# Supplementary material

# DG-LMC: A Turn-key and Scalable Synchronous Distributed MCMC Algorithm via Langevin Monte Carlo within Gibbs

**Vincent Plassier** [1 2 *] **Maxime Vono** [2 *] **Alain Durmus** [3 *] **Eric Moulines** [1]

**Notations and conventions.** We denote by $\mathcal{B}(\mathbb{R}^d)$ the Borel $\sigma$-field of $\mathbb{R}^d$, $\mathbb{M}(\mathbb{R}^d)$ the set of all Borel measurable functions $f$ on $\mathbb{R}^d$, $\|f\|_\infty = \sup_{\mathbf{x}\in\mathbb{R}^d} |f(\mathbf{x})|$ and $\|\cdot\|$ the Euclidean norm on $\mathbb{R}^d$. For $\mu$ a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $f \in \mathbb{M}(\mathbb{R}^d)$ a $\mu$-integrable function, denote by $\mu(f)$ the integral of $f$ with respect to (w.r.t.) $\mu$. Let $\mu$ and $\nu$ be two sigma-finite measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Denote by $\mu \ll \nu$ if $\mu$ is absolutely continuous w.r.t. $\nu$ and $\mathrm{d}\mu/\mathrm{d}\nu$ the associated density. Let $\mu, \nu$ be two probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Define the Kullback-Leibler (KL) divergence of $\mu$ from $\nu$ by

$$\mathrm{KL}(\mu|\nu) = \begin{cases} \int_{\mathbb{R}^d} \frac{\mathrm{d}\mu}{\mathrm{d}\nu}(\mathbf{x}) \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}(\mathbf{x})\right) \mathrm{d}\nu(\mathbf{x}) \;, & \text{if } \mu \ll \nu \\ +\infty & \text{otherwise.} \end{cases}$$

In addition, define the Pearson $\chi^2$-divergence of $\mu$ from $\nu$ by

$$\chi^2(\mu|\nu) = \begin{cases} \int_{\mathbb{R}^d} \left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}(\mathbf{x}) - 1\right)^2 \mathrm{d}\nu(\mathbf{x}) \;, & \text{if } \mu \ll \nu \\ +\infty & \text{otherwise.} \end{cases}$$

We say that $\zeta$ is a transference plan of $\mu$ and $\nu$ if it is a probability measure on $(\mathbb{R}^d \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d))$ such that for all measurable set $\mathsf{A}$ of $\mathbb{R}^d$, $\zeta(\mathsf{A} \times \mathbb{R}^d) = \mu(\mathsf{A})$ and $\zeta(\mathbb{R}^d \times \mathsf{A}) = \nu(\mathsf{A})$. We denote by $\mathcal{T}(\mu,\nu)$ the set of transference plans of $\mu$ and $\nu$. In addition, we say that a couple of $\mathbb{R}^d$-random variables $(X, Y)$ is a coupling of $\mu$ and $\nu$ if there exists $\zeta \in \mathcal{T}(\mu,\nu)$ such that $(X, Y)$ are distributed according to $\zeta$. Let $\mathbf{M}$ be a $d \times d$ symmetric positive definite matrix. Denote $\langle , \rangle_{\mathbf{M}}$ the scalar product corresponding to $\mathbf{M}$, defined for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ by $\langle \mathbf{x}, \mathbf{y}\rangle_{\mathbf{M}} = \mathbf{x}^\top \mathbf{M}\mathbf{y}$. Denote $\|\cdot\|_{\mathbf{M}}$ the corresponding norm. We denote by $\mathcal{P}_2(\mathbb{R}^d)$ the set of probability measures with finite 2-moment: for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\int_{\mathbb{R}^d} \|\mathbf{x}\|^2 \mathrm{d}\mu(\mathbf{x}) < \infty$. We define the Wasserstein distance of order 2 associated with $\|\cdot\|_{\mathbf{M}}$ for any probability measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ by

$$W^2_{\mathbf{M}}(\mu,\nu) = \inf_{\zeta\in\mathcal{T}(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \|\mathbf{x}-\mathbf{y}\|^2_{\mathbf{M}} \,\mathrm{d}\zeta(\mathbf{x},\mathbf{y}) \;.$$

In the case when $\mathbf{M} = \mathbf{I}_d$, we will denote the Wasserstein distance of order 2 by $W_2$. By Villani (2008, Theorem 4.1), for all $\mu$, $\nu$ probability measures on $\mathbb{R}^d$, there exists a transference plan $\zeta^\star \in \mathcal{T}(\mu,\nu)$ such that for any coupling $(X, Y)$ distributed according to $\zeta^\star$, $W_{\mathbf{M}}(\mu,\nu) = \mathbb{E}[\|\mathbf{x}-\mathbf{y}\|^2_{\mathbf{M}}]^{1/2}$. This kind of transference plan (respectively coupling) will be called an optimal transference plan (respectively optimal coupling) associated with $W_{\mathbf{M}}$. By Villani (2008, Theorem 6.16), $\mathcal{P}_2(\mathbb{R}^d)$ equipped with the Wasserstein distance $W_{\mathbf{M}}$ is a complete separable metric space. The total variation norm between two probability measures $\mu$ and $\nu$ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is defined by

$$\|\mu-\nu\|_{\mathrm{TV}} = \sup_{f\in\mathbb{M}(\mathbb{R}^d), \|f\|_\infty\le 1} \left| \int_{\mathbb{R}^d} f(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x}) - \int_{\mathbb{R}^d} f(\mathbf{x})\,\mathrm{d}\nu(\mathbf{x}) \right| \;.$$

[*]Equal contribution [1]CMAP, Ecole Polytechnique, Université Paris-Saclay, Palaiseau, France [2]Lagrange Mathematical and Computing Research Center, Huawei, Paris, France [3]Ecole Normale Supérieure Paris-Saclay, Cachan, France. Correspondence to: Vincent Plassier <vincent.plassier@huawei.com>.

*Proceedings of the $38^{th}$ International Conference on Machine Learning*, PMLR 139, 2021. 

For the sake of simplicity, with little abuse, we shall use the same notations for a probability distribution and its associated probability density function. For a Markov chain with transition kernel $P$ on $\mathbb{R}^d$ and invariant distribution $\pi$, we define the $\varepsilon$-mixing time associated to a statistical distance $D$, precision $\varepsilon > 0$ and initial distribution $\nu$, by

$$t_{\mathrm{mix}}(\varepsilon; \nu) = \min \left\{ t \geq 0 \;\middle|\; D(\nu P^t, \pi) \leq \varepsilon \right\} ,$$

which stands for the minimum number of steps of the Markov chain such that its distribution is at most at an $\varepsilon$ $D$-distance from the invariant distribution $\pi$. For $n \geq 1$, we refer to the set of integers between 1 and $n$ with the notation $[n]$. The $d$-multidimensional Gaussian probability distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is denoted by $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. When $\boldsymbol{\mu} = \mathbf{0}_d$ and $\boldsymbol{\Sigma} = \mathbf{I}_d$, the associated probability density function is denoted by by $\phi_d$. Let $F : \mathbb{R}^d \to \mathbb{R}$ be a twice continuously differentiable function, denote $\vec{\Delta}$ the vector Laplacian of $F$ defined, for all $x \in \mathbb{R}^d$, by $\vec{\Delta}F(x) = \{\sum_{l=1}^d (\partial^2 F_k)(x)/\partial x_l^2\}_{k=1}^d$. For $0 \leq i < j$, we use the notation $\mathbf{u}_{i:j}$ to refer to the vector $[\mathbf{u}_i^\top, \cdots, \mathbf{u}_j^\top]^\top$ built by stacking $j - i + 1$ vectors $(\mathbf{u}_k;\, k \in \{i, \cdots, j\})$. For a given matrix $\mathbf{M} \in \mathbb{R}^{d \times d}$, we denote its smallest and largest eigenvalues by $\lambda_{\min}(\mathbf{M})$ and $\lambda_{\max}(\mathbf{M})$, respectively. Fix $b \in \mathbb{N}^*$ and let $\mathbf{M}_1, \ldots, \mathbf{M}_b$ be $d$-dimensional matrices. We denote $\prod_{\ell=i}^j \mathbf{M}_\ell = \mathbf{M}_j \ldots \mathbf{M}_i$ if $i \leq j$ and with the convention $\prod_{\ell=i}^j \mathbf{M}_\ell = \mathbf{I}_d$ if $i > j$. For any $b \in \mathbb{N}^*$, $(d_i)_{i \in [b]} \in (\mathbb{N}^*)^b$ and $(\mathbf{M}_i)_{i \in [b]} \in \otimes_{i \in [b]} \mathbb{R}^{d_i \times d_i}$, we denote $\mathrm{diag}(\mathbf{M}_1, \ldots, \mathbf{M}_b)$ the unique matrix $\mathbf{M} \in \mathbb{R}^{(\sum_i d_i) \times (\sum_i d_i)}$ satisfying for any $\mathbf{u} = (\mathbf{u}_1, \ldots, \mathbf{u}_b) \in \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_b}$, $\mathbf{M}\mathbf{u} = \sum_{i=1}^b \mathbf{M}_i \mathbf{u}_i$ which corresponds to

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_1 & \mathbf{0}_{d_1,d_2} & \cdots & \mathbf{0}_{d_b,d_b} \\ \mathbf{0}_{d_2,d_1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{d_{b-1},d_b} \\ \mathbf{0}_{d_b,d_1} & \cdots & \mathbf{0}_{d_b,d_{b-1}} & \mathbf{M}_b \end{pmatrix} .$$

For any $\mathbf{v} \in \mathbb{R}^b$, define the block diagonal matrix

$$\mathbf{D}_{\mathbf{v}} = \mathrm{diag}\left(v_1 \cdot \mathbf{I}_{d_1}, \ldots, v_b \cdot \mathbf{I}_{d_b}\right) \in \mathbb{R}^{p \times p} . \tag{S1}$$

For any symmetric matrices $\mathbf{S}_1, \mathbf{S}_2 \in \mathbb{R}^{p \times p}$, we note $\mathbf{S}_1 \preccurlyeq \mathbf{S}_2$ if and only if, for any $\mathbf{u} \in \mathbb{R}^p$, we have $\mathbf{u}^\top(\mathbf{S}_2 - \mathbf{S}_1)\mathbf{u} \geq 0$. Let $(\mathsf{X}, \mathcal{X})$ and $(\mathsf{Y}, \mathcal{Y})$ be two measurable spaces, we say that a transition probability kernel on $(\mathsf{Y} \times \mathsf{X}) \times \mathcal{Y}$ is a conditional Markov kernel. One elementary step in most Gibbs samplers corresponds to a conditional Markov kernel.

