Supplement to High-dimensional Bayesian inference via the Unadjusted Langevin Algorithm

Alain Durmus¹

Éric Moulines²

March 5, 2018

1 Proofs of Section 2

1.1 Proof of Proposition 1

(i) The generator \mathscr{A} associated with $(P_t)_{t\geq 0}$ is given, for all $f\in C^2(\mathbb{R}^d)$ and $y\in\mathbb{R}^d$, by:

$$\mathscr{A}f(y) = -\langle \nabla U(y), \nabla f(y) \rangle + \Delta f(y) .$$
(S1)

Denote for all $y \in \mathbb{R}^d$ by $V(y) = ||y - x^*||^2$. Let $x \in \mathbb{R}^d$ and $(Y_t)_{t \ge 0}$ be a solution of (1) started at x. Under **H1** sup_{t \in [0,T]} $\mathbb{E}[||Y_t||^2] < +\infty$ for all $T \ge 0$. Therefore, the process

$$\left(V(Y_t) - V(x) - \int_0^t \mathscr{A}V(Y_s) \mathrm{d}s\right)_{t \ge 0}$$

is a $(\mathcal{F}_t)_{t\geq 0}$ -martingale. Denote for all $t\geq 0$ and $x\in \mathbb{R}^d$ by $v(t,x)=P_tV(x)$. Then we have, $\partial v(t,x)/\partial t=P_t\mathscr{A}V(x)$.

Since $\nabla U(x^{\star}) = 0$ and by **H**2, $\langle \nabla U(x) - \nabla U(x^{\star}), x - x^{\star} \rangle \ge m ||x - x^{\star}||^2$, we have

$$\mathscr{A}V(x) = 2\left(-\left\langle\nabla U(x) - \nabla U(x^{\star}), x - x^{\star}\right\rangle + d\right) \le 2\left(-mV(x) + d\right) .$$
(S2)

Therefore, we get

$$\frac{\partial v(t,x)}{\partial t} = P_t \mathscr{A} V(x) \le -2mP_t V(x) + 2d = -2mv(t,x) + 2d ,$$

and the proof follows from the Grönwall inequality.

 $^{^1 {\}rm \acute{E}cole}$ Normale Supérieure CMLA 61, Av. du Président Wilson 94235 Cachan Cedex, France Email: alain.durmus@cmla.ens-cachan.fr

 $^{^2 {\}rm Centre}$ de Mathématiques Appliquées, UMR 7641, Ecole Polytechnique, France. eric.moulines@polytechnique.edu

(ii) Set $V(x) = ||x - x^*||^2$. By Jensen's inequality and Lemma 19-(i), for all c > 0 and t > 0, we get

$$\pi(V \wedge c) = \pi P_t(V \wedge c) \le \pi(P_t V \wedge c)$$

$$\le \int \pi(\mathrm{d}x) \, c \wedge \left\{ \|x - x^*\|^2 \mathrm{e}^{-2mt} + \frac{d}{m} (1 - \mathrm{e}^{-2mt}) \right\}$$

$$\le \pi(V \wedge c) \mathrm{e}^{-2mt} + (1 - \mathrm{e}^{-2mt}) d/m .$$

Taking the limit as $t \to +\infty$, we get $\pi(V \land c) \leq d/m$. Using the monotone convergence theorem and taking the limit as $c \to +\infty$ concludes the proof.

(iii) Let $x, y \in \mathbb{R}^d$. Consider the following SDE in $\mathbb{R}^d \times \mathbb{R}^d$:

$$\begin{cases} dY_t = -\nabla U(Y_t)dt + \sqrt{2}dB_t ,\\ d\tilde{Y}_t = -\nabla U(\tilde{Y}_t)dt + \sqrt{2}dB_t ,\end{cases}$$
(S3)

where $(Y_0, \tilde{Y}_0) = (x, y)$. Since ∇U is Lipschitz, then by [4, Theorem 2.5, Theorem 2.9, Chapter 5], this SDE has a unique strong solution $(Y_t, \tilde{Y}_t)_{t\geq 0}$ associated with $(B_t)_{t\geq 0}$. Moreover since $(Y_t, \tilde{Y}_t)_{t\geq 0}$ is a solution of (S3),

$$\left\|Y_t - \tilde{Y}_t\right\|^2 = \left\|Y_0 - \tilde{Y}_0\right\|^2 - 2\int_0^t \left\langle\nabla U(Y_s) - \nabla U(\tilde{Y}_s), Y_s - \tilde{Y}_s\right\rangle \mathrm{d}s ,$$

which implies using H_2 and Grönwall's inequality that

$$\left\|Y_{t} - \tilde{Y}_{t}\right\|^{2} \leq \left\|Y_{0} - \tilde{Y}_{0}\right\|^{2} - 2m \int_{0}^{t} \left\|Y_{s} - \tilde{Y}_{s}\right\|^{2} ds \leq \left\|Y_{0} - \tilde{Y}_{0}\right\|^{2} e^{-2mt}$$

Since for all $t \ge 0$, the law of (Y_t, \tilde{Y}_t) is a coupling between $\delta_x P_t$ and $\delta_y P_t$, by definition of W_2 , $W_2(\delta_x P_t, \delta_y P_t) \le \mathbb{E}[||Y_t - \tilde{Y}_t||^2]^{1/2}$, which concludes the proof.

(iv) The proof is a direct consequence of (ii) and (iii)

1.2 Proof of Proposition 2

(i) Note that the proof is trivial if $\ell < n$. Therefore we only need to consider the case $\ell \geq n$. For any $\gamma \in (0, 2/(m+L))$, we have for all $x \in \mathbb{R}^d$:

$$\int_{\mathbb{R}^d} \|y - x^\star\|^2 R_\gamma(x, \mathrm{d}y) = \|x - \gamma \nabla U(x) - x^\star\|^2 + 2\gamma d \, dx$$

Using that $\nabla U(x^*) = 0$, and (3), we get from the previous inequality:

$$\begin{split} \int_{\mathbb{R}^d} \|y - x^\star\|^2 R_{\gamma}(x, \mathrm{d}y) \\ &\leq (1 - \kappa\gamma) \|x - x^\star\|^2 + \gamma \left(\gamma - \frac{2}{m+L}\right) \|\nabla U(x) - \nabla U(x^\star)\|^2 + 2\gamma d \\ &\leq (1 - \kappa\gamma) \|x - x^\star\|^2 + 2\gamma d , \end{split}$$

where we have used for the last inequality that $\gamma \leq 2/(m+L)$. Then by definition (7) of $Q_{\gamma}^{n,\ell}$ for $\ell, n \geq 1, \ell \geq n$, the proof follows from a straightforward induction.

(ii) By (i), we have for all
$$x \in \mathbb{R}^d$$
 and $n \ge 1$,

$$\int_{\mathbb{R}^d} \|y - x^\star\|^2 R^n_{\gamma}(x, \mathrm{d}y) \le (1 - \kappa\gamma)^n \|x - x^\star\|^2 + 2d \sum_{k=1}^n \gamma (1 - \kappa\gamma)^{n-k} \le (1 - \kappa\gamma)^n \|x - x^\star\|^2 + 2\kappa^{-1} d(1 - (1 - \kappa\gamma)^n) .$$
(S4)

Since any compact set of \mathbb{R}^d is accessible and small for R_{γ} , then [5, Theorem 15.0.1] implies that R_{γ} has a unique stationary distribution π_{γ} . Using (S4), the proof is along the same line as the one of Proposition 1-(ii).

1.3 Proof of Proposition **3**

(i) $(Z_k)_{k\geq 1}$ be a sequence of i.i.d. *d*-dimensional Gaussian random variables. We consider the processes $(X_k^{n,1}, X_k^{n,2})_{k\geq 0}$ with $(X_0^{n,1}, X_0^{n,2}) = (x, y)$ and defined for $k \geq 0$ by

$$X_{k+1}^{n,j} = X_k^{n,j} - \gamma_{k+n} \nabla U(X_k^{n,j}) + \sqrt{2\gamma_{k+n}} Z_{k+1} \quad j = 1, 2.$$
(S5)

Using (S5), we get for any $\ell \ge n \ge 1$. $W_2^2(\delta_x Q_{\gamma}^{n,\ell}, \delta_y Q_{\gamma}^{n,\ell}) \le \mathbb{E}[\|X_{\ell}^{n,1} - X_{\ell}^{n,2}\|^2]$ and (3) implies for $k \ge n-1$,

$$\begin{split} \left\| X_{k+1}^{n,1} - X_{k+1}^{n,2} \right\|^2 &= \left\| X_k^{n,1} - X_k^{n,2} \right\|^2 + \gamma_{n+k}^2 \left\| \nabla U(X_k^{n,1}) - \nabla U(X_k^{n,2}) \right\|^2 \\ &- 2\gamma_{n+k} \left\langle X_k^{n,1} - X_k^{n,2}, \nabla U(X_k^{n,1}) - \nabla U(X_k^{n,2}) \right\rangle \\ &\leq (1 - \kappa \gamma_{n+k}) \left\| X_k^{n,1} - X_k^{n,2} \right\|^2 \,. \end{split}$$

Therefore by a straightforward induction we get for all $\ell \geq n$,

$$\left\|X_{\ell}^{n,1} - X_{\ell}^{n,2}\right\|^{2} \leq \prod_{k=n}^{\ell} (1 - \kappa \gamma_{k}) \left\|X_{0}^{n,1} - X_{0}^{n,2}\right\|^{2}$$

(ii) Let $\mu \in \mathcal{P}_{2p}(\mathbb{R}^d)$ and $p \geq 1$. It is straightforward that for all $n \geq 0$, $\mu R_{\gamma}^n \in \mathcal{P}_{2p}(\mathbb{R}^d)$. Then, by Proposition 3-(i) for $\gamma \leq 2/(m+L)$, R_{γ} is a strict contraction in $(\mathcal{P}_{2p}(\mathbb{R}^d), W_{2p})$ and there is a unique fixed point π_{γ} which is the unique invariant distribution. (ii) follows from Proposition 3-(i).

2 Proofs of Section 3

2.1 Proof of Lemma 7

Let $\ell \in \{1, \ldots, n+1\}$. Since $(\gamma_k)_{k \ge 1}$ is non-increasing,

$$\sum_{i=1}^{n+1} \prod_{k=i+1}^{n+1} (1 - \varpi \gamma_k) \gamma_i^j = \sum_{i=1}^{\ell-1} \prod_{k=i+1}^{n+1} (1 - \varpi \gamma_k) \gamma_i^j + \sum_{i=\ell}^{n+1} \prod_{k=i+1}^{n+1} (1 - \varpi \gamma_k) \gamma_i^j$$
$$\leq \prod_{k=\ell}^{n+1} (1 - \varpi \gamma_k) \sum_{i=1}^{\ell-1} \gamma_i^j + \gamma_\ell^{j-1} \sum_{i=\ell}^{n+1} \prod_{k=i+1}^{n+1} (1 - \varpi \gamma_k) \gamma_i$$
$$\leq \prod_{k=\ell}^{n+1} (1 - \varpi \gamma_k) \sum_{i=1}^{\ell-1} \gamma_i^j + \frac{\gamma_\ell^{j-1}}{\varpi} .$$

2.2 Proof of (28)

Consider the constant sequence $\gamma_k = \gamma$ for all $k \in \mathbb{N}^*$ with $\gamma \in (0, 1/(m+L)]$. By (22), we have for all $n \in \mathbb{N}^*$ and $x \in \mathbb{R}^d$

$$\vartheta_{n,0}^{(1)}(x) \le \gamma \mathsf{D}_1(\gamma, d) + \gamma^3 \mathsf{D}_2(\gamma) \sum_{i=1}^n (1 - \kappa \gamma/2)^{n-i} \delta_{i,n,0}(x) ,$$

where

$$\mathsf{D}_{1}(\gamma, d) = 2L^{2}\kappa^{-1} \left(\kappa^{-1} + \gamma\right) \left(2d + L^{2}\gamma^{2}/6\right) , \quad \mathsf{D}_{2}(\gamma) = L^{4} \left(\kappa^{-1} + \gamma\right) .$$

In addition, using that $\kappa \geq 2m$ and for all $t \geq 0, 1 - t \leq e^{-t}$,

$$\sum_{i=1}^{n} (1 - \kappa \gamma/2)^{n-i} \delta_{i,n,0}(x) = \sum_{i=1}^{n} \left[(1 - \kappa \gamma/2)^{n-i} \left\{ e^{-2m\gamma(i-1)} \|x - x^{\star}\|^{2} + \left(1 - e^{-2m\gamma(i-1)} \right) (d/m) \right\} \right]$$
$$\leq n e^{-m\gamma(n-1)} \|x - x^{\star}\|^{2} + 2d(\kappa\gamma m)^{-1} .$$
(S6)

Therefore for all $n \ge 1$ and $x \in \mathbb{R}^d$ we get

$$\vartheta_{n,0}^{(1)}(x) \le \gamma \mathsf{D}_1(\gamma) + \gamma^3 \mathsf{D}_2(\gamma) \left\{ n \mathrm{e}^{-m\gamma(n-1)} \|x - x^\star\|^2 + 2d(\kappa\gamma m)^{-1} \right\} .$$
(S7)

Let now $\ell \in \mathbb{N}^*$, $\ell \ge \lceil \gamma^{-1} \rceil + 1$ and $n = \ell - \lceil \gamma^{-1} \rceil$. Then,

$$\begin{split} &\sum_{k=n+1}^{\ell} \left\{ (\gamma_k^3 L^2/3) \varrho_{1,k-1}(x) + d\gamma_k^2 \right\} \\ &\leq (L^2 \gamma^3/3) \left\{ (1 - \kappa \gamma)^n (\ell - n - 1) \|x - x^\star\|^2 + 2\kappa^{-1} \gamma d(\ell - n - 1) \right\} + d\gamma^2 (\ell - n - 1) \\ &\leq (L^2 \gamma^3/3) \left\{ (1 + \gamma^{-1}) (1 - \kappa \gamma)^{\ell - \lceil \gamma^{-1} \rceil} \|x - x^\star\|^2 + 2(1 + \gamma)\kappa^{-1}d \right\} + d(1 + \gamma) \,. \end{split}$$

Combining this inequality and (S7) in the bound provided by Theorem 13 shows (28).

