{\gamma_1,\theta_1}-\mathrm{K}_{\gamma_2,\theta_2})V^{1/2}(x)\right\| \\
&\leq 2A_2\sum_{j=0}^{\ell-1}\rho^{(\ell-1-j)\gamma_2/2}\left[\boldsymbol{\Lambda}_1(\gamma_1,\gamma_2)+\boldsymbol{\Lambda}_2(\gamma_1,\gamma_2)\|\theta_1-\theta_2\|\right]\sup_{k\in\mathbb{N}}\mathbb{E}\left[\mathrm{K}_{\gamma_1,\theta_1}^k V(\tilde{X}_0)\right] \\
&\leq 4A_1A_2\log^{-1}(1/\rho)\rho^{-\bar{\gamma}/2}\gamma_2^{-1}\left[\boldsymbol{\Lambda}_1(\gamma_1,\gamma_2)+\boldsymbol{\Lambda}_2(\gamma_1,\gamma_2)\|\theta_1-\theta_2\|\right]\mathbb{E}\left[V(\tilde{X}_0)\right]\ .
\end{aligned}
$$

Taking $\ell \to +\infty$ and using **H** 1-(ii), we obtain that for any $\theta_1, \theta_2 \in \Theta$ and $\gamma_1, \gamma_2 \in (0, \bar{\gamma}]$ with $\gamma_1 > \gamma_2$,

$$\|\pi_{\gamma_1,\theta_1} - \pi_{\gamma_2,\theta_2}\|_{V^{1/2}} \leq C_{c,1}\gamma_2^{-1}\left[\mathbf{\Lambda}_1(\gamma_1,\gamma_2) + \mathbf{\Lambda}_2(\gamma_1,\gamma_2)\|\theta_1 - \theta_2\|\right] , \tag{S26}$$

with $C_{c,1}$ given by(S25).

In what follows we derive an upper bound $\left\|\mathrm{K}_{\gamma_1,\theta_1}\hat{H}_{\gamma_1,\theta_1}(x) - \mathrm{K}_{\gamma_2,\theta_2}\hat{H}_{\gamma_2,\theta_2}(x)\right\|$ for any $\theta_1, \theta_2 \in \Theta$, $\gamma_1, \gamma_2 \in (0, \bar{\gamma}]$ with $\gamma_1 > \gamma_2$ and $x \in \mathbb{R}^d$. By (S18) we have for any $\theta_1, \theta_2 \in \Theta$, $\gamma_1, \gamma_2 \in (0, \bar{\gamma}]$ with $\gamma_1 > \gamma_2$ and $x \in \mathbb{R}^d$,

$$\begin{aligned}
&\left\|\mathrm{K}_{\gamma_1,\theta_1}\hat{H}_{\gamma_1,\theta_1}(x) - \mathrm{K}_{\gamma_2,\theta_2}\hat{H}_{\gamma_2,\theta_2}(x)\right\| \\
&= \left\|\sum_{\ell\in\mathbb{N}^*}\left\{\mathrm{K}^{\ell}_{\gamma_1,\theta_1}H_{\theta_1}(x) - \pi_{\gamma_1,\theta_1}(H_{\theta_1})\right\} - \sum_{\ell\in\mathbb{N}^*}\left\{\mathrm{K}^{\ell}_{\gamma_2,\theta_2}H_{\theta_2}(x) - \pi_{\gamma_2,\theta_2}(H_{\theta_2})\right\}\right\| \\
&\leq \sum_{\ell\in\mathbb{N}^*}\left\|\left\{\mathrm{K}^{\ell}_{\gamma_1,\theta_1}H_{\theta_1}(x) - \pi_{\gamma_1,\theta_1}(H_{\theta_1})\right\} - \left\{\mathrm{K}^{\ell}_{\gamma_2,\theta_2}H_{\theta_2}(x) - \pi_{\gamma_2,\theta_2}(H_{\theta_2})\right\}\right\| .
\end{aligned}$$

We now bound each term of the series in the right hand side. For any measurable functions g_1, g_2 with $g_i : \mathbb{R}^d \to \mathbb{R}^{d_\theta}$ and such that $\sup_{x\in\mathbb{R}^d}\|g_i(x)\|/V^{1/4}(x) < +\infty$ with $i \in \{1,2\}$, $\theta_1, \theta_2 \in \Theta$, $\gamma_1, \gamma_2 \in (0, \bar{\gamma}]$ with $\gamma_1 > \gamma_2$, $x \in \mathbb{R}^d$ and $\ell \in \mathbb{N}^*$, it holds that

$$\begin{aligned}
&\mathrm{K}^{\ell}_{\gamma_1,\theta_1}g_1(x) - \mathrm{K}^{\ell}_{\gamma_2,\theta_2}g_2(x) = \mathrm{K}^{\ell}_{\gamma_1,\theta_1}g_1(x) - \mathrm{K}^{\ell}_{\gamma_2,\theta_2}g_1(x) + \mathrm{K}^{\ell}_{\gamma_2,\theta_2}(g_1(x) - g_2(x)) \\
&= \sum_{j=0}^{\ell-1}\left\{\mathrm{K}^{j}_{\gamma_1,\theta_1} - \pi_{\gamma_1,\theta_1}\right\}(\mathrm{K}_{\gamma_1,\theta_1} - \mathrm{K}_{\gamma_2,\theta_2})\left\{\mathrm{K}^{\ell-1-j}_{\gamma_2,\theta_2}g_1(x) - \pi_{\gamma_2,\theta_2}(g_1)\right\} \\
&\quad + \sum_{j=0}^{\ell-1}\pi_{\gamma_1,\theta_1}\left\{\mathrm{K}^{\ell-1-j}_{\gamma_2,\theta_2}g_1(x) - \mathrm{K}^{\ell-j}_{\gamma_2,\theta_2}g_1(x)\right\} + \mathrm{K}^{\ell}_{\gamma_2,\theta_2}(g_1(x) - g_2(x)) \\
&= \sum_{j=0}^{\ell-1}\left\{\mathrm{K}^{j}_{\gamma_1,\theta_1} - \pi_{\gamma_1,\theta_1}\right\}(\mathrm{K}_{\gamma_1,\theta_1} - \mathrm{K}_{\gamma_2,\theta_2})\left\{\mathrm{K}^{\ell-1-j}_{\gamma_2,\theta_2}g_1(x) - \pi_{\gamma_2,\theta_2}(g_1)\right\} \\
&\quad - \pi_{\gamma_1,\theta_1}(\mathrm{K}^{\ell}_{\gamma_2,\theta_2}g_1(x) - g_1(x)) + \mathrm{K}^{\ell}_{\gamma_2,\theta_2}(g_1(x) - g_2(x)) .
\end{aligned}$$

Setting $g_1 = H_{\theta_1} - \pi_{\gamma_1,\theta_1}(H_{\theta_1})$ and $g_2 = H_{\theta_2} - \pi_{\gamma_2,\theta_2}(H_{\theta_2})$, we obtain that

$$\begin{aligned}
&\mathrm{K}^{\ell}_{\gamma_1,\theta_1}g_1(x) - \mathrm{K}^{\ell}_{\gamma_2,\theta_2}g_2(x) \\
&= \sum_{j=0}^{\ell-1}\left\{\mathrm{K}^{j}_{\gamma_1,\theta_1} - \pi_{\gamma_1,\theta_1}\right\}(\mathrm{K}_{\gamma_1,\theta_1} - \mathrm{K}_{\gamma_2,\theta_2})\left\{\mathrm{K}^{\ell-1-j}_{\gamma_2,\theta_2}H_{\theta_1}(x) - \pi_{\gamma_2,\theta_2}(H_{\theta_1})\right\} \\
&\qquad + \Xi_\ell ,
\end{aligned} \tag{S27}$$

where

$$
\begin{aligned}
\Xi_\ell &= -\pi_{\gamma_1,\theta_1}(\mathrm{K}^\ell_{\gamma_2,\theta_2} H_{\theta_1}(x) - H_{\theta_1}(x)) \\
&\quad + \mathrm{K}^\ell_{\gamma_2,\theta_2}\left[H_{\theta_1}(x) - H_{\theta_2}(x) + \pi_{\gamma_2,\theta_2}(H_{\theta_2}) - \pi_{\gamma_1,\theta_1}(H_{\theta_1})\right] \\
&= -\pi_{\gamma_1,\theta_1}\mathrm{K}^\ell_{\gamma_2,\theta_2} H_{\theta_1}(x) + \mathrm{K}^\ell_{\gamma_2,\theta_2}\left[H_{\theta_1}(x) - H_{\theta_2}(x) + \pi_{\gamma_2,\theta_2}(H_{\theta_2})\right] \\
&= (\pi_{\gamma_2,\theta_2} - \pi_{\gamma_1,\theta_1})(\mathrm{K}^\ell_{\gamma_2,\theta_2} H_{\theta_1}(x) - \pi_{\gamma_2,\theta_2}(H_{\theta_1})) - \pi_{\gamma_2,\theta_2}(H_{\theta_1}) \\
&\qquad\qquad + \mathrm{K}^\ell_{\gamma_2,\theta_2}\left[H_{\theta_1}(x) - H_{\theta_2}(x) + \pi_{\gamma_2,\theta_2}(H_{\theta_2})\right] \\
&= (\pi_{\gamma_2,\theta_2} - \pi_{\gamma_1,\theta_1})(\mathrm{K}^\ell_{\gamma_2,\theta_2} H_{\theta_1}(x) - \pi_{\gamma_2,\theta_2}(H_{\theta_1})) \\
&\qquad\qquad + \mathrm{K}^\ell_{\gamma_2,\theta_2}(H_{\theta_1} - H_{\theta_2})(x) - \pi_{\gamma_2,\theta_2}(H_{\theta_1} - H_{\theta_2}) \; .
\end{aligned} \tag{S28}
$$