## S1. Proof of Proposition 1

Let $b' \in [b-1]$, $p' = \sum_{i=b'+1}^b d_i$ and consider

$$\mathbf{B}_{b'}^\top = [\mathbf{A}_{b'+1}^\top/\rho_{b'+1}^{1/2} \cdots \mathbf{A}_b^\top/\rho_b^{1/2}] \in \mathbb{R}^{d \times p'} , \qquad \bar{\mathbf{B}}_{b'} = \mathbf{B}_{b'}^\top \mathbf{B}_{b'} = \sum_{i=b'+1}^b \{\mathbf{A}_i^\top \mathbf{A}_i/\rho_i\} \in \mathbb{R}^{d \times d} . \tag{S2}$$

Note that under **H**1, $\bar{\mathbf{B}}_{b'}$ is invertible. Indeed, it is a symmetric positive definite matrix since for any $\boldsymbol{\theta} \in \mathbb{R}^d$, $\langle \bar{\mathbf{B}}_{b'}\boldsymbol{\theta}, \boldsymbol{\theta} \rangle \geq [\min_{i \in [b]} \rho_i^{-1}] \langle \sum_{i=b'+1}^b \mathbf{A}_i^\top \mathbf{A}_i \boldsymbol{\theta}, \boldsymbol{\theta} \rangle > 0$ using that $\sum_{i=b'+1}^b \mathbf{A}_i^\top \mathbf{A}_i$ is invertible. Define the orthogonal projection onto the range of $\mathbf{B}_{b'}$ and the diagonal matrix:

$$\mathbf{P}_{b'} = \mathbf{B}_{b'} \bar{\mathbf{B}}_{b'}^{-1} \mathbf{B}_{b'}^\top , \qquad \tilde{\mathbf{D}}_{b'} = \mathrm{diag}(\mathbf{I}_{d_{b'+1}}/\rho_{b'+1}, \ldots, \mathbf{I}_{d_b}/\rho_b) . \tag{S3}$$

### S1.1. Technical lemma

**Lemma S1.** *Assume **H**1. For any $(\boldsymbol{\theta}, \mathbf{z}_{b'+1:b}) \in \mathbb{R}^d \times \mathbb{R}^{p'}$, setting $\mathbf{z} = \mathbf{z}_{b'+1:b}$, we have*

$$\sum_{i=b'+1}^b \left\{ \|\mathbf{z}_i - \mathbf{A}_i \boldsymbol{\theta}\|^2/\rho_i \right\} = (\tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z})^\top \{\mathbf{I}_{p'} - \mathbf{P}_{b'}\} (\tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z}) \\ + (\boldsymbol{\theta} - \bar{\mathbf{B}}_{b'}^{-1}\mathbf{B}_{b'}^\top \tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z})^\top \bar{\mathbf{B}}_{b'} (\boldsymbol{\theta} - \bar{\mathbf{B}}_{b'}^{-1}\mathbf{B}_{b'}^\top \tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z}) .$$

*Proof.* Setting $\mathbf{b} = \mathbf{B}_{b'}^\top \tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z}$ and using the fact that $\bar{\mathbf{B}}_{b'}$ is symmetric, we have

$$
\begin{aligned}
\sum_{i=b'+1}^{b} \left\{\|\mathbf{z}_i - \mathbf{A}_i\boldsymbol{\theta}\|^2/\rho_i\right\} &= \boldsymbol{\theta}^\top\bar{\mathbf{B}}_{b'}\boldsymbol{\theta} - 2\boldsymbol{\theta}^\top\mathbf{b} + \sum_{i=b'+1}^{b}\|\mathbf{z}_i\|^2/\rho_i \\
&= \sum_{i=b'+1}^{b}\|\mathbf{z}_i\|^2/\rho_i - \mathbf{b}^\top\bar{\mathbf{B}}_{b'}^{-1}\mathbf{b} + (\boldsymbol{\theta} - \bar{\mathbf{B}}_{b'}^{-1}\mathbf{b})^\top\bar{\mathbf{B}}_{b'}(\boldsymbol{\theta} - \bar{\mathbf{B}}_{b'}^{-1}\mathbf{b})\,.
\end{aligned}
$$

Using that $\mathbf{b}^\top\bar{\mathbf{B}}_{b'}^{-1}\mathbf{b} = (\tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z})^\top\mathbf{P}_{b'}(\tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z})$ and $\mathbf{P}_{b'}$ is a projection, $\mathbf{P}_{b'}^2 = \mathbf{P}_{b'}$ completes the proof. □

### S1.2. Proof of Proposition 1

**Proposition S2.** *Assume **H**1. Then, the function* $\psi : (\boldsymbol{\theta}, \mathbf{z}_{1:b}) \mapsto \prod_{i=1}^{b}\exp\{-U_i(\mathbf{z}_i) - \|\mathbf{z}_i - \mathbf{A}_i\boldsymbol{\theta}\|^2/(2\rho_i)\}$ *is integrable on* $\mathbb{R}^d \times \mathbb{R}^p$*, where* $p = \sum_{i=1}^{b} d_i$*.*

*Proof.* Using **H**1 and the Fubini theorem, there exists $C_1 > 0$ such that:

$$
\begin{aligned}
&\int_{\mathbb{R}^d}\left[\prod_{i=1}^{b'}\int_{\mathbb{R}^{d_i}} \mathrm{e}^{-U_i(\mathbf{z}_i)}\mathrm{e}^{-\frac{\|\mathbf{z}_i - \mathbf{A}_i\boldsymbol{\theta}\|^2}{2\rho_i}}\,\mathrm{d}\mathbf{z}_i \cdot \prod_{j=b'+1}^{b}\int_{\mathbb{R}^{d_j}} \mathrm{e}^{-U_j(\mathbf{z}_j)}\mathrm{e}^{-\frac{\|\mathbf{z}_j - \mathbf{A}_j\boldsymbol{\theta}\|^2}{2\rho_j}}\,\mathrm{d}\mathbf{z}_j\right]\mathrm{d}\boldsymbol{\theta} \\
&\leq C_1\int_{\mathbb{R}^d}\left[\prod_{i=1}^{b'}\int_{\mathbb{R}^{d_i}} \mathrm{e}^{-\frac{\|\mathbf{z}_i - \mathbf{A}_i\boldsymbol{\theta}\|^2}{2\rho_i}}\,\mathrm{d}\mathbf{z}_i \cdot \prod_{j=b'+1}^{b}\int_{\mathbb{R}^{d_j}} \mathrm{e}^{-U_j(\mathbf{z}_j)}\mathrm{e}^{-\frac{\|\mathbf{z}_j - \mathbf{A}_j\boldsymbol{\theta}\|^2}{2\rho_j}}\,\mathrm{d}\mathbf{z}_j\right]\mathrm{d}\boldsymbol{\theta} \\
&\leq C_1\prod_{i=1}^{b'}(2\pi\rho_i)^{d_i/2}\int_{\mathbb{R}^d}\left[\prod_{j=b'+1}^{b}\int_{\mathbb{R}^{d_j}} \mathrm{e}^{-U_j(\mathbf{z}_j)}\exp\left(-\left\|\mathbf{z}_j - \mathbf{A}_j\boldsymbol{\theta}\right\|^2/(2\rho_j)\right)\mathrm{d}\mathbf{z}_j\right]\mathrm{d}\boldsymbol{\theta} \\
&= C_1\prod_{i=1}^{b'}(2\pi\rho_i)^{d_i/2}\int_{\mathbb{R}^{d_{b'+1}}}\cdots\int_{\mathbb{R}^{d_b}}\left[\prod_{j=b'+1}^{b}\mathrm{e}^{-U_j(\mathbf{z}_j)}\right]\left[\int_{\mathbb{R}^d}\prod_{j=b'+1}^{b}\mathrm{e}^{-\frac{\|\mathbf{z}_j - \mathbf{A}_j\boldsymbol{\theta}\|^2}{2\rho_j}}\mathrm{d}\boldsymbol{\theta}\right]\mathrm{d}\mathbf{z}_{b'+1:b}\,. \qquad \text{(S4)}
\end{aligned}
$$

Using Lemma S1 and the fact that $\mathbf{I}_{p'} - \mathbf{P}_{b'}$ is positive definite, we obtain

$$
\begin{aligned}
&\int_{\mathbb{R}^d}\prod_{j=b'+1}^{b}\exp\left(-\left\|\mathbf{z}_j - \mathbf{A}_j\boldsymbol{\theta}\right\|^2/(2\rho_j)\right)\mathrm{d}\boldsymbol{\theta} \\
&= \exp\left(-(\tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z})^\top\{\mathbf{I}_{p'} - \mathbf{P}_{b'}\}(\tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z})/2\right) \\
&\times\int_{\mathbb{R}^d}\exp\left(-(\boldsymbol{\theta} - \bar{\mathbf{B}}_{b'}^{-1}\mathbf{B}_{b'}^\top\tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z})^\top\bar{\mathbf{B}}_{b'}(\boldsymbol{\theta} - \bar{\mathbf{B}}_{b'}^{-1}\mathbf{B}_{b'}^\top\tilde{\mathbf{D}}_{b'}^{1/2}\mathbf{z})/2\right)\mathrm{d}\boldsymbol{\theta} \\
&\leq \det\left(\bar{\mathbf{B}}_{b'}\right)^{-1/2}(2\pi)^{d/2}.
\end{aligned}
$$

Then, the proof is completed by plugging this expression into (S4) and using from **H**1 that $\mathbf{z}_{b'+1:b} \mapsto \prod_{j=b'+1}^{b}\mathrm{e}^{-U_j(\mathbf{z}_j)}$ is integrable. □

## S2. Proof of Proposition 2

This section aims at proving Proposition 2 in the main paper. To ease the understanding, we dissociate the scenarios where $\max_{i\in[b]} N_i = 1$ and $\max_{i\in[b]} N_i > 1$. In addition, in all this section $\boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$ is assumed to be fixed.