2.3 Proof of (58)

First for all $n \ge 1$ and $x \in \mathbb{R}^d$, we have

$$\vartheta_{n,0}^{(2)}(x) \le \gamma^2 \mathsf{E}_1(\gamma, d) + \gamma^3 \mathsf{E}_2(\gamma) \sum_{i=1}^n \prod_{k=i+1}^n (1 - \kappa \gamma_k/2) \delta_{i,n,0}(x) ,$$

where

$$\mathsf{E}_{1}(\gamma,d) = 2d\kappa^{-1} \left\{ 2L^{2} + 4\kappa^{-1}(d\tilde{L}^{2}/12 + \gamma L^{4}/4) + \gamma^{2}L^{4}/6 \right\} \ , \\ \mathsf{E}_{2}(\gamma) = L^{4}(4\kappa^{-1}/3 + \gamma) \ .$$

$$\begin{split} \mathsf{E}_{3}(\gamma, d, x) &= (\ell - \gamma^{-1}) \mathrm{e}^{-m\gamma(\ell - 2\gamma^{-1} - 1)} \|x - x^{\star}\|^{2} + 2d/(\kappa\gamma m) \\ \mathsf{E}_{4}(\gamma, d, x) &= \mathrm{e}^{-m\gamma(\ell - 2\gamma^{-1} - 1)} \|x - x^{\star}\|^{2} + 2\kappa^{-1}d + d/m \;. \end{split}$$

By (S6), we get for all $n \ge 1$ and $x \in \mathbb{R}^d$,

$$\vartheta_{n,0}^{(2)}(x) \le \gamma^2 \mathsf{E}_1(\gamma, d) + \gamma^3 \mathsf{E}_2(\gamma) \left\{ n \mathrm{e}^{-m\gamma(n-1)} \| x - x^\star \|^2 + 2d(\kappa \gamma m)^{-1} \right\} .$$
(S8)

On the other hand, for all $x \in \mathbb{R}^d$, $\ell, n \in \mathbb{N}$, $n \ge 1$, $\ell > n$ we have using that $\kappa \ge 2m$ and for all $t \ge 0$, $1 - t \le e^{-t}$,

$$\vartheta_{n,\ell}^{(2)}(x) \leq \gamma^{3} n \mathsf{E}_{1}(\gamma) + \gamma^{3} n \mathsf{E}_{2}(\gamma) \left\{ \mathrm{e}^{-m\gamma(n-1)} \varrho_{n,\ell}(x) + d/m \right\} \\
\leq \gamma^{3} n \mathsf{E}_{1}(\gamma) + \gamma^{3} n \mathsf{E}_{2}(\gamma) \left\{ \mathrm{e}^{-m\gamma(n-1)} \left((1 - \kappa\gamma)^{\ell-n} \|x - x^{\star}\|^{2} + 2\kappa^{-1}d \right) + d/m \right\} \\
\leq \gamma^{3} n \mathsf{E}_{1}(\gamma) + \gamma^{3} n \mathsf{E}_{2}(\gamma) \left\{ \mathrm{e}^{-m\gamma(\ell-1)} \|x - x^{\star}\|^{2} + 2\kappa^{-1}d + d/m \right\} .$$
(S9)

Finally, for all $\ell \in \mathbb{N}^*$ and $x \in \mathbb{R}^d$, we have

$$(\gamma^3 L^2/3)\varrho_{1,\ell-1}(x) + d\gamma^2 \le (\gamma^3 L^2/3) \left\{ (1 - \kappa\gamma)^{\ell-1} \|x - x^\star\|^2 + 2d\kappa^{-1} \right\} + d\gamma^2 .$$
(S10)

Combining (S8), (S9) and (S10) in the bound given by Lemma 22, and using that $\gamma^{-1} \leq 2^{n(\gamma)} \leq 2\gamma^{-1}$ we have for all $\ell \in \mathbb{N}^*$, $\ell > 2^{n(\gamma)}$,

$$\begin{split} \|\delta_{x}P_{\ell\gamma} - \delta_{x}R_{\gamma}^{\ell}\|_{\mathrm{TV}} &\leq 2^{-3/2}L\left[(\gamma^{3}L^{2}/3)\left\{(1-\kappa\gamma)^{\ell-1}\|x-x^{\star}\|^{2} + 2d\kappa^{-1}\right\} + d\gamma^{2}\right]^{1/2} \\ &+ (4\pi)^{-1/2}\left[\gamma^{2}\mathsf{E}_{1}(\gamma) + \gamma^{3}\mathsf{E}_{2}(\gamma)\left\{(\ell-\gamma^{-1})\mathrm{e}^{-m\gamma(\ell-2\gamma^{-1}-1)}\|x-x^{\star}\|^{2} + 2d(\kappa\gamma m)^{-1}\right\}\right]^{1/2} \\ &+ \sum_{k=1}^{\mathsf{n}(\gamma)}\left[\frac{\gamma^{3}2^{k-1}\mathsf{E}_{1}(\gamma) + \gamma^{3}2^{k-1}\mathsf{E}_{2}(\gamma)\left\{\mathrm{e}^{-m\gamma(\ell-2^{k}-1)}\|x-x^{\star}\|^{2} + 2\kappa^{-1}d + d/m\right\}}{\pi^{2^{k+1}\gamma}}\right]^{1/2} \\ &\leq 2^{-3/2}L\left\{(\gamma^{3}L^{2}/3)\left\{(1-\kappa\gamma)^{\ell-1}\|x-x^{\star}\|^{2} + 2d\kappa^{-1}\right\} + d\gamma^{2}\right\}^{1/2} \\ &+ (4\pi)^{-1/2}\left[\gamma^{2}\mathsf{E}_{1}(\gamma) + \gamma^{3}\mathsf{E}_{2}(\gamma)\left\{(\ell-\gamma^{-1})\mathrm{e}^{-m\gamma(\ell-2\gamma^{-1}-1)}\|x-x^{\star}\|^{2} + 2d(\kappa\gamma m)^{-1}\right\}\right]^{1/2} \\ &+ (4\pi)^{-1/2}\mathsf{n}(\gamma)\left[\gamma^{2}\mathsf{E}_{1}(\gamma) + \gamma^{2}\mathsf{E}_{2}(\gamma)\left\{\mathrm{e}^{-m\gamma(\ell-2\gamma^{-1}-1)}\|x-x^{\star}\|^{2} + 2\kappa^{-1}d + d/m\right\}\right]^{1/2} \,. \end{split}$$

3 Proof of Theorem 18

Let $N \ge 0$, $n \ge 1$, $x \in \mathbb{R}^d$ and $f \in \mathbb{F}_b(\mathbb{R}^d)$. The main idea of the proof is to consider the decomposition (S31) again but combined with the decomposition of $\Phi_{n,k+1}^N$, for $k \in \{N, \ldots, N+n-1\}$, into a Lipschitz component and a bounded measurable component as it is done in the proof of (39). Let $k \in \{N, \ldots, N+n-1\}$. By definition (33), $\Phi_{n,k}^N = \omega_{k+1,n}^N f + \tilde{\Phi}_{n,k}^N$, where $\tilde{\Phi}_{n,k}^N = \sum_{i=k+2}^{N+n} \omega_{i,n}^N Q_{\gamma}^{k+2,i} f$. Using that f is bounded, we get for all $y \in \mathbb{R}^d$ and $\lambda > 0$,

$$R_{\gamma_{k+1}} \left\{ e^{\lambda \{\Phi_{n,k+1}^{N}(\cdot) - R_{\gamma_{k+1}} \Phi_{n,k+1}^{N}(y)\}} \right\} (y) \\ \leq e^{\lambda \operatorname{osc}(f) \gamma_{k+2}(\Gamma_{N+2,N+n+1})^{-2}} R_{\gamma_{k+1}} \left\{ e^{\lambda \{\tilde{\Phi}_{n,k+1}^{N}(\cdot) - R_{\gamma_{k+1}} \tilde{\Phi}_{n,k+1}^{N}(y)\}} \right\} (y)$$

By (38) and (S32), we obtain for all $y \in \mathbb{R}^d$ and $\lambda > 0$,

$$R_{\gamma_{k+1}} \left\{ e^{\lambda \{\Phi_{n,k+1}^{N}(\cdot) - R_{\gamma_{k+1}} \Phi_{n,k+1}^{N}(y)\}} \right\} (y)$$

$$\leq \exp\left(\lambda \operatorname{osc}(f) \gamma_{k+2} (\Gamma_{N+2,N+n+1})^{-2} + (\lambda \operatorname{osc}(f))^{2} \gamma_{k+1} \left(\sum_{i=k+2}^{N+n} \omega_{i,n}^{N} / (\pi \Lambda_{k+2,i})^{1/2}\right)^{2} \right).$$
(S11)

It remains to control the Laplace transform of Ψ_n^N under $\delta_x Q_{\gamma}^N$. For this, note that by (40) Ψ_n^N is a Lipschitz function. Therefore using Lemma S6, we get an analogue of Corollary 7: for all $y \in \mathbb{R}^d$ and $\lambda > 0$,

$$\mathbb{E}_{y}\left[\mathrm{e}^{\lambda\{\Psi_{n}^{N}(X_{n})-\mathbb{E}_{x}\left[\Psi_{n}^{N}(X_{n})\right]\}}\right] \leq \exp\left(\kappa^{-1}\lambda^{2}\mathrm{osc}(f)^{2}\left(\sum_{i=N+1}^{N+n}\omega_{i,n}^{N}/(\pi\Lambda_{N+1,i})^{1/2}\right)^{2}\right),\tag{S12}$$

Combining (S11) and (S12) in (S31), the Laplace transform of $\hat{\pi}_n^N(f)$ can be explicitly bounded: for all $\lambda > 0$,

$$\mathbb{E}_x\left[\mathrm{e}^{\lambda\{\hat{\pi}_n^N(f)-\mathbb{E}_x[\hat{\pi}_n^N(f)]\}}\right] \le \mathrm{e}^{\lambda\operatorname{osc}(f)(\Gamma_{N+2,N+n+1})^{-1}+(\lambda\operatorname{osc}(f))^2 u_{N,n}^{(5)}(\gamma)}$$

Using this result and the Markov inequality, for all $\lambda > 0$, we have:

$$\mathbb{P}_x\left[\hat{\pi}_n^N(f) \ge \mathbb{E}_x[\hat{\pi}_n^N(f)] + r\right]$$

$$\le \exp\left(-\lambda r + \lambda \operatorname{osc}(f)(\Gamma_{N+2,N+n+1})^{-1} + (\lambda \operatorname{osc}(f))^2 u_{N,n}^{(5)}(\gamma)\right)$$

Then the proof follows from taking

$$\lambda = (r - \operatorname{osc}(f)(\Gamma_{N+2,N+n+1})^{-1}) / (2\operatorname{osc}(f)^2 u_{N,n}^{(5)}(\gamma))$$

4 Proof of Section 8

4.1 Coupling

Consider the Markov kernel K_k on $(\mathbb{R}^d \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^d))$ given for all $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ and $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^d)$ by

$$\begin{aligned} \mathsf{K}_{k}((x,y),\mathbf{A}) &= \frac{\mathbbm{1}_{\mathsf{D}}(h_{k}(x),h_{k}(y))}{(2\pi\sigma_{k}^{2})^{d/2}} \int_{\mathbb{R}^{d}} \mathbbm{1}_{\mathbf{A}}(\tilde{x},\tilde{x})\mathrm{e}^{-\|\tau_{k}(\tilde{x},x)\|^{2}/(2\sigma_{k}^{2})}\mathrm{d}\tilde{x} \\ &+ \frac{\mathbbm{1}_{\mathsf{D}^{c}}(h_{k}(x),h_{k}(y))}{(2\pi\sigma_{k}^{2})^{d/2}} \left[\int_{\mathbb{R}^{d}} \mathbbm{1}_{\mathbf{A}}(\tilde{x},\tilde{x})p_{k}(x,y,\tau_{k}(\tilde{x},x))\mathrm{e}^{-\|\tau_{k}(\tilde{x},x)\|^{2}/(2\sigma_{k}^{2})}\mathrm{d}\tilde{x} \\ &+ \int_{\mathbb{R}^{d}} \mathbbm{1}_{\mathbf{A}}(\tilde{x},\mathsf{F}_{k}(x,y,\tau_{k}(\tilde{x},x))) \left\{ 1 - p_{k}(x,y,\tau_{k}(\tilde{x},x)) \right\} \mathrm{e}^{-\|\tau_{k}(\tilde{x},x)\|^{2}/(2\sigma_{k}^{2})}\mathrm{d}\tilde{x} \right] \,, \end{aligned}$$
(S13)

where for all $\tilde{x} \in \mathbb{R}^d$, $\tau_k(\tilde{x}, x) = \tilde{x} - h_k(x)$ and $\mathsf{D} = \left\{ (\tilde{x}, \tilde{y}) \in \mathbb{R}^d \times \mathbb{R}^d \, \big| \, \tilde{x} = \tilde{y} \right\}$ and

$$p_k(x, y, z) = 1 \land \ \alpha_k(x, y, z)$$

By construction, $\mathsf{K}_k((x,y), \cdot \times \mathbb{R}^d) = \mathsf{P}_k(x, \cdot)$ for all $x, y \in \mathbb{R}^d$ and $\mathsf{K}_k((x,y), \mathbb{R}^d \times \cdot) = \mathsf{P}_k(y, \cdot)$ for all (x, y) such that $h_k(x) = h_k(y)$. In addition, for all $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$, we have for all $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ such that $h_k(x) \neq h_k(y)$, we get

$$\begin{aligned} \mathsf{K}_{k}((x,y), \mathbb{R}^{d} \times \mathsf{A}) &= \frac{1}{(2\pi\sigma_{k}^{2})^{d/2}} \int_{\mathbb{R}^{d}} \mathbb{1}_{\mathsf{A}}(\tilde{x}) p_{k}(x,y,\tilde{x}-h_{k}(x)) \mathrm{e}^{-\|\tilde{x}-h_{k}(x)\|^{2}/(2\sigma_{k}^{2})} \mathrm{d}\tilde{x} \\ &+ \frac{1}{(2\pi\sigma_{k}^{2})^{d/2}} \int_{\mathbb{R}^{d}} \mathbb{1}_{\mathsf{A}}(\mathrm{F}_{k}(x,y,\tilde{x}-h_{k}(x))) \left\{ 1 - p_{k}(x,y,\tilde{x}-h_{k}(x)) \right\} \mathrm{e}^{-\|\tilde{x}-h_{k}(x)\|^{2}/(2\sigma_{k}^{2})} \mathrm{d}\tilde{x} . \end{aligned}$$
(S14)

Since $(\text{Id} - 2e_k(x, y)e_k(x, y)^T)$ is an orthogonal matrix, making the change of variable $\tilde{y} = F_k(x, y, \tilde{x} - h_k(x))$ and using that

$$\langle \mathbf{e}_k(x,y), h_k(y) - \tilde{y} \rangle = \langle \mathbf{e}_k(x,y), \tilde{x} - h_k(x) \rangle$$

we get that

$$\int_{\mathbb{R}^d} \mathbb{1}_{\mathbf{A}}(\mathbf{F}_k(x, y, \tilde{x} - h_k(x))) \left\{ 1 - p_k(x, y, \tilde{x} - h_k(x)) \right\} e^{-\|\tilde{x} - h_k(x)\|^2 / (2\sigma_k^2)} d\tilde{x}$$
$$= \int_{\mathbb{R}^d} \mathbb{1}_{\mathbf{A}}(\tilde{y}) \left\{ 1 - p_k(x, y, h_k(y) - \tilde{y}) \right\} e^{-\|\tilde{y} - h_k(y)\|^2 / (2\sigma_k^2)} d\tilde{y} .$$
(S15)

By definition of $\alpha_k(x, y, (64))$, we have for all $\tilde{x} \in \mathbb{R}^d$,

$$\alpha_k\left(x, y, \tilde{x} - h_k(x)\right) = \frac{\varphi_{\sigma_k^2}\left(\langle e_k(x, y), h_k(y) - \tilde{x} \rangle\right)}{\varphi_{\sigma_k^2}\left(\|E_k(x, y)\| - \langle e_k(x, y), h_k(y) - \tilde{x} \rangle\right)} = \frac{1}{\alpha_k\left(x, y, h_k(y) - \tilde{x}\right)}.$$
(S16)

In addition using that

$$\|\tilde{x} - h_k(x)\|^2 = \|\tilde{x} - h_k(y)\|^2 - 2\langle h_k(y) - \tilde{x}, \mathbf{E}_k(x, y)\rangle + \|\mathbf{E}_k(x, y)\|^2 ,$$

we obtain

$$p_k(x, y, \tilde{x} - h_k(x)) \mathrm{e}^{-\|\tilde{x} - h_k(x)\|^2 / (2\sigma_k^2)} = p_k(x, y, h_k(y) - \tilde{x}) \mathrm{e}^{-\|\tilde{x} - h_k(y)\|^2 / (2\sigma_k^2)} .$$
(S17)

Plugging (S15) and (S17) into (65) implies that $\mathsf{K}_k((x,y), \mathbb{R}^d \times \mathsf{A}) = \mathsf{P}_k(y, \mathsf{A}).$

4.2 Proof of Lemma 24

Let ς , a > 0 and $t \in \mathbb{R}_+$. Let us denote by I the integral on the left hand side in the expression above. Then,

$$I = \int_{-\infty}^{t/2} \left\{ \varphi_{\varsigma^2}(y) - \varphi_{\varsigma^2}(t-y) \right\} \left\{ 1 - 2\Phi\left(\frac{2y-t}{2a}\right) \right\} dy$$

$$= \int_{-\infty}^{t/2} \varphi_{\varsigma^2}(y) \left\{ 1 - 2\Phi\left(\frac{2y-t}{2a}\right) \right\} dy$$

$$- \int_{-\infty}^{-t/2} \varphi_{\varsigma^2}(y) \left\{ 1 - 2\Phi\left(\frac{t+2y}{2a}\right) \right\} dy , \qquad (S18)$$

Now to simplify the proof, we give a probabilistic interpretation of this two integrals. Let X and Y be two real Gaussian random variables with zero mean and variance a^2 and ς^2 respectively. Since for all $u \in \mathbb{R}_+$, $1 - 2\Phi(-u/(2a)) = \mathbb{P}[|X| \le u/2]$, we have by (S18)

$$\begin{split} I &= \mathbb{P} \left(\, Y \leq t/2, X + \, Y \leq t/2, \, Y - X \leq t/2 \right) \\ &\quad - \, \mathbb{P} \left(\, Y \geq t/2, \, X + \, Y \geq t/2, \, Y - X \geq t/2 \right) \, . \end{split}$$

Using that Y and -Y have the same law in the second term, we get

$$I = \mathbb{P} \left(Y \le t/2, X + Y \le t/2, Y - X \le t/2 \right) - \mathbb{P} \left(Y \le -t/2, X - Y \ge t/2, Y + X \le -t/2 \right) = I_1 + I_2 , \qquad (S19)$$

where

$$I_{1} = \mathbb{P}\left(Y \le t/2, X + Y \le t/2, Y - X \le t/2, X \ge 0\right) - \mathbb{P}\left(Y \le -t/2, X - Y \ge t/2, Y + X \le -t/2, X \ge 0\right) = \mathbb{P}\left(|X + Y| \le t/2, X \ge 0\right) ,$$
(S20)

and

$$I_2 = \mathbb{P}\left(Y \le t/2, X + Y \le t/2, Y - X \le t/2, X \le 0\right) \\ - \mathbb{P}\left(Y \le -t/2, X - Y \ge t/2, Y + X \le -t/2, X \le 0\right) .$$

Using again that Y and -Y have the same law in the two terms we have

$$I_{2} = \mathbb{P}\left(Y \ge -t/2, X - Y \le t/2, Y + X \ge -t/2, X \le 0\right) - \mathbb{P}\left(Y \ge t/2, X + Y \ge t/2, X - Y \le -t/2, X \le 0\right) = \mathbb{P}\left(|X + Y| \le t/2, X \le 0\right) .$$
(S21)

Combining (S20), (S21) in (S19), we have $I = \mathbb{P}(|X + Y| \le t/2)$. The proof follows from the fact that X + Y is a real Gaussian random variable with mean zero and variance $a^2 + \varsigma^2$, since X and Y are independent.

