



HAL
open science

Supplement to "Sampling from a strongly log-concave distribution with the Unadjusted Langevin Algorithm"

Alain Durmus, Éric Moulines

► **To cite this version:**

Alain Durmus, Éric Moulines. Supplement to "Sampling from a strongly log-concave distribution with the Unadjusted Langevin Algorithm". 2016. hal-01176084v2

HAL Id: hal-01176084

<https://hal.science/hal-01176084v2>

Preprint submitted on 25 Apr 2016 (v2), last revised 15 Jul 2018 (v5)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Supplement to Sampling from a strongly log-concave distribution with the Unadjusted Langevin Algorithm

Alain Durmus¹ Éric Moulines²

April 25, 2016

1 Discussion of Theorem 3

Note that

$$u_n^{(2)}(\gamma) \leq \sum_{i=1}^n \{A_0 \gamma_i^2 + A_1 \gamma_i^3\} \prod_{k=i+1}^n (1 - \kappa \gamma_k / 2), \quad (\text{S1}) \quad \text{--eq:bound'u'n'2"}$$

where κ is given by (4), and

$$A_0 \stackrel{\text{def}}{=} 2L^2 \kappa^{-1} d, \quad (\text{S2}) \quad \text{--eq:definition-C0"}$$

$$A_1 \stackrel{\text{def}}{=} 2L^2 \kappa^{-1} d(m+L)^{-1} + dL^4 (\kappa^{-1} + (m+L)^{-1})(m^{-1} + 6^{-1}(m+L)^{-1}). \quad (\text{S3}) \quad \text{--eq:definition-C1"}$$

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

Corollary 1 (of Theorem 3). *Assume **H1** and **H2**. Let x^* be the unique minimizer of U . Let $x \in \mathbb{R}^d$ and $\epsilon > 0$. Set for all $k \in \mathbb{N}$, $\gamma_k = \gamma$ with*

$$\begin{aligned} \gamma &= \frac{-2A_0 \kappa^{-1} + (4\kappa^{-1}(A_0^2 \kappa^{-1} + \epsilon^2 A_1))^{1/2}}{4A_1 \kappa^{-1}} \wedge (m+L)^{-1}, \quad (\text{S4}) \quad \text{--eq:const'gamma'target"} \\ n &= \lceil 2(\kappa \gamma)^{-1} \{-\log(\epsilon^2/4) + \log(2d/m)\} \rceil. \end{aligned}$$

Then $W_2(\delta_{x^*} Q_\gamma^n, \pi) \leq \epsilon$.

Note that if γ is given by (S4), and is different from $1/(m+L)$, then $\gamma \leq \epsilon(4A_1 \kappa^{-1})^{-1/2}$ and $2\kappa^{-1}(A_0 \gamma + A_1 \gamma^2) = \epsilon^2/2$. Therefore,

$$\gamma \geq (\epsilon^2 \kappa / 4) \left\{ A_0 + \epsilon(A_1 / (4\kappa))^{1/2} \right\}^{-1}.$$

It is shown in [1, Corollary 1] that under **H2**, for constant step size for any $\epsilon > 0$, we can choose γ and $n \geq 1$ such that if for all $k \geq 1$, $\gamma_k = \gamma$, then $\|\nu^* Q_\gamma^n - \pi\|_{\text{TV}} \leq \epsilon$ where ν^* is the Gaussian measure on \mathbb{R}^d with mean x^* and covariance matrix $L^{-1} \text{I}_d$. We stress that the results in [1,

¹LTCI, Telecom ParisTech 46 rue Barrault, 75634 Paris Cedex 13, France. alain.durmus@telecom-paristech.fr

²Centre de Mathématiques Appliquées, UMR 7641, Ecole Polytechnique, France. eric.moulines@polytechnique.edu

corollary 1] hold only for a particular choice of the initial distribution ν^* , (which might seem a rather artificial assumption) whereas Theorem 3 hold for any initial distribution in $\mathcal{P}_2(\mathbb{R}^d)$.

We compare the optimal value of γ and n obtained from Corollary 1 with those given in [1, Corollary 1]. This comparison is summarized in Table 1 and Table 2; 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 [1, Corollary 1].

Parameter	d	ϵ	L	m
Theorem 3 and (21)	d^{-1}	ϵ^2	L^{-2}	m^2
[1, Corollary 1]	d^{-2}	ϵ^2	L^{-2}	m

Table 1: Dependencies of γ

Parameter	d	ϵ	L	m
Theorem 3 and (21)	$d \log(d)$	$\epsilon^{-2} \log(\epsilon) $	L^2	$ \log(m) m^{-3}$
[1, Corollary 1]	d^3	$\epsilon^{-2} \log(\epsilon) $	L^3	$ \log(m) m^{-2}$

Table 2: Dependencies of n

1.1 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 $t \in \mathbb{R}_+^*$,

$$\psi_\beta(t) = \begin{cases} (t^\beta - 1)/\beta & \text{for } \beta \neq 0 \\ \log(t) & \text{for } \beta = 0. \end{cases} \quad (\text{S5})$$

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

$$\psi_{1-a}(p+1) - \psi_{1-a}(n) \leq \sum_{k=n}^p k^{-a} \leq \psi_{1-a}(p) - \psi_{1-a}(n) + 1, \quad (\text{S6})$$