### S2.1. Single local LMC iteration

In this section, we assume that a single LMC step is performed locally on each worker, that is $\max_{i\in[b]} N_i = 1$. For this, we introduce the conditional Markov transition kernel defined for any $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_b)$, $\boldsymbol{\theta} \in \mathbb{R}^d$, $\mathbf{z} = (\mathbf{z}_1, \cdots, \mathbf{z}_b) \in$

$\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_b}$, and for $i \in [b]$, $\mathsf{B}_i \in \mathcal{B}(\mathbb{R}^{d_i})$, by

$$Q_{\boldsymbol{\rho},\boldsymbol{\gamma}}\left(\mathbf{z}, \mathsf{B}_1 \times \cdots \times \mathsf{B}_b|\boldsymbol{\theta}\right) = \prod_{i=1}^{b} R_{\rho_i,\gamma_i}(\mathbf{z}_i, \mathsf{B}_i|\boldsymbol{\theta}) \;, \tag{S5}$$

where

$$R_{\rho_i,\gamma_i}(\mathbf{z}_i, \mathsf{B}_i|\boldsymbol{\theta}) = \int_{\mathsf{B}_i} \exp\left\{-\frac{1}{4\gamma_i}\left\|\tilde{\mathbf{z}}_i - \left(1 - \frac{\gamma_i}{\rho_i}\right)\mathbf{z}_i - \frac{\gamma_i}{\rho_i}\mathbf{A}_i\boldsymbol{\theta} + \gamma_i \nabla U_i(\mathbf{z}_i)\right\|^2\right\} \frac{\mathrm{d}\tilde{\mathbf{z}}_i}{(4\pi\gamma_i)^{d_i/2}} \;. \tag{S6}$$

Recall that $p = \sum_{i=1}^{b} d_i$. The considered Gibbs sampler in Algorithm 1 defines a homogeneous Markov chain $X_n^\top = (\theta_n^\top, Z_n^\top)_{n\geq 1}$ where $Z_n^\top = ([Z_n^1]^\top, \cdots, [Z_n^b]^\top)$. Indeed, it is easy to show that for any $n \in \mathbb{N}$ and measurable bounded function $f : \mathbb{R}^p \to \mathbb{R}_+$, $\mathbb{E}[f(Z_{n+1})|X_n] = \int_{\mathbb{R}^p} f(\mathbf{z}) Q_{\boldsymbol{\rho},\boldsymbol{\gamma}}(Z_n, \mathrm{d}\mathbf{z}|\theta_n)$ and therefore $(X_n)_{n\in\mathbb{N}}$ is associated with the Markov kernel defined, for any $\mathbf{x}^\top = (\boldsymbol{\theta}^\top, \mathbf{z}^\top) \in \mathbb{R}^d \times \mathbb{R}^p$ and $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$, $\mathsf{B} \in \mathcal{B}(\mathbb{R}^p)$, by

$$P_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathbf{x}, \mathsf{A} \times \mathsf{B}) = \int_{\mathsf{B}} Q_{\boldsymbol{\rho},\boldsymbol{\gamma}}\left(\mathbf{z}, \mathrm{d}\tilde{\mathbf{z}}|\boldsymbol{\theta}\right) \int_{\mathsf{A}} \Pi_{\boldsymbol{\rho}}(\mathrm{d}\tilde{\boldsymbol{\theta}}|\tilde{\mathbf{z}}) \;, \tag{S7}$$

where $\Pi_{\boldsymbol{\rho}}(\cdot|\tilde{\mathbf{z}})$ is defined in (5). Let $(\xi_n)_{n\geq 1}$ be a sequence of i.i.d. $d$-dimensional standard Gaussian random variables independent of the family of independent random variables $\{(\eta_n^i)_{n\geq 1} : i \in [b]\}$ where for any $i \in [b]$ and $n \geq 1$, $\eta_n^i$ is a $d_i$-dimensional standard Gaussian random variable. We define the stochastic processes $(X_n, \tilde{X}_n)_{n\geq 0}$ on $\mathbb{R}^p \times \mathbb{R}^p$ starting from $(X_0, \tilde{X}_0) = (\mathbf{x}, \tilde{\mathbf{x}}) = ((\boldsymbol{\theta}^\top, \mathbf{z}^\top)^\top, (\tilde{\boldsymbol{\theta}}^\top, \tilde{\mathbf{z}}^\top)^\top)$ and following the recursion for $n \geq 0$,

$$X_{n+1} = (\theta_{n+1}^\top, Z_{n+1}^\top)^\top \;, \qquad \tilde{X}_{n+1} = (\tilde{\theta}_{n+1}^\top, \tilde{Z}_{n+1}^\top)^\top \;, \tag{S8}$$

where $Z_{n+1} = ([Z_{n+1}^1]^\top, \ldots, [Z_{n+1}^b]^\top)^\top$, $\tilde{Z}_{n+1} = ([\tilde{Z}_{n+1}^1]^\top, \ldots, [\tilde{Z}_{n+1}^b]^\top)^\top$ are defined, for any $i \in [b]$, by

$$\begin{aligned} Z_{n+1}^i &= \left(1 - \gamma_i/\rho_i\right) Z_n^i + \left(\gamma_i/\rho_i\right) \mathbf{A}_i \theta_n - \gamma_i \nabla U_i(Z_n^i) + \sqrt{2\gamma_i}\eta_{n+1}^i \;, \\ \tilde{Z}_{n+1}^i &= \left(1 - \gamma_i/\rho_i\right) \tilde{Z}_n^i + \left(\gamma_i/\rho_i\right) \mathbf{A}_i \tilde{\theta}_n - \gamma_i \nabla U_i(\tilde{Z}_n^i) + \sqrt{2\gamma_i}\eta_{n+1}^i \;, \end{aligned} \tag{S9}$$

and $\theta_{n+1}, \tilde{\theta}_{n+1}$ by

$$\theta_{n+1} = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2} Z_{n+1} + \bar{\mathbf{B}}_0^{-1/2}\xi_{n+1} \;, \qquad \tilde{\theta}_{n+1} = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2} \tilde{Z}_{n+1} + \bar{\mathbf{B}}_0^{-1/2}\xi_{n+1} \;, \tag{S10}$$

where $\bar{\mathbf{B}}_0$, $\mathbf{B}_0$ and $\tilde{\mathbf{D}}_0$ are given in (S2) and (S3), respectively. Note that $X_n$ and $\tilde{X}_n$ are distributed according to $\updelta_{\mathbf{x}} P_{\boldsymbol{\rho},\boldsymbol{\gamma}}^n$ and $\updelta_{\tilde{\mathbf{x}}} P_{\boldsymbol{\rho},\boldsymbol{\gamma}}^n$, respectively. Hence, by definition of the Wasserstein distance of order 2, it follows that

$$W_2(\updelta_{\mathbf{x}} P_{\boldsymbol{\rho},\boldsymbol{\gamma}}^n, \updelta_{\tilde{\mathbf{x}}} P_{\boldsymbol{\rho},\boldsymbol{\gamma}}^n) \leq \mathbb{E}\left[\|X_n - \tilde{X}_n\|^2\right]^{1/2} . \tag{S11}$$

Thus, in this section we focus on upper bounding the squared norm $\|X_n - \tilde{X}_n\|$ from which we get an explicit bound on the Wasserstein distance thanks to the previous inequality.

#### S2.1.1. Supporting lemmata

Note that **H**1 implies the invertibility of the matrix $\mathbf{B}_0$ defined in (S2) since we have the existence of $b' \in [b-1]$, such that $\sum_{i=b'+1}^{b} \lambda_{\min}(\mathbf{A}_i^\top \mathbf{A}_i)/\rho_i > 0$ and by the semi-positiveness of the symmetric matrices $\{\mathbf{A}_i^\top \mathbf{A}_i\}_{i\in[b]}$, we get that $\lambda_{\min}(\mathbf{B}_0) = \sum_{i=1}^{b} \lambda_{\min}(\mathbf{A}_i^\top \mathbf{A}_i)/\rho_i \geq \sum_{i=b'+1}^{b} \lambda_{\min}(\mathbf{A}_i^\top \mathbf{A}_i)/\rho_i$. To prove Proposition 2 in the case $\max_{i\in[b]} N_i = 1$, we first upper bound (S83) by building upon the following two technical lemmas.

**Lemma S3.** *Assume **H**1 and consider $(X_n, \tilde{X}_n)_{n\in\mathbb{N}}$ defined in (S8). Then, for any $n \in \mathbb{N}$, it holds almost surely that*

$$\|X_{n+1} - \tilde{X}_{n+1}\|^2 \leq (1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2)\|Z_{n+1} - \tilde{Z}_{n+1}\|^2 \;.$$

*Proof.* Let $n \geq 0$. By (S10), we have $\theta_{n+1} - \tilde{\theta}_{n+1} = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}(Z_{n+1} - \tilde{Z}_{n+1})$ which implies that

$$\|X_{n+1} - \tilde{X}_{n+1}\|^2 = \|\theta_{n+1} - \tilde{\theta}_{n+1}\|^2 + \|Z_{n+1} - \tilde{Z}_{n+1}\|^2 \leq (1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2)\|Z_{n+1} - \tilde{Z}_{n+1}\|^2 \;.$$

□

Define the contraction factor

$$\kappa_{\boldsymbol{\gamma}} = \max_{i\in[b]} \left\{ |1-\gamma_i m_i| \vee |1-\gamma_i(M_i + 1/\rho_i)| \right\} \ . \tag{S12}$$

Then, the following result holds.