For the first term in (S27), using **H** 1-(ii), **H** 2, Lemma S2 and and that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/4}(x)$ we obtain for any $\theta_1, \theta_2 \in \Theta$, $\gamma_1, \gamma_2 \in (0, \bar{\gamma}]$ with $\gamma_1 > \gamma_2$, $x \in \mathbb{R}^d$ and $\ell \in \mathbb{N}^*$

$$
\begin{aligned}
&\left\| \sum_{j=0}^{\ell-1} \left\{ \mathrm{K}^j_{\gamma_1,\theta_1} - \pi_{\gamma_1,\theta_1} \right\} (\mathrm{K}_{\gamma_1,\theta_1} - \mathrm{K}_{\gamma_2,\theta_2}) \left\{ \mathrm{K}^{\ell-1-j}_{\gamma_2,\theta_2} H_{\theta_1}(x) - \pi_{\gamma_2,\theta_2}(H_{\theta_1}) \right\} \right\| \\
&\qquad \leq 2A_2 \sum_{j=0}^{\ell-1} \rho^{(\ell-1-j)\gamma_1/2} \left\| \left\{ \mathrm{K}^j_{\gamma_1,\theta_1} - \pi_{\gamma_1,\theta_1} \right\} (\mathrm{K}_{\gamma_1,\theta_1} - \mathrm{K}_{\gamma_2,\theta_2}) V^{1/2}(x) \right\| \\
&\qquad \leq 4A_2^2 \left[\mathbf{\Lambda}_1(\gamma_1, \gamma_2) + \mathbf{\Lambda}_2(\gamma_1, \gamma_2)\|\theta_1 - \theta_2\| \right] \sum_{j=0}^{\ell-1} \rho^{(j + (\ell-1-j))\gamma_2/2} V^{1/2}(x) \\
&\qquad \leq 4A_2^2 \left[\mathbf{\Lambda}_1(\gamma_1, \gamma_2) + \mathbf{\Lambda}_2(\gamma_1, \gamma_2)\|\theta_1 - \theta_2\| \right] \ell \rho^{(\ell-1)\gamma_2/2} V^{1/2}(x) \; .
\end{aligned} \tag{S29}
$$

For the first term in (S28), using **H** 1-(ii), Lemma S2, (S26) and that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/4}(x) \leq V^{1/2}(x)$, we obtain for any $\theta_1, \theta_2 \in \Theta$, $\gamma_1, \gamma_2 \in (0, \bar{\gamma}]$ with $\gamma_1 > \gamma_2$, $x \in \mathbb{R}^d$ and $\ell \in \mathbb{N}^*$

$$
\begin{aligned}
&\left\| (\pi_{\gamma_1,\theta_1} - \pi_{\gamma_2,\theta_2})(\mathrm{K}^\ell_{\gamma_2,\theta_2} H_{\theta_1}(x) - \pi_{\gamma_2,\theta_2}(H_{\theta_1})) \right\| \\
&\qquad \leq 2A_2 \rho^{\ell\gamma_2/2} \|\pi_{\gamma_1,\theta_1} - \pi_{\gamma_2,\theta_2}\|_{V^{1/2}} \\
&\qquad \leq 2A_2 C_{c,1}\, \rho^{\ell\gamma_2/2} \gamma_2^{-1} \left\{ \mathbf{\Lambda}_1(\gamma_1, \gamma_2) + \mathbf{\Lambda}_2(\gamma_1, \gamma_2)\|\theta_1 - \theta_2\| \right\} \; .
\end{aligned} \tag{S30}
$$

For the second term in (S28), using **A** 2, **H** 1-(ii) and Lemma S2, we obtain for any $\theta_1, \theta_2 \in \Theta$, $\gamma_1, \gamma_2 \in (0, \bar{\gamma}]$ with $\gamma_1 > \gamma_2$, $x \in \mathbb{R}^d$ and $\ell \in \mathbb{N}^*$

$$
\left\| \mathrm{K}^\ell_{\gamma_2,\theta_2}(H_{\theta_1} - H_{\theta_2})(x) - \pi_{\gamma_2,\theta_2}(H_{\theta_1} - H_{\theta_2}) \right\| \leq 2A_2 L_H \rho^{\ell\gamma_2/2} \|\theta_1 - \theta_2\| V^{1/2}(x) \; . \tag{S31}
$$

Combining (S28), (S29), (S30), (S31) in (S27) and using Lemma S1, we obtain that for any $\theta_1, \theta_2 \in \Theta$, $\gamma_1, \gamma_2 \in (0, \bar{\gamma}]$ with $\gamma_1 > \gamma_2$, $x \in \mathbb{R}^d$ that

$$
\begin{aligned}
&\left\| \mathrm{K}_{\gamma_1,\theta_1} \hat{H}_{\gamma_1,\theta_1}(x) - \mathrm{K}_{\gamma_2,\theta_2} \hat{H}_{\gamma_2,\theta_2}(x) \right\| \\
&\qquad \leq C_{c,2}\, \gamma_2^{-1} \left[\gamma_2^{-1} \left\{ \mathbf{\Lambda}_1(\gamma_1, \gamma_2) + \mathbf{\Lambda}_2(\gamma_1, \gamma_2)\|\theta_1 - \theta_2\| \right\} + \|\theta_1 - \theta_2\| \right] V^{1/2}(x) \; ,
\end{aligned}
$$

with

$$C_{c,2} = 4A_2 \log^{-1}(1/\rho)\rho^{-\bar{\gamma}/2} \max\left[L_H, C_{c,1} + 2A_2 \log^{-1}(1/\rho)\rho^{-\bar{\gamma}/2}\right] .$$

Since for any $k \in \mathbb{N}$, $\left\|\tilde{\theta}_{k+1} - \tilde{\theta}_k\right\| \leq \delta_{k+1} V^{1/2}(\tilde{X}_{k+1})$ by (15) and the fact that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/2}(x)$ and that Π_Θ is non-expansive, we get that for any $k \in \mathbb{N}$,

$$\begin{aligned}&\left\|\mathrm{K}_{\gamma_k,\tilde{\theta}_k} \hat{H}_{\gamma_k,\tilde{\theta}_k}(\tilde{X}_{k+1}) - \mathrm{K}_{\gamma_{k+1},\tilde{\theta}_{k+1}} \hat{H}_{\gamma_{k+1},\tilde{\theta}_{k+1}}(\tilde{X}_{k+1})\right\| \\ &\qquad \leq C_{c,2}\gamma_{k+1}^{-1}\left[\gamma_{k+1}^{-1}\left\{\boldsymbol{\Lambda}_1(\gamma_k,\gamma_{k+1}) + \boldsymbol{\Lambda}_2(\gamma_k,\gamma_{k+1})\delta_{k+1}\right\} + \delta_{k+1}\right] V(\tilde{X}_{k+1}) ,\end{aligned}$$

which implies by (S21) and using **H**1-(i) that

$$\mathbb{E}\left[\left\|\tilde{\eta}^{(c)}\right\|\right] \leq C_2\gamma_{k+1}^{-1}\left[\gamma_{k+1}^{-1}\left\{\boldsymbol{\Lambda}_1(\gamma_k,\gamma_{k+1}) + \boldsymbol{\Lambda}_2(\gamma_k,\gamma_{k+1})\delta_{k+1}\right\} + \delta_{k+1}\right] ,$$

with C_2 given by (S24). □

Lemma S9. *Assume* **A**1, **H**1 *and that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/4}(x)$. Then we have for any $n \in \mathbb{N}$*

$$\mathbb{E}\left[\left\|\tilde{\eta}_n^{(d)}\right\|\right] \leq \boldsymbol{\Psi}(\gamma_n) .$$

Proof. By a straightforward application of **H**1-(iii) and by (S21) along with the fact that for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\|H_\theta(x)\| \leq V^{1/4}(x)$ we have for any $n \in \mathbb{N}$, $\mathbb{E}\left[\left\|\tilde{\eta}_n^d\right\|\right] \leq \boldsymbol{\Psi}(\gamma_n)$. □

We now turn to the proof of Theorem 5.

Proof. (a) To apply [1, Theorem 2], it is enough to show that the following series converge a.s

$$\sum_{n=0}^{+\infty} \delta_{n+1}\langle \Pi_\Theta(\theta_n - \delta_{n+1}\nabla f(\theta_n)) - \theta^\star, \tilde{\eta}_n^{(i)}\rangle , \quad \sum_{n=0}^{+\infty} \delta_{n+1}^2 \|\tilde{\eta}_n\|^2 ,$$

with $\theta^\star \in \arg\min_{\theta\in\Theta} f(\theta)$. $\sum_{n=0}^{+\infty} \delta_{n+1}^2\|\tilde{\eta}_n\|^2 < +\infty$ a.s by Lemma S5 since $\sum_{n\in\mathbb{N}} \delta_{n+1}^2 < +\infty$. Since $(\langle \Pi_\Theta(\theta_n - \delta_{n+1}\nabla f(\theta_n)) - \theta^\star, \tilde{\eta}_n^{(a)}\rangle)_{n\in\mathbb{N}}$ is a $(\tilde{\mathcal{F}}_n)_{n\in\mathbb{N}}$-martingale increment, see (16) and by Lemma S6 and $\sum_{n\in\mathbb{N}} \delta_{n+1}^2/\gamma_n^2 < +\infty$

$$\mathbb{E}\left[\sum_{n=0}^{+\infty} \delta_{n+1}^2 \langle \Pi_\Theta(\theta_n - \delta_{n+1}\nabla f(\theta_n)) - \theta^\star, \tilde{\eta}_n^{(a)}\rangle^2\right] < +\infty ,$$

we obtain using [19, Section 12.5] that $\sum_{n=0}^{+\infty} \delta_{n+1}\langle \Pi_\Theta(\theta_n - \delta_{n+1}\nabla f(\theta_n)) - \theta^\star, \tilde{\eta}_n^{(a)}\rangle$ converges a.s. Using **A**1, (17) and Lemma S8 and Lemma S9 we get that $\sum_{n=0}^{+\infty} \delta_{n+1}\langle \Pi_\Theta(\theta_n - \delta_{n+1}\nabla f(\theta_n)) - \theta^\star, \tilde{\eta}_n^{(i)}\rangle$ is absolutely convergent a.s for $i \in \{c, d\}$. Finally we obtain that $\sum_{n=0}^{+\infty} \delta_{n+1}\langle \Pi_\Theta(\theta_n - \delta_{n+1}\nabla f(\theta_n)) - \theta^\star, \tilde{\eta}_n^{(b)}\rangle$ converges a.s by Lemma S7-(b).

(b) The proof of is identical to the one of Theorem 3-(b).

□

S1.4 Proof of Theorem 6

The proof is similar to the one of Theorem 4, using Lemma S5, Lemma S7, Lemma S8, Lemma S9 and the fact that $\tilde{\eta}_n^{(a)}$ is a $(\tilde{\mathcal{F}}_n)_{n\in\mathbb{N}}$-martingale increment, see (16).

S1.5 Proof of Theorem 7

In this section, we give the proof of Theorem 7 by showing that **H**1 holds. First of all in Section S1.5.1, we establish under **L**1 and **L**2 stability results uniform in the parameter $\theta \in \Theta$ for the Langevin diffusion (5) and the associated Euler-Maruyama discretization (6) based on a Foster-Lyapunov drift condition with constants independent of θ. Then, in Section S1.5.2, we show that the stability conditions that we derive, are sufficient to prove that **H** 1 holds. The proof of Theorem 7 then consists in combining all these results and is presented in Section S1.5.3.