5 Mean square error and concentration for Lipschitz functions

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz function and $(X_k)_{k\geq 0}$ the Euler discretization of the Langevin diffusion. In this section we study the approximation of $\int_{\mathbb{R}^d} f(y)\pi(\mathrm{d} y)$ by the weighted average estimator

$$\hat{\pi}_{n}^{N}(f) = \sum_{k=N+1}^{N+n} \omega_{k,n}^{N} f(X_{k}) , \quad \omega_{k,n}^{N} = \gamma_{k+1} \Gamma_{N+2,N+n+1}^{-1} .$$
(S22)

where $N \ge 0$ is the length of the burn-in period, $n \ge 1$ is the number of samples, and for $n, p \in \mathbb{N}$, $\Gamma_{n,p}$ is given by (5). In all this section, \mathbb{P}_x and \mathbb{E}_x denote the probability and the expectation respectively, induced on $((\mathbb{R}^d)^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^d)^{\mathbb{N}})$ by the Markov chain $(X_n)_{n\ge 0}$ started at $x \in \mathbb{R}^d$. We first compute an explicit bound for the Mean Squared Error (MSE) of this estimator defined by:

$$MSE_{f}^{N,n} = \mathbb{E}_{x} \left[\left| \hat{\pi}_{n}^{N}(f) - \pi(f) \right|^{2} \right] = \left\{ \mathbb{E}_{x} [\hat{\pi}_{n}^{N}(f)] - \pi(f) \right\}^{2} + Var_{x} \left\{ \hat{\pi}_{n}^{N}(f) \right\} .$$

We first bound the bias. For all $k \in \{N+1, \ldots, N+n\}$, let ξ_k be the optimal transference plan between $\delta_x Q_{\gamma}^k$ and π for W_2 , *i.e.* $W_2^2(\delta_x Q_{\gamma}^k, \pi) = \int_{\mathbb{R}^d \times \mathbb{R}^d} ||x - y||^2 d\xi_k(x, y)$. Then by the Jensen inequality and because f is Lipschitz, we have:

$$\left\{ \mathbb{E}_{x}[\hat{\pi}_{n}^{N}(f)] - \pi(f) \right\}^{2} = \left(\sum_{k=N+1}^{N+n} \omega_{k,n}^{N} \int_{\mathbb{R}^{d} \times \mathbb{R}^{d}} \{f(z) - f(y)\} \xi_{k}(\mathrm{d}z,\mathrm{d}y) \right)^{2} \\ \leq \|f\|_{\mathrm{Lip}}^{2} \sum_{k=N+1}^{N+n} \omega_{k,n}^{N} \int_{\mathbb{R}^{d} \times \mathbb{R}^{d}} \|z - y\|^{2} \xi_{k}(\mathrm{d}z,\mathrm{d}y) \,.$$

Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 1/(m+L)$ and recall that x^* is the unique minimizer of U. Using Theorem 5 and Theorem 9, we end up with the following bound:

$$\left\{\mathbb{E}_{x}[\hat{\pi}_{n}^{N}(f)] - \pi(f)\right\}^{2} \leq \|f\|_{\operatorname{Lip}}^{2} \sum_{k=N+1}^{N+n} \omega_{k,n}^{N} \left\{ (\|x - x^{\star}\|^{2} + d/m)u_{k}^{(1)}(\gamma) + w_{k}(\gamma) \right\} ,$$
(S23)

where $u_n^{(1)}(\gamma)$ is given in (10) and $w_n(\gamma)$ is equal to $u_n^{(2)}(\gamma)$ defined by (11) if **H1-H2** hold, and to $u_n^{(3)}(\gamma)$, defined by (17), if **H1-H2** and **H3** hold.

Consider now the variance term. To control this term, we adapt the proof of [3, Theorem 2] for homogeneous Markov chain to our inhomogeneous setting, and we have:

Theorem S1. Assume H_1 and H_2 . Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L)$. Then for all $N \geq 0$, $n \geq 1$ and Lipschitz functions $f : \mathbb{R}^d \to \mathbb{R}$, we get $\operatorname{Var}_x\{\hat{\pi}_n^N(f)\} \leq 8\kappa^{-2} \|f\|_{\operatorname{Lip}}^2 \Gamma_{N+2,N+n+1}^{-1} v_{N,n}(\gamma)$, where

$$v_{N,n}(\gamma) \stackrel{\text{\tiny def}}{=} \left\{ 1 + \Gamma_{N+2,N+n+1}^{-1}(\kappa^{-1} + 2/(m+L)) \right\} .$$
(S24)

Proof. Our main tool is the Gaussian Poincaré inequality [1, Theorem 3.20] which can be applied to $R_{\gamma}(y, \cdot)$ defined by (6), noticing that $R_{\gamma}(y, \cdot)$ is a Gaussian distribution with mean $y - \gamma \nabla U(y)$ and covariance matrix $2\gamma I_d$: for all Lipschitz function $g : \mathbb{R}^d \to \mathbb{R}$

$$R_{\gamma} \{g(\cdot) - R_{\gamma}g(y)\}^2(y) \le 2\gamma \|g\|_{\text{Lip}}^2$$
 (S25)

To go further, we decompose $\hat{\pi}_n^N(f) - \mathbb{E}_x[\hat{\pi}_n^N(f)]$ as the sum of martingale increments, w.r.t. $(\mathcal{G}_n)_{n\geq 0}$, the natural filtration associated with Euler approximation $(X_n)_{n\geq 0}$, and we get

$$\operatorname{Var}_{x}\left\{\hat{\pi}_{n}^{N}(f)\right\} = \sum_{k=N}^{N+n-1} \mathbb{E}_{x}\left[\left(\mathbb{E}_{x}^{\mathcal{G}_{k+1}}\left[\hat{\pi}_{n}^{N}(f)\right] - \mathbb{E}_{x}^{\mathcal{G}_{k}}\left[\hat{\pi}_{n}^{N}(f)\right]\right)^{2}\right] \\ + \mathbb{E}_{x}\left[\left(\mathbb{E}_{x}^{\mathcal{G}_{N}}\left[\hat{\pi}_{n}^{N}(f)\right] - \mathbb{E}_{x}\left[\hat{\pi}_{n}^{N}(f)\right]\right)^{2}\right]. \quad (S26)$$

Since $\hat{\pi}_n^N(f)$ is an additive functional, the martingale increment $\mathbb{E}_x^{\mathcal{G}_{k+1}} \left[\hat{\pi}_n^N(f) \right] - \mathbb{E}_x^{\mathcal{G}_k} \left[\hat{\pi}_n^N(f) \right]$ has a simple expression. For $k = N + n - 1, \dots, N + 1$, define backward in time the function

$$\Phi_{n,k}^{N} : x_{k} \mapsto \omega_{k,n}^{N} f(x_{k}) + R_{\gamma_{k+1}} \Phi_{n,k+1}^{N}(x_{k}) , \qquad (S27)$$

where $\Phi_{n,N+n}^N : x_{N+n} \mapsto \Phi_{n,N+n}^N(x_{N+n}) = \omega_{N+n,n}^N f(x_{N+n})$. Denote finally

$$\Psi_n^N : x_N \mapsto R_{\gamma_{N+1}} \Phi_{n,N+1}^N(x_N) .$$
(S28)

Note that for $k \in \{N, \ldots, N + n - 1\}$, by the Markov property,

$$\Phi_{n,k+1}^{N}(X_{k+1}) - R_{\gamma_{k+1}}\Phi_{n,k+1}^{N}(X_{k}) = \mathbb{E}_{x}^{\mathcal{G}_{k+1}}\left[\hat{\pi}_{n}^{N}(f)\right] - \mathbb{E}_{x}^{\mathcal{G}_{k}}\left[\hat{\pi}_{n}^{N}(f)\right] , \qquad (S29)$$

and $\Psi_n^N(X_N) = \mathbb{E}_x^{\mathcal{G}_N}[\hat{\pi}_n^N(f)]$. With these notations, (S26) may be equivalently expressed as

$$\operatorname{Var}_{x}\left\{\hat{\pi}_{n}^{N}(f)\right\} = \sum_{k=N}^{N+n-1} \mathbb{E}_{x}\left[R_{\gamma_{k+1}}\left\{\Phi_{n,k+1}^{N}(\cdot) - R_{\gamma_{k+1}}\Phi_{n,k+1}^{N}(X_{k})\right\}^{2}(X_{k})\right] + \operatorname{Var}_{x}\left\{\Psi_{n}^{N}(X_{N})\right\} .$$
(S30)

Now for k = N + n - 1, ..., N, we will use the Gaussian Poincaré inequality (S25) to the sequence of function $\Phi_{n,k+1}^N$ to prove that $x \mapsto R_{\gamma_{k+1}} \{\Phi_{n,k+1}^N(\cdot) - R_{\gamma_{k+1}} \Phi_{n,k+1}^N(x)\}^2(x)$ is uniformly bounded. It is required to bound the Lipschitz constant of $\Phi_{n,k}^N$.

Lemma S2. Assume H1 and H2. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L)$. Let $N \geq 0$ and $n \geq 1$. Then for all $y \in \mathbb{R}^d$, Lipschitz function f and $k \in \{N, \ldots, N+n-1\}$,

$$R_{\gamma_{k+1}} \left\{ \Phi_{n,k+1}^{N}(\cdot) - R_{\gamma_{k+1}} \Phi_{n,k+1}^{N}(y) \right\}^{2}(y) \le 8\gamma_{k+1} \|f\|_{\operatorname{Lip}}^{2} \left(\kappa \Gamma_{N+2,N+n+1}\right)^{-2},$$

where $\Phi_{n,k+1}^N$ is given by (S27).

Proof of Theorem S1. By (S27), $\|\Phi_{n,k}^N\|_{\text{Lip}} \leq \sum_{i=k+1}^{N+n} \omega_{i,n}^N \|Q_{\gamma}^{k+2,i}f\|_{\text{Lip}}$. Using Corollary 4, the bound $(1-t)^{1/2} \leq 1-t/2$ for $t \in [0,1]$ and the definition of $\omega_{i,n}^N$ given by (30), we have

$$\left\|\Phi_{n,k}^{N}\right\|_{\text{Lip}} \le \|f\|_{\text{Lip}} \sum_{i=k+1}^{N+n} \omega_{i,n}^{N} \prod_{j=k+2}^{i} (1-\kappa\gamma_{j}/2) \le 2 \|f\|_{\text{Lip}} (\kappa\Gamma_{N+2,N+n+1})^{-1}.$$

Finally, the proof follows from (S25).

Also to control the last term in right hand side of (S30), we need to control the variance of $\Psi_n^N(X_N)$ under $\delta_x Q_\gamma^N$. But similarly to the sequence of functions $\Phi_{n,k}^N, \Psi_n^N$ is Lipschitz by Corollary 4 by definition, see (S28). Therefore it suffices to find some bound for the variance of g under $\delta_y Q_\gamma^{n,p}$, for $g : \mathbb{R}^d \to \mathbb{R}$ a Lipschitz function, $y \in \mathbb{R}^d$ and $\gamma > 0$, which is done in Lemma 16.

Corollary 3. Assume **H1** and **H2**. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L)$. Then for all Lipschitz function f and $x \in \mathbb{R}^d$, $\operatorname{Var}_x\{\Psi_n^N(X_N)\} \leq 8\kappa^{-3} \|f\|_{\operatorname{Lip}}^2 \Gamma_{N+2,N+n+1}^{-2}$, where Ψ_n^N is given by (S28).

Proof. By (S28) and Corollary 4, Ψ_n^N is Lipschitz function with $\|\Psi_n^N\|_{\text{Lip}} \leq \sum_{i=N+1}^{N+n} \omega_{i,n}^N \|Q_{\gamma}^{N+1,i}f\|_{\text{Lip}}$. Using Corollary 4, the bound $(1-t)^{1/2} \leq 1-t/2$ for $t \in [0,1]$ and the definition of $\omega_{i,n}^N$ given by (30), we have

$$\left\|\Psi_{n}^{N}\right\|_{\text{Lip}} \leq \|f\|_{\text{Lip}} \sum_{i=N+1}^{N+n} \omega_{i,n}^{N} \prod_{j=N+2}^{i} (1-\kappa\gamma_{j}/2) \leq 2 \|f\|_{\text{Lip}} (\kappa\Gamma_{N+2,N+n+1})^{-1}$$

The proof follows from Lemma 16.

Plugging the bounds given by Lemma S2 and Corollary 3 in (S30), we have

$$\operatorname{Var}_{x}\left\{\hat{\pi}_{n}^{N}(f)\right\} \leq 8\kappa^{-2} \left\|f\right\|_{\operatorname{Lip}}^{2}\left\{\Gamma_{N+2,N+n+1}^{-2}\Gamma_{N+1,N+n}+\kappa^{-1}\Gamma_{N+2,N+n+1}^{-2}\right\} \\ \leq 8\kappa^{-2} \left\|f\right\|_{\operatorname{Lip}}^{2}\left\{\Gamma_{N+2,N+n+1}^{-1}+\Gamma_{N+2,N+n+1}^{-2}(\gamma_{N+1}+\kappa^{-1})\right\}.$$

Using that $\gamma_{N+1} \leq 2/(m+L)$ concludes the proof.

It is worth to observe that the bound for the variance is independent from the dimension.

We may now discuss the bounds on the MSE (obtained by combining the bounds for the squared bias (S23) and the variance Theorem S1) for step sizes given for $k \ge 1$ by $\gamma_k = \gamma_1 k^{-\alpha}$ where $\alpha \in [0, 1]$ and $\gamma_1 < 1/(m + L)$. Details of these calculations are postponed to Section 9.1-Section 9.2. The order of the bounds (up to numerical constants) of the MSE are summarized in Table 1 as a function of γ_1 , n and N. Note that in the infinite horizon setting, it is optimal to take $\alpha = 1/2$ under H1 and H2, and $\alpha = 1/3$ under H1, H2 and H3.