**Lemma S4.** *Assume **H**1-**H**2 and let $\boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$. Then for any $\mathbf{x} = (\mathbf{z}^\top, \boldsymbol{\theta}^\top)^\top, \tilde{\mathbf{x}} = (\tilde{\mathbf{z}}^\top, \tilde{\boldsymbol{\theta}}^\top)^\top$, with $(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}) \in (\mathbb{R}^d)^2$ and $(\mathbf{z}, \tilde{\mathbf{z}}) \in (\mathbb{R}^p)^2$, for any $n \geq 1$, we have*

$$W_2(\delta_{\mathbf{x}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma}}, \delta_{\tilde{\mathbf{x}}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma}}) \leq \kappa_{\boldsymbol{\gamma}}^{n-1} \cdot \left( (1 + \|\bar{\mathbf{B}}_0^{-1} \mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2) \cdot \frac{\max_{i\in[b]}\{\gamma_i\}}{\min_{i\in[b]}\{\gamma_i\}} \right)^{1/2} \times \left[ \kappa_{\boldsymbol{\gamma}} \|\mathbf{z} - \tilde{\mathbf{z}}\| + \|\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}} \mathbf{B}_0\| \|\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\| \right] \ ,$$

*where $\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}$ is defined as in (S1) with $\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}} = (\gamma_1/\rho_1^{1/2}, \ldots, \gamma_b/\rho_b^{1/2})$, $\bar{\mathbf{B}}_0$, $\mathbf{B}_0$, $P_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ and $\kappa_{\boldsymbol{\gamma}}$ are given in (S2), (S7), (S12), respectively.*

*Proof.* Consider $(X_k, \tilde{X}_k)_{k\in\mathbb{N}}$ defined in (S8). By (S83) and Lemma S3, we need to bound $(\|Z_k - \tilde{Z}_k\|)_{k\in\mathbb{N}}$. Let $n \in \mathbb{N}^*$. For any $i \in [b]$, we have by (S9), that

$$Z^i_{n+1} - \tilde{Z}^i_{n+1} = \left(1 - \frac{\gamma_i}{\rho_i}\right)(Z^i_n - \tilde{Z}^i_n) + \frac{\gamma_i}{\rho_i} \mathbf{A}_i(\theta_n - \tilde{\theta}_n) - \gamma_i \left( \nabla U_i(Z^i_n) - \nabla U_i(\tilde{Z}^i_n) \right) \ . \tag{S13}$$

Since $U_i$ is twice differentiable, we have

$$\nabla U_i(Z^i_n) - \nabla U_i(\tilde{Z}^i_n) = \int_0^1 \nabla^2 U_i(\tilde{Z}^i_n + t(Z^i_n - \tilde{Z}^i_n))\,\mathrm{d}t \cdot (Z^i_n - \tilde{Z}^i_n) \ .$$

Using $\theta_n - \tilde{\theta}_n = \bar{\mathbf{B}}_0^{-1} \mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}(Z_n - \tilde{Z}_n)$, it follows that

$$Z^i_{n+1} - \tilde{Z}^i_{n+1} = \left( \left[1 - \frac{\gamma_i}{\rho_i}\right] \mathbf{I}_{d_i} - \gamma_i \int_0^1 \nabla^2 U_i(\tilde{Z}^i_n + t(Z^i_n - \tilde{Z}^i_n))\,\mathrm{d}t \right) (Z^i_n - \tilde{Z}^i_n) + \frac{\gamma_i}{\rho_i} \mathbf{A}_i \bar{\mathbf{B}}_0^{-1} \mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2} (Z_n - \tilde{Z}_n) \ .$$

Consider the $p \times p$ block diagonal matrix defined by

$$\mathbf{D}_{U,n} = \mathrm{diag}\left( \gamma_1 \int_0^1 \nabla^2 U_1(\tilde{Z}^1_n + t(Z^1_n - \tilde{Z}^1_n))\,\mathrm{d}t, \cdots, \gamma_b \int_0^1 \nabla^2 U_b(\tilde{Z}^b_n + t(Z^b_n - \tilde{Z}^b_n))\,\mathrm{d}t \right) \ .$$

With the projection matrix $\mathbf{P}_0$ defined in (S3), the difference $Z_{n+1} - \tilde{Z}_{n+1}$ can be rewritten as

$$Z_{n+1} - \tilde{Z}_{n+1} = \left( \mathbf{I}_p - \mathbf{D}_{U,n} - \mathbf{D}_{\boldsymbol{\gamma}}^{1/2} \mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2} (\mathbf{I}_p - \mathbf{P}_0) \tilde{\mathbf{D}}_0^{1/2} \right) (Z_n - \tilde{Z}_n) \ ,$$

where $\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}$ is defined as in (S1) with $\boldsymbol{\gamma}/\boldsymbol{\rho} = (\gamma_1/\rho_1, \ldots, \gamma_b/\rho_b)$. Since $\mathbf{D}_{U,n}$ commutes with $\mathbf{D}_{\boldsymbol{\gamma}}$ and $\mathbf{P}_0$ is an orthogonal projection matrix, using **H**2-(i)-(ii), we get

$$\begin{aligned}
&\|Z_{n+1} - \tilde{Z}_{n+1}\|_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} \\
&= \|\mathbf{D}_{\boldsymbol{\gamma}}^{-1/2}(\mathbf{D}_{\boldsymbol{\gamma}}^{1/2}\mathbf{D}_{\boldsymbol{\gamma}}^{-1/2} - \mathbf{D}_{\boldsymbol{\gamma}}^{1/2}\mathbf{D}_{U,n}\mathbf{D}_{\boldsymbol{\gamma}}^{-1/2} - \mathbf{D}_{\boldsymbol{\gamma}}^{1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}(\mathbf{I}_p - \mathbf{P}_0)\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{\gamma}}^{-1/2})(Z_n - \tilde{Z}_n)\| \\
&\leq \|\mathbf{I}_p - \mathbf{D}_{U,n} - \mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2} \left(\mathbf{I}_p - \mathbf{P}_0\right) \mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\| \|Z_n - \tilde{Z}_n\|_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} \ .
\end{aligned}$$

Note that **H**1 and **H**2 and the fact that $\mathbf{P}_0$ is an orthogonal projector, so $\mathbf{0}_p \preccurlyeq \mathbf{I}_p - \mathbf{P}_0$, imply that

$$\begin{aligned}
\mathrm{diag}(\{1 - \gamma_1(M_1 + 1/\rho_1)\}\mathbf{I}_{d_1}, \cdots, \{1 - \gamma_b(M_b + 1/\rho_b)\}\mathbf{I}_{d_b}) &\preccurlyeq \mathbf{I}_p - \mathbf{D}_{U,n} - \mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2} \left(\mathbf{I}_p - \mathbf{P}_0\right) \mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2} \\
&\preccurlyeq \mathrm{diag}\left(\{1 - \gamma_1 m_1\}\mathbf{I}_{d_1}, \ldots, \{1 - \gamma_b m_b\}\mathbf{I}_{d_b}\right) \ .
\end{aligned}$$

Therefore, we get

$$\begin{aligned}\|Z_{n+1}-\tilde{Z}_{n+1}\|_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} &\leq \max_{i\in[b]}\left\{\max(|1-\gamma_i m_i|, |1-\gamma_i(M_i+1/\rho_i)|)\right\}\|Z_n-\tilde{Z}_n\|_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} \\ &= \kappa_{\boldsymbol{\gamma}}\|Z_n-\tilde{Z}_n\|_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} \ . \end{aligned} \tag{S14}$$

An immediate induction shows, for any $n \geq 1$,

$$\|Z_n-\tilde{Z}_n\|_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} \leq \kappa_{\boldsymbol{\gamma}}^{n-1}\|Z_1-\tilde{Z}_1\|_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} \ . \tag{S15}$$

In addition, by (S13), we have for any $i \in [b]$,

$$Z_1^i - \tilde{Z}_1^i = \left(1-\frac{\gamma_i}{\rho_i}\right)(\mathbf{z}_i-\tilde{\mathbf{z}}_i) + \frac{\gamma_i}{\rho_i}\mathbf{A}_i(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}) - \gamma_i(\nabla U_i(\mathbf{z}_i) - \nabla U_i(\tilde{\mathbf{z}}_i)) \ .$$

It follows that $Z_1 - \tilde{Z}_1 = (\mathbf{I}_p - \mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}} - \mathbf{D}_{U,0})(\mathbf{z}-\tilde{\mathbf{z}}) + \mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}\tilde{\mathbf{D}}_0^{-1/2}\mathbf{B}_0(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}})$. Using the triangle inequality and **H**2 gives

$$\begin{aligned}\|Z_1-\tilde{Z}_1\|_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} &\leq (\min_{i\in[b]}\{\gamma_i\})^{-1/2}\|(\mathbf{I}_p - \mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}} - \mathbf{D}_{U,0})(\mathbf{z}-\tilde{\mathbf{z}}) + (\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}})\| \\ &\leq (\min_{i\in[b]}\{\gamma_i\})^{-1/2}\left[\|\mathbf{I}_p - \mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}} - \mathbf{D}_{U,0}\|\|\mathbf{z}-\tilde{\mathbf{z}}\| + \|\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\|\|\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}\|\right] \\ &\leq (\min_{i\in[b]}\{\gamma_i\})^{-1/2}\Big[\max_{i\in[b]}\{|1-\gamma_i(m_i+1/\rho_i)|, |1-\gamma_i(M_i+1/\rho_i)|\}\|\mathbf{z}-\tilde{\mathbf{z}}\| \\ &\qquad + \|\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\|\|\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}\|\Big] \\ &\leq (\min_{i\in[b]}\{\gamma_i\})^{-1/2}\left[\kappa_\gamma\|\mathbf{z}-\tilde{\mathbf{z}}\| + \|\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\|\|\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}\|\right] \ .\end{aligned}$$

Combining (S15) and the previous inequality and using Lemma S3, we get for $n \geq 1$,

$$\begin{aligned}\|X_n-\tilde{X}_n\|^2 \leq \kappa_{\boldsymbol{\gamma}}^{2(n-1)}\left(1+\|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2\right)\frac{\max_{i\in[b]}\{\gamma_i\}}{\min_{i\in[b]}\{\gamma_i\}} \\ \times\left[\kappa_\gamma\|\mathbf{z}-\tilde{\mathbf{z}}\| + \|\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\|\|\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}\|\right]^2 \ .\end{aligned}$$

The proof is concluded by (S83). □

### S2.1.2. Specific case of Proposition 2

Based on the previous lemmata, we provide in what follows a specific instance of Proposition 2 in the scenario where $\max_{i\in[b]} N_i = 1$.