Under **L**1 and **L**2, for any $\theta \in \Theta$, (5) defines a Markov semi-group $(\mathrm{P}_{t,\theta})_{t\geq 0}$ for any $x \in \mathbb{R}^d$ and $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$ by $\mathrm{P}_{t,\theta}(x, \mathsf{A}) = \mathbb{P}(Y_t^\theta \in \mathsf{A})$ where $(Y_t^\theta)_{t\geq 0}$ is the solution of (5) with $Y_0^\theta = x$. Consider now the generator of $(\mathrm{P}_{t,\theta})_{t\geq 0}$ for any $\theta \in \Theta$, defined for any $f \in \mathrm{C}^2(\mathbb{R}^d)$ by

$$\mathcal{A}_\theta f = -\langle \nabla_x f, \nabla_x U_\theta(x)\rangle + \Delta_x f \ . \tag{S32}$$

We say that a Markov semi-group $(\mathrm{P}_t)_{t\geq 0}$ on $\mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d)$ with extended infinitesimal generator $(\mathcal{A}, \mathrm{D}(\mathcal{A}))$ (see e.g. [15] for the definition of $(\mathcal{A}, \mathrm{D}(\mathcal{A}))$) satisfies a continuous drift condition $\mathbf{D}_{\mathrm{c}}(V, \zeta, \beta)$ if there exist $\zeta > 0$, $\beta \geq 0$ and a measurable function $V : \mathbb{R}^d \to [1, +\infty)$ with $V \in \mathrm{D}(\mathcal{A})$ such that for all $x \in \mathbb{R}^d$

$$\mathcal{A}V(x) \leq -\zeta V(x) + \beta \ .$$

S1.5.1 Foster-Lyapunov drift conditions uniform on θ

Define $V_{\mathrm{e}} : \mathbb{R}^d \to [1, +\infty)$ for all $x \in \mathbb{R}^d$ by

$$V_{\mathrm{e}}(x) = \exp(\tilde{\mathtt{m}}_1 \phi(x)) \ , \quad \text{with } \phi(x) = \sqrt{1 + \|x\|^2} \text{ and } \tilde{\mathtt{m}}_1 = \mathtt{m}_1/4 \ . \tag{S33}$$

Proposition S10. *Assume* **L** 1 *and* **L** 2*. Let* $\bar{\gamma} < \min(1, 2\mathtt{m}_2)$*. Then there exist* $\lambda_{\mathrm{e}} \in (0,1)$ *and* $b_{\mathrm{e}} \geq 0$ *such that for all* $\gamma \in (0, \bar{\gamma}]$ *and* $\theta \in \Theta$ *the Markov kernel* $\mathrm{R}_{\gamma,\theta}$ *associated with the recursion* (6) *satisfies the discrete drift condition* $\mathbf{D}_{\mathrm{d}}(V, \lambda^\gamma, b\gamma)$*, i.e. for all* $x \in \mathbb{R}^d$

$$\mathrm{R}_{\gamma,\theta} V_{\mathrm{e}}(x) \leq \lambda_{\mathrm{e}}^\gamma V_{\mathrm{e}}(x) + b_{\mathrm{e}} \gamma \mathbb{1}_{\mathrm{B}(0, r_{\mathrm{e}})}(x) \ , \tag{S34}$$

with

$$\lambda_{\mathrm{e}} = \mathrm{e}^{-\tilde{\mathtt{m}}_1^2 (2^{1/2} - 1)} \ , \quad r_{\mathrm{e}} = \max(1, 2(d + \mathtt{c})/\mathtt{m}_1, R_1) \ ,$$
$$b_{\mathrm{e}} = \tilde{\mathtt{m}}_1 (d + c + 2^{1/2} \tilde{\mathtt{m}}_1) \exp\left[\tilde{\mathtt{m}}_1 \left\{(d + c + \tilde{\mathtt{m}}_1)\bar{\gamma} + \sqrt{1 + r_{\mathrm{e}}^2}\right\}\right] \ .$$

Proof. Since ϕ is 1-Lipschitz, by the log-Sobolev inequality [3, Proposition 5.4.1], we have for any $x \in \mathbb{R}^d$ and $\theta \in \Theta$,

$$\begin{aligned} \mathrm{R}_{\gamma,\theta} V_{\mathrm{e}}(x) &\leq \exp\left[\tilde{\mathtt{m}}_1 \mathrm{R}_{\gamma,\theta}\phi(x) + \tilde{\mathtt{m}}_1^2\gamma\right] \\ &\leq \exp\left[\tilde{\mathtt{m}}_1 \sqrt{\|x - \gamma\nabla_x U_\theta(x)\|^2 + 2\gamma d + 1} + \tilde{\mathtt{m}}_1^2\gamma\right] \;, \end{aligned} \tag{S35}$$

where we have used Jensen's inequality in the last line. Second, using **L**2 and $\gamma < 2\mathtt{m}_2$, we obtain that for any $x \in \mathbb{R}^d$ and $\theta \in \Theta$,

$$\begin{aligned} \|x - \gamma\nabla_x U_\theta(x)\|^2 &\leq \|x\|^2 - 2\gamma\langle x, \nabla_x U_\theta(x)\rangle + \gamma^2\|\nabla_x U_\theta(x)\|^2 \\ &\leq \|x\|^2 - 2\mathtt{m}_1\gamma\|x\|\mathbb{1}_{\mathrm{B}(0,R_1)^{\mathrm{c}}}(x) + \gamma(\gamma - 2\mathtt{m}_2)\|\nabla_x U_\theta(x)\|^2 + 2\gamma\mathtt{c} \\ &\leq \|x\|^2 - 2\mathtt{m}_1\gamma\|x\|\mathbb{1}_{\mathrm{B}(0,R_1)^{\mathrm{c}}}(x) + 2\gamma\mathtt{c} \;. \end{aligned}$$

Therefore, using for any $a > 0$, $\sqrt{1+a} - 1 \leq a/2$, we get for any $x \in \mathbb{R}^d$ and $\theta \in \Theta$,

$$\begin{aligned} &\sqrt{\|x - \gamma\nabla_x U_\theta(x)\|^2 + 2\gamma d + 1} - \phi(x) \\ &\leq \phi(x)\left\{\sqrt{1 + 2\gamma\phi^{-2}(x)(d + \mathtt{c} - \mathtt{m}_1\|x\|\mathbb{1}_{\mathrm{B}(0,R_1)^{\mathrm{c}}}(x))} - 1\right\} \\ &\leq \gamma\phi^{-1}(x)(d + \mathtt{c} - \mathtt{m}_1\|x\|\mathbb{1}_{\mathrm{B}(0,R_1)^{\mathrm{c}}}(x)) \;. \end{aligned} \tag{S36}$$

Therefore, combining this result with (S35) and using that for any $\tilde{x} \in \overline{\mathrm{B}}(0, r_{\mathrm{e}})^{\mathrm{c}}$, $\phi(\tilde{x})^2/\left\|\tilde{x}\right\|^2 \leq 2$ and $d+c \leq \mathtt{m}_1 \left\|x\right\|/2$, we obtain for any $x \in \overline{\mathrm{B}}(0, r_{\mathrm{e}})^{\mathrm{c}}$ and $\theta \in \Theta$,

$$\begin{aligned} \mathrm{R}_{\gamma,\theta} V_{\mathrm{e}}(x) &\leq \exp\left[\tilde{\mathtt{m}}_1\phi^{-1}(x)(d + \mathtt{c} - \mathtt{m}_1\|x\|) + \tilde{\mathtt{m}}_1^2\gamma\right] V_{\mathrm{e}}(x) \\ &\leq \exp\left[-2\tilde{\mathtt{m}}_1^2\gamma\phi^{-1}(x)\|x\| + \tilde{\mathtt{m}}_1^2\gamma\right] V_{\mathrm{e}}(x) \leq \lambda_{\mathrm{e}}^\gamma V_{\mathrm{e}}(x) \;. \end{aligned}$$

Using (S35), (S36), and the fact that $\phi(\tilde{x}) \geq 1$ for any $\tilde{x} \in \mathbb{R}^d$, we have for any $x \in \mathrm{B}(0, r_{\mathrm{e}})$ and $\theta \in \Theta$,

$$\mathrm{R}_{\gamma,\theta} V_{\mathrm{e}}(x) \leq \lambda_{\mathrm{e}}^\gamma V_{\mathrm{e}}(x) + \left(\mathrm{e}^{\tilde{\mathtt{m}}_1(d+c+\tilde{\mathtt{m}}_1)\gamma} - \lambda_{\mathrm{e}}^\gamma\right)\exp\left[\tilde{\mathtt{m}}_1\sqrt{1 + r_{\mathrm{e}}^2}\right] \;.$$

The proof of (S34) for $x \in \mathrm{B}(0, r_{\mathrm{e}})$ and $\theta \in \Theta$ is then completed upon using that $\mathrm{e}^a - \mathrm{e}^b \leq (a-b)\mathrm{e}^a$ for all $a, b \in \mathbb{R}$ with $a \geq b$. $\square$

Proposition S11. *Assume* **L**1 *and* **L**2*. Then for any $\theta \in \Theta$, $(\mathrm{P}_{t,\theta})_{t\geq 0}$ associated with* (5) *satisfies the continuous drift condition $\mathbf{D}_{\mathrm{c}}(V_{\mathrm{e}}, \zeta_{\mathrm{e}}, \beta_{\mathrm{e}})$ for V_{e} defined in* (S33) *and*

$$\zeta_{\mathrm{e}} = 3\tilde{\mathtt{m}}_1^2/2^{1/2} \;, \quad \beta_{\mathrm{e}} = \tilde{\mathtt{m}}_1 \exp\left[\tilde{\mathtt{m}}_1\sqrt{1 + \tilde{r}_{\mathrm{e}}^2}\right](1 + \tilde{\mathtt{m}}_1 + \mathtt{c} + d) \;, \quad \tilde{r}_{\mathrm{e}} = \max(1, R_1) \;.$$

Proof. First, by definition, for any $x \in \mathbb{R}^d$, we have

$$\begin{aligned} \nabla_x V(x) &= \tilde{\mathtt{m}}_1 x V(x)/\phi(x) \\ \Delta_x V(x) &= \{\tilde{\mathtt{m}}_1 V(x)/\phi(x)\}\{\tilde{\mathtt{m}}_1 \left\|x\right\|^2/\phi(x) + d - \left\|x\right\|^2/\phi^2(x)\} \;. \end{aligned}$$

Therefore, by (S32) and **L**2, we get for any $\theta \in \Theta$ and $x \in \mathbb{R}^d$,

$$
\begin{aligned}
\mathcal{A}_\theta V(x) &= \{\tilde{\mathtt{m}}_1 V(x)/\phi(x)\}\left[-\langle \nabla_x U_\theta(x), x\rangle + \tilde{\mathtt{m}}_1 \|x\|^2/\phi(x) + d - \|x\|^2/\phi^2(x)\right] \\
&\leq \{\tilde{\mathtt{m}}_1 V(x)/\phi(x)\}\left[-\mathtt{m}_1 \|x\| \mathbb{1}_{\mathrm{B}(0,R_1)^{\mathrm{c}}}(x) + \mathtt{c} + \tilde{\mathtt{m}}_1 \|x\|^2/\phi(x) + d - \|x\|^2/\phi^2(x)\right] \\
&\leq \{\tilde{\mathtt{m}}_1 V(x)/\phi(x)\}\left[-(3\mathtt{m}_1/4) \|x\| \mathbb{1}_{\mathrm{B}(0,R_1)^{\mathrm{c}}}(x) + \mathtt{c} + \tilde{\mathtt{m}}_1 \|x\| \mathbb{1}_{\mathrm{B}(0,R_1)}(x) + d\right] .
\end{aligned}
$$

The proof is then complete upon using that for any $x \in \mathrm{B}(0, \tilde{r}_{\mathrm{e}})^{\mathrm{c}}$, $\|x\|/\phi(x) \geq 2^{-1/2}$, for any $y \in \mathbb{R}^d$, $\|y\|/\phi(y) \leq 1$. $\square$

S1.5.2 Checking H1

Lemma S12. *Assume* **L**1 *and let* $V : \mathbb{R}^d \to [1, +\infty)$ *measurable and satisfying* $\lim_{\|x\|\to+\infty} V(x) = +\infty$ *and* $V \in \mathrm{D}(\mathcal{A}_\theta)$*, for any* $\theta \in \Theta$*, where* $\mathcal{A}_\theta$ *is defined by* (S32). .