	Bound for the MSE		
$\alpha = 0$	$\gamma_1 + (\gamma_1 n)^{-1} \{1 + \exp(-\kappa \gamma_1 N/2)\}$		
$\alpha \in (0, 1/2)$	$\gamma_1 n^{-\alpha} + (\gamma_1 n^{1-\alpha})^{-1} \left\{ 1 + \exp(-\kappa \gamma_1 N^{1-\alpha} / (2(1-\alpha))) \right\}$		
$\alpha = 1/2$	$\gamma_1 \log(n) n^{-1/2} + (\gamma_1 n^{1/2})^{-1} \left\{ 1 + \exp(-\kappa \gamma_1 N^{1/2}/4) \right\}$		
$\alpha \in (1/2, 1)$	$n^{\alpha-1} \left[\gamma_1 + \gamma_1^{-1} \left\{ 1 + \exp(-\kappa \gamma_1 N^{1-\alpha} / (2(1-\alpha))) \right\} \right]$		
$\alpha = 1$	$\log(n)^{-1} \left\{ \gamma_1 + \gamma_1^{-1} (1 + N^{-\gamma_1 \kappa/2}) \right\}$		

Table 1: Bound for the MSE for $\gamma_k = \gamma_1 k^{-\alpha}$ for fixed γ_1 and N under H1 and H2

	Bound for the MSE
$\alpha = 0$	$\gamma_1^2 + (\gamma_1 n)^{-1} \{ 1 + \exp(-\kappa \gamma_1 N/2) \}$
$\alpha \in (0, 1/3)$	$\gamma_1^2 n^{-2\alpha} + (\gamma_1 n^{1-\alpha})^{-1} \{ 1 + \exp(-\kappa \gamma_1 N^{1-\alpha} / (2(1-\alpha))) \}$
$\alpha = 1/3$	$\gamma_1^2 \log(n) n^{-2/3} + (\gamma_1 n^{2/3})^{-1} \{1 + \exp(-\kappa \gamma_1 N^{1/2}/4)\}$
$\alpha \in (1/3, 1)$	$n^{\alpha-1} \left[\gamma_1^2 + \gamma_1^{-1} \{ 1 + \exp(-\kappa \gamma_1 N^{1-\alpha} / (2(1-\alpha))) \} \right]$
$\alpha = 1$	$\log(n)^{-1} \left\{ \gamma_1^2 + \gamma_1^{-1} (1 + N^{-\gamma_1 \kappa/2}) \right\}$

Table 2: Bound for the MSE for $\gamma_k = \gamma_1 k^{-\alpha}$ for fixed γ_1 and N under H1, H2 and H3

If the total number of iterations n + N is held fixed (fixed horizon setting), we may optimize the value of the step size γ_1 but also of the burn-in period N to minimize the upper bound on the MSE. The order (in n) for different values of $\alpha \in [0, 1]$ are summarized in Table 3 (we display the order in n but not the constants, which are quite involved).

Let us discuss first the bounds based on Theorem 5. This time for any $\alpha \in [0, 1/2)$, we can always achieve a MSE of order $n^{-1/2}$ by choosing appropriately γ_1 and N (for $\alpha = 1/2$ we have only $\log(n)n^{-1/2}$). For $\alpha \in (1/2, 1]$, the best strategy is to take N = 0and the largest possible value for $\gamma_1 = 1/(m+L)$, which leads to a MSE of order $n^{\alpha-1}$ for $\alpha \in (0, 1/2)$ and $\log(n)$ for $\alpha = 1$. We now discuss the bounds provided by Theorem 9. It appears that, for any $\alpha \in [0, 1/3)$, we can always achieved the order $n^{-2/3}$ by choosing appropriately γ_1 and N (for $\alpha = 1/3$ we have only $\log^{1/3}(n)n^{-2/3}$). The worst case is for $\alpha \in (1/3, 1]$, where in fact the best strategy is to take N = 0 and the largest possible value for $\gamma_1 = 1/(m + L)$.

We can also follow the proof of [3, Theorem 5] to establish an exponential deviation

	H1, $H2$ and $H3$	H 1, H 2 and H 3		
$\alpha = 0$	$n^{-1/2}$	$n^{-2/3}$		
$\alpha \in (0, 1/2)$	$n^{-1/2}$	$n^{-2/3}$		
$\alpha = 1/2$	$\log(n)n^{-1/2}$	$\log^{1/3}(n)n^{-2/3}$		
$\alpha \in (1/2, 1)$	$n^{\alpha-1}$	$n^{\alpha-1}$		
$\alpha = 1$	$\log(n)$	$\log(n)$		

Table 3: Optimal bound for the MSE by choosing γ_1

inequality for $\hat{\pi}_n^N(f) - \mathbb{E}_x[\hat{\pi}_n^N(f)]$ given by (30).

Theorem S4. Assume H_1 and H_2 . Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L)$. Then for all $N \geq 0$, $n \geq 1$, r > 0 and Lipschitz functions $f : \mathbb{R}^d \to \mathbb{R}$:

$$\mathbb{P}_x\left[\hat{\pi}_n^N(f) \ge \mathbb{E}_x[\hat{\pi}_n^N(f)] + r\right] \le \exp\left(-\frac{r^2 \kappa^2 \Gamma_{N+2,N+n+1}}{16 \left\|f\right\|_{\operatorname{Lip}}^2 v_{N,n}(\gamma)}\right) ,$$

where $v_{N,n}(\gamma)$ is defined by (S24).

Proof. Let $N \ge 0$, $n \ge 1$, $x \in \mathbb{R}^d$ and f be a Lipschitz function. To prove Theorem S4, we derive an upper bound of the Laplace transform of $\hat{\pi}_n^N(f) - \mathbb{E}_x[\hat{\pi}_n^N(f)]$. Consider the decomposition by martingale increments

$$\mathbb{E}_{x}\left[\mathrm{e}^{\lambda\{\hat{\pi}_{n}^{N}(f)-\mathbb{E}_{x}[\hat{\pi}_{n}^{N}(f)]\}}\right] = \mathbb{E}_{x}\left[\mathrm{e}^{\lambda\{\mathbb{E}_{x}^{\mathcal{G}_{N}}[\hat{\pi}_{n}^{N}(f)]-\mathbb{E}_{x}[\hat{\pi}_{n}^{N}(f)]\}+\sum_{k=N}^{N+n-1}\lambda\{\mathbb{E}_{x}^{\mathcal{G}_{k+1}}[\hat{\pi}_{n}^{N}(f)]-\mathbb{E}_{x}^{\mathcal{G}_{k}}[\hat{\pi}_{n}^{N}(f)]\}}\right]$$

Now using (S29) with the sequence of functions $(\Phi_{n,k}^N)$ and Ψ_n^N given by (S27) and (S28), respectively, we have by the Markov property

$$\mathbb{E}_{x} \left[e^{\lambda \{ \hat{\pi}_{n}^{N}(f) - \mathbb{E}_{x}[\hat{\pi}_{n}^{N}(f)] \}} \right] = \mathbb{E}_{x} \left[e^{\lambda \{ \Psi_{n}^{N}(X_{n}) - \mathbb{E}_{x}[\Psi_{n}^{N}(X_{n})] \}} \prod_{k=N}^{N+n-1} R_{\gamma_{k+1}} \left[e^{\lambda \{ \Phi_{n,k+1}^{N}(\cdot) - R_{\gamma_{k+1}} \Phi_{n,k+1}^{N}(X_{k}) \}} \right] (X_{k}) \right],$$
(S31)

where R_{γ} is given by (6) for $\gamma > 0$. We use the same strategy to get concentration inequalities than to bound the variance term in the previous section, replacing the Gaussian Poincaré inequality by the log-Sobolev inequality to get uniform bound on

$$R_{\gamma_{k+1}}\{\exp(\lambda\{\Phi_{n,k+1}^{N}(\cdot) - R_{\gamma_{k+1}}\Phi_{n,k+1}^{N}(X_{k})\})\}(X_{k})$$

w.r.t. X_k , for all $k \in \{N + 1, ..., N + n\}$. Indeed for all $x \in \mathbb{R}^d$ and $\gamma > 0$, recall that $R_{\gamma}(x, \cdot)$ is a Gaussian distribution with mean $x - \gamma \nabla U(x)$ and covariance matrix

 $2\gamma I_d$. The log-Sobolev inequality [1, Theorem 5.5] shows that for all Lipschitz function $g: \mathbb{R}^d \to \mathbb{R}, x \in \mathbb{R}^d, \gamma > 0$ and $\lambda > 0$,

$$\int R_{\gamma}(x, \mathrm{d}y) \left\{ \exp\left(\lambda \{g(y) - R_{\gamma}g(x)\}\right) \right\} \le \exp\left(\gamma \lambda^2 \left\|g\right\|_{\mathrm{Lip}}^2\right) \,. \tag{S32}$$

We deduced from this result, (S29) and Corollary 4, an equivalent of Lemma S2 for the Laplace transform of $\Phi_{n,k+1}^N$ under $\delta_y R_{\gamma_{k+1}}$ for $k \in \{N+1,\ldots,N+n\}$ and all $y \in \mathbb{R}^d$.

Corollary 5. Assume **H** 1 and **H** 2. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L)$. Let $N \geq 0$ and $n \geq 1$. Then for all $k \in \{N, \ldots, N+n-1\}$, $y \in \mathbb{R}^d$ and $\lambda > 0$,

$$R_{\gamma_{k+1}}\left\{e^{\lambda\{\Phi_{n,k+1}^{N}(\cdot)-R_{\gamma_{k+1}}\Phi_{n,k+1}^{N}(y)\}}\right\}(y) \le \exp\left(4\gamma_{k+1}\lambda^{2} \|f\|_{\operatorname{Lip}}^{2} (\kappa\Gamma_{N+2,N+n+1})^{-2}\right),$$

where $\Phi_{n,k}^N$ is given by (33).

It remains to control the Laplace transform of Ψ_n^N under $\delta_x Q_\gamma^N$, where $\delta_x Q_\gamma^N$ is defined by (7). For this, using again that by (S28) and Corollary 4, Ψ_n^N is a Lipschitz function, we iterate (S32) to get bounds on the Laplace transform of Lipschitz function g under $Q_\gamma^{n,\ell}(y,\cdot)$ for all $y \in \mathbb{R}^d$ and $n, \ell \geq 1$, since for all $n, \ell \geq 1$, $Q_\gamma^{n,\ell}g$ is a Lipschitz function by Corollary 4.

Lemma S6. Assume H1 and H2. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L)$. Let $g : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz function, then for all $n, p \geq 1$, $n \leq p, y \in \mathbb{R}^d$ and $\lambda > 0$:

$$Q_{\gamma}^{n,p}\left\{\exp\left(\lambda\{g(\cdot) - Q_{\gamma}^{n,p}g(y)\}\right)\right\}(y) \le \exp\left(\kappa^{-1}\lambda^2 \|g\|_{\operatorname{Lip}}^2\right) , \qquad (S33)$$

where $Q_{n,p}^{\gamma}$ is given by (7).

Proof. Let $(X_n)_{n\geq 0}$ the Euler approximation given by (2) and started at $y \in \mathbb{R}^d$. By decomposing $g(X_p) - \mathbb{E}_{y}^{\mathcal{G}_n}[g(X_p)] = \sum_{k=n+1}^p \{\mathbb{E}_y^{\mathcal{G}_k}[g(X_p)] - \mathbb{E}_y^{\mathcal{G}_{k-1}}[g(X_p)]\}$, and using $\mathbb{E}_y^{\mathcal{G}_k}[g(X_p)] = Q_{\gamma}^{k+1,p}g(X_k)$, we get

$$\mathbb{E}_{y}^{\mathcal{G}_{n}}\left[\exp\left(\lambda\left\{g(X_{p})-\mathbb{E}_{y}^{\mathcal{G}_{n}}\left[g(X_{p})\right]\right\}\right)\right]$$

$$=\mathbb{E}_{y}^{\mathcal{G}_{n}}\left[\prod_{k=n+1}^{p}\mathbb{E}_{y}^{\mathcal{G}_{k-1}}\left[\exp\left(\lambda\left\{\mathbb{E}_{y}^{\mathcal{G}_{k}}\left[g(X_{p})\right]-\mathbb{E}_{y}^{\mathcal{G}_{k-1}}\left[g(X_{p})\right]\right\}\right)\right]\right]$$

$$=\mathbb{E}_{y}^{\mathcal{G}_{n}}\left[\prod_{k=n+1}^{p}R_{\gamma_{k}}\exp\left(\lambda\left\{Q_{\gamma}^{k+1,p}g(\cdot)-R_{\gamma_{k}}Q_{\gamma}^{k+1,p}g(X_{k-1})\right\}\right)(X_{k-1})\right]$$

By the Gaussian log-Sobolev inequality (S32), we get

$$\mathbb{E}_{y}^{\mathcal{G}_{n}}\left[\exp\left(\lambda\left\{g(X_{p})-\mathbb{E}_{y}^{\mathcal{G}_{n}}\left[g(X_{p})\right]\right\}\right)\right] \leq \exp\left(\lambda^{2}\sum_{k=n+1}^{p}\gamma_{k}\left\|Q_{\gamma}^{k+1,p}g\right\|_{\mathrm{Lip}}^{2}\right) \,.$$

The proof follows from Corollary 4 and Lemma 7, using the bound $(1-t)^{1/2} \le 1-t/2$ for $t \in [0,1]$.

Combining this result and $\|\Psi_n^N\|_{\text{Lip}} \leq 2\kappa^{-1} \|f\|_{\text{Lip}} \Gamma_{N+2,N+n+1}^{-1}$ by Corollary 4, we get an analogue of Corollary 3 for the Laplace transform of Ψ_n^N :

Corollary 7. Assume **H**¹ and **H**². Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L)$. Let $N \geq 0$ and $n \geq 1$. Then for all $\lambda > 0$ and $x \in \mathbb{R}^d$,

$$\mathbb{E}_{x}\left[\mathrm{e}^{\lambda\{\Psi_{n}^{N}(X_{n})-\mathbb{E}_{x}\left[\Psi_{n}^{N}(X_{n})\right]\}}\right] \leq \exp\left(4\kappa^{-3}\lambda^{2}\left\|f\right\|_{\mathrm{Lip}}^{2}\Gamma_{N+2,N+n+1}^{-2}\right),$$

where Ψ_n^N is given by (S28).

The Laplace transform of $\hat{\pi}_n^N(f)$ can be explicitly bounded using Corollary 5 and Corollary 7 in (S31).

Proposition S8. Assume H1 and H2. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L)$. Then for all $N \geq 0$, $n \geq 1$, Lipschitz functions $f : \mathbb{R}^d \to \mathbb{R}$, $\lambda > 0$ and $x \in \mathbb{R}^d$:

$$\mathbb{E}_{x}\left[e^{\lambda\{\hat{\pi}_{n}^{N}(f)-\mathbb{E}_{x}[\hat{\pi}_{n}^{N}(f)]\}}\right] \leq \exp\left(4\kappa^{-2}\lambda^{2}\|f\|_{\operatorname{Lip}}^{2}\Gamma_{N+2,N+n+1}^{-1}u_{N,n}^{(3)}(\gamma)\right)$$

where $u_{N,n}^{(3)}(\gamma)$ is given by (S24).

Proof of Theorem S4. Using the Markov inequality and Proposition S8, for all $\lambda > 0$, we have:

$$\mathbb{P}_x\left[\hat{\pi}_n^N(f) \ge \mathbb{E}_x[\hat{\pi}_n^N(f)] + r\right] \le \exp\left(-\lambda r + 4\kappa^{-2}\lambda^2 \left\|f\right\|_{\operatorname{Lip}}^2 \Gamma_{N+2,N+n+1}^{-1} v_{N,n}(\gamma)\right) \ .$$

Then the result follows from taking $\lambda = (r\kappa^2 \Gamma_{N+2,N+n+1})/(8 \|f\|_{\text{Lip}}^2 v_{N,n}(\gamma)).$

If we apply this result to the sequence $(\gamma_k)_{k\geq 1}$ defined for all $k \geq 1$ by $\gamma_k = \gamma_1 k^{-\alpha}$, for $\alpha \in [0,1]$, we end up with a concentration of order $\exp(-Cr^2\gamma_1n^{1-\alpha})$ for $\alpha \in [0,1)$, for some constant $C \geq 0$ independent of γ_1 and n.

6 Discussion on Theorem 5

Note that

$$u_n^{(2)}(\gamma) \le \sum_{i=1}^n \left\{ \mathsf{A}_0 \gamma_i^2 + \mathsf{A}_1 \gamma_i^3 \right\} \prod_{k=i+1}^n (1 - \kappa \gamma_k/2) , \qquad (S34)$$

where κ is given by (3), and

$$A_0 = 2L^2 \kappa^{-1} d, \qquad (S35)$$

$$A_1 = 2L^2 \kappa^{-1} d(m+L)^{-1} + dL^4 (\kappa^{-1} + (m+L)^{-1}) (m^{-1} + 6^{-1} (m+L)^{-1}). \qquad (S36)$$

If $(\gamma_k)_{k\geq 1}$ is a constant step size $\gamma_k = \gamma$ for all $k \geq 1$, then a straightforward consequence of Theorem 5 and (S34) is the following result, which gives the minimal number of iterations n_{ε} and a step-size γ_{ε} to get $W_2(\delta_{x*}Q_{\gamma}^n, \pi)$ smaller than $\varepsilon > 0$.

