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

Alain Durmus, Éric Moulines

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

Alain Durmus[1] and Éric Moulines[2]

[1]CMLA - École normale supérieure Paris-Saclay, CNRS, Université Paris-Saclay, 94235 Cachan, France.
[2]Centre de Mathématiques Appliquées, UMR 7641, Ecole Polytechnique, France.

July 15, 2018

1 Discussions on Theorem 5

In this section, we give explicit bounds for ULA which are consequences of Theorem 5. First note that, if $(\gamma_k)_{k\geq 1}$ is a non-increasing sequence of step sizes with $\gamma_1 < (m+L)^{-1}$, we have by (9) that

$$u_n^{(2)}(\gamma) \leq \sum_{i=1}^{n} \left\{ \mathsf{A}_0 \gamma_i^2 + \mathsf{A}_1 \gamma_i^3 \right\} \prod_{k=i+1}^{n} (1 - \kappa\gamma_k/2) \; , \tag{S1}$$

where κ is given by (42), and

$$\mathsf{A}_0 = 2L^2 \kappa^{-1} d \; , \tag{S2}$$

$$\mathsf{A}_1 = 2L^2 d + dL^4 (\kappa^{-1} + (m+L)^{-1})(m^{-1} + 6^{-1}(m+L)^{-1}) \; . \tag{S3}$$

1.1 Explicit bounds for fixed step size and fixed precision

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 (S1) is the following result, which gives the minimal number of iterations n_ε and a step size γ_ε to get $W_2(\delta_{x^\star} Q_\gamma^n, \pi)$ smaller than $\varepsilon > 0$.

[1]Email: alain.durmus@cmla.ens-cachan.fr
[2]eric.moulines@polytechnique.edu

Corollary S1 (of Theorem 5)**.** *Assume* **H** *1 and* **H** *2. Let $x^\star$ 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*

$$\begin{aligned} \gamma &= \frac{-\mathsf{A}_0 + (\mathsf{A}_0^2 + \varepsilon^2 \kappa \mathsf{A}_1)^{1/2}}{2\mathsf{A}_1} \wedge (m+L)^{-1} \,, \\ n &= \left\lceil \log^{-1}(1-\kappa\gamma/2) \left\{ -\log(\varepsilon^2/2) + \log(2d/m) \right\} \right\rceil \,. \end{aligned} \tag{S4}$$

Then $W_2(\delta_{x^\star} R_\gamma^n, \pi) \le \varepsilon$.

Note that if γ is given by (S4), and is different from $1/(m+L)$, then $\gamma \le \varepsilon(4\mathsf{A}_1\kappa^{-1})^{-1/2}$ and $2\kappa^{-1}(\mathsf{A}_0\gamma + \mathsf{A}_1\gamma^2) = \varepsilon^2/2$. Therefore,

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

It is shown in [1, Corollary 1, Proposition 2] that under **H** 2, 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 $\| \nu^\star Q_n^\gamma - \pi \|_{\mathrm{TV}} \le \varepsilon$ where $\nu^\star$ is the Gaussian measure on $\mathbb{R}^d$ with mean $x^\star$ and covariance matrix $L^{-1} \mathrm{I}_d$ or a warm start. We stress that the results in [1, Corollary 1, Proposition 2] hold only for particular choices of the initial distribution $\nu^\star$, (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 S1 with those given in [1, Corollary 1, Proposition 2] and [2, Table 2] for the total variation distance and established under the same conditions as Theorem 5. This comparison is summarized in Table 1; for simplicity, we provide only the dependencies of the 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 [1, Corollary 1]. The dependency on the dimension and the precision is same compared to [1, Proposition 2] (up to logarithmic terms), but this result only holds if the initial distribution is a warm start. In addition, the dependency on L and m is not explicit in [1, Proposition 2], and that is why we do indicate it in Table 1. On the other hand, we get the same dependency on ε and d as [2, Table 2]. Note that however the result of [2] holds for the total variation and not the Wasserstein distance.

Parameter	d	ε	L	m
Corollary S1	$d\log(d)$	$\varepsilon^{-2}\lvert\log(\varepsilon)\rvert$	L^2	$\lvert\log(m)\rvert\, m^{-3}$
[1, Corollary 1] Gaussian start	d^3	$\varepsilon^{-2}\lvert\log(\varepsilon)\rvert$	L^3	$\lvert\log(m)\rvert\, m^{-2}$
[1, Proposition 2] warm start	d	$\varepsilon^{-2}\lvert\log(\varepsilon)\rvert$	–	–
[2, Table 2]	$d\log(d)$	$\varepsilon^{-2}\lvert\log(\varepsilon)\rvert$	L^2	m^{-2}

Table 1: Dependencies of n

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

We give here a bound on the sequences $(u_n^{(1)}(\gamma))_{n\ge 0}$ and $(u_n^{(2)}(\gamma))_{n\ge 0}$ for $(\gamma_k)_{k\ge 1}$ 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 $t \in \mathbb{R}_+^*$,

$$\boldsymbol{\psi}_\beta(t) = \begin{cases} (t^\beta - 1)/\beta & \text{for } \beta \neq 0 \\ \log(t) & \text{for } \beta = 0 \,. \end{cases} \tag{S5}$$

We easily get for $a \geq 0$ that for all $n, p \geq 1$, $n \leq p$

$$\boldsymbol{\psi}_{1-a}(p+1) - \boldsymbol{\psi}_{1-a}(n) \leq \sum_{k=n}^{p} k^{-a} \leq \boldsymbol{\psi}_{1-a}(p) - \boldsymbol{\psi}_{1-a}(n) + 1 \,, \tag{S6}$$

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

$$\sum_{k=n}^{p} k^{-a} \leq \boldsymbol{\psi}_{1-a}(p+1) - \boldsymbol{\psi}_{1-a}(n) + 1 \,. \tag{S7}$$

1. For $\alpha = 1$, using that for all $t \in \mathbb{R}$, $(1+t) \leq \mathrm{e}^t$ and by (S6) and (S7), we have

$$u_n^{(1)}(\gamma) \leq (n+1)^{-\kappa\gamma_1/2} \;,\; u_n^{(2)}(\gamma) \leq (n+1)^{-\kappa\gamma_1/2} \sum_{j=0}^{1} \mathsf{A}_j (\boldsymbol{\psi}_{\kappa\gamma_1/2-1-j}(n+1) + 1) \,.$$