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

$$\sum_{k=n}^p k^{-a} \leq \psi_{1-a}(p+1) - \psi_{1-a}(n) + 1. \quad (\text{S7})$$

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

$$u_n^{(1)}(\gamma) \leq (n+1)^{-\kappa\gamma_1/2}, \quad u_n^{(2)}(\gamma) \leq (n+1)^{-\kappa\gamma_1/2} \sum_{j=0}^1 A_j (\psi_{\kappa\gamma_1/2-1-j}(n+1) + 1).$$

2. For $\alpha \in (0, 1)$, by (S6) and Lemma 14 applied with $\ell = \lceil n/2 \rceil$, where $\lceil \cdot \rceil$ is the ceiling function, we have

$$\begin{aligned} u_n^{(1)}(\gamma) &\leq \exp(-\kappa\gamma_1 \psi_{1-\alpha}(n+1)/2) \\ u_n^{(2)}(\gamma) &\leq \sum_{j=0}^1 A_j \left(2\kappa^{-1} \gamma_1^{j+1} (n/2)^{-\alpha(j+1)} + \gamma_1^{j+2} (\psi_{1-\alpha(j+2)}(\lceil n/2 \rceil) + 1) \right) \\ &\quad \times \exp\{-(\kappa\gamma_1/2)(\psi_{1-\alpha}(n+1) - \psi_{1-\alpha}(\lceil n/2 \rceil))\}. \end{aligned} \quad (\text{S8})$$

1.2 Optimal strategy with a fixed number of iterations

Corollary 2. 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

$$\begin{aligned}\gamma^+ &\stackrel{\text{def}}{=} 2(\kappa n)^{-1}(\log(\kappa n/2) + \log(2(\|x - x^*\|^2 + d/m))) - \log(2\kappa^{-1}\mathbf{A}_0) \\ \gamma_- &\stackrel{\text{def}}{=} 2(\kappa n)^{-1}(\log(\kappa n/2) + \log(2(\|x - x^*\|^2 + d/m))) - \log(2\kappa^{-1}(\mathbf{A}_0 + 2\mathbf{A}_1(m+L)^{-1})).\end{aligned}$$

Assume $\gamma^+ \in (0, (m+L)^{-1})$. Then, the optimal choice of γ to minimize the bound on $W_2(\delta_x Q_n^\gamma, \pi)$ given by Theorem 3 belongs to $[\gamma_-, \gamma^+]$. Moreover if $\gamma = \gamma_+$, then the bound on $W_2^2(\delta_x Q_n^\gamma, \pi)$ is equivalent to $4\mathbf{A}_0(\kappa^2 n)^{-1} \log(\kappa n/2)$ as $n \rightarrow +\infty$.

Similarly, we have the following result.

Corollary 3. Assume **H1** and **H2**. Let $(\gamma_k)_{k \geq 1}$ be the decreasing sequence, defined by $\gamma_k = \gamma_\alpha \kappa^{-\alpha}$, with $\alpha \in (0, 1)$. Let $n \geq 1$ and set

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

Assume $\gamma_\alpha \in (0, (m+L)^{-1})$. Then the bound on $W_2^2(\delta_x Q_n^\gamma, \pi)$ given by Theorem 3 is smaller for large n than $8(1-\alpha)\mathbf{A}_0(\kappa^2 n)^{-1} \log(\kappa n/(2(1-\alpha)))$.

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

2 Discussion of Theorem 6

Based on Theorem 6, we can follow the same discussion than for Theorem 3. Note that

$$u_n^{(3)}(\gamma) \leq \sum_{i=1}^n \{B_0 \gamma_i^3 + B_1 \gamma_i^4\} \prod_{k=i+1}^n (1 - \kappa \gamma_k/2), \quad (\text{S9})$$

where κ is given by (4), and

$$B_0 \stackrel{\text{def}}{=} d \left(2L^2 + \kappa^{-1}(\tilde{L}^2/3 + 4L^4/(3m)) \right), \quad (\text{S10})$$

$$B_1 \stackrel{\text{def}}{=} d \left(\kappa^{-1}L^4 + L^4(1/(6(m+L)) + m^{-1}) \right). \quad (\text{S11})$$

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

Corollary 4 (of Theorem 6). Assume **H1**, **H2** and **H3**. Let x^* be the unique minimizer of U . Let $x \in \mathbb{R}^d$ and $\epsilon > 0$. Set for all $k \in \mathbb{N}$, $\gamma_k = \gamma$ with

$$\gamma = (\epsilon/2)\kappa^{-1}(\mathbf{B}_0 + \mathbf{B}_1(m+L)^{-1})^{-1/2}, \quad (\text{S12})$$

$$n = \lceil 2(\kappa\gamma)^{-1} \{-\log(\epsilon^2/4) + \log(2d/m)\} \rceil. \quad (\text{S13})$$

Then $W_2(\delta_{x^*} Q_n^\gamma, \pi) \leq \epsilon$.

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, \tilde{L} in Table 3 and Table 4.

Parameter	d	ϵ	L	m
Theorem 6 and (21)	$d^{-1/2}$	ϵ	$L^{-7/2}$	m^2

Table 3: Dependencies of γ

tab:comparison'gammaD

Parameter	d	ϵ	L	m
Theorem 6 and (21)	$d^{1/2} \log(d)$	$\epsilon^{-1} \log(\epsilon) $	$L^{9/2}$	$ \log(m) m^{-3}$

Table 4: Dependencies of n

tab:comparison'nD

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

explicit'bound'k'alphaD

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 1.1. Recall that the function ψ is defined by (S5).