Corollary 9 (of Theorem 5). Assume H1 and H2. Let x^* be the unique minimizer of U. Let $x \in \mathbb{R}^d$ and $\varepsilon > 0$. Set for all $k \in \mathbb{N}$, $\gamma_k = \gamma$ with

$$\gamma = \frac{-\mathsf{A}_0 + (\mathsf{A}_0^2 + \varepsilon^2 \kappa \mathsf{A}_1)^{1/2}}{2\mathsf{A}_1} \wedge (m+L)^{-1} , \qquad (S37)$$
$$n = \left\lceil \log^{-1}(1 - \kappa \gamma/2) \left\{ -\log(\varepsilon^2/2) + \log(2d/m) \right\} \right\rceil .$$

Then $W_2(\delta_{x^{\star}} R^n_{\gamma}, \pi) \leq \varepsilon.$

Note that if γ is given by (S37), and is different from 1/(m+L), then $\gamma \leq \varepsilon (4A_1\kappa^{-1})^{-1/2}$ and $2\kappa^{-1}(A_0\gamma + A_1\gamma^2) = \varepsilon^2/2$. Therefore,

$$\gamma \ge \left(\varepsilon^2 \kappa/4\right) \left\{ \mathsf{A}_0 + \varepsilon (\mathsf{A}_1/(4\kappa))^{1/2} \right\}^{-1} \,,$$

It is shown in [2, Corollary 1] that under **H2**, for constant step size for any $\varepsilon > 0$, we can choose γ and $n \ge 1$ such that if for all $k \ge 1$, $\gamma_k = \gamma$, then $\|\nu^* Q_n^{\gamma} - \pi\|_{\text{TV}} \le \varepsilon$ where ν^* is the Gaussian measure on \mathbb{R}^d with mean x^* and covariance matrix $L^{-1} \mathbf{I}_d$. We stress that the results in [2, corollary 1] hold only for a particular choice of the initial distribution ν^* , (which might seem a rather artificial assumption) whereas Theorem 5 holds for any initial distribution in $\mathcal{P}_2(\mathbb{R}^d)$.

We compare the optimal value of γ and n obtained from Corollary 9 with those given in [2, Corollary 1]. This comparison is summarized in Table 4 and Table 5; for simplicity, we provide only the dependencies of the optimal stepsize γ and minimal number of simulations n as a function of the dimension d, the precision ε and the constants m, L. It can be seen that the dependency on the dimension is significantly better than those in [2, Corollary 1].

Parameter	d	ε	L	m
Corollary 9	d^{-1}	ε^2	L^{-2}	m^2
[2, Corollary 1]	d^{-2}	ε^2	L^{-2}	m

Table 4: Dependencies of γ

Parameter	d	ε	L	m
Corollary 9	$d\log(d)$	$\varepsilon^{-2} \left \log(\varepsilon) \right $	L^2	$\left \log(m)\right m^{-3}$
[2, Corollary 1]	d^3	$\varepsilon^{-2} \left \log(\varepsilon) \right $	L^3	$\left \log(m)\right m^{-2}$

Table 5: Dependencies of n

6.1 Explicit bounds for $\gamma_k = \gamma_1 k^{\alpha}$ with $\alpha \in (0, 1]$

We give here a bound on the sequences $(u_n^{(1)}(\gamma))_{n\geq 0}$ and $(u_n^{(2)}(\gamma))_{n\geq 0}$ for $(\gamma_k)_{k\geq 1}$ defined by $\gamma_1 < 1/(m+L)$ and $\gamma_k = \gamma_1 k^{-\alpha}$ for $\alpha \in (0,1]$. Also for that purpose we introduce for $t \in \mathbb{R}^*_+$,

$$\boldsymbol{\psi}_{\beta}(t) = \begin{cases} (t^{\beta} - 1)/\beta & \text{for } \beta \neq 0\\ \log(t) & \text{for } \beta = 0 \end{cases}$$
(S38)

We easily get for $a \ge 0$ that for all $n, p \ge 1, n \le p$

$$\boldsymbol{\psi}_{1-a}(p+1) - \boldsymbol{\psi}_{1-a}(n) \le \sum_{k=n}^{p} k^{-a} \le \boldsymbol{\psi}_{1-a}(p) - \boldsymbol{\psi}_{1-a}(n) + 1 , \qquad (S39)$$

and for $a \in \mathbb{R}$

$$\sum_{k=n}^{p} k^{-a} \le \psi_{1-a}(p+1) - \psi_{1-a}(n) + 1.$$
 (S40)

1. For $\alpha = 1$, using that for all $t \in \mathbb{R}$, $(1 + t) \leq e^t$ and by (S39) and (S40), we have

$$u_n^{(1)}(\gamma) \le (n+1)^{-\kappa\gamma_1/2}$$
, $u_n^{(2)}(\gamma) \le (n+1)^{-\kappa\gamma_1/2} \sum_{j=0}^1 \mathsf{A}_j(\psi_{\kappa\gamma_1/2-1-j}(n+1)+1)$.

2. For $\alpha \in (0,1)$, by (S39) and Lemma 7 applied with $\ell = \lceil n/2 \rceil$, where $\lceil \cdot \rceil$ is the ceiling function, we have

$$u_{n}^{(1)}(\gamma) \leq \exp\left(-\kappa\gamma_{1}\psi_{1-\alpha}(n+1)/2\right)$$
$$u_{n}^{(2)}(\gamma) \leq \sum_{j=0}^{1} \mathsf{A}_{j} \left[2\kappa^{-1}\gamma_{1}^{j+1}(n/2)^{-\alpha(j+1)} + \gamma_{1}^{j+2} \left\{ \psi_{1-\alpha(j+2)}(\lceil n/2 \rceil) + 1 \right\} \times \exp\left(-(\kappa\gamma_{1}/2) \left\{ \psi_{1-\alpha}(n+1) - \psi_{1-\alpha}(\lceil n/2 \rceil) \right\} \right) \right].$$
(S41)

6.2 Optimal strategy with a fixed number of iterations

Corollary 10. Let $n \in \mathbb{N}^*$ be a fixed number of iteration. Assume H1, H2, and $(\gamma_k)_{k\geq 1}$ is a constant sequence, $\gamma_k = \gamma$ for all $k \geq 1$. Set

$$\gamma^{+} = 2(\kappa n)^{-1} \left[\log(\kappa n/2) + \log(2(\|x - x^{\star}\|^{2} + d/m)) - \log(2\kappa^{-1}\mathsf{A}_{0}) \right]$$

$$\gamma_{-} = 2(\kappa n)^{-1} \left[\log(\kappa n/2) + \log(2(\|x - x^{\star}\|^{2} + d/m)) - \log\left\{ 2\kappa^{-1}(\mathsf{A}_{0} + 2\mathsf{A}_{1}(m + L)^{-1}) \right\} \right] .$$

Assume $\gamma^+ \in (0, (m+L)^{-1})$. Then, the optimal choice of γ to minimize the bound on $W_2(\delta_x R^n_{\gamma}, \pi)$ given by Theorem 5 belongs to $[\gamma_-, \gamma^+]$. Moreover if $\gamma = \gamma_+$, then there exists $C \geq 0$ independent of the dimension such that the bound on $W_2^2(\delta_x R^n_{\gamma}, \pi)$ is equivalent to $Cdn^{-1}\log(n)$ as n goes to $+\infty$.

Similarly, we have the following result.

Corollary 11. Assume **H1** and **H2**. Let $(\gamma_k)_{k\geq 1}$ be the decreasing sequence, defined by $\gamma_k = \gamma_\alpha k^{-\alpha}$, with $\alpha \in (0, 1)$. Let $n \geq 1$ and set

$$\gamma_{\alpha} = 2(1-\alpha)\kappa^{-1}(2/n)^{1-\alpha}\log(\kappa n/(2(1-\alpha)))$$
.

Assume $\gamma_{\alpha} \in (0, (m+L)^{-1})$. Then there exists $C \geq 0$ independent of the dimension such that the bound on $W_2^2(\delta_x Q_{\gamma}^n, \pi)$ is equivalent to $Cdn^{-1}\log(n)$ as n goes to $+\infty$.

Proof. Follows from (S34), (S41) and the choice of γ_{α} .

7 Discussion on Theorem 9

Based on Theorem 9, we can follow the same discussion as for Theorem 5. Note that

$$u_n^{(3)}(\gamma) \le \sum_{i=1}^n \left\{ \mathsf{B}_0 \gamma_i^3 + \mathsf{B}_1 \gamma_i^4 \right\} \prod_{k=i+1}^n (1 - \kappa \gamma_k/2) , \qquad (S42)$$

where κ is given by (3), and

$$\mathsf{B}_{0} = d \left[2L^{2} + \kappa^{-1} \left\{ d\tilde{L}^{2}/3 + 4L^{4}/(3m) \right\} \right] , \qquad (S43)$$

$$\mathsf{B}_1 = d \left[\kappa^{-1} L^4 + L^4 / (6(m+L)) + m^{-1} \right] \,. \tag{S44}$$

The following result gives for a target precision $\varepsilon > 0$, the minimal number of iterations n_{ε} and a step-size γ_{ε} to get $W_2(\delta_{x^*}Q_{\gamma}^n, \pi)$ smaller than ε , when $(\gamma_k)_{k\geq 1}$ is a constant step size, $\gamma_k = \gamma_{\varepsilon}$ for all $k \geq 1$.

Corollary 12. Assume H_1 , H_2 and H_3 . Let x^* be the unique minimizer of U. Let $x \in \mathbb{R}^d$ and $\varepsilon > 0$. Set for all $k \in \mathbb{N}$, $\gamma_k = \gamma$ with

$$\gamma = \left[(\varepsilon/2)\kappa^{-1} \left\{ \mathsf{B}_0 + \mathsf{B}_1(m+L)^{-1} \right\}^{-1/2} \right] \wedge (1/(m+L)) ,$$

$$n = \left\lceil \log^{-1}(1-\kappa\gamma/2) \left\{ -\log(\varepsilon^2/2) + \log(2d/m) \right\} \right\rceil .$$

Then $W_2(\delta_{x^*} R^n_{\gamma}, \pi) \leq \varepsilon.$

We provide the dependencies of the optimal stepsize γ_{ε} and minimal number of simulations n_{ε} as a function of the dimension d, the precision ε and the constants m, L in Table 6 and Table 7.

Parameter	d	ε	L	m
Corollary 12	d^{-1}	ε	L^{-2}	m

Table 6: Dependencies of γ_{ε}

Parameter	d	ε	L	m
Corollary 12	$d\log(d)$	$\varepsilon^{-1} \left \log(\varepsilon) \right $	L^2	$\left \log(m)\right m^{-2}$

Table 7: Dependencies of n_{ε}

7.1 Explicit bounds for $\gamma_k = \gamma_1 k^{\alpha}$ with $\alpha \in (0, 1]$

We give here a bound on the sequence $(u_n^{(3)}(\gamma))_{n\geq 0}$ for $(\gamma_k)_{k\geq 1}$ defined by $\gamma_1 < 1/(m+L)$ and $\gamma_k = \gamma_1 k^{-\alpha}$ for $\alpha \in (0,1]$. Bounds for $(u_n^{(1)}(\gamma))_{n\geq 0}$ have already been given in Section 6.1. Recall that the function $\boldsymbol{\psi}$ is defined by (S38). For $\alpha \in (0,1]$, by (S39) and Lemma 7 applied with $\ell = \lceil n/2 \rceil$, where $\lceil \cdot \rceil$ is the ceiling function, we have

$$u_{n}^{(3)}(\gamma) \leq \sum_{j=1}^{2} \mathsf{B}_{j-1} \left[2\kappa^{-1} \gamma_{1}^{j+1} (n/2)^{-\alpha(j+1)} + \gamma_{1}^{j+2} \left(\psi_{1-\alpha(j+2)}(\lceil n/2 \rceil) + 1 \right) \right. \\ \left. \times \exp\left(-(\kappa \gamma_{1}/2) \left\{ \psi_{1-\alpha}(n+1) - \psi_{1-\alpha}(\lceil n/2 \rceil) \right\} \right) \right] .$$
(S45)

7.2 Optimal strategy with a fixed number of iterations

Corollary 13. Let $n \in \mathbb{N}^*$ be a fixed number of iteration. Assume H1, H2, H3 and $(\gamma_k)_{k\geq 1}$ is a constant sequence, $\gamma_k = \gamma^*$ for all $k \geq 1$, with

$$\gamma^{\star} = 4(\kappa n)^{-1} \left\{ \log(\kappa n/2) + \log(2(\|x - x^{\star}\|^2 + d/m)) \right\} .$$

Assume $\gamma^* \in (0, (m+L)^{-1})$. Then there exists $C \geq 0$ independent of the dimension such that the bound on $W_2^2(\delta_x R_{\gamma}^n, \pi)$ is equivalent to $Cd^2n^{-2}\log^2(n)$ as n goes to $+\infty$.

Similarly, we have the following result.

Corollary 14. Assume H1, H2 and H3. Let $(\gamma_k)_{k\geq 1}$ be the decreasing sequence, defined by $\gamma_k = \gamma_{\alpha} k^{-\alpha}$, with $\alpha \in (0, 1)$. Let $n \geq 1$ and set

$$\gamma_{\alpha} = 2(1-\alpha)\kappa^{-1}(2/n)^{1-\alpha}\log(\kappa n/(2(1-\alpha)))$$

Assume $\gamma_{\alpha} \in (0, (m+L)^{-1})$. Then there exists $C \geq 0$ independent of the dimension such that the bound on $W_2^2(\delta_x R_{\gamma}^n, \pi)$ is equivalent to $Cd^2n^{-2}\log^2(n)$ as n goes to $+\infty$.

Proof. Follows from (S42), (S45) and the choice of γ_{α} .

Note that in Corollary 13 and Corollary 14, we do not find the optimal convergence rates obtained for the sequence of step-sizes $\gamma_k = \gamma_1/k$, for $k \ge 1$ and $\gamma_1 > 0$, up to a logarithmic factor $\log(n)$. This most likely due to the fact that the bounds (for example (S45)) used to compute the optimal parameters γ^* and γ_{α} are not the most appropriate.

8 Generalization of Theorem 5

In this section, we weaken the assumption $\gamma_1 \leq 1/(m+L)$ of Theorem 5. We assume now:

G1. The sequence $(\gamma_k)_{k\geq 1}$ is non-increasing, and there exists $\rho > 0$ and n_1 such that $(1+\rho)\gamma_{n_1} \leq 2/(m+L)$.