(a) Assume that there exist $\lambda \in (0,1)$*,* $b \geq 0$ *and* $\bar{\gamma} > 0$ *such that for any* $\theta \in \Theta$ *and* $\gamma \in (0, \bar{\gamma}]$*,* $\mathrm{R}_{\gamma,\theta}$ *associated with the recursion* (18)*, satisfies* $\mathbf{D}_{\mathrm{d}}(V, \lambda^\gamma, b\gamma)$*. Then for any* $\theta \in \Theta$ *and* $\gamma \in (0, \bar{\gamma}]$*,* $\mathrm{R}_{\gamma,\theta}$ *admits an invariant probability measure* $\pi_{\gamma,\theta}$ *on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ *and there exists* $D_3 \geq 0$ *such that for any* $x \in \mathbb{R}^d$ *and* $k \in \mathbb{N}$

$$
\delta_x \mathrm{R}^k_{\gamma,\theta} V \leq D_3 + V(x) , \qquad \pi_{\gamma,\theta}(V) \leq D_3 , \qquad D_3 = b\lambda^{-\bar{\gamma}}/\log(1/\lambda) .
$$

In addition, for all $\theta \in \Theta$ *and* $x \in \mathbb{R}^d$*,* $\lim_{k\to+\infty} \|\delta_x \mathrm{R}^k_{\gamma,\theta} - \pi_{\gamma,\theta}\|_V = 0$.

(b) Assume that there exist $\zeta > 0$ *and* $\beta \geq 0$ *such that for any* $\theta \in \Theta$*,* $(\mathrm{P}_{t,\theta})_{t\geq 0}$ *associated with* (5) *satisfies* $\mathbf{D}_{\mathrm{c}}(V, \zeta, \beta)$*. Then for any* $\theta \in \Theta$*, the diffusion is non-explosive,* $\mathcal{A}_\theta$ *admits* π_θ *as an invariant probability measure and*

$$
\pi_\theta(V) \leq D_0 , \qquad D_0 = \beta/\zeta .
$$

In addition, for all $\theta \in \Theta$ *and* $x \in \mathbb{R}^d$*,* $\lim_{t\to+\infty} \|\delta_x \mathrm{P}_{t,\theta} - \pi_\theta\|_V = 0$.

Proof. (a) for any $\gamma \in (0, \bar{\gamma}]$ and $\theta \in \Theta$, $\mathrm{R}_{\gamma,\theta}$ is irreducible with respect to the Lebesgue measure on $\mathbb{R}^d$, has the Feller property and satisfies $\mathbf{D}_{\mathrm{d}}(V, \lambda^\gamma, b\gamma)$ then [13, Section 4.4] applies and $\mathrm{R}_{\gamma,\theta}$ admits an invariant probability measure $\pi_{\gamma,\theta}$. The discrete drift condition and [6, Lemma 1] give that for any $\gamma \in (0, \bar{\gamma}]$ and $\theta \in \Theta$

$$
\mathrm{R}^k_{\gamma,\theta} V(x) \leq V(x) + b\lambda^{-\bar{\gamma}}/\log(1/\lambda) , \qquad \pi_{\gamma,\theta}(V) \leq b\lambda^{-\bar{\gamma}}/\log(1/\lambda) .
$$

We obtain that for all $\theta \in \Theta$ and $x \in \mathbb{R}^d$, $\lim_{k\to+\infty} \|\delta_x \mathrm{R}^k_{\gamma,\theta} - \pi_{\gamma,\theta}\|_V = 0$ using [14, Theorem 16.0.1].

(b) Using $\mathbf{D}_{\mathrm{c}}(V, \zeta, \beta)$ and [15, Theorem 2.1] we get that the diffusion process is non-explosive and thus $(\mathrm{P}_{t,\theta})_{t\geq 0}$ is defined for any $\theta \in \Theta$ and $t \geq 0$. Using [18, Corollary 10.1.4] for any $\theta \in \Theta$, $(\mathrm{P}_{t,\theta})_{t\geq 0}$ is strongly Feller continuous, therefore any compact sets is petite for the Markov kernel $\mathrm{P}_{h,\theta}$, for any $h > 0$ and $\theta \in \Theta$, by [14, Theorem 6.0.1]. Using [17, Chapter 7, Proposition 1.5], [7, Chapter 4, Theorem 9.17], and the fact that $\pi_\theta(\mathcal{A}_\theta f) = 0$ for any $\theta \in \Theta$ and $f \in \mathrm{C}_c^2(\mathbb{R}^d)$, we obtain that for any $\theta \in \Theta$, π_θ is an invariant measure for $(\mathrm{P}_{t,\theta})_{t\geq 0}$. Using $\mathbf{D}_\mathrm{c}(V,\zeta,\beta)$ and [15, Theorem 4.5] we get that for all $\theta \in \Theta$, $\pi_\theta(V) \leq \beta/\zeta$. Finally, the convergence is ensured using [16, Theorem 5.1].

□

As an immediate corollary we obtain that under the conditions of Lemma S12 for any $\theta \in \Theta$, $\gamma \in (0,\bar{\gamma}]$ and $k \in \mathbb{N}$,

$$\pi_\theta \mathrm{R}_{\gamma,\theta}^k V \leq \beta/\zeta + b\lambda^{-\bar{\gamma}}/\log(1/\lambda) \,. \tag{S37}$$

Lemma S13. *Let $V : \mathbb{R}^d \to [1,+\infty)$. Assume there exist $\lambda \in (0,1)$, $b \geq 0$ and $\bar{\gamma} > 0$ such that for any $\theta \in \Theta$ and $\gamma \in (0,\bar{\gamma}]$, $\mathrm{R}_{\gamma,\theta}$ associated with the recursion (6) satisfies $\mathbf{D}_\mathrm{d}(V,\lambda^\gamma,b\gamma)$. Let $(\gamma_n)_{n\in\mathbb{N}}$, $(\delta_n)_{n\in\mathbb{N}}$ be sequences of non-increasing positive real numbers and $(m_n)_{n\in\mathbb{N}}$ be a sequence of positive integers satisfying $\sup_{n\in\mathbb{N}} \gamma_n < \bar{\gamma}$. Then, $(X_k^n)_{n\in\mathbb{N},k\in\{0,\ldots,m_n\}}$ given by (10) with $\{\mathrm{K}_{\gamma,\theta} : \gamma \in (0,\bar{\gamma}], \theta \in \Theta\} = \{\mathrm{R}_{\gamma,\theta} : \gamma \in (0,\bar{\gamma}], \theta \in \Theta\}$ satisfies for all $p, n \in \mathbb{N}$ and $k \in \{0,\ldots,m_n\}$*

$$\mathbb{E}\left[\mathrm{R}_{\gamma_n,\theta_n}^p V(X_k^n)\middle|X_0^0\right] \leq D_1 V(X_0^0)\,, \qquad D_1 = 1 + 2b\lambda^{-\bar{\gamma}}/\log(1/\lambda)\,.$$

Proof. By induction we obtain that

$$\begin{aligned}\mathbb{E}\left[V(X_k^{n+1})\middle|\mathcal{F}_n\right] &= \mathrm{R}_{\gamma_{n+1},\theta_{n+1}}^k V(X_0^{n+1}) \\ &\leq \lambda^{k\gamma_{n+1}} V(X_0^{n+1}) + b\gamma_{n+1}\sum_{i=1}^k \lambda^{\gamma_{n+1}(k-i)}\,,\end{aligned} \tag{S38}$$

where $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is defined by (7). Similarly, we obtain for any $k \in \{0,\ldots,m_0\}$,

$$\mathbb{E}\left[V(X_k^0)\middle|X_0^0\right] = \mathrm{R}_{\gamma_0,\theta_0}^k V(X_0^0) \leq \lambda^{k\gamma_0}V(X_0^0) + b\gamma_0\sum_{i=1}^k \lambda^{\gamma_0(k-i)}\,. \tag{S39}$$

Define for $k,\ell \in \mathbb{N}$ and $i \in \mathbb{N}^*$, $\mathrm{q}_{\ell,k} = \sum_{j=0}^{\ell-1} m_j + k$, $\mathrm{q}_n = \mathrm{q}_{\ell,0}$ and $\tilde{\gamma}_i = \sum_{j=0}^{+\infty} \gamma_j \mathbb{1}_{(\mathrm{q}_j,\mathrm{q}_{j+1}]}(i)$. In addition, consider for any $\mathrm{p},\mathrm{q} \in \mathbb{N}^*$, $\Gamma_{\mathrm{p},\mathrm{q}} = \sum_{i=\mathrm{p}}^{\mathrm{q}} \tilde{\gamma}_i$ and $\Gamma_\mathrm{p} = \Gamma_{1,\mathrm{p}}$. Combining (S38), (S39) and Lemma S1 we get for any $n \in \mathbb{N}$ and $k \in \{0,\ldots,m_n\}$

$$\begin{aligned}\mathbb{E}\left[\mathrm{R}_{\gamma_n,\theta_n}^p V(X_k^n)\middle|X_0^0\right] &\leq \lambda^{\gamma_n p}\mathbb{E}\left[V(X_k^n)\middle|X_0^0\right] + b\log(1/\lambda)\lambda^{-\bar{\gamma}} \\ &\leq \lambda^{\Gamma_{\mathrm{q}_{n,k}}} V(X_0^0) + b\sum_{i=1}^{\mathrm{q}_{n,k}} \tilde{\gamma}_i \lambda^{\Gamma_{i+1,\mathrm{q}_{n,k}}} + b\log(1/\lambda)\lambda^{-\bar{\gamma}}\,.\end{aligned} \tag{S40}$$