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

$$u_n^{(3)}(\gamma) \leq (n+1)^{-\kappa\gamma_1/2} \sum_{j=1}^2 \mathbf{B}_{j-1}(\psi_{\kappa\gamma_1/2-1-j}(n+1) + 1).$$

2. For $\alpha \in (0, 1)$, by (S6) and Lemma 14 applied with $\ell = \lceil n/2 \rceil$, where $\lceil \cdot \rceil$ is the ceiling function, we have

$$u_n^{(3)}(\gamma) \leq \sum_{j=1}^2 \mathbf{B}_{j-1} \left(2\kappa^{-1} \gamma_1^{j+1} (n/2)^{-\alpha(j+1)} + \gamma_1^{j+2} (\psi_{1-\alpha(j+2)}(\lceil n/2 \rceil) + 1) \right. \\ \left. \times \exp\{-(\kappa\gamma_1/2)(\psi_{1-\alpha}(n+1) - \psi_{1-\alpha}(\lceil n/2 \rceil))\} \right). \quad (\text{S14})$$

-eq:bound'n'alphaD

2.2 Optimal strategy with a fixed number of iterations

sapp:optimal'strategyD

coro:fixed'iterationD

Corollary 5. *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*

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

Assume $\gamma^ \in (0, (m+L)^{-1})$. Then the bound on $W_2^2(\delta_x Q_n^\gamma, \pi)$ given by Theorem 6 is of order $\mathcal{O}(\mathbf{B}_0(\kappa^2 n)^{-2} \log^2(n))$ as $n \rightarrow +\infty$.*

Similarly, we have the following result.

Corollary 6. *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*

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

Assume $\gamma_\alpha \in (0, (m+L)^{-1})$. Then the bound on $W_2^2(\delta_x Q_n^\gamma, \pi)$ given by Theorem 3 is smaller for large n than $\mathcal{O}(\mathbf{B}_0(\kappa^2 n)^{-2} \log^2(n))$.

Proof. Follows from (S9), (S14) and the choice of γ_α . □

3 Generalization of Theorem 3

In this section, we weaken the assumption $\gamma_1 \leq 1/(m+L)$ of Theorem 3. We assume now:

assum:gamma'sum

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

$$n_0 \stackrel{\text{def}}{=} \min \{k \in \mathbb{N} \mid \gamma_k \leq 2/(m+L)\} \quad (\text{S15})$$

-eq:def:index'n'0"

Denote in the sequel $(\cdot)_+ = \max(\cdot, 0)$. Recall that under **H2**, x^* is the unique minimizer of U , and κ is defined in (6)

theo:kind'drift'annexe

Theorem S7. *Assume **H1**, **H2** and **G1**. Then for all $n, p \in \mathbb{N}^*$, $n \leq p$*

$$\int_{\mathbb{R}^d} \|x - x^*\|^2 \mu_0 Q_n^p(dx) \leq E_{n,p}(\mu_0, \gamma),$$

where

$$\begin{aligned} E_{n,p}(\mu_0, \gamma) &\stackrel{\text{def}}{=} \exp\left(-\sum_{k=n}^p \gamma_k \kappa + \sum_{k=n}^{n_0-1} L^2 \gamma_k^2\right) \int_{\mathbb{R}^d} \|x - x^*\|^2 \mu_0(dx) \\ &+ 2d\kappa^{-1} + 2d \left\{ \prod_{k=n}^{n_0-1} (\gamma_{n_0-1} L^2)^{-1} (1 + L^2 \gamma_k^2) \right\} \exp\left(-\sum_{k=n}^p \kappa \gamma_k + \sum_{k=n}^{n_0-1} \gamma_k^2 mL\right). \end{aligned} \quad (\text{S16})$$

-eq:def borneKindDriftA

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

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

Using that $\nabla U(x^*) = 0$, (4) and **H1**, we get from the previous inequality:

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

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

$$\int_{\mathbb{R}^d} \|x - x^*\|^2 \mu_0 Q_n^p(dx) \leq \prod_{k=n}^p \eta_k \int_{\mathbb{R}^d} \|x - x^*\|^2 \mu_0(dx) + (2d) \sum_{i=n}^p \prod_{k=i+1}^p \eta_k \gamma_i, \quad (\text{S17})$$

-eq:lem'kind'drift'first'es

with the convention that for $n, p \in \mathbb{N}$, $n < p$, $\prod_p^n = 1$. 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 **G1**

$$\prod_{k=n}^p \eta_k \leq \exp\left(-\sum_{k=n}^p \kappa \gamma_k + \sum_{k=n}^{n_0-1} L^2 \gamma_k^2\right), \quad (\text{S18})$$

-eq:lem'kind'drift'second

where n_0 is defined in (S15). Consider now the second term in the right hand side of (S17).