**Proposition S5.** *Assume **H**1-**H**2 and let $\boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $\gamma_i \leq 2(m_i + M_i + 1/\rho_i)^{-1}$. Then, $P_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ defined in (S7) admits a unique stationary distribution $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ and for any $\mathbf{x} = (\mathbf{z}^\top, \boldsymbol{\theta}^\top)^\top$ with $\boldsymbol{\theta} \in \mathbb{R}^d$, $\mathbf{z} \in \mathbb{R}^p$ and any $n \in \mathbb{N}^*$, we have*

$$\begin{aligned}W_2^2(\delta_{\mathbf{x}}P_{\boldsymbol{\rho},\boldsymbol{\gamma}}^n, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}) \leq \left(1-\min_{i\in[b]}\{\gamma_i m_i\}\right)^{2(n-1)}\left((1+\|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2)\cdot\frac{\max_{i\in[b]}\{\gamma_i\}}{\min_{i\in[b]}\{\gamma_i\}}\right) \\ \times\int_{\mathbb{R}^d\times\mathbb{R}^p}\left[(1-\min_{i\in[b]}\{\gamma_i m_i\})\|\mathbf{z}-\tilde{\mathbf{z}}\| + \|\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\|\|\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}}\|\right]^2 \mathrm{d}\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\tilde{\mathbf{x}}) \ ,\end{aligned}$$

*where $\bar{\mathbf{B}}_0, \mathbf{B}_0, \tilde{\mathbf{D}}_0$, $P_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ are defined in (S2) and (S3).*

*Proof.* For any $i \in [b]$, note that the condition $0 < \gamma_i \leq 2(m_i+M_i+1/\rho_i)^{-1}$ ensures that $\kappa_{\boldsymbol{\gamma}} = 1-\min_{i\in[b]}\{\gamma_i m_i\} \in (0,1)$ and the proof follows from Lemma S4 combined with Douc et al. (2018, Lemma 20.3.2, Theorem 20.3.4). □

### S2.2. Multiple local LMC iterations

In this section, we consider the general case $\max_{i\in[b]} N_i \geq 1$. For this, we introduce the conditional Markov transition kernel defined for any $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_b)$, $\boldsymbol{N} = (N_1, \dots, N_b)$, $\boldsymbol{\theta} \in \mathbb{R}^d$, $\mathbf{z} = (\mathbf{z}_1, \cdots, \mathbf{z}_b) \in \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_b}$, for $i \in [b]$ and $\mathsf{B}_i \in \mathcal{B}(\mathbb{R}^{d_i})$, by

$$Q_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}\left(\mathbf{z}, \mathsf{B}_1 \times \cdots \times \mathsf{B}_b | \boldsymbol{\theta}\right) = \prod_{i=1}^{b} R_{\rho_i,\gamma_i}^{N_i}(\mathbf{z}_i, \mathsf{B}_i|\boldsymbol{\theta}) \ , \tag{S16}$$

where $R_{\rho_i,\gamma_i}$ is defined by (S6). Then, as in the case $\max_{i\in[b]} N_i = 1$, the Gibbs sampler presented in Algorithm 1 defines a homogeneous Markov chain $X_n^\top = (\theta_n^\top, Z_n^\top)_{n\geq 1}$ where $Z_n^\top = ([Z_n^1]^\top, \cdots, [Z_n^b]^\top)$. Indeed, it is easy to show that for any $n \in \mathbb{N}$ and measurable function $f : \mathbb{R}^p \to \mathbb{R}_+$, $\mathbb{E}[f(Z_{n+1})|X_n] = \int_{\mathbb{R}^p} f(\mathbf{z}) Q_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(Z_n, \mathrm{d}\mathbf{z}|\theta_n)$. Therefore, $(X_n)_{n\in\mathbb{N}}$ is associated with the Markov kernel defined, for any $\mathbf{x}^\top = (\boldsymbol{\theta}^\top, \mathbf{z}^\top)$ and $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$, $\mathsf{B} \in \mathcal{B}(\mathbb{R}^p)$, by

$$P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathbf{x}, \mathsf{A} \times \mathsf{B}) = \int_{\mathsf{B}} Q_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}\left(\mathbf{z}, \mathrm{d}\tilde{\mathbf{z}}|\boldsymbol{\theta}\right) \int_{\mathsf{A}} \Pi_{\boldsymbol{\rho}}(\mathrm{d}\tilde{\boldsymbol{\theta}}|\tilde{\mathbf{z}}) \ , \tag{S17}$$

where $\Pi_{\boldsymbol{\rho}}(\cdot|\tilde{\mathbf{z}})$ is defined in (5). We now define a coupling between $\delta_{\mathbf{x}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ and $\delta_{\tilde{\mathbf{x}}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ for any $n \geq 1$ and $\mathbf{x}, \tilde{\mathbf{x}} \in \mathbb{R}^d \times \mathbb{R}^p$. Let $(\xi_n)_{n\geq 1}$ be a sequence of i.i.d. $d$-dimensional standard Gaussian random variables independent of the family of independent random variables $\{(\eta_n^i)_{n\geq 1} : i \in [b]\}$ where for any $i \in [b]$ and $n \geq 1$, $\eta_n^i$ is a $d_i$-dimensional standard Gaussian random variable. Define by induction the synchronous coupling $(\theta_n, Z_n)_{n\geq 0}$, $(\tilde{\theta}_n, \tilde{Z}_n)_{n\geq 0}$, for any $i \in [b]$ starting from $(\theta_0, Z_0) = \mathbf{x} = (\boldsymbol{\theta}, \mathbf{z})$, $(\tilde{\theta}_0, \tilde{Z}_0) = \tilde{\mathbf{x}} = (\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{z}})$ and for any $n \geq 0$ by

$$\begin{aligned} \tilde{Z}^i_{n+1} &= \tilde{Y}^{(i,n)}_{N_i} \ , \qquad \tilde{\theta}_{n+1} = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2} \tilde{Z}_{n+1} + \bar{\mathbf{B}}_0^{-1/2}\xi_{n+1} \ , \\ Z^i_{n+1} &= Y^{(i,n)}_{N_i} \ , \qquad \theta_{n+1} = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2} Z_{n+1} + \bar{\mathbf{B}}_0^{-1/2}\xi_{n+1} \ , \end{aligned} \tag{S18}$$

where $\bar{\mathbf{B}}_0, \mathbf{B}_0, \tilde{\mathbf{D}}_0$ are given by (S2)-(S3) and $\tilde{Y}^{(i,n)}_0 = \tilde{Z}^i_n$, $Y^{(i,n)}_0 = Z^i_n$, and for any $k \in \mathbb{N}$

$$\begin{aligned} \tilde{Y}^{(i,n)}_{k+1} &= \tilde{Y}^{(i,n)}_k - \gamma_i \nabla V_i(\tilde{Y}^{(i,n)}_k) + (\gamma_i/\rho_i)\mathbf{A}_i\tilde{\theta}_n + \sqrt{2\gamma_i}\eta^{(i,n)}_{k+1} \ , \\ Y^{(i,n)}_{k+1} &= Y^{(i,n)}_k - \gamma_i \nabla V_i(Y^{(i,n)}_k) + (\gamma_i/\rho_i)\mathbf{A}_i\theta_n + \sqrt{2\gamma_i}\eta^{(i,n)}_{k+1} \ , \end{aligned} \tag{S19}$$

where, for any $\mathbf{z}_i \in \mathbb{R}^{d_i}$, $V_i$ is defined by

$$V_i(\mathbf{z}_i) = U_i(\mathbf{z}_i) + (2\rho_i)^{-1}\|\mathbf{z}_i\|^2 \ . \tag{S20}$$

For any $n, k \in \mathbb{N}$ consider the $p \times p$ matrices defined by

$$\begin{aligned} \mathbf{H}^{(n)}_{U,k} &= \mathrm{diag}\bigg(\gamma_1 \int_0^1 \nabla^2 U_1((1-s)Y^{(1,n)}_k + s\tilde{Y}^{(1,n)}_k)\,\mathrm{d}s, \\ &\qquad\qquad \dots, \gamma_b \int_0^1 \nabla^2 U_b((1-s)Y^{(b,n)}_k + s\tilde{Y}^{(b,n)}_k)\,\mathrm{d}s\bigg) \ , \\ \mathbf{J}(k) &= \mathrm{diag}\left(\mathbb{1}_{[N_1]}(k+1)\cdot\mathbf{I}_{d_1}, \cdots, \mathbb{1}_{[N_b]}(k+1)\cdot\mathbf{I}_{d_b}\right) \ , \end{aligned} \tag{S21}$$

$$\mathbf{C}^{(n)}_k = \mathbf{J}(k)(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}} + \mathbf{H}^{(n)}_{U,k}) \ , \tag{S22}$$

$$\mathbf{M}^{(n)}_{k+1} = (\mathbf{I}_p - \mathbf{C}^{(n)}_0)^{-1} \dots (\mathbf{I}_p - \mathbf{C}^{(n)}_k)^{-1} \ , \qquad \text{with } \mathbf{M}^{(n)}_0 = \mathbf{I}_p \ . \tag{S23}$$

Under **H**2, we have $\|\mathbf{C}^{(n)}_k\| \leq \max_{i\in[b]}\{\gamma_i(M_i + 1/\rho_i)\}$, thus if we suppose that for any $i \in [b], 0 < \gamma_i < (M_i + 1/\rho_i)^{-1}$, the matrix $(\mathbf{I}_p - \mathbf{C}^{(n)}_k)$ is invertible. In addition, for any $n \in \mathbb{N}, k \geq \max_{i\in[b]}\{N_i\}, \mathbf{C}^{(n)}_k = \mathbf{0}_{p\times p}$, hence the sequence $(\mathbf{M}^{(n)}_k)_{k\in\mathbb{N}}$ is stationary and we denote its limit by $\mathbf{M}^{(n)}_\infty$ which is equal to $\mathbf{M}^{(n)}_{\max_{i\in[b]}\{N_i\}}$.