Since $(\tilde{\gamma}_i)_{i\in\mathbb{N}}$ is nonincreasing and for all $t \geq 0$, $1-\lambda^t \geq -t\lambda^t\log(\lambda)$, we have for all $\mathrm{q} \in \mathbb{N}^*$,

$$\begin{aligned}
&\sum_{i=1}^{\mathrm{q}} \tilde{\gamma}_i \lambda^{\Gamma_{i+1,\mathrm{q}}} \leq \sum_{i=1}^{\mathrm{q}} \tilde{\gamma}_i \prod_{j=i+1}^{\mathrm{q}} (1+\lambda^{\tilde{\gamma}_1}\log(\lambda)\tilde{\gamma}_j) \\
&\leq (-\lambda^{\tilde{\gamma}_1}\log(\lambda))^{-1} \sum_{i=1}^{\mathrm{q}} \left\{ \prod_{j=i+1}^{\mathrm{q}} (1+\lambda^{\tilde{\gamma}_1}\log(\lambda)\tilde{\gamma}_j) - \prod_{j=i}^{\mathrm{q}} (1+\lambda^{\tilde{\gamma}_1}\log(\lambda)\tilde{\gamma}_j) \right\} \\
&\leq (-\lambda^{\tilde{\gamma}_1}\log(\lambda))^{-1} .
\end{aligned}$$

Combining this result and (S40) completes the proof. □

Lemma S14. *Let $V : \mathbb{R}^d \to [1,+\infty)$ with $\sup_{x\in\mathbb{R}^d}\left\{(1+\|x\|)^2/V(x)\right\} \leq M_V$ where $M_V \geq 0$ and V measurable. Assume* **L**1 *and that for any $\theta \in \Theta$, $\gamma \in (0,\bar{\gamma}]$ and $k \in \mathbb{N}$,*

$$\pi_\theta \mathrm{R}_{\gamma,\theta}^k(V) \leq \tilde{D}_1 , \qquad \pi_\theta \mathrm{P}_{\gamma m_\gamma,\theta} V \leq \tilde{D}_1 , \tag{S41}$$

with $m_\gamma = \lceil 1/\gamma \rceil$. Then for any $\theta \in \Theta$ and $\gamma \in (0,\bar{\gamma}]$

$$\begin{aligned}
&\|\pi_\theta \mathrm{R}_{\gamma,\theta}^{m_\gamma} - \pi_\theta \mathrm{P}_{\gamma m_\gamma,\theta}\|_{V^{1/2}}^2 \\
&\qquad \leq 2\tilde{D}_1 \mathtt{L}^2 \gamma(1+\bar{\gamma}) \left\{ d + 2\bar{\gamma}(\sup_{\theta\in\Theta}\|\nabla_x U_\theta(0)\|^2 + \mathtt{L}^2 M_V \tilde{D}_1) \right\} ,
\end{aligned}$$

Proof. The proof follows the lines of [6, Theorem 10]. Let $\theta \in \Theta$ and $\gamma \in (0,\bar{\gamma}]$. We have, using a generalized Pinsker inequality [6, Lemma 24], that

$$\begin{aligned}
\|\pi_\theta \mathrm{R}_{\gamma,\theta}^{m_\gamma} - \pi_\theta \mathrm{P}_{\gamma m_\gamma,\theta}\|_{V^{1/2}}^2 &\leq 2(\pi_\theta \mathrm{R}_{\gamma,\theta}^{m_\gamma} V + \pi_\theta \mathrm{P}_{\gamma m_\gamma,\theta} V)\mathrm{KL}\left(\pi_\theta \mathrm{R}_{\gamma,\theta}^{m_\gamma} | \pi_\theta \mathrm{P}_{\gamma m_\gamma,\theta}\right) . \\
&\leq 4\tilde{D}_1 \mathrm{KL}\left(\pi_\theta \mathrm{R}_{\gamma,\theta}^{m_\gamma} | \pi_\theta \mathrm{P}_{\gamma m_\gamma,\theta}\right) .
\end{aligned}$$

Using **L**1, [6, Equation (15)], [9, Theorem 4.1, Chapter 2], (S41) and that for any $a, b \in \mathbb{R}$, $(a+b)^2 \leq 2(a^2+b^2)$ we obtain that

$$\begin{aligned}
\mathrm{KL}\left(\pi_\theta \mathrm{R}_{\gamma,\theta}^{m_\gamma} | \pi_\theta \mathrm{P}_{\gamma m_\gamma,\theta}\right) &\leq \mathtt{L}^2 m_\gamma \gamma^2 (d + \bar{\gamma} \sup_{k\in\mathbb{N}} \pi_\theta \mathrm{R}_{\gamma,\theta}^k \|\nabla_x U_\theta(x)\|^2) \\
&\leq \mathtt{L}^2(1+\bar{\gamma})\gamma(d + 2\bar{\gamma}(\sup_{\theta\in\Theta}\|\nabla_x U_\theta(0)\|^2 + \mathtt{L}^2 M_V \tilde{D}_1)) ,
\end{aligned}$$

which concludes the proof. □

Proposition S15. *Let $V : \mathbb{R}^d \to [1,+\infty)$ measurable and $M_V \geq 0$ satisfying $\sup_{x\in\mathbb{R}^d}\left\{(1+\|x\|)^2/V(x)\right\} \leq M_V$. Assume* **L**1 *and that there exist $\lambda \in (0,1)$, $b \geq 0$ and $\bar{\gamma} > 0$ such that for any $\theta \in \Theta$ and $\gamma \in (0,\bar{\gamma}]$ $\mathrm{R}_{\gamma,\theta}$ satisfies $\mathbf{D}_{\mathrm{d}}(V, \lambda^\gamma, b\gamma)$. Assume that there exists $D_0 \geq 0$. Then there exists $D_4 \geq 0$ such that for any $\theta \in \Theta$ and $\gamma \in (0,\bar{\gamma}]$*

$$\|\pi_{\gamma,\theta} - \pi_\theta\|_{V^{1/2}} \leq D_4 \gamma^{1/2} . \tag{S42}$$

Proof. Using Lemma S12 we obtain that for any $\theta \in \Theta$

$$\lim_{k \to +\infty} \|\pi_\theta \mathrm{R}^k_{\gamma,\theta} - \pi_\theta \mathrm{P}_{\gamma k,\theta}\|_{V^{1/2}} = \|\pi_{\gamma,\theta} - \pi_\theta\|_{V^{1/2}} \ . \tag{S43}$$

We now give an upper bound on $\|\pi_\theta \mathrm{R}^k_{\gamma,\theta} - \pi_\theta \mathrm{P}_{\gamma k,\theta}\|_{V^{1/2}}$ for $k = q_\gamma m_\gamma$ with $m_\gamma = \lceil 1/\gamma \rceil$ and $q_\gamma \in \mathbb{N}$. Using [4, Theorem 6] and that π_θ is invariant for $\mathrm{P}_{t,\theta}$ with $t \geq 0$, see Lemma S12, we obtain for all $\theta \in \Theta$, $\gamma \in (0, \bar{\gamma}]$ and $k \in \mathbb{N}$

$$\begin{aligned}
&\|\pi_\theta \mathrm{R}^k_{\gamma,\theta} - \pi_\theta \mathrm{P}_{\gamma k,\theta}\|_{V^{1/2}} \\
&\leq \sum_{\ell=0}^{q_\gamma - 1} \|\pi_\theta \mathrm{P}_{\gamma(\ell+1)m_\gamma,\theta} \mathrm{R}^{(q_\gamma-(\ell+1))m_\gamma}_{\gamma,\theta} - \pi_\theta \mathrm{P}_{\gamma \ell m_\gamma,\theta} \mathrm{R}^{(q_\gamma-\ell)m_\gamma}_{\gamma,\theta}\|_{V^{1/2}} \\
&\leq \sum_{\ell=0}^{q_\gamma - 1} C \xi^{\gamma m_\gamma (q_\gamma - (\ell+1))} \|\pi_\theta \mathrm{P}_{\gamma \ell m_\gamma,\theta} \mathrm{P}_{m_\gamma \gamma,\theta} - \pi_\theta \mathrm{P}_{\gamma \ell m_\gamma,\theta} \mathrm{R}^{m_\gamma}_{\gamma,\theta}\|_{V^{1/2}} \\
&\leq \|\pi_\theta \mathrm{P}_{m_\gamma \gamma,\theta} - \pi_\theta \mathrm{R}^{m_\gamma}_{\gamma,\theta}\|_{V^{1/2}} \sum_{\ell=1}^{q_\gamma} C \xi^{\ell \gamma m_\gamma} \ ,
\end{aligned} \tag{S44}$$

where $C \geq 0, \xi \in (0,1)$ are the constants given by [4, Theorem 6] with minorization condition given by [4, Proposition 8a] with $\mathtt{m} = -\mathtt{L}$ since **L**1 holds and drift condition $\mathbf{D}_\mathrm{d}(V^{1/2}, \lambda^\gamma, b\lambda^{-\bar{\gamma}/2}\gamma/2)$, since for all $\theta \in \Theta$ and $\gamma \in (0, \bar{\gamma}]$ we have that $\mathrm{R}_{\gamma,\theta}$ satisfies $\mathbf{D}_\mathrm{d}(V, \lambda^\gamma, b\gamma)$ and therefore using Jensen's inequality that $\mathrm{R}_{\gamma,\theta}$ satisfies the drift condition $\mathbf{D}_\mathrm{d}(V^{1/2}, \lambda^{\gamma/2}, b\lambda^{-\bar{\gamma}/2}\gamma/2)$.