$$\begin{aligned}
\sum_{i=n}^p \prod_{k=i+1}^p \eta_k \gamma_i &\leq \sum_{i=n_0}^p \prod_{k=i+1}^p (1 - \kappa \gamma_k) \gamma_i + \sum_{i=n}^{n_0-1} \prod_{k=i+1}^p \eta_k \gamma_i \\
&\leq \kappa^{-1} \sum_{i=n_0}^p \left\{ \prod_{k=i+1}^p (1 - \kappa \gamma_k) - \prod_{k=i}^p (1 - \kappa \gamma_k) \right\} \\
&\quad + \left\{ \sum_{i=n}^{n_0-1} \prod_{k=i+1}^{n_0-1} (1 + L^2 \gamma_k^2) \gamma_i \right\} \prod_{k=n_0}^p (1 - \kappa \gamma_k)
\end{aligned} \tag{S19}$$

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

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

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

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

Using the two previous inequalities in (S19), we get

$$\sum_{i=n}^p \prod_{k=i+1}^p \eta_k \gamma_i \leq \kappa^{-1} + \left\{ \prod_{k=n}^{n_0-1} (\gamma_{n_0-1} L^2)^{-1} (1 + L^2 \gamma_k^2) \right\} \exp \left(- \sum_{k=n}^p \kappa \gamma_k + \sum_{k=n}^{n_0-1} \gamma_k^2 m L \right) . \tag{S20}$$

Combining (S18) and (S20) in (S17) concluded the proof. \square

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 S8. *Assume **H1** and **H2**. Let $\zeta_0 \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, $(Y_t, \bar{Y}_t)_{t \geq 0}$ such that (Y_0, \bar{Y}_0) is distributed according to ζ_0 and given by (12). 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, \bar{Y}_0) . Then for all $n \geq 0$, $\epsilon_1 > 0$ and $\epsilon_2 > 0$,*

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

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

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

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

$$\begin{aligned} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\|\Delta_{n+1}\|^2 \right] &\leq \{1 - \gamma_{n+1}(\kappa - 2\epsilon_1)\} \|\Delta_n\|^2 - 2\gamma_{n+1}/(m+L) \|\nabla U(Y_{\Gamma_n}) - \nabla U(\bar{Y}_{\Gamma_n})\|^2 \\ &\quad + \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\left\| \int_{\Gamma_n}^{\Gamma_{n+1}} \{ \nabla U(Y_s) - \nabla U(\bar{Y}_{\Gamma_n}) \} ds \right\|^2 \right] + \frac{1}{2\epsilon_1} \int_{\Gamma_n}^{\Gamma_{n+1}} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\|\nabla U(Y_s) - \nabla U(Y_{\Gamma_n})\|^2 \right] ds . \end{aligned} \tag{S21}$$

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

$$\begin{aligned} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\left\| \int_{\Gamma_n}^{\Gamma_{n+1}} \{ \nabla U(Y_s) - \nabla U(\bar{Y}_{\Gamma_n}) \} ds \right\|^2 \right] &\leq (1 + \epsilon_2) \gamma_{n+1}^2 \|\nabla U(Y_{\Gamma_n}) - \nabla U(\bar{Y}_{\Gamma_n})\|^2 \\ &\quad \gamma_{n+1} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\int_{\Gamma_n}^{\Gamma_{n+1}} \|\nabla U(Y_s) - \nabla U(Y_{\Gamma_n})\|^2 ds \right] . \end{aligned}$$

This result and **H1** imply,

$$\begin{aligned} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\|\Delta_{n+1}\|^2 \right] &\leq \{1 - \gamma_{n+1}(\kappa - 2\epsilon_1) + \gamma_{n+1}L((1 + \epsilon_2)\gamma_{n+1} - 2/(m+L))_+\} \|\Delta_n\|^2 \\ &\quad + ((1 + \epsilon_2^{-1})\gamma_{n+1} + (2\epsilon_1)^{-1}) \int_{\Gamma_n}^{\Gamma_{n+1}} \mathbb{E}^{\mathcal{F}_{\Gamma_n}} \left[\|\nabla U(Y_s) - \nabla U(Y_{\Gamma_n})\|^2 \right] ds . \end{aligned} \tag{S22}$$

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

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

Plugging this bound in (S22) concludes the proof. \square

Lemma S9. 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 $n \geq 0$, $j \geq 2$

(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 $\ell \in \{N, \dots, n\}$,

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

Proof. By definition of N we have

$$\begin{aligned} & \sum_{i=1}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi + \gamma_k \beta (\gamma_k - \tau)_+) \gamma_i^j \\ & \leq \sum_{i=N}^{n+1} \prod_{k=i+1}^{n+1} (1 - \gamma_k \varpi) \gamma_i^j + \left\{ \sum_{i=1}^{N-1} \prod_{k=i+1}^{N-1} (1 + \gamma_k^2 \beta) \gamma_i^j \right\} \prod_{k=N}^{n+1} (1 - \gamma_k \varpi). \end{aligned} \quad (\text{S23})$$

-eq:suite' recurrence' 2'1'

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

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

Plugging this inequality in (S23) concludes the proof of (i). Let $\ell \in \{N, \dots, n+1\}$. Since $(\gamma_k)_{k \geq 1}$ is nonincreasing and for every $x \in \mathbb{R}$, $(1+x) \leq e^x$, we get

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

□

Lemma S10. 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 S9. Let $\mathbf{P} \in \mathbb{N}^*$, $C_i \geq 0$, $i = 0, \dots, \mathbf{P}$ be positive constants and $(u_n)_{n \geq 0}$ be a sequence of real numbers with $u_0 \geq 0$ satisfying for all $n \geq 0$