2. For $\alpha \in (0,1)$, by (S6) and Lemma 23 applied with $\ell = \lceil n/2 \rceil$, where $\lceil \cdot \rceil$ is the ceiling function, we have

$$\begin{aligned} u_n^{(1)}(\gamma) &\leq \exp\left(-\kappa\gamma_1 \boldsymbol{\psi}_{1-\alpha}(n+1)/2\right) \\ u_n^{(2)}(\gamma) &\leq \sum_{j=0}^{1} \mathsf{A}_j \left[2\kappa^{-1}\gamma_1^{j+1}(n/2)^{-\alpha(j+1)} + \gamma_1^{j+2}\left\{\boldsymbol{\psi}_{1-\alpha(j+2)}(\lceil n/2 \rceil) + 1\right\} \right. \\ &\qquad\qquad \left. \times \exp\left(-(\kappa\gamma_1/2)\left\{\boldsymbol{\psi}_{1-\alpha}(n+1) - \boldsymbol{\psi}_{1-\alpha}(\lceil n/2 \rceil)\right\}\right) \right] \,. \end{aligned} \tag{S8}$$

1.3 Optimal strategy with a fixed number of iterations

Corollary S2. *Let $n \in \mathbb{N}^*$ be a fixed number of iteration. Assume **H**1, **H**2, and $(\gamma_k)_{k\geq 1}$ is a constant sequence, $\gamma_k = \gamma$ for all $k \geq 1$. Set*

$$\begin{aligned} \gamma^+ &= 2(\kappa n)^{-1}\left[\log(\kappa n/2) + \log(2(\|x - x^\star\|^2 + d/m)) - \log(2\kappa^{-1}\mathsf{A}_0)\right] \\ \gamma_- &= 2(\kappa n)^{-1}\left[\log(\kappa n/2) + \log(2(\|x - x^\star\|^2 + d/m)) \right. \\ &\qquad\qquad \left. - \log\left\{2\kappa^{-1}(\mathsf{A}_0 + 2\mathsf{A}_1(m+L)^{-1})\right\}\right] \,. \end{aligned}$$

Assume $\gamma^+ \in \left(0, (m+L)^{-1}\right)$. Then, the optimal choice of γ to minimize the bound on $W_2(\delta_x R_\gamma^n, \pi)$ 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 $W_2^2(\delta_x R_\gamma^n, \pi)$ given by Theorem 5 is equivalent to $Cdn^{-1}\log(n)$ as n goes to $+\infty$.

Similarly, we have the following result.

Corollary S3. *Assume* ***H**1 and **H**2. 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 $\gamma_\alpha \in \left(0, (m+L)^{-1}\right)$. Then there exists $C \geq 0$ independent of the dimension such that the bound on $W_2^2(\delta_x Q_\gamma^n, \pi)$ given by Theorem 5 is equivalent to $Cdn^{-1}\log(n)$ as n goes to $+\infty$.

Proof. Follows from (S1), (S8) and the choice of γ_α. □

2 Discussion on Theorem 8

Based on Theorem 8, we can follow the same discussion as for Theorem 5. Note that

$$u_n^{(3)}(\gamma) \leq \sum_{i=1}^{n} \left\{ \mathsf{B}_0 \gamma_i^3 + \mathsf{B}_1 \gamma_i^4 \right\} \prod_{k=i+1}^{n} (1 - \kappa\gamma_k/2) \; , \tag{S9}$$

where κ is given by (42), and

$$\mathsf{B}_0 = d\left[2L^2 + (4/(3\kappa))\left\{d\tilde{L}^2 + L^4/m\right\}\right] \; , \tag{S10}$$

$$\mathsf{B}_1 = d\left[\kappa^{-1}L^4 + L^4/(6(m+L)) + m^{-1}\right] \; . \tag{S11}$$

2.1 Explicit bounds for fixed step size and fixed precision

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 $W_2(\delta_{x^\star} Q_\gamma^n, \pi)$ smaller than ε, when $(\gamma_k)_{k\geq 1}$ is constant, $\gamma_k = \gamma_\varepsilon$ for all $k \geq 1$.

Corollary S4. *Assume* ***H**1, **H**2 and **H**3. Let $x^\star$ 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*

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

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

We provide the dependencies of t minimal number of simulations n_ε as a function of the dimension d, the precision ε and the constants $m, L, \tilde{L}$ in Table 2.

Parameter	d	ε	L	m
Corollary S4	$d\log(d)$	$\varepsilon^{-1}\lvert\log(\varepsilon)\rvert$	L^2	$\lvert\log(m)\rvert\, m^{-2}$

Table 2: Dependencies of n

2.2 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 1.2. Recall that the function $\boldsymbol{\psi}$ is defined by (S5). For $\alpha \in (0,1]$, by (S6) and Lemma 23 applied with $\ell = \lceil n/2 \rceil$, where $\lceil \cdot \rceil$ is the ceiling function, we have

$$u_n^{(3)}(\gamma) \leq \sum_{j=1}^{2} \mathsf{B}_{j-1} \left[2\kappa^{-1}\gamma_1^{j+1}(n/2)^{-\alpha(j+1)} + \gamma_1^{j+2}\left(\boldsymbol{\psi}_{1-\alpha(j+2)}(\lceil n/2 \rceil) + 1\right) \right.$$
$$\left. \times \exp\left(-(\kappa\gamma_1/2)\left\{\boldsymbol{\psi}_{1-\alpha}(n+1) - \boldsymbol{\psi}_{1-\alpha}(\lceil n/2 \rceil)\right\}\right)\right] . \tag{S12}$$

2.3 Optimal strategy with a fixed number of iterations

Corollary S5. *Let $n \in \mathbb{N}^*$ be a fixed number of iteration. Assume* ***H**1*, ***H**2*, ***H**3* *and $(\gamma_k)_{k\geq 1}$ is a constant sequence, $\gamma_k = \gamma^\star$ 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 $\gamma^\star \in \left(0, (m+L)^{-1}\right)$. 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 $Cd^2 n^{-2}\log^2(n)$ as n goes to $+\infty$.