We now give an upper bound on error $\|\pi_\theta \mathrm{P}_{m_\gamma \gamma,\theta} - \pi_\theta \mathrm{R}^{m_\gamma}_{\gamma,\theta}\|_{V^{1/2}}$. Indeed, since $\mathcal{A}_\theta$ satisfies a $\mathbf{D}_\mathrm{c}(V, \zeta, \beta)$ and $\mathrm{R}_{\gamma,\theta}$ satisfies $\mathbf{D}_\mathrm{d}(V, \lambda^\gamma, b\gamma)$ for any $\theta \in \Theta$ and $\gamma \in (0, \bar{\gamma}]$, we obtain using (S37) that for any $\theta \in \Theta$ and $\gamma \in (0, \bar{\gamma}]$

$$\pi_\theta \mathrm{P}_{\gamma m_\gamma,\theta}(V) \leq D_0 \ , \qquad \pi_\theta \mathrm{R}^{m_\gamma}_{\gamma,\theta}(V) \leq \tilde{D}_1 \ , \qquad \tilde{D}_1 = D_0 + b\lambda^{-\bar{\gamma}} \log(1/\lambda)^{-1} \ ,$$

Combining this result and Lemma S14 we have for any $\theta \in \Theta$ and $\gamma \in (0, \bar{\gamma}]$

$$\|\pi_\theta \mathrm{P}_{\gamma m_\gamma,\theta} - \pi_\theta \mathrm{R}^{m_\gamma}_{\gamma,\theta}\|_{V^{1/2}} \leq \tilde{D}_2 \gamma^{1/2} \ , \tag{S45}$$

with

$$\tilde{D}_2 = 2\tilde{D}_1^{1/2}(1+\bar{\gamma})^{1/2} \left\{ d + 2\bar{\gamma}(\mathtt{L}^2 M_V + \sup_{\theta \in \Theta} \|\nabla_x U_\theta(0)\|^2) \tilde{D}_1 \right\}^{1/2} \mathtt{L} \ .$$

Combining (S44) and (S45) we get for any $k \in \mathbb{N}$, $\theta \in \Theta$ and $\gamma \in (0, \bar{\gamma}]$

$$\begin{aligned}
\|\pi_\theta \mathrm{R}^k_{\gamma,\theta} - \pi_\theta \mathrm{P}_{\gamma k,\theta}\|_{V^{1/2}} &\leq C\tilde{D}_2 \sum_{\ell=1}^{q_\gamma} \xi^{\gamma m_\gamma \ell} \gamma^{1/2} \\
&\leq C\tilde{D}_2 (1-\xi)^{-1} \gamma^{1/2} \ ,
\end{aligned}$$

where we used that $\xi^{\gamma m_\gamma} \leq \xi$. The conclusion follows from this result and (S43). □

S1.5.3 Proof of Theorem 7

Combining Proposition S10 and Lemma S13 we get that **H**1-(i) is satisfied with constant $A_1 \leftarrow D_1$. **L**1, **L**2, Proposition S10 and Lemma S12-(a) ensure that **H**1-(ii) is satisfied by [4, Theorem 14] with $A_3 \leftarrow D_3$. **H**1-(iii) is satisfied combining Proposition S10, Proposition S11 and Proposition S15 with $\boldsymbol{\Psi}(\gamma) \leftarrow D_4\gamma^{1/2}$.

S1.6 Proof of Theorem 8

We preface the proof by a technical lemma.

Proposition S16. *Let $V : \mathbb{R}^d \to [1,+\infty)$ measurable and $M_{V,4} \geq 0$ satisfying $\sup_{x\in\mathbb{R}^d} \left\{(1+\|x\|^4)/V(x)\right\} \leq M_{V,4}$. Assume that there exists $M \geq 1$ such that for any $\theta \in \Theta$, $\gamma \in (0,\bar{\gamma}]$, with $\bar{\gamma} > 0$ and $x \in \mathbb{R}^d$, $\mathrm{R}_{\gamma,\theta}V(x) \leq MV(x)$. Assume **L** 1 and **L** 3, then we have for any $\theta_1,\theta_2 \in \Theta$, $\gamma_1,\gamma_2 \in (0,\bar{\gamma}]$ with $\gamma_2 < \gamma_1$, $a \in [1/4, 1/2]$ and $x \in \mathbb{R}^d$*

$$\|\delta_x \mathrm{R}_{\gamma_1,\theta_1} - \delta_x \mathrm{R}_{\gamma_2,\theta_2}\|_{V^a} \leq D_5 \left[\gamma_1/\gamma_2 - 1 + \gamma_2^{1/2}\|\theta_1-\theta_2\|\right] V(x)^{2a} , \tag{S46}$$

where $\{\mathrm{R}_{\gamma,\theta}, \gamma \in (0,\bar{\gamma}], \theta \in \Theta\}$ is the sequence of Markov kernels associated with the recursion (6) and

$$D_5 = \max\left(2M^{1/2}\left[d/4 + \sup_{\theta\in\Theta}\|\nabla_x U_\theta(0)\|^2 + \mathrm{L}^2 M_{4,V}^{1/2}\right]^{1/2}, (2M)^{1/2} L_U\right) .$$

Proof. Let $x \in \mathbb{R}^d$, $\theta_1,\theta_2 \in \Theta$ and $\gamma_1,\gamma_2 \in (0,\bar{\gamma}]$, $\gamma_2 < \gamma_1$. Using [6, Lemma 24] we have that

$$\begin{aligned}
&\|\delta_x \mathrm{R}_{\gamma_1,\theta_1} - \delta_x \mathrm{R}_{\gamma_2,\theta_2}\|_{V^a} \\
&\qquad \leq \sqrt{2}\left(\mathrm{R}_{\gamma_1,\theta_1}V^{2a}(x) + \mathrm{R}_{\gamma_2,\theta_2}V^{2a}(x)\right)^{1/2} \mathrm{KL}\left(\delta_x \mathrm{R}_{\gamma_1,\theta_1}|\delta_x \mathrm{R}_{\gamma_2,\theta_2}\right)^{1/2} \\
&\qquad \leq 2M^a V^a(x) \mathrm{KL}\left(\delta_x \mathrm{R}_{\gamma_1,\theta_1}|\delta_x \mathrm{R}_{\gamma_2,\theta_2}\right)^{1/2}
\end{aligned} \tag{S47}$$

Denote for any $\mu \in \mathbb{R}^d$ and $\sigma^2 > 0$, γ_{μ,σ^2} the d-dimensional Gaussian distribution with mean μ and covariance matrix $\sigma^2\,\mathrm{Id}$. Using that for any $\mu_1,\mu_2 \in \mathbb{R}^d$ and $\sigma_1,\sigma_2 > 0$,

$$\mathrm{KL}\left(\Upsilon_{\mu_1,\sigma_1\,\mathrm{Id}}|\Upsilon_{\mu_2,\sigma_2\,\mathrm{Id}}\right) = \|\mu_1-\mu_2\|^2/(2\sigma_2^2) + (d/2)\left\{-\log(\sigma_1^2/\sigma_2^2) - 1 + \sigma_1^2/\sigma_2^2\right\} .$$

In addition, if $\sigma_1 \geq \sigma_2$

$$\mathrm{KL}\left(\Upsilon_{\mu_1,\sigma_1\,\mathrm{Id}}|\Upsilon_{\mu_2,\sigma_2\,\mathrm{Id}}\right) \leq \|\mu_1-\mu_2\|^2/(2\sigma_2^2) + (d/2)(1-\sigma_1^2/\sigma_2^2)^2 .$$

Therefore, we obtain that

$$\mathrm{KL}\left(\delta_x \mathrm{R}_{\gamma_1,\theta_1}|\delta_x \mathrm{R}_{\gamma_2,\theta_2}\right) \leq \Xi/(4\gamma_2) + (d/2)(1-\gamma_1/\gamma_2)^2 , \tag{S48}$$

where Ξ satisfies

$$\begin{aligned}\Xi &= \|\gamma_1 \nabla_x U_{\theta_1}(x) - \gamma_2 \nabla_x U_{\theta_2}(x)\|^2 \\ &= \|\gamma_1 \nabla_x U_{\theta_1}(x) - \gamma_2 \nabla_x U_{\theta_1}(x) + \gamma_2 \nabla_x U_{\theta_1}(x) - \gamma_2 \nabla_x U_{\theta_2}(x)\|^2 \\ &\leq 2\|\gamma_1 \nabla_x U_{\theta_1}(x) - \gamma_2 \nabla_x U_{\theta_1}(x)\|^2 + 2\|\gamma_2 \nabla_x U_{\theta_1}(x) - \gamma_2 \nabla_x U_{\theta_2}(x)\|^2 \\ &\leq 2(\gamma_1-\gamma_2)^2\|\nabla_x U_{\theta_1}(x)\|^2 + 2\gamma_2^2\|\nabla_x U_{\theta_1}(x) - \nabla_x U_{\theta_2}(x)\|^2 \\ &\leq 2(\gamma_1-\gamma_2)^2\|\nabla_x U_{\theta_1}(x)\|^2 + 2\gamma_2^2 L_U^2\|\theta_1-\theta_2\|^2 V^{2a}(x) \;, \end{aligned} \tag{S49}$$

where we have used **L**3 in the last line. Using **L**3 again and that $\sup_{\theta\in\Theta}\|\nabla_x U_\theta(0)\| < +\infty$ by **L**1, we get for any $a \in [1/4, 1/2]$

$$\|\nabla_x U_\theta(x)\|^2 \leq 2(\|\nabla_x U_\theta(x) - \nabla_x U_\theta(0)\|^2 + \sup_{\theta\in\Theta}\|\nabla_x U_\theta(0)\|^2) \leq C_\Theta V^{2a}(x) \;,$$

with $C_\Theta = 2\sup_{\theta\in\Theta}\|\nabla_x U_\theta(0)\|^2 + 2\mathrm{L}^2 M_{4,V}^{1/2}$. Combining this result, $\log(\gamma_2/\gamma_1) \leq 0$ and and (S49) in (S48), it follows that

$$\begin{aligned}&\mathrm{KL}\left(\delta_x \mathrm{R}_{\gamma_1,\theta_1} | \delta_x \mathrm{R}_{\gamma_2,\theta_2}\right) \leq d(1-\gamma_1/\gamma_2)^2/2 \\ &+ \gamma_2^{-1}(\gamma_1-\gamma_2)^2\|\nabla_x U(\theta_1,x)\|^2/2 + \gamma_2 L_U^2\|\theta_1-\theta_2\|^2 V^{2a}(x)/2 \\ &\leq \left[d\gamma_2^{-1}(1-\gamma_2/\gamma_1)/4 + \gamma_2^{-1}(\gamma_1-\gamma_2)^2 C_\Theta/2 + \gamma_2 L_U^2\|\theta_1-\theta_2\|^2/2\right] V^{2a}(x) \;.\end{aligned}$$

This result substituted in (S47) completes the proof with the fact that for any $a, b \in \mathbb{R}_+$, $(a+b)^{1/2} \leq a^{1/2} + b^{1/2}$. □

Proof of Theorem 8. **L**1 and **L**2 ensure a uniform drift condition on $\mathrm{R}_{\gamma,\theta}$, see Proposition S10 . Note that the Lyapunov function V defined by Proposition S10 satisfies $\sup_{x\in\mathbb{R}^d}(1+\|x\|^4)/V(x) < +\infty$. is then a direct consequence of Proposition S16 □

S1.7 Proof of Theorem 9

We preface the proof with the following lemmas.

Lemma S17. *Assume* **L**1 *and* **L**4*.*

(a) There exist $b \geq 0$ and $\lambda \in [0,1)$ such that for any $\gamma \in (0,\bar{\gamma}]$ with $\bar{\gamma} < \mathtt{m}_3(2\mathrm{L}^2 + 4\sup_{\theta\in\Theta}\|\nabla_x U_\theta(0)\|^2/R_3^2)^{-1}$ and $\theta \in \Theta$, $\mathrm{R}_{\gamma,\theta}$ associated with the recursion (6) satisfies $\mathbf{D}_{\mathrm{d}}(V, \lambda^\gamma, b\gamma)$ with $V : \mathbb{R}^d \to [1,+\infty)$ given for any $x \in \mathbb{R}^d$ by $V(x) = 1 + \|x\|^4$.