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

Then for all $n \geq 1$,

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

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

Proposition S11. Assume **H 1**, **H 2** and **G 1**. Let x^* be the unique minimizer of U . Let $\zeta_0 \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, $(Y_t, \bar{Y}_t)_{t \geq 0}$ such that (Y_0, \bar{Y}_0) is distributed according to ζ_0 and given by (12). Then for all $n \geq 0$ and $t \in [\Gamma_n, \Gamma_{n+1}]$:

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

where

$$\tilde{u}_n^{(1)}(\gamma) \stackrel{\text{def}}{=} \left\{ \prod_{k=1}^{n_1-1} (1 + L^2(1 + \rho)\gamma_k^2) \right\} \prod_{k=n_1}^n (1 - \kappa\gamma_k/2), \quad (\text{S24})$$

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

where for $\gamma > 0$,

$$\begin{aligned} f(\gamma) &= \{2\kappa^{-1} + (1 + \rho^{-1})\gamma\} (dL^2 + \delta L^4\gamma/2 + dL^4\gamma^2/12), \\ \delta &= \max_{i \geq 1} \left\{ e^{-2m\Gamma_{i-1}} \mathbb{E} \left[\|Y_0 - x^*\|^2 \right] + (1 - e^{-2m\Gamma_{i-1}})(d/m) \right\}, \end{aligned}$$

and

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

where $E_{1,n}(\mu_0, \gamma)$ is given by (S16) and μ_0 is the initial distribution of \bar{Y}_0 .

Proof. Lemma S8 with $\epsilon_1 = \kappa/4$ and $\epsilon_2 = \rho$, **G 1**, Lemma S9, ((i)) imply for all $n \geq 0$

$$\mathbb{E} \left[\|Y_{\Gamma_n} - \bar{Y}_{\Gamma_n}\|^2 \right] \leq \tilde{u}_n^{(1)}(\gamma) \mathbb{E} \left[\|Y_0 - \bar{Y}_0\|^2 \right] + \tilde{u}_n^{(4)}(\gamma). \quad (\text{S25})$$

Now let $n \geq 0$ and $t \in [\Gamma_n, \Gamma_{n+1}]$. By (12),

$$\|Y_t - \bar{Y}_t\|^2 = \|Y_{\Gamma_n} - \bar{Y}_{\Gamma_n}\|^2 - 2 \int_{\Gamma_n}^t \langle \nabla U(Y_s) - \nabla U(\bar{Y}_{\Gamma_n}), Y_s - \bar{Y}_s \rangle ds. \quad (\text{S26})$$

Moreover for all $s \in [\Gamma_n, \Gamma_{n+1}]$, by (4) we get

$$\begin{aligned} \langle \nabla U(Y_s) - \nabla U(\bar{Y}_{\Gamma_n}), Y_s - \bar{Y}_s \rangle &= \langle \nabla U(Y_s) - \nabla U(\bar{Y}_{\Gamma_n}), Y_s - \bar{Y}_{\Gamma_n} + \bar{Y}_{\Gamma_n} - \bar{Y}_s \rangle \\ &\geq (m+L)^{-1} \|\nabla U(Y_s) - \nabla U(\bar{Y}_{\Gamma_n})\|^2 + \langle \nabla U(Y_s) - \nabla U(\bar{Y}_{\Gamma_n}), \bar{Y}_{\Gamma_n} - \bar{Y}_s \rangle. \end{aligned} \quad (\text{S27})$$

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

$$\langle \nabla U(Y_s) - \nabla U(\bar{Y}_{\Gamma_n}), \bar{Y}_{\Gamma_n} - \bar{Y}_s \rangle \geq -(m+L)^{-1} \|\nabla U(Y_s) - \nabla U(\bar{Y}_{\Gamma_n})\|^2 - (m+L) \|\bar{Y}_s - \bar{Y}_{\Gamma_n}\|^2/4.$$

Using this inequality in (S27), we get

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

and (S26) becomes

$$\|Y_t - \bar{Y}_t\|^2 \leq \|Y_{\Gamma_n} - \bar{Y}_{\Gamma_n}\|^2 + ((m+L)/2) \int_{\Gamma_n}^t \|\bar{Y}_s - \bar{Y}_{\Gamma_n}\|^2 ds \quad (\text{S28})$$

Therefore, by the previous inequality, it remains to bound the expectation of $\|\bar{Y}_s - \bar{Y}_{\Gamma_n}\|^2$. By (12) and using $\nabla U(x^*) = 0$,

$$\|\bar{Y}_s - \bar{Y}_{\Gamma_n}\|^2 = \left\| -(s - \Gamma_n)(\nabla U(\bar{Y}_{\Gamma_n}) - \nabla U(x^*)) + \sqrt{2}(B_s - B_{\Gamma_n}) \right\|^2.$$

Then taking the expectation, using the Markov property of $(B_t)_{t \geq 0}$ and **H1**, we have

$$\mathbb{E} \left[\|\bar{Y}_s - \bar{Y}_{\Gamma_n}\|^2 \right] \leq (s - \Gamma_n)^2 L^2 \mathbb{E} \left[\|\bar{Y}_{\Gamma_n} - x^*\|^2 \right] + 2(s - \Gamma_n)d. \quad (\text{S29})$$