#### S2.2.1. Technical lemmata

Similarly to Lemma S3, the following result shows that it is enough to consider the marginal process $(Z_n, \tilde{Z}_n)_{n\geq 0}$ to control

$$W_2(\delta_{\mathbf{x}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \delta_{\tilde{\mathbf{x}}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}) \leq \mathbb{E}\left[\|X_n - \tilde{X}_n\|^2\right]^{1/2} \ . \tag{S24}$$

**Lemma S6.** *Assume **H**1 and let* $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$. *Then, for any* $n \in \mathbb{N}$, *the random variables* $X_n = (\theta_n^\top, Z_n^\top)^\top, \tilde{X}_n = (\tilde{\theta}_n^\top, \tilde{Z}_n^\top)^\top$ *defined in* (S18) *satisfy*

$$\|\tilde{X}_{n+1} - X_{n+1}\|^2 \leq (1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2)\|\tilde{Z}_{n+1} - Z_{n+1}\|^2 \ ,$$

*where* $\bar{\mathbf{B}}_0, \mathbf{B}_0, \tilde{\mathbf{D}}_0$ *are defined in* (S2)-(S3).

*Proof.* The proof is similar to the proof of Lemma S3 and is omitted. □

To ease notation, for any $i \in [b]$, we consider all along this section the quantities

$$\tilde{m}_i = m_i + 1/\rho_i \ , \qquad \tilde{M}_i = M_i + 1/\rho_i \ . \tag{S25}$$

The following lemma provides an explicit expression for $\|\tilde{Z}_{n+1} - Z_{n+1}\|$ with respect to $\|\tilde{Z}_n - Z_n\|$.

**Lemma S7.** *Assume **H**1-**H**2 and let* $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ *such that, for any* $i \in [b], \gamma_i < \tilde{M}_i^{-1}$. *Then, for any* $n \geq 1$, *we have*

$$\|\tilde{Z}_{n+1} - Z_{n+1}\|_{\mathbf{D}_{\boldsymbol{N\gamma}}^{-1}} \leq \left\| [\mathbf{M}_\infty^{(n)}]^{-1} + \sum_{k=0}^{\infty} [\mathbf{M}_\infty^{(n)}]^{-1} \mathbf{M}_{k+1}^{(n)} \mathbf{J}(k) \mathbf{D}_{\boldsymbol{N}}^{-1/2} \mathbf{D}_{\boldsymbol{\gamma/\rho}}^{1/2} \mathbf{P}_0 \mathbf{D}_{\boldsymbol{\gamma/\rho}}^{1/2} \mathbf{D}_{\boldsymbol{N}}^{1/2} \right\| \times \|\tilde{Z}_n - Z_n\|_{\mathbf{D}_{\boldsymbol{N\gamma}}^{-1}} \ , \tag{S26}$$

*where* $(\mathbf{M}_k^{(n)})_{k\in\mathbb{N}}$ *is defined in* (S23), $(\tilde{Z}_k, Z_k)_{k\in\mathbb{N}}$ *in* (S18), $\boldsymbol{N\gamma} = (\gamma_1 N_1, \ldots, \gamma_b N_b)$ *and* $\boldsymbol{\gamma/\rho} = (\gamma_1/\rho_1, \ldots, \gamma_b/\rho_b)$.

*Proof.* Let $n \geq 1$. By (S19), for any $i \in [b], k \in \mathbb{N}$, we obtain

$$\tilde{Y}_{k+1}^{(i,n)} - Y_{k+1}^{(i,n)} = \left(\mathbf{I}_{d_i} - \gamma_i \int_0^1 \nabla^2 V_i((1-s)Y_k^{(i,n)} + s\tilde{Y}_k^{(i,n)})\, \mathrm{d}s\right)(\tilde{Y}_k^{(i,n)} - Y_k^{(i,n)}) + (\gamma_i/\rho_i)\mathbf{A}_i(\tilde{\theta}_n - \theta_n) \ .$$

Consider the process $((\tilde{\mathrm{Y}}_k^{(n)}, \mathrm{Y}_k^{(n)}) = \{\tilde{\mathrm{Y}}_k^{(i,n)}, \mathrm{Y}_k^{(i,n)}\}_{i=1}^b)_{k\in\mathbb{N}}$ with values in $\mathbb{R}^p \times \mathbb{R}^p$ defined for any $i \in [b]$, $k \geq 0$, by

$$\tilde{\mathrm{Y}}_k^{(i,n)} = \tilde{Y}_{\min(k,N_i)}^{(i,n)} \ , \qquad \mathrm{Y}_k^{(i,n)} = Y_{\min(k,N_i)}^{(i,n)} \ . \tag{S27}$$

By (S18), we have $\mathbf{A}_i(\tilde{\theta}_n - \theta_n) = \mathbf{A}_i\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}(\tilde{Z}_n - Z_n)$. Since $\mathbf{B}_0^\top = [\mathbf{A}_1^\top/\rho_1^{1/2} \cdots \mathbf{A}_b^\top/\rho_b^{1/2}]$ and $\mathbf{P}_0 = \mathbf{B}_0\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top$ is the orthogonal projection matrix defined in (S3), it follows that

$$\tilde{\mathrm{Y}}_{k+1}^{(n)} - \mathrm{Y}_{k+1}^{(n)} = \big(\mathbf{I}_p - \mathbf{C}_k^{(n)}\big)(\tilde{\mathrm{Y}}_k^{(n)} - \mathrm{Y}_k^{(n)}) + \mathbf{J}(k)\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{P}_0\tilde{\mathbf{D}}_0^{1/2}(\tilde{\mathrm{Y}}_0^{(n)} - \mathrm{Y}_0^{(n)}) \ . \tag{S28}$$

Since $\mathbf{D}_{\boldsymbol{N\gamma}}$ commutes with $\mathbf{C}_k^{(n)}$ and $\mathbf{J}(k)$, multiplying (S28) by $\mathbf{M}_{k+1}^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}$, yields

$$\mathbf{M}_{k+1}^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}(\tilde{\mathrm{Y}}_{k+1}^{(n)} - \mathrm{Y}_{k+1}^{(n)}) = \mathbf{M}_k^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}(\tilde{\mathrm{Y}}_k^{(n)} - \mathrm{Y}_k^{(n)}) + \mathbf{M}_{k+1}^{(n)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma/\rho}}^{1/2}\mathbf{P}_0\tilde{\mathbf{D}}_0^{1/2}(\tilde{\mathrm{Y}}_0^{(n)} - \mathrm{Y}_0^{(n)}) \ . \tag{S29}$$

By definition of the processes in (S18)-(S19) and (S27), we have for $k \geq \max_{i\in[b]}\{N_i\}$, $(\tilde{\mathrm{Y}}_k^{(n)}, \mathrm{Y}_k^{(n)}) = (\tilde{Z}_{n+1}, Z_{n+1})$ and $\mathbf{J}(k) = \mathbf{0}_{p\times p}$. Therefore summing the previous equality (S29) yields

$$\mathbf{M}_\infty^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}(\tilde{Z}_{n+1} - Z_{n+1}) = \big[\mathbf{M}_0^{(n)} + \textstyle\sum_{k=0}^{\infty}\mathbf{M}_{k+1}^{(n)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma/\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma/\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}\big] \times \mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}(\tilde{\mathrm{Y}}_0^{(n)} - \mathrm{Y}_0^{(n)}) \ .$$

Multiplying this last equality by $[\mathbf{M}_\infty^{(n)}]^{-1}$ and applying the norm $\|\cdot\|_{\mathbf{D}_{\boldsymbol{N\gamma}}^{-1}}$ concludes the proof. □

The three following lemmata aim at providing an explicit upper bound on (S26). To this end, for $n, k \in \mathbb{N}$ and $i \in [b]$, consider $\mathbf{C}_k^{(i,n)}$ corresponding to the $i$-th diagonal block of $\mathbf{C}_k^{(n)}$ defined in (S22), *i.e.*

$$\mathbf{C}_k^{(i,n)} = \mathbb{1}_{[N_i]}(k+1)\gamma_i \left\{ \rho_i^{-1}\mathbf{I}_{d_i} + \int_0^1 \nabla^2 U_i((1-s)Y_k^{(i,n)} + s\tilde{Y}_k^{(i,n)})\,\mathrm{d}s \right\} \in \mathbb{R}^{d_i \times d_i} , \tag{S30}$$

where, for any $n \in \mathbb{N}$ and $i \in [b]$, $(Y_k^{(i,n)}, \tilde{Y}_k^{(i,n)})_{k\in\mathbb{N}}$ is defined in (S19). Thus, using the definition (S23) of $\mathbf{M}_k^{(n)}$, we can write $[\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_k^{(n)}$ as a block-diagonal matrix $\mathrm{diag}(([\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_k^{(n)})^1, \ldots, ([\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_k^{(n)})^b)$ where for any $i \in [b]$, $([\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_k^{(n)})^i = \prod_{l=k}^{N_i-1}(\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)}) \in \mathbb{R}^{d_i \times d_i}$.