Similarly, we have the following result.

Corollary S6. *Assume* ***H**1*, ***H**2* *and* ***H**3*. *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 $\gamma_\alpha \in \left(0, (m+L)^{-1}\right)$. 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 $Cd^2 n^{-2}\log^2(n)$ as n goes to $+\infty$.

Proof. Follows from (S9), (S12) and the choice of γ_α. □

Note that in Corollary S5 and Corollary S6, we do not find the optimal convergence rates obtained for the sequence of step sizes $\gamma_k = \gamma_1/k$, for $k \geq 1$ and $\gamma_1 > 0$, up to a logarithmic factor $\log(n)$. This most likely due to the fact that the bounds (for example (S12)) used to compute the optimal parameters $\gamma^\star$ and γ_α are not the most appropriate.

3 Generalization of Theorem 5 and Theorem 8

In this section, we weaken the assumption $\gamma_1 \leq 1/(m+L)$ of Theorem 5 and Theorem 8. We assume now:

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

Under **G**1, we denote by

$$n_0 = \min\left\{k \in \mathbb{N} \mid \gamma_k \le 2/(m+L)\right\} \tag{S13}$$

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

Theorem S7. *Assume* ***H****1,* ***H****2 and* ***G****1. Then for all $n, p \in \mathbb{N}^*$, $n \le p$*

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

where

$$\begin{aligned}\mathsf{G}_{n,p}(\mu_0,\gamma) &= \exp\left(-\sum_{k=n}^{p} \gamma_k \kappa + \sum_{k=n}^{n_0-1} L^2\gamma_k^2\right) \int_{\mathbb{R}^d} \|x - x^\star\|^2 \, \mu_0(\mathrm{d}x) \\ &+ 2d\kappa^{-1} + 2d \left\{\prod_{k=n}^{n_0-1} (\gamma_{n_0-1}L^2)^{-1}\left(1 + L^2\gamma_k^2\right)\right\} \exp\left(-\sum_{k=n}^{p} \kappa\gamma_k + \sum_{k=n}^{n_0-1} \gamma_k^2 m L\right) \,. \end{aligned} \tag{S14}$$

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 $\nabla U(x^\star) = 0$, (42) and **H**1, we get from the previous inequality:

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

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 $p, n \in \mathbb{N}$, $p \le n$,

$$\int_{\mathbb{R}^d} \|x - x^\star\|^2 \, \mu_0 Q_n^p(\mathrm{d}x) \le \prod_{k=n}^{p} \eta_k \int_{\mathbb{R}^d} \|x - x^\star\| \, \mu_0(\mathrm{d}x) + (2d) \sum_{i=n}^{p} \prod_{k=i+1}^{p} \eta_k \gamma_i \,. \tag{S15}$$

For the first term of the right hand side, we simply use the bound, for all $x \in \mathbb{R}$, $(1 + x) \le \mathrm{e}^x$, and we get by **G**1

$$\prod_{k=n}^{p} \eta_k \le \exp\left(-\sum_{k=n}^{p} \kappa\gamma_k + \sum_{k=n}^{n_0-1} L^2\gamma_k^2\right) \,, \tag{S16}$$

where n_0 is defined in (S13). Consider now the second term in the right hand side of (S15).

$$\sum_{i=n}^{p} \prod_{k=i+1}^{p} \eta_k \gamma_i \leq \sum_{i=n_0}^{p} \prod_{k=i+1}^{p} (1-\kappa\gamma_k)\,\gamma_i + \sum_{i=n}^{n_0-1} \prod_{k=i+1}^{p} \eta_k \gamma_i$$
$$\leq \kappa^{-1} \sum_{i=n_0}^{p} \left\{ \prod_{k=i+1}^{p} (1-\kappa\gamma_k) - \prod_{k=i}^{p} (1-\kappa\gamma_k) \right\}$$
$$+ \left\{ \sum_{i=n}^{n_0-1} \prod_{k=i+1}^{n_0-1} \left(1+L^2\gamma_k^2\right) \gamma_i \right\} \prod_{k=n_0}^{p} (1-\kappa\gamma_k) \tag{S17}$$

Since $(\gamma_k)_{k\geq 1}$ is non-increasing, 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 $(1+x) \leq \mathrm{e}^x$ on $\mathbb{R}$:

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

Using the two previous inequalities in (S17), we get

$$\sum_{i=n}^{p} \prod_{k=i+1}^{p} \eta_k \gamma_i$$
$$\leq \kappa^{-1} + \left\{ \prod_{k=n}^{n_0-1} (\gamma_{n_0-1} L^2)^{-1} \left(1+L^2\gamma_k^2\right) \right\} \exp\left(-\sum_{k=n}^{p} \kappa\gamma_k + \sum_{k=n}^{n_0-1} \gamma_k^2 m L\right) . \tag{S18}$$

Combining (S16) and (S18) in (S15) concluded the proof. □

We now deal with bounds on $W_2(\mu_0 Q_\gamma^n, \pi)$ under **G** 1. But before we preface our result by some techical lemmas.