The proof follows from taking the expectation in (S28), combining (S25)-(S29), and using Theorem S7. \square

Theorem S12. Assume **H1**, **H2** and **G1**. Then for all $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ and $n \geq 1$,

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

where $(\tilde{u}_n^{(1)})_{n \geq 0}$ is given by (S24) and

$$\begin{aligned} \tilde{u}_n^{(2)}(\gamma) \stackrel{\text{def}}{=} & \sum_{i=n_1}^n \gamma_i^2 \phi(\gamma_i) \prod_{k=i+1}^n (1 - \kappa \gamma_k / 2) \\ & + \phi(\gamma_1) (L^2(1 + \rho))^{-1} \left\{ \prod_{k=1}^{n_1-1} (1 + \gamma_k^2(1 + \rho)L^2) \right\} \prod_{k=n_1}^n (1 - \kappa \gamma_k / 2), \end{aligned} \quad (\text{S31})$$

where

$$\phi(\gamma) = \{2\kappa^{-1} + (1 + \rho^{-1})\gamma\} (dL^2 + dL^4\gamma/(2m) + dL^4\gamma^2/12).$$

Proof. Let ζ_0 be an optimal transference plan of μ_0 and π . Let $(Y_t, \bar{Y}_t)_{t \geq 0}$ with (Y_0, \bar{Y}_0) distributed according to ζ_0 and defined by (12). By definition of W_2 and since for all $t \geq 0$, π is invariant for P_t , $W_2^2(\mu_0 Q_\gamma^n, \pi) \leq \mathbb{E}[\|Y_{\Gamma_n} - X_{\Gamma_n}\|^2]$. Then the proof follows from Proposition S11 since Y_0 is distributed according to π and by (7), which shows that $\delta \leq d/m$. \square

3.1 Explicit bound based on Theorem S12 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 (S5). First note, since $(\gamma_k)_{k \geq 1}$ is nonincreasing, for all $n \geq 1$, we have

$$\begin{aligned} \tilde{u}_n^{(2)}(\gamma) \leq & \sum_{j=0}^1 C_j \sum_{i=n_1}^n \gamma_i^{j+2} \prod_{k=i+1}^n (1 - \kappa \gamma_k / 2) \\ & + \sum_{j=0}^1 C_j (L^2(1 + \rho))^{-1} \gamma_1^j \left\{ \prod_{k=1}^{n_1-1} (1 + \gamma_k^2(1 + \rho)L^2) \right\} \prod_{k=n_1}^n (1 - \kappa \gamma_k / 2), \end{aligned} \quad (\text{S32})$$

where

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

1. For $\alpha = 1$ and $n_1 = 1$, by (S6) and (S7), we have

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

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

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

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 (S6) and Lemma S9-(ii) applied with $\ell = \lceil n/2 \rceil$, we have

$$\begin{aligned}\tilde{u}_n^{(1)}(\gamma) &\leq \exp(-(\kappa\gamma_1/2) \boldsymbol{\psi}_{\alpha-1}(n+1)) \\ \tilde{u}_n^{(2)}(\gamma) &\leq \sum_{j=0}^1 C_j \left\{ \gamma_1^{j+2} (\boldsymbol{\psi}_{1-\alpha(j+2)}(\lceil n/2 \rceil) + 1) \exp \left\{ -(\kappa\gamma_1/2) (\boldsymbol{\psi}_{1-\alpha}(n+1) - \boldsymbol{\psi}_{1-\alpha}(\lceil n/2 \rceil)) \right\} \right. \\ &\quad \left. + 2\kappa^{-1} \gamma_1^{j+1} (n/2)^{-\alpha(j+1)} + \gamma_1^j \exp \left\{ -(\kappa\gamma_1/2) \boldsymbol{\psi}_{\alpha-1}(n+1) \right\} \right\}.\end{aligned}$$

For $n_1 > 1$ and $\lceil n/2 \rceil \geq n_1$, since $(\gamma_k)_{k \geq 0}$ is non increasing, using again (S6), and Lemma S9-(ii) applied with $\ell = \lceil n/2 \rceil$, and the bound for $t \in \mathbb{R}$, $(1+t) \leq e^t$, we get

$$\begin{aligned}\tilde{u}_n^{(1)}(\gamma) &\leq \exp \left\{ -\kappa\gamma_1 (\boldsymbol{\psi}_{\alpha-1}(n+1) - \boldsymbol{\psi}_{1-\alpha}(n_1))/2 + L^2(1+\rho) \gamma_1^2 (\boldsymbol{\psi}_{1-2\alpha}(n_1-1) + 1) \right\} \\ \tilde{u}_n^{(2)}(\gamma) &\leq \sum_{j=0}^1 C_j \left\{ 2\kappa^{-1} \gamma_1^{j+1} (n/2)^{-\alpha(j+1)} \right. \\ &\quad \left. + \gamma_1^{j+2} (\boldsymbol{\psi}_{1-\alpha(j+2)}(\lceil n/2 \rceil) - \boldsymbol{\psi}_{1-\alpha(j+2)}(n_1) + 1) \exp \left\{ -(\kappa\gamma_1/2) (\boldsymbol{\psi}_{1-\alpha}(n+1) - \boldsymbol{\psi}_{1-\alpha}(\lceil n/2 \rceil)) \right\} \right. \\ &\quad \left. + (\gamma_1^j / (L^2(1+\rho))) \exp \left\{ -\kappa\gamma_1 (\boldsymbol{\psi}_{\alpha-1}(n+1) - \boldsymbol{\psi}_{1-\alpha}(n_1))/2 + L^2(1+\rho) \gamma_1^2 (\boldsymbol{\psi}_{1-2\alpha}(n_1-1) + 1) \right\} \right\}.\end{aligned}$$