**Lemma S8.** *Assume **H**1-**H**2 and let $\boldsymbol{N} \in (\mathbb{R}_+^*)^b$, $\boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $\gamma_i < \tilde{M}_i^{-1}$. Then, for any $i \in [b]$, $n \in \mathbb{N}$ and $k \in [N_i]$, we have*

$$\|([\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_k^{(n)})^i - \mathbf{I}_{d_i} - \textstyle\sum_{l=k}^{\infty} \mathbf{C}_l^{(i,n)}\| \leq \exp\{(N_i - k)\gamma_i \tilde{M}_i\} - 1 - (N_i - k)\gamma_i \tilde{M}_i ,$$

*where $\mathbf{M}_k^{(n)}, \tilde{M}_i$ are defined in (S23), (S25) respectively, and $\mathbf{M}_\infty^{(n)}$ is the limit of the stationnary sequence $(\mathbf{M}_k^{(n)})_{k\in\mathbb{N}}$.*

*Proof.* Let $n \in \mathbb{N}$, $i \in [b]$ and $k \in [N_i]$. The approximation error between $\prod_{l=k}^{\infty}(\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)})$ and its linear approximation can be upper bounded as

$$\begin{aligned}
\left\| \prod_{l=k}^{\infty}(\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)}) - \mathbf{I}_{d_i} - \sum_{l=k}^{\infty} \mathbf{C}_l^{(i,n)} \right\| &= \left\| \sum_{m=2}^{\infty} (-1)^m \sum_{k \leq l_1 < \cdots < l_m} \mathbf{C}_{l_1}^{(i,n)} \cdots \mathbf{C}_{l_m}^{(i,n)} \right\| \\
&\leq \sum_{m=2}^{\infty} \sum_{k \leq l_1 < \cdots < l_m} \|\mathbf{C}_{l_1}^{(i,n)}\| \cdots \|\mathbf{C}_{l_m}^{(i,n)}\| = \prod_{l=k}^{\infty}(1 + \|\mathbf{C}_l^{(i,n)}\|) - 1 - \sum_{l \geq k} \|\mathbf{C}_l^{(i,n)}\| \\
&\leq \exp\left( \sum_{l=k}^{\infty} \|\mathbf{C}_l^{(i,n)}\| \right) - 1 - \sum_{l=k}^{\infty} \|\mathbf{C}_l^{(i,n)}\| ,
\end{aligned}$$

where the products and the sums are well defined since for any $l \geq N_i$, we have $\mathbf{C}_l^{(i,n)} = \mathbf{0}_{d_i}$. Finally, the proof is concluded using that $x \mapsto \exp(x) - 1 - x$ is increasing on $\mathbb{R}$ and for $l \in \mathbb{N}$, $\|\mathbf{C}_l^{(i,n)}\| \leq \gamma_i \tilde{M}_i \mathbb{1}_{[N_i]}(l+1)$ from **H**2-(i). □

For any $\boldsymbol{N} = (N_1, \ldots, N_b) \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_b) \in (\mathbb{R}_+^*)^b$, define the $p \times p$ block matrices

$$\begin{aligned}
\mathbf{S}_1 &= \mathrm{diag}(\{1 - N_1\gamma_1\tilde{M}_1\}\mathbf{I}_{d_1}, \cdots, \{1 - N_b\gamma_b\tilde{M}_b\}\mathbf{I}_{d_b}) , \\
\mathbf{S}_2 &= \mathbf{I}_p - \sum_{l=0}^{\infty} \mathbf{J}(l)\mathbf{H}_{U,l}^{(n)} - (\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2}(\mathbf{I}_p - \mathbf{P}_0)(\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2} , \\
\mathbf{S}_3 &= \mathrm{diag}\left(\{1 - N_1\gamma_1 m_1\}\mathbf{I}_{d_1}, \ldots, \{1 - N_b\gamma_b m_b\}\mathbf{I}_{d_b}\right) ,
\end{aligned} \tag{S31}$$

where for any $i \in [b]$, $\tilde{M}_i$ is defined in (S25) and $\mathbf{P}_0, \mathbf{J}(l), \mathbf{H}_{U,l}^{(n)}$ are defined in (S3), (S86), (S87), respectively.

**Lemma S9.** *Assume **H**1-**H**2. Then, for any $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$, we have*

$$\mathbf{S}_1 \preccurlyeq \mathbf{S}_2 \preccurlyeq \mathbf{S}_3 .$$

*As a result, under the additional assumption, for any $i \in [b]$, $\gamma_i N_i \leq 2/(m_i + M_i + 1/\rho_i)$, we get*

$$\|\mathbf{S}_2\| \leq 1 - \min_{i\in[b]}\{N_i\gamma_i m_i\} . \tag{S32}$$

*Proof.* Since $\mathbf{P}_0$ is an orthogonal projection defined in (S3), we have $\mathbf{P}_0 \preccurlyeq \mathbf{I}_p$, therefore we easily get

$$\mathbf{0}_{p\times p} \preccurlyeq (\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2}(\mathbf{I}_p - \mathbf{P}_0)(\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2} \preccurlyeq \mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}$$

and **H**2-(i)-(ii) imply

$$\operatorname{diag}(N_1\gamma_1 m_1\mathbf{I}_{d_1},\cdots,N_b\gamma_b m_b\mathbf{I}_{d_b}) \preccurlyeq \sum_{l=0}^{\infty}\mathbf{J}(l)\mathbf{H}_{U,l}^{(n)} \preccurlyeq \operatorname{diag}(N_1\gamma_1 M_1\mathbf{I}_{d_1},\cdots,N_b\gamma_b M_b\mathbf{I}_{d_b}) \ .$$

Substracting these previous inequalities and adding $\mathbf{I}_p$ complete the first part of the proof. The additional condition, for any $i\in[b]$, $\gamma_i N_i \le 2/(m_i+M_i+1/\rho_i)$, ensures that $\mathbf{S}_1$ is definite-positive. Since $\mathbf{S}_1 \preceq \mathbf{S}_2$, we deduce that $\mathbf{S}_2$ is symmetric positive-definite as well. Then, $\|\mathbf{S}_2\|$ is equal to the largest eigenvalue of $\mathbf{S}_2$. The inequality $\mathbf{S}_2 \preceq \mathbf{S}_3$ concludes the second part of the proof. □

For any $\boldsymbol{N}=(N_1,\ldots,N_b)\in(\mathbb{N}^*)^b, \boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_b)\in(\mathbb{R}_+^*)^b$, define

$$r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} = \max_{i\in[b]}\{N_i\gamma_i/\rho_i\}\max_{i\in[b]}\{N_i\gamma_i\tilde{M}_i\}\Big(1/2 + \max_{i\in[b]}\{N_i\gamma_i\tilde{M}_i\}\Big) + 4\max_{i\in[b]}\{N_i\gamma_i\tilde{M}_i\}^2 \quad , \tag{S33}$$

where $\tilde{M}_i$ is defined in (S25).

**Lemma S10.** *Assume **H**1-**H**2. Let $\boldsymbol{N}\in(\mathbb{N}^*)^b, \boldsymbol{\gamma}\in(\mathbb{R}_+^*)^b$ such that, for any $i\in[b]$, $N_i\gamma_i\le 2/(m_i+\tilde{M}_i)$ and $\gamma_i<\tilde{M}_i^{-1}$. Then, for any $n\in\mathbb{N}$, we have*

$$\|[\mathbf{M}_\infty^{(n)}]^{-1}+\textstyle\sum_{k=0}^{\infty}[\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_{k+1}^{(n)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}\| \le 1-\min_{i\in[b]}\{N_i\gamma_i m_i\}+r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \ ,$$

*where $\mathbf{P}_0$, $\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}$, $\mathbf{J}(k)$, $\mathbf{M}_k^{(n)}$ and $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ are defined in (S3), (S21), (S23) and (S33), respectively.*

*Proof.* Let $n\in\mathbb{N}$. For any $k\in\mathbb{N}$, define

$$\mathbf{R}_k^{(n)}=\prod_{l=k}^{\infty}(\mathbf{I}_p-\mathbf{C}_l^{(n)})-\mathbf{I}_p+\sum_{l=k}^{\infty}\mathbf{C}_l^{(n)} \ , \qquad \mathbf{R}_k^{(i,n)}=\prod_{l=k}^{\infty}(\mathbf{I}_{d_i}-\mathbf{C}_l^{(i,n)})-\mathbf{I}_{d_i}+\sum_{l=k}^{\infty}\mathbf{C}_l^{(i,n)} \ , \quad i\in[b] \ , \tag{S34}$$

where $(\mathbf{C}_l^{(i,n)})_{l\in\mathbb{N}}$ is defined in (S30) and remark that the products and the sums are well defined since for any $l\ge N_i$, we have $\mathbf{C}_l^{(i,n)}=\mathbf{0}_{d_i}$. By noting, for any $k\in[\max_{i\in[b]}N_i]$, that $[\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_k^{(n)}=\prod_{l=k}^{\infty}(\mathbf{I}_p-\mathbf{C}_l^{(n)})$, it follows that $[\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_k^{(n)}=\mathbf{I}_p-\sum_{l=k}^{\infty}\mathbf{C}_l^{(n)}+\mathbf{R}_k^{(n)}$. Since for any $i\in[b], l\ge N_i$, $\mathbf{R}_k^{(i,n)}=\mathbf{0}_{d_i}$, thus we have $\mathbf{J}(k)\mathbf{R}_{k+1}^{(n)}=\mathbf{R}_{k+1}^{(n)}$. In addition, using that $\mathbf{M}_0^{(n)}=\mathbf{I}_p$, $\mathbf{C}_l^{(n)}=\mathbf{J}(l)(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}+\mathbf{H}_{U,l}^{(n)})$, $\mathbf{D}_{\boldsymbol{N}}=\sum_{k=0}^{\infty}\mathbf{J}(k)$, $\mathbf{D}_{\boldsymbol{N}}\mathbf{C}_l^{(n)}=\mathbf{C}_l^{(n)}\mathbf{D}_{\boldsymbol{N}}$, we get