Lemma S8. *Assume **H** 1 and **H** 2. 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 (50). 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 $(Y_0, \overline{Y}_0)$.*

(i) For all $n \geq 0$, $\epsilon_1 > 0$ and $\epsilon_2 > 0$,

$$\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\left\|Y_{\Gamma_{n+1}} - \overline{Y}_{\Gamma_{n+1}}\right\|^2\right]$$
$$\leq \{1 - \gamma_{n+1}(\kappa - 2\epsilon_1) + \gamma_{n+1}L((1+\epsilon_2)\gamma_{n+1} - 2/(m+L))_+\}\left\|Y_{\Gamma_n} - \overline{Y}_{\Gamma_n}\right\|^2$$
$$+\gamma_{n+1}^2(1/(2\epsilon_1)+(1+\epsilon_2^{-1})\gamma_{n+1})\left(dL^2 + (L^4\gamma_{n+1}/2)\left\|Y_{\Gamma_n} - x^\star\right\|^2 + dL^4\gamma_{n+1}^2/12\right).$$

*(ii) If in addition **H**3 holds then for all $n \geq 0$, $\epsilon_1 > 0$ and $\epsilon_2 > 0$,*

$$\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\left\|Y_{\Gamma_{n+1}} - \overline{Y}_{\Gamma_{n+1}}\right\|^2\right]$$
$$\leq \{1 - \gamma_{n+1}(\kappa - 2\epsilon_1) + \gamma_{n+1}L((1+\epsilon_2)\gamma_{n+1} - 2/(m+L))_+\}\left\|Y_{\Gamma_n} - \overline{Y}_{\Gamma_n}\right\|^2$$
$$+ (2\epsilon_1)^{-1}\gamma_{n+1}^3\left\{(2L^4/3)\left\|Y_{\Gamma_n} - x^\star\right\|^2 + L^4 d\gamma_{n+1}/2 + 2d^2\tilde{L}^2/3\right\}$$
$$+ \gamma_{n+1}^3(1+\epsilon_2^{-1})\left(dL^2 + (L^4\gamma_{n+1}/2)\left\|Y_{\Gamma_n} - x^\star\right\|^2 + dL^4\gamma_{n+1}^2/12\right).$$

Proof. (i) Let $n \geq 0$ and $\epsilon_1 > 0$, and set $\Delta_n = Y_{\Gamma_n} - \overline{Y}_{\Gamma_n}$ by definition we have:

$$\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\|\Delta_{n+1}\|^2\right] = \|\Delta_n\|^2 + \mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\left\|\int_{\Gamma_n}^{\Gamma_{n+1}}\left\{\nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n})\right\}\mathrm{d}s\right\|^2\right] \tag{S19}$$
$$- 2\gamma_{n+1}\left\langle \Delta_n, \nabla U(Y_{\Gamma_n}) - \nabla U(\overline{Y}_{\Gamma_n})\right\rangle - 2\int_{\Gamma_n}^{\Gamma_{n+1}}\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\langle \Delta_n, \{\nabla U(Y_s) - \nabla U(Y_{\Gamma_n})\}\rangle\,\mathrm{d}s\right].$$

Using the two inequalities $|\langle a, b\rangle| \leq \epsilon_1\|a\|^2 + (4\epsilon_1)^{-1}\|b\|^2$ and (42), we get

$$\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\|\Delta_{n+1}\|^2\right] \leq \{1 - \gamma_{n+1}(\kappa - 2\epsilon_1)\}\|\Delta_n\|^2$$
$$- 2\gamma_{n+1}/(m+L)\left\|\nabla U(Y_{\Gamma_n}) - \nabla U(\overline{Y}_{\Gamma_n})\right\|^2 \tag{S20}$$
$$+ \mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\left\|\int_{\Gamma_n}^{\Gamma_{n+1}}\{\nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n})\}\mathrm{d}s\right\|^2\right]$$
$$+ (2\epsilon_1)^{-1}\int_{\Gamma_n}^{\Gamma_{n+1}}\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\|\nabla U(Y_s) - \nabla U(Y_{\Gamma_n})\|^2\right]\mathrm{d}s. \tag{S21}$$

Using $\|a+b\|^2 \leq (1+\epsilon_2)\|a\|^2 + (1+\epsilon_2^{-1})\|b\|^2$ and the Jensen's inequality, we have

$$\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\left\|\int_{\Gamma_n}^{\Gamma_{n+1}}\left\{\nabla U(Y_s) - \nabla U(\overline{Y}_{\Gamma_n})\right\}\mathrm{d}s\right\|^2\right] \leq (1+\epsilon_2)\gamma_{n+1}^2\left\|\nabla U(Y_{\Gamma_n}) - \nabla U(\overline{Y}_{\Gamma_n})\right\|^2$$
$$+ (1+\epsilon_2^{-1})\gamma_{n+1}\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\int_{\Gamma_n}^{\Gamma_{n+1}}\|\nabla U(Y_s) - \nabla U(Y_{\Gamma_n})\|^2\,\mathrm{d}s\right]. \tag{S22}$$

This result and **H**1 imply,

$$\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\|\Delta_{n+1}\|^2\right] \leq \{1-\gamma_{n+1}(\kappa-2\epsilon_1)+\gamma_{n+1}L((1+\epsilon_2)\gamma_{n+1}-2/(m+L))_+\}\|\Delta_n\|^2$$
$$+((1+\epsilon_2^{-1})\gamma_{n+1}+(2\epsilon_1)^{-1})\int_{\Gamma_n}^{\Gamma_{n+1}}\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\|\nabla U(Y_s)-\nabla U(Y_{\Gamma_n})\|^2\right]\mathrm{d}s\,. \quad \text{(S23)}$$

By **H**1, the Markov property of $(Y_t)_{t\geq 0}$ and Lemma 21, we have

$$\int_{\Gamma_n}^{\Gamma_{n+1}}\mathbb{E}^{\mathcal{F}'_{\Gamma_n}}\left[\|\nabla U(Y_s)-\nabla U(Y_{\Gamma_n})\|^2\right]\mathrm{d}s$$
$$\leq L^2\left(d\gamma_{n+1}^2+dL^2\gamma_{n+1}^4/12+(L^2\gamma_{n+1}^3/2)\|Y_{\Gamma_n}-x^\star\|^2\right)\,. \quad \text{(S24)}$$

Plugging this bound in (S23) concludes the proof.