4 Explicit bounds on the MSE

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

5 Explicit bounds based on Theorem 3

1. First for $\alpha = 0$, recall by (S1), (20) and (21) we have for all $p \geq 1$,

$$u_p^{(1)}(\gamma) W_2^2(\delta_x, \pi) + u_p^{(2)}(\gamma) \leq 2\Omega(x) (1 - \kappa\gamma_1/2)^p + 2\kappa^{-1} (A_0\gamma_1 + A_1\gamma_1^2),$$

where \mathbf{A}_0 and \mathbf{A}_1 are given by (S2) and (S3) respectively. So by Proposition 8 and Lemma 14, we have the following bound for the bias

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

Therefore plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\kappa^{-1} \mathbf{A}_0 \gamma_1 + \frac{\kappa^{-2} + \kappa^{-1} \exp(-\kappa N \gamma_1/2) \Omega(x)}{n \gamma_1} \right). \quad (\text{S33})$$

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

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

and (S33) becomes if $\gamma_1 \leftarrow \gamma_{\star,0}(n)$,

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

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 $\alpha \in (0, 1/2)$, recall by (S1), (20) and (21) we have for all $p \geq 1$,

$$u_p^{(1)}(\gamma) W_2^2(\delta_x, \pi) \leq 2\Omega(x) \prod_{i=1}^p (1 - \kappa \gamma_i/2). \quad (\text{S34})$$

So by Proposition 8, Lemma 14, (S6) and (S8), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\frac{\kappa^{-1} \mathbf{A}_0 \gamma_1}{(1 - 2\alpha) n^\alpha} + \frac{\kappa^{-1} \exp\{-\kappa \gamma_1 N^{1-\alpha}/(2(1-\alpha))\} \Omega(x) + \kappa^{-2}}{\gamma_1 n^{1-\alpha}} \right). \quad (\text{S35})$$

At fixed γ_1 , this bound is of order $n^{-\alpha}$, and is better than (S33) 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

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

(S35) becomes with $\gamma_1 \leftarrow \gamma_{\star,\alpha}(n)$,

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

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

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

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 $\alpha \rightarrow 1/2$.

3. If $\alpha = 1/2$, (S34) still holds. By Lemma 14, (S6) and (S8), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\frac{\kappa^{-1} \mathbf{A}_0 \gamma_1 \log(n)}{n^{1/2}} + \frac{\kappa^{-1} \exp\{-\kappa \gamma_1 N^{1/2}/4\} \Omega(x) + \kappa^{-2}}{\gamma_1 n^{1/2}} \right). \quad (\text{S36})$$

—eq:bound'MSE'gamma1

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

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

and (S36) becomes with $\gamma_1 \leftarrow \gamma_{*,1/2}(n)$,

$$\begin{aligned} \text{MSE}_f(N, n) \\ \leq C \left(\frac{\log(n)}{nC_{\text{MSE},1/2}} \right)^{1/2} \left(\kappa^{-1} \exp\left\{-\kappa N^{1/2} \gamma_{*,1/2}(n)/4\right\} \Omega(x) + \frac{C_{\text{MSE},1/2}}{\log(n)} \right). \end{aligned}$$

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

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

We can see that we obtain a worse bound than for $\alpha = 0$ and $\alpha \in (0, 1/2)$.

4. For $\alpha \in (1/2, 1]$, (S34) still holds. By Lemma 14, (S6) and (S8), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\frac{\kappa^{-1} \mathbf{A}_0 \gamma_1}{n^{1-\alpha}} + \frac{\kappa^{-1} \exp\{-\kappa \gamma_1 N^{1-\alpha}/(2(1-\alpha))\} \Omega(x) + \kappa^{-2}}{\gamma_1 n^{1-\alpha}} \right).$$

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 $\gamma_1 \rightarrow +\infty$ as $n \rightarrow +\infty$. Therefore, in that case, the best choice of γ_1 is the largest possible value $1/(m+L)$.

5. For $\alpha = 1$, (S34) still holds. By Lemma 14, (S6) and (S8), the bias is upper bounded by

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

Plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\frac{\kappa^{-1} \mathbf{A}_0 \gamma_1}{\log(n)} + \frac{\kappa^{-1} N^{-\kappa \gamma_1/2} \Omega(x) + \kappa^{-2}}{\gamma_1 \log(n)} \right).$$

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 $\alpha \in (1/2, 1)$.