Under G1, we denote by

$$n_0 = \min\left\{k \in \mathbb{N} \mid \gamma_k \le 2/(m+L)\right\}$$
(S46)

We first give an extension of Proposition 2. Denote in the sequel $(\cdot)_+ = \max(\cdot, 0)$. Recall that under **H2**, x^* is the unique minimizer of U, and κ is defined in (4)

Theorem S15. Assume H1, H2 and G1. Then for all $n, p \in \mathbb{N}^*$, $n \leq p$

$$\int_{\mathbb{R}^d} \|x - x^\star\|^2 \,\mu_0 Q_n^p(\mathrm{d} x) \le \mathsf{G}_{n,p}(\mu_0,\gamma) \;,$$

where

$$G_{n,p}(\mu_0,\gamma) = \exp\left(-\sum_{k=n}^p \gamma_k \kappa + \sum_{k=n}^{n_0-1} L^2 \gamma_k^2\right) \int_{\mathbb{R}^d} \|x - x^\star\|^2 \mu_0(\mathrm{d}x) + 2d\kappa^{-1} + 2d\left\{\prod_{k=n}^{n_0-1} (\gamma_{n_0-1}L^2)^{-1} \left(1 + L^2 \gamma_k^2\right)\right\} \exp\left(-\sum_{k=n}^p \kappa \gamma_k + \sum_{k=n}^{n_0-1} \gamma_k^2 m L\right) .$$
 (S47)

Proof. For any $\gamma > 0$, we have for all $x \in \mathbb{R}^d$:

$$\int_{\mathbb{R}^d} \|y - x^{\star}\|^2 R_{\gamma}(x, \mathrm{d}y) = \|x - \gamma \nabla U(x) - x^{\star}\|^2 + 2\gamma d \,.$$

Using that $\nabla U(x^*) = 0$, (3) and H1, we get from the previous inequality:

$$\begin{split} &\int_{\mathbb{R}^d} \|y - x^\star\|^2 R_{\gamma}(x, \mathrm{d}y) \\ &\leq (1 - \kappa \gamma) \|x - x^\star\|^2 + \gamma \left(\gamma - \frac{2}{m+L}\right) \|\nabla U(x) - \nabla U(x^\star)\|^2 + 2\gamma d \\ &\leq \eta(\gamma) \|x - x^\star\|^2 + 2\gamma d \;, \end{split}$$

where $\eta(\gamma) = (1 - \kappa \gamma + \gamma L(\gamma - 2/(m+L))_+)$. Denote for all $k \ge 1$, $\eta_k = \eta(\gamma_k)$. By a straightforward induction, we have by definition of Q_n^p for $p, n \in \mathbb{N}$, $p \le n$,

$$\int_{\mathbb{R}^d} \|x - x^\star\|^2 \,\mu_0 Q_n^p(\mathrm{d}x) \le \prod_{k=n}^p \eta_k \int_{\mathbb{R}^d} \|x - x^\star\| \,\mu_0(\mathrm{d}x) + (2d) \sum_{i=n}^p \prod_{k=i+1}^p \eta_k \gamma_i \,, \quad (S48)$$

with the convention that for $n, p \in \mathbb{N}$, n < p, $\prod_{p=1}^{n} = 1$. For the first term of the right hand side, we simply use the bound, for all $x \in \mathbb{R}$, $(1 + x) \leq e^x$, and we get by **G**1

$$\prod_{k=n}^{p} \eta_k \le \exp\left(-\sum_{k=n}^{p} \kappa \gamma_k + \sum_{k=n}^{n_0-1} L^2 \gamma_k^2\right) , \qquad (S49)$$

where n_0 is defined in (S46). Consider now the second term in the right hand side of (S48).

$$\sum_{i=n}^{p} \prod_{k=i+1}^{p} \eta_{k} \gamma_{i} \leq \sum_{i=n_{0}}^{p} \prod_{k=i+1}^{p} (1 - \kappa \gamma_{k}) \gamma_{i} + \sum_{i=n}^{n_{0}-1} \prod_{k=i+1}^{p} \eta_{k} \gamma_{i}$$
$$\leq \kappa^{-1} \sum_{i=n_{0}}^{p} \left\{ \prod_{k=i+1}^{p} (1 - \kappa \gamma_{k}) - \prod_{k=i}^{p} (1 - \kappa \gamma_{k}) \right\}$$
$$+ \left\{ \sum_{i=n}^{n_{0}-1} \prod_{k=i+1}^{n_{0}-1} (1 + L^{2} \gamma_{k}^{2}) \gamma_{i} \right\} \prod_{k=n_{0}}^{p} (1 - \kappa \gamma_{k})$$
(S50)

Since $(\gamma_k)_{k\geq 1}$ is nonincreasing, we have

$$\sum_{i=n}^{n_0-1} \prod_{k=i+1}^{n_0-1} \left(1 + L^2 \gamma_k^2\right) \gamma_i = \sum_{i=n}^{n_0-1} (\gamma_i L^2)^{-1} \left\{ \prod_{k=i}^{n_0-1} \left(1 + L^2 \gamma_k^2\right) - \prod_{k=i+1}^{n_0-1} \left(1 + L^2 \gamma_k^2\right) \right\}$$
$$\leq \prod_{k=n}^{n_0-1} (\gamma_{n_0-1} L^2)^{-1} \left(1 + L^2 \gamma_k^2\right) .$$

Furthermore for $k < n_0 \ \gamma_k > 2/(m+L)$. This implies with the bound $(1+x) \leq e^x$ on \mathbb{R} :

$$\prod_{k=n_0}^{p} (1 - \kappa \gamma_k) \le \exp\left(-\sum_{k=n}^{p} \kappa \gamma_k\right) \exp\left(\sum_{k=n}^{n_0-1} \kappa \gamma_k\right)$$
$$\le \exp\left(-\sum_{k=n}^{p} \kappa \gamma_k\right) \exp\left(\sum_{k=n}^{n_0-1} \gamma_k^2 m L\right) \ .$$

Using the two previous inequalities in (S50), we get

$$\sum_{i=n}^{p} \prod_{k=i+1}^{p} \eta_k \gamma_i \\ \leq \kappa^{-1} + \left\{ \prod_{k=n}^{n_0-1} (\gamma_{n_0-1} L^2)^{-1} \left(1 + L^2 \gamma_k^2\right) \right\} \exp\left(-\sum_{k=n}^{p} \kappa \gamma_k + \sum_{k=n}^{n_0-1} \gamma_k^2 m L\right) .$$
(S51)

Combining (S49) and (S51) in (S48) concluded the proof.

We now deal with bounds on $W_2(\mu_0 Q_{\gamma}^n, \pi)$ under **G1**. But before we preface our result by some technical lemmas.

Lemma S16. Assume H1 and H2. Let $\zeta_0 \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, $(Y_t, \overline{Y}_t)_{t\geq 0}$ such that (Y_0, \overline{Y}_0) is distributed according to ζ_0 and given by (9). Let $(\mathcal{F}_t)_{t\geq 0}$ be the filtration associated with $(B_t)_{t\geq 0}$ with \mathcal{F}_0 , the σ -field generated by (Y_0, \overline{Y}_0) . Then for all $n \geq 0$, $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$,

$$\mathbb{E}^{\mathcal{F}_{\Gamma_{n}}} \left[\left\| Y_{\Gamma_{n+1}} - \overline{Y}_{\Gamma_{n+1}} \right\|^{2} \right]$$

$$\leq \{ 1 - \gamma_{n+1} \left(\kappa - 2\varepsilon_{1} \right) + \gamma_{n+1} L ((1 + \varepsilon_{2})\gamma_{n+1} - 2/(m+L))_{+} \} \left\| Y_{\Gamma_{n}} - \overline{Y}_{\Gamma_{n}} \right\|^{2}$$

$$+ \gamma_{n+1}^{2} (1/(2\varepsilon_{1}) + (1 + \varepsilon_{2}^{-1})\gamma_{n+1}) \left(dL^{2} + (L^{4}\gamma_{n+1}/2) \left\| Y_{\Gamma_{n}} - x^{\star} \right\|^{2} + dL^{4}\gamma_{n+1}^{2}/12 \right) .$$

Proof. Let $n \ge 0$ and $\varepsilon_1 > 0$, and set $\Delta_n = Y_{\Gamma_n} - \overline{Y}_{\Gamma_n}$ by definition we have:

$$\mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\left\| \Delta_{n+1} \right\|^2 \right] = \left\| \Delta_n \right\|^2 + \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\left\| \int_{\Gamma_n}^{\Gamma_{n+1}} \left\{ \nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}) \right\} ds \right\|^2 \right] - 2\gamma_{n+1} \left\langle \Delta_n, \nabla U(Y_{\Gamma_n}) - \nabla U(\overline{Y}_{\Gamma_n}) \right\rangle - 2 \int_{\Gamma_n}^{\Gamma_{n+1}} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\left\langle \Delta_n, \{\nabla U(Y_s) - \nabla U(Y_{\Gamma_n})\} \right\rangle ds \right] .$$

Using the two inequalities $|\langle a, b \rangle| \leq \varepsilon_1 ||a||^2 + (4\varepsilon_1)^{-1} ||b||^2$ and (3), we get

$$\mathbb{E}^{\mathcal{F}_{\Gamma_{n}}}\left[\|\Delta_{n+1}\|^{2}\right] \leq \left\{1 - \gamma_{n+1}(\kappa - 2\varepsilon_{1})\right\} \|\Delta_{n}\|^{2} - 2\gamma_{n+1}/(m+L) \left\|\nabla U(Y_{\Gamma_{n}}) - \nabla U(\overline{Y}_{\Gamma_{n}})\right\|^{2} \\ + \mathbb{E}^{\mathcal{F}_{\Gamma_{n}}}\left[\left\|\int_{\Gamma_{n}}^{\Gamma_{n+1}} \left\{\nabla U(Y_{s}) - \nabla U(\overline{Y}_{\Gamma_{n}})\right\} \mathrm{d}s\right\|^{2}\right] \\ + (2\varepsilon_{1})^{-1} \int_{\Gamma_{n}}^{\Gamma_{n+1}} \mathbb{E}^{\mathcal{F}_{\Gamma_{n}}}\left[\left\|\nabla U(Y_{s}) - \nabla U(Y_{\Gamma_{n}})\right\|^{2}\right] \mathrm{d}s .$$
(S52)

Using $||a+b||^2 \le (1+\varepsilon_2) ||a||^2 + (1+\varepsilon_2^{-1}) ||b||^2$ and the Jensen's inequality, we have

$$\mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\left\| \int_{\Gamma_n}^{\Gamma_{n+1}} \left\{ \nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}) \right\} \mathrm{d}s \right\|^2 \right] \le (1 + \varepsilon_2) \gamma_{n+1}^2 \left\| \nabla U(Y_{\Gamma_n}) - \nabla U(\overline{Y}_{\Gamma_n}) \right\|^2 + \gamma_{n+1} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\int_{\Gamma_n}^{\Gamma_{n+1}} \left\| \nabla U(Y_s) - \nabla U(Y_{\Gamma_n}) \right\|^2 \mathrm{d}s \right] .$$

This result and H1 imply,

$$\mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\|\Delta_{n+1}\|^2 \right] \leq \left\{ 1 - \gamma_{n+1} (\kappa - 2\varepsilon_1) + \gamma_{n+1} L ((1+\varepsilon_2)\gamma_{n+1} - 2/(m+L))_+ \right\} \|\Delta_n\|^2 + \left((1+\varepsilon_2^{-1})\gamma_{n+1} + (2\varepsilon_1)^{-1} \right) \int_{\Gamma_n}^{\Gamma_{n+1}} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\|\nabla U(Y_s) - \nabla U(Y_{\Gamma_n})\|^2 \right] \mathrm{d}s \,.$$
(S53)

By **H1**, the Markov property of $(Y_t)_{t\geq 0}$ and Lemma 19, we have

$$\begin{split} \int_{\Gamma_n}^{\Gamma_{n+1}} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\|\nabla U(Y_s) - \nabla U(Y_{\Gamma_n})\|^2 \right] \mathrm{d}s \\ & \leq L^2 \left(d\gamma_{n+1}^2 + dL^2 \gamma_{n+1}^4 / 12 + (L^2 \gamma_{n+1}^3 / 2) \|Y_{\Gamma_n} - x^\star\|^2 \right) \,. \end{split}$$

Plugging this bound in (S53) concludes the proof.

Lemma S17. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence of positive numbers. Let $\varpi, \beta > 0$ be positive constants satisfying $\varpi^2 \leq 4\beta$ and $\tau > 0$. Assume there exists $N \geq 1$, $\gamma_N \leq \tau$ and $\gamma_N \varpi \leq 1$. Then for all $n \geq 0$, $j \geq 2$

$$\sum_{i=1}^{n+1} \prod_{k=i+1}^{n+1} \left(1 - \gamma_k \varpi + \gamma_k \beta(\gamma_k - \tau)_+\right) \gamma_i^j \le \sum_{i=N}^{n+1} \prod_{k=i+1}^{n+1} \left(1 - \gamma_k \varpi\right) \gamma_i^j + \left\{\beta^{-1} \gamma_1^{j-2} \prod_{k=1}^{N-1} \left(1 + \gamma_k^2 \beta\right)\right\} \prod_{k=N}^{n+1} \left(1 - \varpi \gamma_k\right) .$$

(ii) For all $\ell \in \{N, \ldots, n\}$,

$$\sum_{i=N}^{n+1} \prod_{k=i+1}^{n+1} \left(1 - \gamma_k \varpi\right) \gamma_i^j \le \exp\left(-\sum_{k=\ell}^{n+1} \varpi \gamma_k\right) \sum_{i=N}^{\ell-1} \gamma_i^j + \frac{\gamma_\ell^{j-1}}{\varpi} \,.$$

Proof. By definition of N we have

$$\sum_{i=1}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi + \gamma_k \beta (\gamma_k - \tau)_+) \gamma_i^j$$

$$\leq \sum_{i=N}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi) \gamma_i^j + \left\{ \sum_{i=1}^{N-1} \prod_{k=i+1}^{N-1} (1 + \gamma_k^2 \beta) \gamma_i^j \right\} \prod_{k=N}^{n+1} (1 - \gamma_k \varpi) .$$
(S54)

Using that $(\gamma_k)_{k\geq 1}$ is nonincreasing, we have

$$\begin{split} \sum_{i=1}^{N-1} \prod_{k=i+1}^{N-1} \left(1 + \gamma_k^2 \beta \right) \gamma_i^j &\leq \sum_{i=1}^{N-1} \frac{\gamma_i^{j-2}}{\beta} \left\{ \prod_{k=i}^{N-1} \left(1 + \gamma_k^2 \beta \right) - \prod_{k=i+1}^{N-1} \left(1 + \gamma_k^2 \beta \right) \right\} \\ &\leq \beta^{-1} \gamma_1^{j-2} \prod_{k=1}^{N-1} \left(1 + \gamma_k^2 \beta \right) \;. \end{split}$$

Plugging this inequality in (S54) concludes the proof of (i). Let $\ell \in \{N, \ldots, n+1\}$. Since $(\gamma_k)_{k\geq 1}$ is nonincreasing and for every $x \in \mathbb{R}$, $(1+x) \leq e^x$, we get

$$\sum_{i=N}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi) \gamma_i^j = \sum_{i=N}^{\ell-1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi) \gamma_i^j + \sum_{i=\ell}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi) \gamma_i^j$$

$$\leq \sum_{i=N}^{\ell-1} \exp\left(-\sum_{k=i+1}^{n+1} \varpi \gamma_k\right) \gamma_i^j + \gamma_\ell^{j-1} \sum_{i=\ell}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi) \gamma_i$$

$$\leq \exp\left(-\sum_{k=\ell}^{n+1} \varpi \gamma_k\right) \sum_{i=N}^{\ell-1} \gamma_i^j + \frac{\gamma_\ell^{j-1}}{\varpi} .$$

Lemma S18. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence of positive numbers, $\varpi, \beta, \tau > 0$ be positive real numbers, and $N \geq 1$ satisfying the assumptions of Lemma S17. Let $\mathsf{P} \in \mathbb{N}^*, C_i \geq 0, i = 0, \dots, \mathsf{P}$ be positive constants and $(u_n)_{n\geq 0}$ be a sequence of real numbers with $u_0 \geq 0$ satisfying for all $n \geq 0$

$$u_{n+1} \le (1 - \gamma_{n+1}\varpi + \beta\gamma_{n+1}(\gamma_{n+1} - \tau)_+)u_n + \sum_{i=0}^{\mathsf{P}} C_j \gamma_{n+1}^{j+2}.$$

Then for all $n \geq 1$,

$$\begin{split} u_n &\leq \left\{ \prod_{k=1}^{N-1} \left(1 + \beta \gamma_k^2 \right) \right\} \prod_{k=N}^n \left(1 - \gamma_k \varpi \right) u_0 + \sum_{j=0}^{\mathsf{P}} C_j \sum_{i=N}^n \prod_{k=i+1}^n \left(1 - \gamma_k \varpi \right) \gamma_i^{j+2} \\ &+ \left\{ \sum_{j=0}^{\mathsf{P}} C_j \beta^{-1} \gamma_1^j \prod_{k=1}^{N-1} \left(1 + \gamma_k^2 \beta \right) \right\} \prod_{k=N}^n \left(1 - \varpi \gamma_k \right) \;. \end{split}$$

Proof. This is a consequence of a straightforward induction and Lemma S17-(i). \Box

Proposition S19. Assume H1, H2 and G1. Let x^* be the unique minimizer of U. Let $\zeta_0 \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, $(Y_t, \overline{Y}_t)_{t \geq 0}$ such that (Y_0, \overline{Y}_0) is distributed according to ζ_0 and given by (9). Then for all $n \geq 0$ and $t \in [\Gamma_n, \Gamma_{n+1}]$:

$$\mathbb{E}\left[\left\|Y_t - \overline{Y}_t\right\|^2\right] \le \tilde{u}_n^{(1)} \mathbb{E}\left[\left\|Y_0 - \overline{Y}_0\right\|^2\right] + \tilde{u}_n^{(4)} + \tilde{u}_{t,n}^{(5)},$$

where

$$\tilde{u}_{n}^{(1)}(\gamma) = \left\{ \prod_{k=1}^{n_{1}-1} \left(1 + L^{2}(1+\rho)\gamma_{k}^{2} \right) \right\} \prod_{k=n_{1}}^{n} \left(1 - \kappa\gamma_{k}/2 \right) , \qquad (S55)$$

$$\begin{split} \tilde{u}_n^{(4)}(\gamma) &= \sum_{i=n_1}^n \gamma_i^2 \mathbf{a}(\gamma_i) \prod_{k=i+1}^n (1 - \kappa \gamma_k/2) \\ &+ \mathbf{a}(\gamma_1) (L^2 (1+\rho))^{-1} \left\{ \prod_{k=1}^{n_1-1} (1 + \gamma_k^2 (1+\rho)L^2) \right\} \prod_{k=n_1}^n (1 - \kappa \gamma_k/2) \,, \end{split}$$

where for $\gamma > 0$,

$$\mathbf{a}(\gamma) = \left\{ 2\kappa^{-1} + (1+\rho^{-1})\gamma \right\} (dL^2 + \delta L^4 \gamma/2 + dL^4 \gamma^2/12) ,$$

$$\delta = \max_{i \ge 1} \left\{ e^{-2m\Gamma_{i-1}} \mathbb{E} \left[\|Y_0 - x^\star\|^2 \right] + (1 - e^{-2m\Gamma_{i-1}})(d/m) \right\} ,$$

and

$$\tilde{u}_{t,n}^{(5)}(\gamma) = \frac{m+L}{2} \left(\frac{(t-\Gamma_n)^3 L^2}{3} \mathsf{G}_{1,n}(\mu_0,\gamma) + (t-\Gamma_n)^2 d \right) ,$$

where $\mathsf{G}_{1,n}(\mu_0,\gamma)$ is given by (S47) and μ_0 is the initial distribution of \overline{Y}_0 .