(ii) Let $n\geq 0$ and $\epsilon>0$, and set $\Theta_n=Y_{\Gamma_n}-\overline{Y}_{\Gamma_n}$. Using Itô's formula, we have for all $s\in[\Gamma_n,\Gamma_{n+1})$,

$$\nabla U(Y_s)-\nabla U(Y_{\Gamma_n})=\int_{\Gamma_n}^{s}\left\{\nabla^2U(Y_u)\nabla U(Y_u)+\vec{\Delta}(\nabla U)(Y_u)\right\}\mathrm{d}u$$
$$+\sqrt{2}\int_{\Gamma_n}^{s}\nabla^2U(Y_u)\mathrm{d}B_u\,. \quad \text{(S25)}$$

Since Θ_n is $\mathcal{F}_{\Gamma_n}$-measurable and $(\int_0^s\nabla^2U(Y_u)\mathrm{d}B_u)_{s\in[0,\Gamma_{n+1}]}$ is a $(\mathcal{F}_s)_{s\in[0,\Gamma_{n+1}]}$-martingale under **H**1, by (S25) we have:

$$\left|\mathbb{E}^{\mathcal{F}_{\Gamma_n}}\left[\langle\Theta_n,\nabla U(Y_s)-\nabla U(Y_{\Gamma_n})\rangle\right]\right|$$
$$=\left|\left\langle\Theta_n,\mathbb{E}^{\mathcal{F}_{\Gamma_n}}\left[\int_{\Gamma_n}^{s}\left\{\nabla^2U(Y_u)\nabla U(Y_u)+\vec{\Delta}(\nabla U)(Y_u)\right\}\mathrm{d}u\right]\right\rangle\right|$$

Combining this equality, (S22) and $|\langle a,b\rangle|\leq\epsilon_1\|a\|^2+(4\epsilon_1)^{-1}\|b\|^2$ in we have

$$\mathbb{E}^{\mathcal{F}_{\Gamma_n}}\left[\|\Theta_{n+1}\|^2\right]\leq\{1-\gamma_{n+1}(\kappa-2\epsilon_1)+\gamma_{n+1}L((1+\epsilon_2)\gamma_{n+1}-2/(m+L))_+\}\|\Theta_n\|^2$$
$$+(2\epsilon_1)^{-1}A+(1+\epsilon_2^{-1})\gamma_{n+1}\mathbb{E}^{\mathcal{F}_{\Gamma_n}}\left[\int_{\Gamma_n}^{\Gamma_{n+1}}\|\nabla U(Y_s)-\nabla U(Y_{\Gamma_n})\|^2\mathrm{d}s\right]\,, \quad \text{(S26)}$$

where

$$A=\int_{\Gamma_n}^{\Gamma_{n+1}}\left\|\mathbb{E}^{\mathcal{F}_{\Gamma_n}}\left[\int_{\Gamma_n}^{s}\nabla^2U(Y_u)\nabla U(Y_u)+\vec{\Delta}(\nabla U)(Y_u)\mathrm{d}u\right]\right\|^2\mathrm{d}s\,.$$

We now separately bound the two last terms of the right hand side. By **H**1, the Markov property of $(Y_t)_{t\geq 0}$ and Lemma 21, we have

$$\int_{\Gamma_n}^{\Gamma_{n+1}}\mathbb{E}^{\mathcal{F}_{\Gamma_n}}\left[\|\nabla U(Y_s)-\nabla U(Y_{\Gamma_n})\|^2\right]\mathrm{d}s$$
$$\leq L^2\left(d\gamma_{n+1}^2+dL^2\gamma_{n+1}^4/12+(1/2)L^2\gamma_{n+1}^3\|Y_{\Gamma_n}-x^\star\|^2\right)\,. \quad \text{(S27)}$$

We now bound A. We get using Jensen's inequality, Fubini's theorem, $\nabla U(x^\star) = 0$ and (10)

$$
\begin{aligned}
A &\leq 2 \int_{\Gamma_n}^{\Gamma_{n+1}} (s-\Gamma_n) \int_{\Gamma_n}^{s} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\left\| \nabla^2 U(Y_u) \nabla U(Y_u) \right\|^2 \right] \mathrm{d}u \, \mathrm{d}s \\
&\qquad + 2 \int_{\Gamma_n}^{\Gamma_{n+1}} (s-\Gamma_n) \int_{\Gamma_n}^{s} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\left\| \vec{\Delta}(\nabla U)(Y_u) \right\|^2 \right] \mathrm{d}u \, \mathrm{d}s \\
&\leq 2 \int_{\Gamma_n}^{\Gamma_{n+1}} (s-\Gamma_n) L^4 \int_{\Gamma_n}^{s} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\|Y_u - x^\star\|^2 \right] \mathrm{d}u \, \mathrm{d}s + 2\gamma_{n+1}^3 d^2 \tilde{L}^2 / 3 \,. \qquad \text{(S28)}
\end{aligned}
$$

By Lemma 21-(i), the Markov property and for all $t \geq 0$, $1 - \mathrm{e}^{-t} \leq t$, we have for all $s \in [\Gamma_n, \Gamma_{n+1}]$,

$$
\int_{\Gamma_n}^{s} \mathbb{E}^{\mathcal{F}_{\Gamma_n} c\gamma_{n+1}^3} \left[\|Y_u - x^\star\|^2 \right] \mathrm{d}u \leq (2m)^{-1} (1 - \mathrm{e}^{-2m(s-\Gamma_n)}) \|Y_{\Gamma_n} - x^\star\|^2 + d(s-\Gamma_n)^2 \,.
$$

Using this inequality in (S28) and for all $t \geq 0$, $1 - \mathrm{e}^{-t} \leq t$, we get

$$
A \leq (2L^4 \gamma_{n+1}^3 / 3) \|Y_{\Gamma_n} - x^\star\|^2 + L^4 d \gamma_{n+1}^4 / 2 + 2\gamma_{n+1}^3 d^2 \tilde{L}^2 / 3 \,.
$$

Combining this bound and (S27) in (S26) concludes the proof.

□

Lemma S9. *Let $(\gamma_k)_{k \geq 1}$ be a non-increasing 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 $n \geq 0$, $j \geq 2$*

(i)

$$
\begin{aligned}
\sum_{i=1}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi + \gamma_k \beta (\gamma_k - \tau)_+) \gamma_i^j &\leq \sum_{i=N}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi) \gamma_i^j \\
&\quad + \left\{ \beta^{-1} \gamma_1^{j-2} \prod_{k=1}^{N-1} (1 + \gamma_k^2 \beta) \right\} \prod_{k=N}^{n+1} (1 - \varpi \gamma_k) \,.
\end{aligned}
$$

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

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

Proof. (i) By definition of N we have

$$
\begin{aligned}
&\sum_{i=1}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi + \gamma_k \beta (\gamma_k - \tau)_+) \gamma_i^j \\
&\leq \sum_{i=N}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi) \gamma_i^j + \left\{ \sum_{i=1}^{N-1} \prod_{k=i+1}^{N-1} (1 + \gamma_k^2 \beta) \gamma_i^j \right\} \prod_{k=N}^{n+1} (1 - \gamma_k \varpi) \,. \qquad \text{(S29)}
\end{aligned}
$$

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

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

Plugging this inequality in (S29) concludes the proof of (i).