5.1 Explicit bound based on Theorem 6

1. First for $\alpha = 0$, recall by (S9), (20) and (21) we have for all $p \geq 1$,

$$u_p^{(1)}(\gamma)W_2^2(\delta_x, \pi) + u_p^{(3)}(\gamma) \leq 2\Omega(x)(1 - \kappa\gamma_1/2)^p + 2\kappa^{-1}(\mathbf{B}_0\gamma_1 + \mathbf{B}_1\gamma_1^2),$$

where \mathbf{B}_0 and \mathbf{B}_1 are given by (S10) and (S11) respectively. So by Proposition 8 and Lemma 14, we have the following bound for the bias

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

Therefore plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\kappa^{-1}\mathbf{B}_0\gamma_1^2 + \frac{\kappa^{-2} + \kappa^{-1} \exp(-\kappa N\gamma_1/2)\Omega(x)}{n\gamma_1} \right). \quad (\text{S37})$$

-eq:bound'MSE'gamma

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

$$\gamma_{*,0}(n) = (\kappa\mathbf{B}_0n)^{-1/3},$$

and (S33) becomes if $\gamma_1 \leftarrow \gamma_{*,0}(n)$,

$$\text{MSE}_f(N, n) \leq C(\mathbf{B}_0^{-1/2}n)^{-2/3} \left(\kappa^{-4/3} \exp(-\kappa N\gamma_{*,0}(n)/2)\Omega(x) + \kappa^{-5/3} \right).$$

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

$$\text{MSE}_f(N_0(n), n) \leq C(\mathbf{B}_0^{-1/2}\kappa^{5/2}n)^{-2/3}.$$

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

2. For $\alpha \in (0, 1/3)$, recall by (S9), (20) and (21) we have for all $p \geq 1$,

$$u_p^{(1)}(\gamma)W_2^2(\delta_x, \pi) \leq 2\Omega(x) \prod_{i=1}^p (1 - \kappa\gamma_i/2). \quad (\text{S38})$$

-eq:bound'for'MSE'u'n'

So by Proposition 8, Lemma 14, (S6) and (S14), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\frac{\kappa^{-1}\mathbf{B}_0\gamma_1^2}{(1 - 3\alpha)n^{2\alpha}} + \frac{\kappa^{-1} \exp\{-\kappa\gamma_1 N^{1-\alpha}/(2(1 - \alpha))\}\Omega(x) + \kappa^{-2}}{\gamma_1 n^{1-\alpha}} \right). \quad (\text{S39})$$

-eq:bound'MSE'gamma

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

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

(S39) becomes with $\gamma_1 \leftarrow \gamma_{*,\alpha}(n)$,

$$\text{MSE}_f(N, n) \leq C(\mathbf{B}_0^{-1/2}n)^{-2/3} \left(\kappa^{-4/3} \exp(-\kappa N \gamma_{*,0}(n)/(2(1-\alpha))) \Omega(x) + \kappa^{-5/3} (1-3\alpha)^{-1/3} \right).$$

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

$$\text{MSE}_f(N_\alpha(n), n) \leq C(\mathbf{B}_0^{-1/2} \kappa^{5/2} n)^{-2/3}.$$

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

3. If $\alpha = 1/3$, (S38) still holds. By Lemma 14, (S6) and (S14), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\frac{\kappa^{-1} \mathbf{B}_0 \gamma_1^2 \log(n)}{n^{2/3}} + \frac{\kappa^{-1} \exp\{-\kappa \gamma_1 N^{2/3}/4\} \Omega(x) + \kappa^{-2}}{\gamma_1 n^{2/3}} \right). \quad (\text{S40})$$

—eq:bound'MSE'gamma

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

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

and (S40) becomes with $\gamma_1 \leftarrow \gamma_{*,1/2}(n)$,

$$\text{MSE}_f(N, n) \leq C \left(\frac{\log(n) \mathbf{B}_0}{n^2} \right)^{1/3} \left(\kappa^{-4/3} \exp\{-\kappa N^{1/2} \gamma_{*,1/2}(n)/4\} \Omega(x) + \kappa^{-5/3} \right).$$

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

$$\text{MSE}_f(N_{1/2}(n), n) \leq C \left(\frac{\log(n) \mathbf{B}_0}{\kappa^5 n^2} \right)^{1/3}.$$

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

4. For $\alpha \in (1/3, 1]$, (S38) still holds. By Lemma 14, (S6) and (S14), we have the following bound for the bias

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

Plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\frac{\kappa^{-1} \mathbf{B}_0 \gamma_1}{n^{1-\alpha}} + \frac{\kappa^{-1} \exp\{-\kappa \gamma_1 N^{1-\alpha}/(2(1-\alpha))\} \Omega(x) + \kappa^{-2}}{\gamma_1 n^{1-\alpha}} \right).$$

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 $\gamma_1 \rightarrow +\infty$ as $n \rightarrow +\infty$. Therefore, in that case, the best choice of γ_1 is the largest possible value $1/(m+L)$.

5. For $\alpha = 1$, (S34) still holds. By Lemma 14, (S6) and (S8), the bias is upper bounded by

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

Plugging this inequality and the one given by Theorem 9 in (23) implies:

$$\text{MSE}_f(N, n) \leq C \left(\frac{\kappa^{-1} \mathbf{B}_0 \gamma_1}{\log(n)} + \frac{\kappa^{-1} N^{-\kappa \gamma_1 / 2} \Omega(x) + \kappa^{-2}}{\gamma_1 \log(n)} \right).$$

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

- [1] A. Dalalyan. Theoretical guarantees for approximate sampling from a smooth and log-concave density. submitted 1412.7392, arXiv, December 2014.