(b) There exist $\zeta > 0$ and $b \geq 0$ such that for $\theta \in \Theta$, $(\mathrm{P}_{t,\theta})_{t\geq 0}$ associated with (5) satisfies $\mathbf{D}_{\mathrm{c}}(V, \zeta, \beta)$.

(c) In addition, there exist $\varpi > 0$, $\tilde{b}, \tilde{\beta} \geq 0$ independent of the dimension d such that

$$b \leq \tilde{b}(1+d^\varpi) \;, \qquad \beta \leq \tilde{\beta}(1+d^\varpi) \;,$$

and λ and ζ are independent of the dimension d.

Proof. Let $\theta \in \Theta$ and $x \in \mathbb{R}^d$

(a) Denote $T_{\gamma,\theta}(x) = x - \gamma \nabla_x U_\theta(x)$ and $X = T_{\gamma,\theta}(x) + \sqrt{2\gamma} Z$ where Z is a d-dimensional Gaussian random variable with zero mean and identity covariance matrix. Using **L**1 and **L**4, if $\|x\| \geq R_3$ we have for any $\gamma \in (0, \bar{\gamma}]$

$$
\begin{aligned}
\|T_{\gamma,\theta}(x)\|^2 &\leq (1 - 2\gamma \mathtt{m}_3 + 2\gamma^2 \mathtt{L}^2) \|x\|^2 + 2\gamma^2 M^2 \ , \\
\|T_{\gamma,\theta}(x)\|^4 &\leq (1 - 2\gamma \mathtt{m}_3 + 2\gamma^2 \mathtt{L}^2) \|x\|^4 + 4\gamma^2 M^2 \|x\|^2 + 4\gamma^4 M^4 \\
&\leq (1 - 2\gamma \mathtt{m}_3 + 2\gamma^2 \mathtt{L}^2 + 4\gamma^2 M) \|x\|^4 + 4\gamma^2 M^2 (1 + M^2 \gamma^2) \ ,
\end{aligned}
$$

where $M = \sup_{\theta \in \Theta} \|\nabla_x U_\theta(0)\|$. In addition, $\|T_{\gamma,\theta}(x)\| \leq \|x\| + \gamma M$ and

$$
\begin{aligned}
&\|X\|^4 = \|T_{\gamma,\theta}(x)\|^4 + 4\gamma \|T_{\gamma,\theta}(x)\|^2 \|Z\|^2 + 4\gamma^2 \|Z\|^4 \\
&+ 2^{5/2} \gamma^{1/2} \|T_{\gamma,\theta}(x)\|^2 \langle T_{\gamma,\theta}(x), Z \rangle + 8\gamma \|T_{\gamma,\theta}(x)\|^2 \|Z\|^2 + 2\gamma^{3/2} \|Z\|^2 \langle T_{\gamma,\theta}(x), Z \rangle \ .
\end{aligned}
$$

Therefore, we obtain

$$
\begin{aligned}
\mathbb{E}[\|X\|^4] &= \|T_{\gamma,\theta}(x)\|^4 + 12\gamma d \|T_{\gamma,\theta}(x)\|^2 + 4d(d+2)\gamma^2 \\
&\leq \|T_{\gamma,\theta}(x)\|^4 + 24\gamma d \|x\|^2 + 24\gamma^2 M^2 + 4d(d+2)\gamma^2 \ .
\end{aligned}
$$

Let $R_4 \geq R_3$ First, assume that $\|x\| \geq R_4$ then

$$
\begin{aligned}
\mathbb{E}[\|X\|^4] \leq (1 - 2\gamma \mathtt{m}_3 + 2\gamma^2 \mathtt{L}^2 + 4\gamma^2 M^2 + 24\gamma d / R_4^2) \|x\|^4 \\
+ 24\gamma^2 M^2 + 4d(d+2)\gamma^2 + 4\gamma^2 M^2 (1 + \gamma^2 M^2) \ .
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
\mathbb{E}[1 + \|X\|^4] \leq (1 - 2\gamma \mathtt{m}_3 + 2\gamma^2 \mathtt{L}^2 + 4\gamma^2 M^2 + 24\gamma d (1 + \gamma \mathtt{L}) / R_4^2)(1 + \|x\|^4) \\
+ 24\gamma^2 M^2 + 4d(d+2)\gamma^2 + 4\gamma^2 M^2 (1 + \gamma^2 M^2) + 2\gamma \mathtt{m}_3 \ .
\end{aligned}
$$

Second assume that $\|x\| \leq R_4$, we have

$$
\|T_{\gamma,\theta}(x)\|^4 \leq \|x\|^4 + 4\gamma M R_4^3 + 6\gamma^2 M^2 R_4^2 + 4\gamma^3 M^3 R_4 + \gamma^4 M^4 \ .
$$

Therefore,

$$
\begin{aligned}
\mathbb{E}[1 + \|X\|^4] \leq (1 - 2\gamma \mathtt{m}_3 + 2\gamma^2 \mathtt{L}^2 + 4\gamma^2 M^2 + 24\gamma d R_4^2)(1 + \|x\|^4) \\
+ 4\gamma M R_4^3 + 6\gamma^2 M^2 R_4^2 + 4\gamma^3 M^3 R_4 + \gamma^4 M^4 + 2\gamma \mathtt{m}_3 (1 + R_4^4) \\
+ 24\gamma d R_4^2 + 24\gamma^2 M^2 + 4d(d+2)\gamma^2 \ .
\end{aligned}
$$

Let $R_4 = \max\{(24d/\mathtt{m}_3^2)^{1/2}, R_3\}$. Then $\mathbf{D}_{\mathrm{d}}(V, \lambda^\gamma, b\gamma)$ holds with $V : \mathbb{R}^d \to [1, +\infty)$ and

$$
\begin{aligned}
\lambda &= \exp[-\mathtt{m}_3 + 2\bar{\gamma} \mathtt{L}^2 + 4\bar{\gamma} M^2] \\
b &= 4M R_4^3 + 6\bar{\gamma} M^2 R_4^2 + 4\bar{\gamma}^2 M^3 R_4 + \bar{\gamma}^3 M^4 + 2\mathtt{m}_3 (1 + R_4^4) \\
&\quad + 24 d R_4^2 + 24\bar{\gamma} M^2 + 4d(d+2)\bar{\gamma} + 4\bar{\gamma} M^2 (1 + \bar{\gamma}^2 M^2) \ .
\end{aligned}
\tag{S50}
$$

(b) We have for any $x \in \mathbb{R}^d$

$$\nabla V(x) = 4 \|x\|^2 x \;, \qquad \Delta V(x) = 4(d+2) \|x\|^2 \;.$$

Therefore we have

$$\mathcal{A}_\theta V(x) = -4 \|x\|^2 \langle \nabla_x U_\theta(x), x \rangle + 4(d+2) \|x\|^2 \;.$$

Let $R_4 = \max(R_3, (2(d+2)\mathtt{m}_3^{-1})^{1/2})$. If $\|x\| \geq R_4$ then

$$\mathcal{A}_\theta V(x) \leq -4\mathtt{m}_3 \|x\|^4 + 4(d+2) \|x\|^2 \leq -2\mathtt{m}_3 V(x) + 2\mathtt{m}_3 \;.$$

If $\|x\| \leq R_4$ then

$$\mathcal{A}_\theta V(x) \leq 4R_4^3 M + 4\mathtt{L} R_4^4 + 4d(d+2)R_4^2 \;.$$

Hence, we obtain that $\mathbf{D}_\mathrm{c}(V, \zeta, \beta)$ with

$$\zeta = 2\mathtt{m}_3 \;, \qquad \beta = 4R_4^3 M + 4\mathtt{L} R_4^4 + 4d(d+2)R_4^2 + 4(d+2)R_4^2 + 2\mathtt{m}_3 \;, \tag{S51}$$

which concludes the proof.

(c) The last point follows from (S50), (S51) and **L**4.

$\square$

Proof of Theorem 9. Combining Lemma S17 and Lemma S13 we get that **H** 1-(i) is satisfied with constant $A_1 \leftarrow \tilde{A}_1(1 + d^{\varpi_1})$ with $\tilde{A}_1 \geq 0$ and $\varpi_1 \in \mathbb{N}^*$ independent of the dimension. Using **L**1, **L**4 and a straightforward adaptation of [4, Theorem 16] (where $V(x) = 1 + \|x\|^2$ is replaced by $V(x) = 1 + \|x\|^4$) and the discussion that follows we obtain **H**1-(ii) is satisfied with $A_2 = \tilde{A}_2(1 + d^{\varpi_2})$, $\log^{-1}(\rho^{-1}) = \log^{-1}(\tilde{\rho}^{-1})(1 + d^{\varpi_2})$ and $A_3 = \tilde{A}_3(1 + d^{\varpi_3})$ with $\varpi_2, \varpi_3 \in \mathbb{N}^*$, $\tilde{\rho} \in [0, 1)$ and $\tilde{A}_2, \tilde{A}_3 \geq 0$ independent of the dimension d. Using **L**1, **L**4 and a straightforward adaptation of [4, Theorem 23] (where $V(x) = 1 + \|x\|^2$ is replaced by $V(x) = 1 + \|x\|^4$) we obtain that for any $x \in \mathbb{R}^d$, $\|\delta_x \mathrm{P}_{t,\theta} - \pi_\theta\|_V \leq C\xi^t V(x)$ with $C = \tilde{C}(1 + d^{\varpi_c})$ and $\log^{-1}(\xi^{-1}) = \log^{-1}(\tilde{\xi}^{-1})(1 + d^{\varpi_c})$ with $\tilde{C}, \varpi_c \geq 0$ and $\tilde{\xi} \in [0, 1)$ independent of the dimension d. Using this result and Proposition S15 we obtain that **H**1-(iii) holds with the constant D_4 in (S42) which satisfies $D_4 = \tilde{A}_4(1 + d^{\varpi_4})$ with $\tilde{A}_4 \geq 0$ and $\varpi_4 \in \mathbb{N}^*$. Using Proposition S16 we obtain that **H**2 holds with the constant D_5 in (S46) which satisfies $D_5 = \tilde{A}_5(1 + d^{\varpi_5})$ with $\varpi_5 \in \mathbb{N}^*$. Finally note that $B_1, B_2, C_1, \tilde{C}_1, C_3, C_2$ in Lemma S3, Lemma S4, Lemma S5, Lemma S6, Lemma S7 and Lemma S8 depend polynomially on A_i with $i \in \{1, \ldots, 5\}$ and $\log^{-1}(\rho^{-1})$. Therefore Theorem 4 and Theorem 6 apply and the conclusion follows using the explicit expressions for the constants appearing in these results. $\square$