(ii) Let $\ell\in\{N,\ldots,n+1\}$. Since $(\gamma_k)_{k\geq 1}$ is non-increasing and for every $x\in\mathbb{R}$, $(1+x)\leq \mathrm{e}^x$, we get

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

(iii)

□

Lemma S10. *Let $(\gamma_k)_{k\geq 1}$ be a non-increasing sequence of positive numbers, $\varpi,\beta,\tau>0$ be positive real numbers, and $N\geq 1$ satisfying the assumptions of Lemma S9. Let $\mathsf{P}\in\mathbb{N}^*$, $C_i\geq 0$, $i=0,\ldots,\mathsf{P}$ be positive constants and $(u_n)_{n\geq 0}$ be a sequence of real numbers with $u_0\geq 0$ satisfying for all $n\geq 0$*

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

Then for all $n\geq 1$,

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

Proof. This is a consequence of a straightforward induction and Lemma S9-(i). □

Theorem S11. *Assume **H1**, **H2** and **G1**.*

(i) For all $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ and $n \geq 1$,

$$W_2^2(\mu_0 Q_\gamma^n, \pi) \leq \tilde{u}_n^{(1)}(\gamma) W_2^2(\mu_0, \pi) + \tilde{u}_n^{(2)}(\gamma) \,, \tag{S30}$$

where

$$\tilde{u}_n^{(1)}(\gamma) = \left\{ \prod_{k=1}^{n_1-1} (1 + 2L^2\gamma_k^2) \right\} \prod_{k=n_1}^{n} (1 - \kappa\gamma_k/2) \,, \tag{S31}$$

$$\begin{aligned} \tilde{u}_n^{(2)}(\gamma) = & \sum_{i=n_1}^{n} \gamma_i^2 \mathsf{b}(\gamma_i) \prod_{k=i+1}^{n} (1 - \kappa\gamma_k/2) \\ & + \mathsf{b}(\gamma_1)(2L^2)^{-1} \left\{ \prod_{k=1}^{n_1-1} (1 + 2\gamma_k^2 L^2) \right\} \prod_{k=n_1}^{n} (1 - \kappa\gamma_k/2) \,, \end{aligned} \tag{S32}$$

with

$$\mathsf{b}(\gamma) = L^2 d \left\{ \kappa^{-1} + \gamma \right\} (2 + L^2\gamma/m + L^2\gamma^2/6) \,.$$

*(ii) If in addition **H**3 holds, for all $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ and $n \geq 1$,*

$$W_2^2(\mu_0 Q_\gamma^n, \pi) \leq \tilde{u}_n^{(1)}(\gamma) W_2^2(\mu_0, \pi) + \tilde{u}_n^{(3)}(\gamma) \,, \tag{S33}$$

where

$$\begin{aligned} \tilde{u}_n^{(3)}(\gamma) = & \sum_{i=n_1}^{n} \gamma_i^3 \mathsf{c}(\gamma_i) \prod_{k=i+1}^{n} (1 - \kappa\gamma_k/2) \\ & + \gamma_1 \mathsf{c}(\gamma_1)(2L^2)^{-1} \left\{ \prod_{k=1}^{n_1-1} (1 + 2\gamma_k^2 L^2) \right\} \prod_{k=n_1}^{n} (1 - \kappa\gamma_k/2) \,, \end{aligned} \tag{S34}$$

with

$$\mathsf{c}(\gamma) = d \left\{ 2L^2 + \gamma_i L^4 \left(\frac{\gamma_i}{6} + m^{-1} \right) + \kappa^{-1} \left(\frac{4d\tilde{L}^2}{3} + \gamma_i L^4 + \frac{4L^4}{3m} \right) \right\} \,.$$

Proof. (i) 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 (50). By definition of W_2 and since for all $t \geq 0$, π 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 Lemma S8-(i) and Lemma S10 using that for all $k \in \mathbb{N}$, $\mathbb{E}[\|Y_{\Gamma_k} - x^\star\|] \leq d/m$ by since Y_0 is distributed according to π.

(ii) The proof follows the same line as the first statement using Lemma S8-(ii) instead of Lemma S8-(i).

□

3.1 Explicit bound based on Theorem S11 for $\gamma_k = \gamma_1/k$

We give here a bound on the sequences $(\tilde{u}_n^{(1)}(\gamma))_{n\geq 1}$ and $(\tilde{u}_n^{(2)}(\gamma))_{n\geq 1}$, $(\tilde{u}_n^{(3)}(\gamma))_{n\geq 1}$ for $(\gamma_k)_{k\geq 1}$ defined by $\gamma_1 > 0$ and $\gamma_k = \gamma_1/k$. Recall that $\boldsymbol{\psi}_\beta$ is given by (S5). First note, since $(\gamma_k)_{k\geq 1}$ is non-increasing, for all $n \geq 1$, we have

$$\tilde{u}_n^{(2)}(\gamma) \leq \sum_{j=0}^{1} \mathsf{C}_j \sum_{i=n_1}^{n} \gamma_i^{j+2} \prod_{k=i+1}^{n} (1-\kappa\gamma_k/2) \\ + \sum_{j=0}^{1} \mathsf{C}_j (2L^2)^{-1} \gamma_1^j \left\{ \prod_{k=1}^{n_1-1} (1+2\gamma_k^2 L^2) \right\} \prod_{k=n_1}^{n} (1-\kappa\gamma_k/2) \ , \quad \text{(S35)}$$

where

$$\mathsf{C}_1 = 2\mathsf{b} d L^2 \ , \mathsf{C}_2 = \mathsf{b}(dL^4/m + \gamma_1 d L^4/6) \ , \mathsf{b} = \kappa^{-1} + \gamma_1 \ .$$