Proof. Lemma S16 with $\varepsilon_1 = \kappa/4$ and $\varepsilon_2 = \rho$, G1, Lemma S17, ((i)) imply for all $n \ge 0$

$$\mathbb{E}\left[\left\|Y_{\Gamma_n} - \overline{Y}_{\Gamma_n}\right\|^2\right] \le \tilde{u}_n^{(1)}(\gamma) \mathbb{E}\left[\left\|Y_0 - \overline{Y}_0\right\|^2\right] + \tilde{u}_n^{(4)}(\gamma) .$$
(S56)

Now let $n \ge 0$ and $t \in [\Gamma_n, \Gamma_{n+1}]$. By (9),

$$\left\|Y_t - \overline{Y}_t\right\|^2 = \left\|Y_{\Gamma_n} - \overline{Y}_{\Gamma_n}\right\|^2 - 2\int_{\Gamma_n}^t \left\langle\nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}), Y_s - \overline{Y}_s\right\rangle \mathrm{d}s \,. \tag{S57}$$

Moreover for all $s \in [\Gamma_n, \Gamma_{n+1}]$, by (3) we get

$$\left\langle \nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}), Y_s - \overline{Y}_s \right\rangle = \left\langle \nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}), Y_s - \overline{Y}_{\Gamma_n} + \overline{Y}_{\Gamma_n} - \overline{Y}_s \right\rangle$$

$$\geq (m+L)^{-1} \left\| \nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}) \right\|^2 + \left\langle \nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}), \overline{Y}_{\Gamma_n} - \overline{Y}_s \right\rangle .$$
 (S58)

Since $|\langle a, b \rangle| \le (m+L)^{-1} ||a||^2 + (m+L) ||b||^2 /4$, we have

$$\left\langle \nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}), \overline{Y}_{\Gamma_n} - \overline{Y}_s \right\rangle$$

$$\geq -(m+L)^{-1} \left\| \nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}) \right\|^2 - (m+L) \left\| \overline{Y}_s - \overline{Y}_{\Gamma_n} \right\|^2 / 4.$$

Using this inequality in (S58), we get

$$\left\langle \nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n}), Y_s - \overline{Y}_s \right\rangle \ge -(m+L) \left\| \overline{Y}_s - \overline{Y}_{\Gamma_n} \right\|^2 / 4,$$

and (S57) becomes

$$\left\|Y_t - \overline{Y}_t\right\|^2 \le \left\|Y_{\Gamma_n} - \overline{Y}_{\Gamma_n}\right\|^2 + \left((m+L)/2\right) \int_{\Gamma_n}^t \left\|\overline{Y}_s - \overline{Y}_{\Gamma_n}\right\|^2 \mathrm{d}s \tag{S59}$$

Therefore, by the previous inequality, it remains to bound the expectation of $\|\overline{Y}_s - \overline{Y}_{\Gamma_n}\|^2$. By (9) and using $\nabla U(x^*) = 0$,

$$\left\|\overline{Y}_{s} - \overline{Y}_{\Gamma_{n}}\right\|^{2} = \left\|-(s - \Gamma_{n})(\nabla U(\overline{Y}_{\Gamma_{n}}) - \nabla U(x^{\star})) + \sqrt{2}(B_{s} - B_{\Gamma_{n}})\right\|^{2}.$$

Then taking the expectation, using the Markov property of $(B_t)_{t\geq 0}$ and **H**1, we have

$$\mathbb{E}\left[\left\|\overline{Y}_{s} - \overline{Y}_{\Gamma_{n}}\right\|^{2}\right] \leq (s - \Gamma_{n})^{2} L^{2} \mathbb{E}\left[\left\|\overline{Y}_{\Gamma_{n}} - x^{\star}\right\|^{2}\right] + 2(s - \Gamma_{n})d.$$
(S60)

The proof follows from taking the expectation in (S59), combining (S56)-(S60), and using Theorem S15.

Theorem S20. Assume H1, H2 and G1. Then for all $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ and $n \ge 1$,

$$W_2^2(\mu_0 Q_\gamma^n, \pi) \le \tilde{u}_n^{(1)}(\gamma) W_2^2(\mu_0, \pi) + \tilde{u}_n^{(2)}(\gamma) , \qquad (S61)$$

where $(\tilde{u}_n^{(1)})_{n\geq 0}$ is given by (S55) and

$$\tilde{u}_{n}^{(2)}(\gamma) = \sum_{i=n_{1}}^{n} \gamma_{i}^{2} \mathsf{b}(\gamma_{i}) \prod_{k=i+1}^{n} (1 - \kappa \gamma_{k}/2) + \mathsf{b}(\gamma_{1}) (L^{2}(1+\rho))^{-1} \left\{ \prod_{k=1}^{n_{1}-1} (1 + \gamma_{k}^{2}(1+\rho)L^{2}) \right\} \prod_{k=n_{1}}^{n} (1 - \kappa \gamma_{k}/2) , \quad (S62)$$

where

$$\mathsf{b}(\gamma) = \left\{ 2\kappa^{-1} + (1+\rho^{-1})\gamma \right\} \left(dL^2 + dL^4\gamma/(2m) + dL^4\gamma^2/12 \right) \,.$$

Proof. Let ζ_0 be an optimal transference plan of μ_0 and π . Let $(Y_t, \overline{Y}_t)_{t\geq 0}$ with (Y_0, \overline{Y}_0) distributed according to ζ_0 and defined by (9). By definition of W_2 and since for all $t \geq 0, \pi$ is invariant for $P_t, W_2^2(\mu_0 Q^n, \pi) \leq \mathbb{E}[||Y_{\Gamma_n} - X_{\Gamma_n}||^2]$. Then the proof follows from Proposition S19 since Y_0 is distributed according to π and by Proposition 1-(ii), which shows that $\delta \leq d/m$.

8.1 Explicit bound based on Theorem S20 for $\gamma_k = \gamma_1 k^{\alpha}$ with $\alpha \in (0, 1]$

We give here a bound on the sequences $(\tilde{u}_n^{(1)}(\gamma))_{n\geq 1}$ and $(\tilde{u}_n^{(2)}(\gamma))_{n\geq 1}$ for $(\gamma_k)_{k\geq 1}$ defined by $\gamma_1 > 0$ and $\gamma_k = \gamma_1 k^{-\alpha}$ for $\alpha \in (0, 1]$. Recall that ψ_β is given by (S38). First note, since $(\gamma_k)_{k\geq 1}$ is nonincreasing, for all $n \geq 1$, we have

$$\tilde{u}_{n}^{(2)}(\gamma) \leq \sum_{j=0}^{1} \mathsf{C}_{j} \sum_{i=n_{1}}^{n} \gamma_{i}^{j+2} \prod_{k=i+1}^{n} (1 - \kappa \gamma_{k}/2) + \sum_{j=0}^{1} \mathsf{C}_{j} (L^{2}(1+\rho))^{-1} \gamma_{1}^{j} \left\{ \prod_{k=1}^{n_{1}-1} (1 + \gamma_{k}^{2}(1+\rho)L^{2}) \right\} \prod_{k=n_{1}}^{n} (1 - \kappa \gamma_{k}/2) , \quad (S63)$$

where

$$C_1 = bdL^2$$
, $C_2 = b(dL^4/(2m) + \gamma_1 dL^4/12)$, $b = 2\kappa^{-1} + (1 + \rho^{-1})\gamma_1$.

1. For $\alpha = 1$ and $n_1 = 1$, by (S39) and (S40), we have

$$\begin{split} \tilde{u}_n^{(1)}(\gamma) &\leq (n+1)^{-\kappa\gamma_1/2} \\ \tilde{u}_n^{(2)}(\gamma) &\leq (n+1)^{-\kappa\gamma_1/2} \sum_{j=0}^1 \mathsf{C}_j \left\{ \gamma_1^{j+2} (\boldsymbol{\psi}_{\kappa\gamma_1/2-1-j}(n+1)+1) + (L^2(1+\rho))^{-1} \gamma_1^j \right\} \,. \end{split}$$

For $n_1 > 1$, since $(\gamma_k)_{k \ge 0}$ is non increasing, using again (S39), (S40), and the bound for $t \in \mathbb{R}$, $(1+t) \le e^t$

$$\begin{split} \tilde{u}_n^{(1)}(\gamma) &\leq (n+1)^{-\kappa\gamma_1/2} \exp\left\{\kappa\gamma_1 \psi_0(n_1)/2 + L^2(1+\rho)\gamma_1^2(\psi_{-1}(n_1-1)+1)\right\} \\ \tilde{u}_n^{(2)}(\gamma) &\leq (n+1)^{-\kappa\gamma_1/2} \sum_{j=0}^1 \mathsf{C}_j\left(\gamma_1^{j+2}(\psi_{\kappa\gamma_1/2-1-j}(n+1) - \psi_{\kappa\gamma_1/2-1-j}(n_1)+1) \right. \\ &\left. + (\gamma_1^j/(L^2(1+\rho)) \exp\left\{\kappa\gamma_1\psi_0(n_1)/2 + L^2(1+\rho)\gamma_1^2(\psi_{-1}(n_1-1)+1)\right\}\right) \,. \end{split}$$

Thus, for $\gamma_1 > \kappa/2$, we get a bound in $\mathcal{O}(n^{-1})$.

2. For $\alpha \in (0,1)$ and $n_1 = 1$, by (S39) and Lemma S17-(ii) applied with $\ell = \lceil n/2 \rceil$, we have

$$\begin{split} \tilde{u}_{n}^{(1)}(\gamma) &\leq \exp\left(-(\kappa\gamma_{1}/2)\psi_{\alpha-1}(n+1)\right) \\ \tilde{u}_{n}^{(2)}(\gamma) &\leq \sum_{j=0}^{1} \mathsf{C}_{j} \left\{ \gamma_{1}^{j+2} \left(\psi_{1-\alpha(j+2)}(\lceil n/2 \rceil) + 1\right) \mathrm{e}^{-(\kappa\gamma_{1}/2)(\psi_{1-\alpha}(n+1)-\psi_{1-\alpha}(\lceil n/2 \rceil))} \right. \\ &\left. + 2\kappa^{-1}\gamma_{1}^{j+1}(n/2)^{-\alpha(j+1)} + \gamma_{1}^{j} \mathrm{e}^{-(\kappa\gamma_{1}/2)\psi_{\alpha-1}(n+1)} \right\} \,. \end{split}$$

For $n_1 > 1$ and $\lceil n/2 \rceil \ge n_1$, since $(\gamma_k)_{k \ge 0}$ is non increasing, using again (S39), and Lemma S17-(ii) applied with $\ell = \lceil n/2 \rceil$, and the bound for $t \in \mathbb{R}$, $(1 + t) \le e^t$, we get

$$\begin{split} \tilde{u}_{n}^{(1)}(\gamma) &\leq e^{-\kappa\gamma_{1}(\boldsymbol{\psi}_{\alpha-1}(n+1)-\boldsymbol{\psi}_{1-\alpha}(n_{1}))/2+L^{2}(1+\rho)\gamma_{1}^{2}(\boldsymbol{\psi}_{1-2\alpha}(n_{1}-1)+1)} \\ \tilde{u}_{n}^{(2)}(\gamma) &\leq \sum_{j=0}^{1} \mathsf{C}_{j} \left\{ 2\kappa^{-1}\gamma_{1}^{j+1}(n/2)^{-\alpha(j+1)} \right. \\ &+ \gamma_{1}^{j+2} \left(\boldsymbol{\psi}_{1-\alpha(j+2)}(\lceil n/2 \rceil) - \boldsymbol{\psi}_{1-\alpha(j+2)}(n_{1}) + 1 \right) e^{-(\kappa\gamma_{1}/2)(\boldsymbol{\psi}_{1-\alpha}(n+1)-\boldsymbol{\psi}_{1-\alpha}(\lceil n/2 \rceil))} \\ &+ \left(\gamma_{1}^{j}/(L^{2}(1+\rho)) e^{-\kappa\gamma_{1}(\boldsymbol{\psi}_{\alpha-1}(n+1)-\boldsymbol{\psi}_{1-\alpha}(n_{1}))/2+L^{2}(1+\rho)\gamma_{1}^{2}(\boldsymbol{\psi}_{1-2\alpha}(n_{1}-1)+1)} \right\} \,. \end{split}$$

9 Explicit bounds on the MSE

Without loss of generality, assume that $||f||_{\text{Lip}} = 1$. In the following, denote by $\Omega(x) = ||x - x^*||^2 + d/m$ and C a constant (which may take different values upon each appearance), which does not depend on m, L, γ_1, α and $||x - x^*||$.

9.1 Explicit bounds based on Theorem 5

1. First for $\alpha = 0$, recall by Theorem 5 and (S34) we have for all $p \ge 1$,

$$W_2^2(\delta_x R_{\gamma}^p, \pi) \le 2\Omega(x)(1-\kappa\gamma_1/2)^p + 2\kappa^{-1}(\mathsf{A}_0\gamma_1 + \mathsf{A}_1\gamma_1^2),$$

where A_0 and A_1 are given by (S35) and (S36) respectively. So by (S23) and Lemma 7, we have the following bound for the bias

$$\left\{\mathbb{E}_x[\hat{\pi}_n^N(f)] - \pi(f)\right\}^2 \le C\left(\frac{\kappa^{-1}\exp(-\kappa N\gamma_1/2)\Omega(x)}{\gamma_1 n} + \kappa^{-1}\mathsf{A}_0\gamma_1\right) \ .$$

Therefore plugging this inequality and the one given by Theorem S1 implies:

$$\mathrm{MSE}_{f}^{N,n} \leq C\left(\kappa^{-1}\mathsf{A}_{0}\gamma_{1} + \frac{\kappa^{-2} + \kappa^{-1}\exp(-\kappa N\gamma_{1}/2)\Omega(x)}{n\gamma_{1}}\right) \,. \tag{S64}$$

So with fixed γ_1 this bound is of order γ_1 . If we fix the number of iterations n, we can optimize the choice of γ_1 . Set

$$\gamma_{\star,0}(n) = (\kappa^{-1} \mathsf{A}_0)^{-1} (C_{\text{MSE},0}/n)^{1/2}$$
, where $C_{\text{MSE},0} = \kappa^{-3} \mathsf{A}_0$

and (S64) becomes if $\gamma_1 \leftarrow \gamma_{\star,0}(n)$,

$$MSE_{f}^{N,n} \le C(C_{MSE,0}n)^{-1/2} \left(\kappa^{-1} \exp(-\kappa N\gamma_{\star,0}(n)/2)\Omega(x) + C_{MSE,0}\right)$$

Setting $N_0(n) = 2(\kappa \gamma_{\star,0}(n))^{-1} \log(\Omega(x))$, we end up with

$$\text{MSE}_{f}^{N_{0}(n),n} \leq C (C_{\text{MSE},0}/n)^{1/2}$$
.

Note that $N_0(n)$ is of order $n^{1/2}$.