S2 Posterior convexity

Lemma S18. *For any $y \in \{0,1\}^{d_y}$, $\theta \mapsto p(y|\theta)$ given by (19) is log-concave.*

Proof. Let $\theta \in \mathbb{R}$, then by (19), for any $y \in \mathbb{R}$ we have $p(y|\theta) = \int_{\mathbb{R}^d} p(y,\beta|\theta)\mathrm{d}\beta$ with

$$p(y,\beta|\theta) = (2\pi\sigma^2)^{-d/2}\left\{\prod_{i=1}^{d_y} s(x_i^{\mathrm{T}}\beta)^{y_i}(1-s(x_i^{\mathrm{T}}\beta))^{1-y_i}\right\}\mathrm{e}^{-\frac{\|\beta-\theta\mathbf{1}_d\|^2}{2\sigma^2}} .$$

Therefore we have using that for any $t \in \mathbb{R}$, $1-s(t) = s(-t)$

$$\log p(y,\beta|\theta) = (-d/2)\log(2\pi\sigma^2) + \left\{\sum_{i=1}^{d_y} y_i \log(s(x_i^{\mathrm{T}}\beta)) + (1-y_i)\log(s(-x_i^{\mathrm{T}}\beta))\right\} - \frac{\|\beta-\theta\mathbf{1}_d\|^2}{2\sigma^2} .$$

Since $y_i \geq 0$, $1-y_i \geq 0$, $(\beta,\theta) \mapsto \|\beta-\theta\mathbf{1}_d\|^2$, $t \mapsto \log(s(t))$ and $t \mapsto \log(s(-t))$ are convex, we obtain that $(\beta,\theta) \mapsto p(y,\beta|\theta)$ is log-concave. Using the Prékopa–Leindler inequality [8, Theorem 7.1] we obtain that $\theta \mapsto p(y|\theta)$ is log-concave which concludes the proof. □

S3 Non-convex objective function

In this section we turn to the case where f is non-convex. We recall that the normal cone of a sub-manifold $\mathcal{M} \subset \mathbb{R}^{d_\Theta}$ (see [2]) at point θ is given by

$$\mathrm{N}(\theta,\mathcal{M}) = \begin{cases} \left\{u \in \mathbb{R}^{d_\Theta} \,:\, \langle u, \tilde{\theta}-\theta\rangle \leq 0\,, \text{ for any } \tilde{\theta} \in \Theta\right\} & \text{if } \theta \in \mathcal{M}\,; \\ \{0\} \quad \text{otherwise}\,, \end{cases}$$

Note that for any $\theta \in \mathrm{int}(\mathcal{M})$, $\mathrm{N}(\theta,\mathcal{M}) = \{0\}$, where $\mathrm{int}(\mathcal{M})$ denotes the interior of $\mathcal{M}$. For any $R \geq 0$, we have $\mathrm{N}(\theta,\overline{\mathrm{B}}(0,R)) = \{r\theta \,:\, r \geq 0\}$ if $\|\theta\| = R$. We consider the following assumption on the set Θ.

A 2. $\Theta = \{\theta \in \mathbb{R}^{d_\Theta} \,:\, q_i(\theta) \leq 0\,,$ *for any* $i \in \{1,\dots,p\}\}$ *with* $p \in \mathbb{N}^*$ *and for any* $i \in \{1,\dots,p\}$, $q_i \in \mathrm{C}^1(\mathbb{R}^{d_\Theta})$ *is convex and such that for any* $\theta \in \mathbb{R}^{d_\Theta}$, $\nabla q_i(\theta) \neq 0$ *if* $q_i(\theta) = 0$.

Under **A**2, for any $\theta \in \Theta\backslash\mathrm{int}(\Theta)$ we have that $\mathsf{I}(\theta) = \{i \in \{1,\dots,p\} \,:\, q_i(\theta) = 0\}$ is not empty. In this case we obtain that $\mathrm{N}(\theta,\Theta)$ is the convex cone generated by $\{\nabla q_i(\theta) \,:\, i \in \mathsf{I}(\theta)\}$, see [11].

Theorem S19. *Assume* **A**1 *and* **A**2. *Let* $\bar{\gamma} > 0$, $(\gamma_n)_{n\in\mathbb{N}}$, $(\delta_n)_{n\in\mathbb{N}}$ *be sequences of non-increasing positive real numbers and* $(m_n)_{n\in\mathbb{N}}$ *be a sequence of positive integers such that* $\sup_{n\in\mathbb{N}}\delta_n < 1/L_f$, $\sup_{n\in\mathbb{N}}\gamma_n < \bar{\gamma}$ *and* (11) *are satisfied. Let* $\{(X_k^n)_{k\in\{0,\dots,m_n\}} \,:\, n \in \mathbb{N}\}$ *be given by* (10). *Assume in addition that* **H**1 *is satisfied. Then* $(\theta_n)_{n\in\mathbb{N}}$ *defined by* (10) *satisfies* $\lim_{n\to+\infty} d(\theta_n,\mathsf{A}) = 0$ *a.s with* $\mathsf{A} = \{\theta \in \Theta \,:\, \nabla f(\theta) + \mathbf{n} = 0,\ \mathbf{n} \in \mathrm{N}(\theta,\Theta)\}$. *In addition, if* $|\mathsf{A}| < +\infty$ *then* $\lim_{n\to+\infty}\theta_n = \theta^\star \in \mathsf{A}$ *a.s.*

Proof. The proof is an application of [11, Chapter 5, Theorem 2.3] using the decomposition of the error term considered in the proof of Theorem 3 and Theorem 5. Indeed we decompose the error term η_n defined by (S1) as $\eta_n = \delta M_n + B_n$, where δM_n is a martingale increment. Then, we only need to show that the following sums converge

$$\sum_{k=0}^{n} \delta_{k+1}^2 \mathbb{E}\left[\|\delta M_k\|^2\right] , \qquad \sum_{k=0}^{n} \delta_{k+1} \mathbb{E}\left[\|B_k\|\right] ,$$

which is established in Lemma S3 and Lemma S4. Denote, B the limit set (see [11] for a definition) of the projected ODE $(\mathrm{d}\theta_t/\mathrm{d}t) = -\bar{\Pi}(\theta_t, \nabla f(\theta_t))$, where $\bar{\Pi}$ is the projection operator defined in [10, p.191]. For any $t \geq 0$, denote $u(t) = f(\theta_t)$. Then u satisfies $u'(t) \leq 0$, see [11, p.108]. Using Lasalle's invariance principle [12, Theorem 1] we obtain that $\mathsf{B} \subset \mathsf{A}$ which concludes the first part of the proof using [11, Chapter 5, Theorem 2.3]. Assume that $\mathsf{A} = \{\theta_1^\star, \dots, \theta_M^\star\}$ with $M \in \mathbb{N}^*$. Using Lemma S3 and Lemma S4 and that for any $a, b \geq 0$, $(a+b)^2 \leq 2a^2 + 2b^2$ we obtain that

$$\mathbb{E}\left[\sum_{n \in \mathbb{N}} \|\theta_{n+1} - \theta_n\|^2\right] \leq 2 \sum_{n \in \mathbb{N}} \delta_n^2 \left\{\mathbb{E}\left[\left\|\nabla f(\theta_n)^2\right\|\right] + \mathbb{E}[\|\eta_n\|^2]\right\} < +\infty .$$

Hence, we obtain that $\lim_{k \to +\infty} \|\theta_{n+1} - \theta_n\| = 0$ a.s. Denote $r = \min\{\left\|\theta_i^\star - \theta_j^\star\right\| : i, j \in \{1, \dots, M\},\ i \neq j\}$. For any $n \in \mathbb{N}$ let $i_n \in \{1, \dots, M\}$ such that $d(\theta_n, \mathsf{A}) = \left\|\theta_n - \theta_{i_n}^\star\right\|$. Then, since $\lim_{n \to +\infty} d(\theta_n, \mathsf{A}) = 0$, there exists $n_0 \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ with $n \geq n_0$, i_n is uniquely defined, $d(\theta_n, \mathsf{A}) \leq r/4$ and $\|\theta_{n+1} - \theta_n\| \leq r/4$. Assume that there exists $n_1 \in \mathbb{N}$ with $n_1 \geq n_0$ such that $i_{n_1} \neq i_{n_1+1}$. Then $\|\theta_{n_1+1} - \theta_{i_{n_1+1}}^\star\| \leq r/4$ and

$$\|\theta_{n_1+1} - \theta_{i_{n_1+1}}^\star\| \geq \|\theta_{i_{n_1}}^\star - \theta_{i_{n_1+1}}^\star\| - \|\theta_{n_1} - \theta_{i_{n_1}}^\star\| - \|\theta_{n_1} - \theta_{i_{n_1+1}}\| \geq r/2 ,$$

which is absurd. Therefore, for any $n \in \mathbb{N}$ with $n \geq n_0$, $i_n = i_{n_0}$ which concludes the proof. □

References

[1] Y. F. Atchadé, G. Fort, and E. Moulines. On perturbed proximal gradient algorithms. *J. Mach. Learn. Res*, 18(1):310–342, 2017.

[2] T. Aubin. *A course in differential geometry*. Graduate Studies in Mathematics. AMS, 2000.

[3] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[4] V. De Bortoli and A. Durmus. Convergence of diffusions and their discretizations:from continuous to discrete processes and back. *arXiv preprint arXiv:1904.09808*, 2019.

[5] R. Douc, E. Moulines, P. Priouret, and P. Soulier. *Markov Chains*. Springer, 2018. to be published.

[6] A. Durmus and E. Moulines. Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *Ann. Appl. Probab.*, 27(3):1551–1587, 2017.

[7] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[8] R. J. Gardner. The Brunn-Minkowski inequality. *Bull. Amer. Math. Soc. (N.S.)*, 39(3):355–405, 2002.

[9] S. Kullback. *Information theory and statistics*. Dover Publications, Inc., Mineola, NY, 1997. Reprint of the second (1968) edition.

[10] H. J. Kushner and D. S. Clark. *Stochastic approximation methods for constrained and unconstrained systems*, volume 26. Springer Science & Business Media, 2012.

[11] H. J. Kushner and G. G. Yin. *Stochastic approximation and recursive algorithms and applications*, volume 35 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2003. Stochastic Modelling and Applied Probability.

[12] J. LaSalle. Some extensions of liapunov's second method. *IRE Transactions on circuit theory*, 7(4):520–527, 1960.

[13] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. I. Criteria for discrete-time chains. *Adv. in Appl. Probab.*, 24(3):542–574, 1992.

[14] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London, Ltd., London, 1993.

[15] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548, 1993.

[16] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548, 1993.

[17] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[18] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes.* Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

[19] D. Williams. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.