and

$$\tilde{u}_n^{(3)}(\gamma) \leq \sum_{j=0}^{1} \mathsf{D}_j \sum_{i=n_1}^{n} \gamma_i^{j+3} \prod_{k=i+1}^{n} (1-\kappa\gamma_k/2) \\ + \sum_{j=0}^{1} \mathsf{D}_j (2L^2)^{-1} \gamma_1^{j+1} \left\{ \prod_{k=1}^{n_1-1} (1+2\gamma_k^2 L^2) \right\} \prod_{k=n_1}^{n} (1-\kappa\gamma_k/2) \ , \quad \text{(S36)}$$

where

$$\mathsf{D}_1 = d\left[2L^2 + (4/(3\kappa))\left\{d\tilde{L}^2 + L^4/m\right\}\right] \ , \mathsf{D}_2 = d\left[\kappa^{-1}L^4 + L^4\gamma_1/(m+L) + m^{-1}\right] \ .$$

1. We first give explicit bound based on Theorem S11-(i). For $n_1 = 1$, by (S6) and (S7), we have

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

For $n_1 > 1$, since $(\gamma_k)_{k\geq 0}$ is non increasing, using again (S6), (S7), and the bound for $t \in \mathbb{R}$, $(1+t) \leq \mathrm{e}^t$, we get

$$\tilde{u}_n^{(1)}(\gamma) \leq (n+1)^{-\kappa\gamma_1/2} \exp\left\{\kappa\gamma_1\boldsymbol{\psi}_0(n_1)/2 + 2L^2\gamma_1^2(\boldsymbol{\psi}_{-1}(n_1-1)+1)\right\}$$
$$\tilde{u}_n^{(2)}(\gamma) \leq (n+1)^{-\kappa\gamma_1/2} \sum_{j=0}^{1} \mathsf{C}_j \Big(\gamma_1^{j+2}(\boldsymbol{\psi}_{\kappa\gamma_1/2-1-j}(n+1) - \boldsymbol{\psi}_{\kappa\gamma_1/2-1-j}(n_1)+1) \\ + (\gamma_1^j/(2L^2)) \exp\left\{\kappa\gamma_1\boldsymbol{\psi}_0(n_1)/2 + 2L^2\gamma_1^2(\boldsymbol{\psi}_{-1}(n_1-1)+1)\right\} \Big) \ .$$

Thus, for $\gamma_1 > 2\kappa^{-1}$, the bound given by Theorem S11-(i) is of order $\mathcal{O}(n^{-1})$.

2. We first give explicit bound based on Theorem S11-(ii). Note that bounds on $(\tilde{u}_n^{(1)}(\gamma))_{n\geq 1}$ have been obtained below. We just need to give some bounds on $(\tilde{u}_n^{(3)}(\gamma))_{n\geq 1}$. For $n_1 = 1$, by (S6), (S7), we have

$$\tilde{u}_n^{(3)}(\gamma) \leq (n+1)^{-\kappa\gamma_1/2} \sum_{j=0}^{1} \mathsf{D}_j \left\{ \gamma_1^{j+3}(\boldsymbol{\psi}_{\kappa\gamma_1/2-2-j}(n+1)+1) + (2L^2)^{-1}\gamma_1^{j+1} \right\} .$$

For $n_1 > 1$, since $(\gamma_k)_{k\geq 0}$ is non increasing, using again (S6), (S7), and the bound for $t \in \mathbb{R}$, $(1+t) \leq \mathrm{e}^t$, we get

$$\begin{aligned}\tilde{u}_n^{(3)}(\gamma) \leq (n+1)^{-\kappa\gamma_1/2} \sum_{j=0}^{1} \mathsf{D}_j \Big(& \gamma_1^{j+3}(\boldsymbol{\psi}_{\kappa\gamma_1/2-2-j}(n+1) - \boldsymbol{\psi}_{\kappa\gamma_1/2-2-j}(n_1)+1) \\ & + (\gamma_1^{j+1}/(2L^2)) \exp\left\{\kappa\gamma_1\boldsymbol{\psi}_0(n_1)/2 + 2L^2\gamma_1^2(\boldsymbol{\psi}_{-1}(n_1-1)+1)\right\}\Big) .\end{aligned}$$

Thus, for $\gamma_1 > 4\kappa^{-1}$, the bound given by Theorem S11-(i) is of order $\mathcal{O}(n^{-1})$.

4 Explicit bounds on the MSE

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

4.1 Explicit bounds based on Theorem 5

1. First for $\alpha = 0$, recall by Theorem 5 and (S1) we have for all $p \geq 1$,

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

where A and A_1 are given by (S2) and (S3) respectively. Set

$$\mathsf{A} = \mathsf{A} \vee (\mathsf{A}_1/(m+L))$$

So by (27) and Lemma 23, we have the following bound for the bias

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

Therefore plugging this inequality and the one given by Theorem 15 implies:

$$\mathrm{MSE}_f^{N,n} \leq C\left(\kappa^{-1}\mathsf{A}\gamma_1 + \frac{\kappa^{-2} + \kappa^{-1}\exp(-\kappa N\gamma_1/2)\Omega(x)}{n\gamma_1}\right) . \tag{S37}$$

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

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

and (S37) becomes if $\gamma_1 \leftarrow \gamma_{\star,0}(n)$,

$$\mathrm{MSE}_f^{N,n} \leq C(C_{\mathrm{MSE},0}n)^{-1/2}\left(\kappa^{-1}\exp(-\kappa N\gamma_{\star,0}(n)/2)\Omega(x) + C_{\mathrm{MSE},0}\right) \;.$$

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

$$\mathrm{MSE}_f^{N_0(n),n} \leq C(C_{\mathrm{MSE},0}/n)^{1/2} \;.$$

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

2. For $\alpha \in (0, 1/2)$ by Theorem 5, Lemma 23, (S6) and (S8), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 15 implies:

$$\mathrm{MSE}_f^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{A}\gamma_1}{(1-2\alpha)n^\alpha} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_1 N^{1-\alpha}/(2(1-\alpha))\right\}\Omega(x) + \kappa^{-2}}{\gamma_1 n^{1-\alpha}}\right) \;. \tag{S38}$$