2. For $\alpha \in (0, 1/2)$ by Theorem 5, Lemma 7, (S39) and (S41), we have the following bound for the bias

$$\left\{\mathbb{E}_{x}[\hat{\pi}_{n}^{N}(f)] - \pi(f)\right\}^{2} \leq C\left(\frac{\kappa^{-1}\mathsf{A}_{0}\gamma_{1}}{(1-2\alpha)n^{\alpha}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_{1}N^{1-\alpha}/(2(1-\alpha))\right\}\Omega(x)}{\gamma_{1}n^{1-\alpha}}\right)$$

Plugging this inequality and the one given by Theorem S1 implies:

$$\mathrm{MSE}_{f}^{N,n} \leq C \left(\frac{\kappa^{-1} \mathsf{A}_{0} \gamma_{1}}{(1-2\alpha)n^{\alpha}} + \frac{\kappa^{-1} \exp\left\{ -\kappa \gamma_{1} N^{1-\alpha} / (2(1-\alpha)) \right\} \Omega(x) + \kappa^{-2}}{\gamma_{1} n^{1-\alpha}} \right).$$
(S65)

At fixed γ_1 , this bound is of order $n^{-\alpha}$, and is better than (S64) for $(\gamma_k)_{k\geq 1}$ constant. If we fix the number of iterations n, we can optimize the choice of γ_1 again. Set

$$\gamma_{\star,\alpha}(n) = (\kappa^{-1} \mathsf{A}_0 / (1 - 2\alpha))^{-1} (C_{\mathrm{MSE},\alpha} / n^{1 - 2\alpha})^{1/2}$$
, where $C_{\mathrm{MSE},\alpha} = \kappa^{-3} \mathsf{A}_0 / (1 - 2\alpha)$,

(S65) becomes with $\gamma_1 \leftarrow \gamma_{\star,\alpha}(n)$,

$$\operatorname{MSE}_{f}^{N,n} \leq C(C_{\mathrm{MSE},\alpha}n)^{-1/2} \left(\kappa^{-1} \exp\left\{ -\kappa N^{1-\alpha} \gamma_{\star,\alpha}(n)/(2(1-\alpha)) \right\} \Omega(x) + C_{\mathrm{MSE},\alpha} \right) .$$

Setting $N_{\alpha}(n) = \{2(1-\alpha)(\kappa\gamma_{\star,\alpha}(n))^{-1}\log(\Omega(x))\}^{1/(1-\alpha)}$, we end up with

$$\operatorname{MSE}_{f}^{N_{\alpha}(n),n} \leq C(C_{\mathrm{MSE},\alpha}/n)^{1/2}$$
.

It is worthwhile to note that the order of $N_{\alpha}(n)$ in n is $n^{(1-2\alpha)/(2(1-\alpha))}$, and $C_{\text{MSE},\alpha}$ goes to infinity as $\alpha \to 1/2$.

3. If $\alpha = 1/2$, by Theorem 5, Lemma 7, (S39) and (S41), we have the following bound for the bias

$$\left\{\mathbb{E}_x[\hat{\pi}_n^N(f)] - \pi(f)\right\}^2 \le C\left(\frac{\kappa^{-1}\mathsf{A}_0\gamma_1\log(n)}{n^{1/2}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_1N^{1/2}/4\right\}\Omega(x)}{\gamma_1n^{1/2}}\right) \ .$$

Plugging this inequality and the one given by Theorem S1 implies:

$$\mathrm{MSE}_{f}^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{A}_{0}\gamma_{1}\log(n)}{n^{1/2}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_{1}N^{1/2}/4\right\}\Omega(x) + \kappa^{-2}}{\gamma_{1}n^{1/2}}\right) .$$
 (S66)

At fixed γ_1 , the order of this bound is $\log(n)n^{-1/2}$, and is the best bound for the MSE. Fix the number of iterations n, and we now optimize the choice of γ_1 . Set

$$\gamma_{\star,1/2}(n) = (\kappa^{-1}\mathsf{A}_0)^{-1} (C_{\text{MSE},1/2}/\log(n))^{1/2}$$
, where $C_{\text{MSE},1/2} = \kappa^{-3}\mathsf{A}_0$,

and (S66) becomes with $\gamma_1 \leftarrow \gamma_{\star,1/2}(n)$,

$$MSE_{f}^{N,n} \leq C \left(\frac{\log(n)}{nC_{MSE,1/2}}\right)^{1/2} \left(\kappa^{-1} \exp\left\{-\kappa N^{1/2} \gamma_{\star,1/2}(n)/4\right\} \Omega(x) + \frac{C_{MSE,1/2}}{\log(n)}\right) .$$

Setting $N_{1/2}(n) = (4(\kappa \gamma_{\star,1/2}(n))^{-1} \log(\Omega(x)))^2$, we end up with

$$\mathrm{MSE}_f^{N_{1/2}(n),n} \leq C \left(\frac{\log(n)C_{\mathrm{MSE},1/2}}{n}\right)^{1/2}$$

.

4. For $\alpha \in (1/2, 1]$, by Theorem 5, Lemma 7, (S39) and (S41), we have the following bound for the bias

$$\left\{\mathbb{E}_x[\hat{\pi}_n^N(f)] - \pi(f)\right\}^2 \le C\left(\frac{\kappa^{-1}\mathsf{A}_0\gamma_1}{n^{1-\alpha}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_1N^{1-\alpha}/(2(1-\alpha))\right\}\Omega(x)}{\gamma_1n^{1-\alpha}}\right)$$

Plugging this inequality and the one given by Theorem S1 implies:

$$\mathrm{MSE}_{f}^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{A}_{0}\gamma_{1}}{n^{1-\alpha}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_{1}N^{1-\alpha}/(2(1-\alpha))\right\}\Omega(x) + \kappa^{-2}}{\gamma_{1}n^{1-\alpha}}\right)$$

For fixed γ_1 , the MSE is of order $n^{1-\alpha}$, and is worse than for $\alpha \in [0, 1/2]$. For a fixed number of iteration n, optimizing γ_1 would imply to choose $\gamma_1 \to +\infty$ as $n \to +\infty$. Therefore, in that case, the best choice of γ_1 is the largest possible value 1/(m+L).

5. For $\alpha = 1$, by Theorem 5, Lemma 7, (S39) and (S41), the bias is upper bounded by

$$\left\{\mathbb{E}_x[\hat{\pi}_n^N(f)] - \pi(f)\right\}^2 \le C\left(\frac{\kappa^{-1}\mathsf{A}_0\gamma_1}{\log(n)} + \frac{\kappa^{-1}N^{-\kappa\gamma_1/2}\Omega(x)}{\gamma_1\log(n)}\right)$$

Plugging this inequality and the one given by Theorem S1 implies:

$$\mathrm{MSE}_{f}^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{A}_{0}\gamma_{1}}{\log(n)} + \frac{\kappa^{-1}N^{-\kappa\gamma_{1}/2}\Omega(x) + \kappa^{-2}}{\gamma_{1}\log(n)}\right)$$

For fixed γ_1 , the order of the MSE is $(\log(n))^{-1}$. For a fixed number of iterations, the conclusions are the same than for $\alpha \in (1/2, 1)$.

9.2 Explicit bound based on Theorem 9

1. First for $\alpha = 0$, recall by Theorem 9 and (S42) we have for all $p \ge 1$,

$$W_2^2(\delta_x R_{\gamma}^p, \pi) \le 2\Omega(x)(1-\kappa\gamma_1/2)^p + 2\kappa^{-1}(\mathsf{B}_0\gamma_1 + \mathsf{B}_1\gamma_1^2),$$

where B_0 and B_1 are given by (S43) and (S44) respectively. So by and Lemma 7, we have the following bound for the bias

$$\left\{\mathbb{E}_x[\hat{\pi}_n^N(f)] - \pi(f)\right\}^2 \le C\left(\frac{\kappa^{-1}\exp(-\kappa N\gamma_1/2)\Omega(x)}{\gamma_1 n} + \kappa^{-1}\mathsf{B}_0\gamma_1^2\right) \,.$$

Therefore plugging this inequality and the one given by Theorem S1 implies:

$$\mathrm{MSE}_{f}^{N,n} \leq C\left(\kappa^{-1}\mathsf{B}_{0}\gamma_{1}^{2} + \frac{\kappa^{-2} + \kappa^{-1}\exp(-\kappa N\gamma_{1}/2)\Omega(x)}{n\gamma_{1}}\right) \,. \tag{S67}$$

So with fixed γ_1 this bound is of order γ_1 . If we fix the number of iterations n, we can optimize the choice of γ_1 . Set

$$\gamma_{\star,0}(n) = (\kappa \mathsf{B}_0 n)^{-1/3}$$

and (S64) becomes if $\gamma_1 \leftarrow \gamma_{\star,0}(n)$,

$$\mathrm{MSE}_{f}^{N,n} \leq C(\mathsf{B}_{0}^{-1/2}n)^{-2/3} \left(\kappa^{-4/3} \exp(-\kappa N \gamma_{\star,0}(n)/2) \Omega(x) + \kappa^{-5/3} \right) \,.$$

Setting $N_0(n) = 2(\kappa \gamma_{\star,0}(n))^{-1} \log(\Omega(x))$, we end up with

$$MSE_f^{N_0(n),n} \le C(\mathsf{B}_0^{-1/2}\kappa^{5/2}n)^{-2/3}$$

Note that $N_0(n)$ is of order $n^{1/3}$.

2. For $\alpha \in (0, 1/3)$ by Theorem 9, Lemma 7, (S39) and (S45), we have the following bound for the bias

$$\left\{\mathbb{E}_{x}[\hat{\pi}_{n}^{N}(f)] - \pi(f)\right\}^{2} \leq C\left(\frac{\kappa^{-1}\mathsf{B}_{0}\gamma_{1}^{2}}{(1 - 3\alpha)n^{2\alpha}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_{1}N^{1 - \alpha}/(2(1 - \alpha))\right\}\Omega(x)}{\gamma_{1}n^{1 - \alpha}}\right).$$

Plugging this inequality and the one given by Theorem S1 implies:

$$\operatorname{MSE}_{f}^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{B}_{0}\gamma_{1}^{2}}{(1-3\alpha)n^{2\alpha}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_{1}N^{1-\alpha}/(2(1-\alpha))\right\}\Omega(x) + \kappa^{-2}}{\gamma_{1}n^{1-\alpha}}\right).$$
(S68)

If we fix the number of iterations n, we can optimize the choice of γ_1 again. Set

$$\gamma_{\star,\alpha}(n) = (n^{1-3\alpha} \kappa \mathsf{B}_0 / (1-3\alpha))^{-1/3}$$

(S68) becomes with $\gamma_1 \leftarrow \gamma_{\star,\alpha}(n)$,

$$\mathrm{MSE}_{f}^{N,n} \le C(\mathsf{B}_{0}^{-1/2}n)^{-2/3} \left(\kappa^{-4/3} \exp(-\kappa N \gamma_{\star,0}(n)/(2(1-\alpha)))\Omega(x) + \kappa^{-5/3}(1-3\alpha)^{-1/3} \right)$$

Setting $N_{\alpha}(n) = \{(\kappa \gamma_{\star,\alpha}(n))^{-1} \log(\Omega(x))\}^{1/(1-\alpha)}$, we end up with

$$\operatorname{MSE}_{f}^{N_{\alpha}(n),n} \leq C(\mathsf{B}_{0}^{-1/2}\kappa^{5/2}n)^{-2/3}$$

It is worthwhile to note that the order of $N_{\alpha}(n)$ in n is $n^{(1-3\alpha)/(3(1-\alpha))}$.

3. If $\alpha = 1/3$, by Theorem 9, Lemma 7, (S39) and (S45), we have the following bound for the bias

$$\left\{\mathbb{E}_x[\hat{\pi}_n^N(f)] - \pi(f)\right\}^2 \le C\left(\frac{\kappa^{-1}\mathsf{B}_0\gamma_1^2\log(n)}{n^{2/3}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_1N^{2/3}/4\right\}\Omega(x)}{\gamma_1n^{2/3}}\right).$$

Plugging this inequality and the one given by Theorem S1 implies:

$$\mathrm{MSE}_{f}^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{B}_{0}\gamma_{1}^{2}\log(n)}{n^{2/3}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_{1}N^{2/3}/4\right\}\Omega(x) + \kappa^{-2}}{\gamma_{1}n^{2/3}}\right) .$$
 (S69)

At fixed γ_1 , the order of this bound is $\log(n)n^{-2/3}$, and is the best bound for the MSE. Fix the number of iterations n, and we now optimize the choice of γ_1 . Set

$$\gamma_{\star,1/2}(n) = (\kappa \mathsf{B}_0 \log(n))^{-1/3}$$

and (S69) becomes with $\gamma_1 \leftarrow \gamma_{\star,1/2}(n)$,

$$\mathrm{MSE}_{f}^{N,n} \leq C \left(\frac{\log(n)\mathsf{B}_{0}}{n^{2}}\right)^{1/3} \left(\kappa^{-4/3} \exp\left\{-\kappa N^{1/2} \gamma_{\star,1/2}(n)/4\right\} \Omega(x) + \kappa^{-5/3}\right) \ .$$

Setting $N_{1/2}(n) = (4(\kappa \gamma_{\star,1/2}(n))^{-1} \log(\Omega(x)))^{3/2}$, we end up with

$$\mathrm{MSE}_{f}^{N_{1/2}(n),n} \leq C \left(\frac{\log(n)\mathsf{B}_{0}}{\kappa^{5}n^{2}}\right)^{1/3}$$

We can see that we obtain a worse bound than for $\alpha = 0$ and $\alpha \in (0, 1/3)$.

4. For $\alpha \in (1/3, 1]$, by Theorem 9, Lemma 7, (S39) and (S45), we have the following bound for the bias

$$\left\{\mathbb{E}_{x}[\hat{\pi}_{n}^{N}(f)] - \pi(f)\right\}^{2} \leq C\left(\frac{\kappa^{-1}\mathsf{B}_{0}\gamma_{1}}{n^{1-\alpha}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_{1}N^{1-\alpha}/(2(1-\alpha))\right\}\Omega(x)}{\gamma_{1}n^{1-\alpha}}\right).$$

Plugging this inequality and the one given by Theorem S1 implies:

$$\mathrm{MSE}_{f}^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{B}_{0}\gamma_{1}}{n^{1-\alpha}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_{1}N^{1-\alpha}/(2(1-\alpha))\right\}\Omega(x) + \kappa^{-2}}{\gamma_{1}n^{1-\alpha}}\right)$$

For fixed γ_1 , the MSE is of order $n^{1-\alpha}$, and is worse than for $\alpha = 1/2$. For a fixed number of iteration n, optimizing γ_1 would imply to choose $\gamma_1 \to +\infty$ as $n \to +\infty$. Therefore, in that case, the best choice of γ_1 is the largest possible value 1/(m+L).

5. For $\alpha = 1$, by Theorem 9, Lemma 7, (S39) and (S41), the bias is upper bounded by

$$\left\{\mathbb{E}_x[\hat{\pi}_n^N(f)] - \pi(f)\right\}^2 \le C\left(\frac{\kappa^{-1}\mathsf{B}_0\gamma_1}{\log(n)} + \frac{\kappa^{-1}N^{-\kappa\gamma_1/2}\Omega(x)}{\gamma_1\log(n)}\right)$$

Plugging this inequality and the one given by Theorem S1 implies:

$$\mathrm{MSE}_{f}^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{B}_{0}\gamma_{1}}{\log(n)} + \frac{\kappa^{-1}N^{-\kappa\gamma_{1}/2}\Omega(x) + \kappa^{-2}}{\gamma_{1}\log(n)}\right)$$

For fixed γ_1 , the order of the MSE is $(\log(n))^{-1}$. For a fixed number of iterations, the conclusions are the same than for $\alpha \in (1/2, 1)$.

References

S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

- [2] A. S. Dalalyan. Theoretical guarantees for approximate sampling from smooth and log-concave densities. Journal of the Royal Statistical Society: Series B (Statistical Methodology), pages n/a-n/a, 2016.
- [3] A. Joulin and Y. Ollivier. Curvature, concentration and error estimates for Markov chain Monte Carlo. Ann. Probab., 38(6):2418–2442, 2010.
- [4] I. Karatzas and S.E. Shreve. Brownian Motion and Stochastic Calculus. Graduate Texts in Mathematics. Springer New York, 1991.
- [5] S. Meyn and R. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, New York, NY, USA, 2nd edition, 2009.