At fixed γ_1, this bound is of order $n^{-\alpha}$, and is better than (S37) 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

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

(S38) becomes with $\gamma_1 \leftarrow \gamma_{\star,\alpha}(n)$,

$$\begin{aligned}&\mathrm{MSE}_f^{N,n} \\ &\leq C(C_{\mathrm{MSE},\alpha}n)^{-1/2}\left(\kappa^{-1}\exp\left\{-\kappa N^{1-\alpha}\gamma_{\star,\alpha}(n)/(2(1-\alpha))\right\}\Omega(x) + C_{\mathrm{MSE},\alpha}\right) \;.\end{aligned}$$

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

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

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

3. If $\alpha = 1/2$, by Theorem 5, Lemma 23, (S6) and (S8), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 15 implies:

$$\mathrm{MSE}_f^{N,n} \le C\left(\frac{\kappa^{-1}\mathsf{A}\gamma_1\log(n)}{n^{1/2}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_1 N^{1/2}/4\right\}\Omega(x) + \kappa^{-2}}{\gamma_1 n^{1/2}}\right) . \tag{S39}$$

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

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

and (S39) becomes with $\gamma_1 \leftarrow \gamma_{\star,1/2}(n)$,

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

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

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

4. For $\alpha \in (1/2, 1]$, by Theorem 5, Lemma 23, (S6) and (S8), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 15 implies:

$$\mathrm{MSE}_f^{N,n} \le C\left(\frac{\kappa^{-1}\mathsf{A}\gamma_1}{n^{1-\alpha}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_1 N^{1-\alpha}/(2(1-\alpha))\right\}\Omega(x) + \kappa^{-2}}{\gamma_1 n^{1-\alpha}}\right) .$$

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 Section 3.1, for $\gamma_1 > 2\kappa^{-1}$ there exists $\tilde{C}_1 \ge 0$, independent of d and n such that the bias is upper bounded by

$$\left\{\mathbb{E}_x[\hat{\pi}_n^N(f)] - \pi(f)\right\}^2 = \tilde{C}_1/\log(n) .$$

Plugging this inequality and the one given by Theorem 15 implies there exists $\tilde{C} \ge 0$, independent of d and n such that the bias is upper bounded by

$$\mathrm{MSE}_f^{N,n} = \tilde{C}_2/\log(n) .$$

4.2 Explicit bound based on Theorem 8

1. First for $\alpha = 0$, recall by Theorem 8 and (S9) we have for all $p \geq 1$,

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

where B_0 and B_1 are given by (S10) and (S11) respectively. Set

$$\mathsf{B} = \mathsf{B}_0 \vee (\mathsf{B}_1/(m + L))$$

So by and Lemma 23, we have the following bound for the bias

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

Therefore plugging this inequality and the one given by Theorem 15 implies:

$$\mathrm{MSE}_f^{N,n} \leq C\left(\kappa^{-1}\mathsf{B}\gamma_1^2 + \frac{\kappa^{-2} + \kappa^{-1}\exp(-\kappa N\gamma_1/2)\Omega(x)}{n\gamma_1}\right) . \tag{S40}$$

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

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

and (S37) becomes if $\gamma_1 \leftarrow \gamma_{\star,0}(n)$,

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

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

$$\mathrm{MSE}_f^{N_0(n),n} \leq C(\mathsf{B}^{-1/2}\kappa^{5/2}n)^{-2/3} .$$

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

2. For $\alpha \in (0, 1/3)$ by Theorem 8, Lemma 23, (S6) and (S12), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 15 implies:

$$\mathrm{MSE}_f^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{B}\gamma_1^2}{(1-3\alpha)n^{2\alpha}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_1 N^{1-\alpha}/(2(1-\alpha))\right\}\Omega(x) + \kappa^{-2}}{\gamma_1 n^{1-\alpha}}\right) . \tag{S41}$$

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

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

(S41) becomes with $\gamma_1 \leftarrow \gamma_{\star,\alpha}(n)$,

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

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

$$\mathrm{MSE}_f^{N_\alpha(n),n} \leq C(\mathsf{B}^{-1/2}\kappa^{5/2}n)^{-2/3} .$$

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 8, Lemma 23, (S6) and (S12), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 15 implies:

$$\mathrm{MSE}_f^{N,n} \leq C\left(\frac{\kappa^{-1}\mathsf{B}\gamma_1^2\log(n)}{n^{2/3}} + \frac{\kappa^{-1}\exp\left\{-\kappa\gamma_1 N^{2/3}/4\right\}\Omega(x) + \kappa^{-2}}{\gamma_1 n^{2/3}}\right) . \tag{S42}$$

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

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

and (S42) becomes with $\gamma_1 \leftarrow \gamma_{\star,1/2}(n)$,

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

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

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

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 8, Lemma 23, (S6) and (S12), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 15 implies:

$$\mathrm{MSE}_f^{N,n} \leq C \left(\frac{\kappa^{-1} \mathsf{B} \gamma_1}{n^{1-\alpha}} + \frac{\kappa^{-1} \exp\left\{-\kappa\gamma_1 N^{1-\alpha}/(2(1-\alpha))\right\} \Omega(x) + \kappa^{-2}}{\gamma_1 n^{1-\alpha}} \right) .$$

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 Section 3.1, for $\gamma_1 > 2\kappa^{-1}$ there exists $\tilde{C}_1 \geq 0$, independent of d and n such that the bias is upper bounded by

$$\left\{\mathbb{E}_x[\hat{\pi}_n^N(f)] - \pi(f)\right\}^2 = \tilde{C}_1/\log(n) .$$

Plugging this inequality and the one given by Theorem 15 implies there exists $\tilde{C} \geq 0$, independent of d and n such that the bias is upper bounded by

$$\mathrm{MSE}_f^{N,n} = \tilde{C}_2/\log(n) .$$

References

[1] A. S. Dalalyan. Theoretical guarantees for approximate sampling from smooth and log-concave densities. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 79(3):651–676, 2017.

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