$$\begin{aligned}
&[\mathbf{M}_\infty^{(n)}]^{-1}+\sum_{k=0}^{\infty}[\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_{k+1}^{(n)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}\\
&=\mathbf{I}_p-\sum_{l=0}^{\infty}\mathbf{C}_l^{(n)}+\sum_{k=0}^{\infty}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}-\sum_{k=0}^{\infty}\sum_{l=k+1}^{\infty}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{C}_l^{(n)}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}\\
&\quad+\mathbf{R}_0^{(n)}+\sum_{k=0}^{\infty}\mathbf{R}_{k+1}^{(n)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}\\
&=\mathbf{I}_p-\sum_{l=0}^{\infty}\mathbf{J}(l)\mathbf{H}_{U,l}^{(n)}-\Big(\sum_{k=0}^{\infty}\mathbf{J}(k)\Big)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}(\mathbf{I}_p-\mathbf{P}_0)\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}\\
&\quad-\sum_{l=1}^{\infty}\Big(\sum_{k=0}^{l-1}\mathbf{J}(k)\Big)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{C}_l^{(n)}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}+\mathbf{R}_0^{(n)}+\sum_{k=0}^{\infty}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{R}_{k+1}^{(n)}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}\\
&=\mathbf{S}_2-\sum_{l=1}^{\infty}\Big(\sum_{k=0}^{l-1}\mathbf{J}(k)\Big)\mathbf{D}_{\boldsymbol{N}}^{-1}\mathbf{C}_l^{(n)}(\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2}\mathbf{P}_0(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}\mathbf{D}_{\boldsymbol{N}})^{1/2}\\
&\quad+\mathbf{R}_0^{(n)}+\sum_{k=1}^{\infty}\mathbf{D}_{\boldsymbol{N}}^{-1}\mathbf{R}_k^{(n)}(\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2}\mathbf{P}_0(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}\mathbf{D}_{\boldsymbol{N}})^{1/2} \ ,
\end{aligned} \tag{S35}$$

where $\mathbf{S}_2$ is defined in (S31). We now bound the different terms of (S35) separately. First, using (S32), we have

$$\|\mathbf{S}_2\| \leq 1 - \min_{i\in[b]}\{N_i\gamma_i m_i\} \ . \tag{S36}$$

By recalling $\mathbf{R}_0^{(n)}$ defined in (S34), Lemma S8 shows that

$$\begin{aligned}
\|\mathbf{R}_0^{(n)}\| \leq \max_{i\in[b]}\|\mathbf{R}_0^{(i,n)}\| &= \max_{i\in[b]}\left\{\left\|\prod_{l=0}^{\infty}\left(\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)}\right) - \mathbf{I}_{d_i} - \sum_{l=0}^{\infty}\mathbf{C}_l^{(i,n)}\right\|\right\} &\text{(S37)}\\
&\leq \max_{i\in[b]}\left\{\exp\left(\sum_{l=0}^{\infty}\|\mathbf{C}_l^{(i,n)}\|\right) - 1 - \sum_{l=0}^{\infty}\|\mathbf{C}_l^{(i,n)}\|\right\} &\text{(S38)}\\
&\leq \max_{i\in[b]}\left\{\exp\{(N_i-1)\gamma_i\tilde{M}_i\} - 1 - (N_i-1)\gamma_i\tilde{M}_i\right\} &\text{(S39)}\\
&\leq \max_{i\in[b]}\{((N_i-1)\gamma_i\tilde{M}_i)^2\mathrm{e}^{(N_i-1)\gamma_i\tilde{M}_i}\}/2 &\text{(S40)}\\
&\leq 4\max_{i\in[b]}\{(N_i-1)\gamma_i\tilde{M}_i\}^2 \ , &\text{(S41)}
\end{aligned}$$

where, in the penultimate line, we used for any $t \geq 0$, that $\exp(t) - 1 - t \leq t^2\exp(t)/2$. Regarding the second term of (S35), using that $\mathbf{P}_0$ is an orthogonal projector, we get

$$\left\|\sum_{l=1}^{\infty}\Big(\sum_{k=0}^{l-1}\mathbf{J}(k)\Big)\mathbf{D}_{\boldsymbol{N}}^{-1}\mathbf{C}_l^{(n)}(\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2}\mathbf{P}_0(\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2}\right\| \leq \max_{i\in[b]}\left(\frac{N_i\gamma_i}{\rho_i}\right)\left\|\sum_{l=1}^{\infty}\Big(\sum_{k=0}^{l-1}\mathbf{J}(k)\Big)\mathbf{D}_{\boldsymbol{N}}^{-1}\mathbf{C}_l^{(n)}\right\| \ .$$

Combining the following upper bound

$$\left\|\sum_{l=1}^{\infty}\Big(\sum_{k=0}^{l-1}\mathbf{J}(k)\Big)\mathbf{D}_{\boldsymbol{N}}^{-1}\mathbf{C}_l^{(n)}\right\| \leq \max_{i\in[b]}\left\{\frac{1}{N_i}\sum_{l=1}^{\infty}l\|\mathbf{C}_l^{(i,n)}\|\right\}$$

with the fact, for any $i \in [b]$, that $\|\mathbf{C}_l^{(i,n)}\| \leq \gamma_i\tilde{M}_i\mathbb{1}_{[N_i]}(l+1)$, we get that

$$\left\|\sum_{l=1}^{\infty}\Big(\sum_{k=0}^{l-1}\mathbf{J}(k)\Big)\mathbf{D}_{\boldsymbol{N}}^{-1}\mathbf{C}_l^{(n)}(\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2}\mathbf{P}_0(\mathbf{D}_{\boldsymbol{N}}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}})^{1/2}\right\| \leq \max_{i\in[b]}\left(\frac{N_i\gamma_i}{\rho_i}\right)\max_{i\in[b]}\left\{\frac{N_i\gamma_i\tilde{M}_i}{2}\right\} . \tag{S42}$$

To upper bound the last term of (S35), we start from the following inequality

$$\left\|\sum_{k=1}^{\infty}\mathbf{D}_{\boldsymbol{N}}^{-1}\mathbf{R}_k^{(n)}\right\| \leq \max_{i\in[b]}\left\{\frac{1}{N_i}\sum_{k=1}^{N_i-1}\|\mathbf{R}_k^{(i,n)}\|\right\} .$$

Lemma S8 shows that for any $k \in [N_i - 1]$ and $i \in [b]$, $\|\mathbf{R}_k^{(i,n)}\| \leq \exp\{(N_i-k)\gamma_i\tilde{M}_i\} - 1 - (N_i-k)\gamma_i\tilde{M}_i$. Then, for any $i \in [b]$, we have

$$\begin{aligned}
\frac{1}{N_i}\sum_{k=1}^{N_i-1}\|\mathbf{R}_k^{(i,n)}\| &\leq \frac{1}{N_i}\sum_{k=1}^{N_i-1}[\exp\{(N_i-k)\gamma_i\tilde{M}_i\} - 1 - (N_i-k)\gamma_i\tilde{M}_i]\\
&\leq (N_i\gamma_i\tilde{M}_i)^{-1}\int_0^{N_i\gamma_i\tilde{M}_i}(\mathrm{e}^t - 1 - t)\,\mathrm{d}t \leq \frac{(N_i\gamma_i\tilde{M}_i)^2}{12}\left(\mathrm{e}^{N_i\gamma_i\tilde{M}_i} + 1\right)\\
&\leq \max_{i\in[b]}\{(N_i\gamma_i\tilde{M}_i)^2\} \ , &\text{(S43)}
\end{aligned}$$

where we have used $\mathrm{e}^2+1 \leq 12$. Plugging (S43), (S42), (S41) into (S32), we get

$$\left\|[\mathbf{M}_\infty^{(n)}]^{-1} + \sum_{k\in\mathbb{N}} [\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_{k+1}^{(n)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}\right\| \leq 1 - \min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \;,$$

where $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ is defined in (S33). □

**Lemma S11.** *Assume **H**1-**H**2. Let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $N_i\gamma_i \leq 2/(m_i + \tilde{M}_i)$ and $\gamma_i < \tilde{M}_i^{-1}$. Then, for any $\mathbf{x} = (\mathbf{z}^\top, \boldsymbol{\theta}^\top)^\top, \tilde{\mathbf{x}} = (\tilde{\mathbf{z}}^\top, \tilde{\boldsymbol{\theta}}^\top)^\top \in \mathbb{R}^{p+d}$, with $(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}) \in (\mathbb{R}^d)^2, (\mathbf{z}, \tilde{\mathbf{z}}) \in (\mathbb{R}^p)^2$ and any $n \geq 1$ we have*

$$\begin{aligned} W_2^2(\delta_{\tilde{\mathbf{x}}}P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \delta_{\mathbf{x}}P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}) \leq (1 - \min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^{2n-2}(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2) \\ \times \frac{\max_{i\in[b]}\{N_i\gamma_i\}}{\min_{i\in[b]}\{N_i\gamma_i\}}\left[\|[\mathbf{M}_\infty^{(0)}]^{-1}\|\|\tilde{\mathbf{z}} - \mathbf{z}\| + (\textstyle\sum_{i\in[b]}\|\mathbf{A}_i\|/\rho_i)\|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}\|\right]^2 \;, \end{aligned}$$

*where $\mathbf{B}_0, \bar{\mathbf{B}}_0, \tilde{\mathbf{D}}_0, P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \mathbf{M}_\infty^{(0)}, r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ are defined in (S2), (S3), (S17), (S23), (S33), respectively.*

*Proof.* Combining Lemma S7 and Lemma S10, we have for $n \geq 1$,

$$\|\tilde{Z}_{n+1} - Z_{n+1}\|_{\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1}} \leq (1 - \min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})\|\tilde{Z}_n - Z_n\|_{\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1}} \;.$$

Thereby, for any $n \geq 1$, we obtain by induction

$$\|\tilde{Z}_n - Z_n\|_{\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1}} \leq (1 - \min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^{n-1}\|\tilde{Z}_1 - Z_1\|_{\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1}} \;. \tag{S44}$$

Define the process $((\tilde{\mathrm{Y}}_k^{(0)}, \mathrm{Y}_k^{(0)}) = \{\tilde{\mathrm{Y}}_k^{(i,0)}, \mathrm{Y}_k^{(i,0)}\}_{i=1}^b)_{k\in\mathbb{N}}$ with values in $\mathbb{R}^p \times \mathbb{R}^p$ defined for any $i \in [b]$, $k \geq 0$ by

$$\tilde{\mathrm{Y}}_k^{(i,0)} = \tilde{Y}_{\min(k,N_i)}^{(i,0)} \;, \qquad\qquad \mathrm{Y}_k^{(i,0)} = Y_{\min(k,N_i)}^{(i,0)} \;.$$

By (S18), it follows that for any $i \in [b]$, $(\tilde{Z}_1^i, Z_1^i) = (\tilde{Y}_{N_i}^{(i,0)}, Y_{N_i}^{(i,0)})$ where $(\tilde{Y}_0^{(i,0)}, Y_0^{(i,0)}) = (\tilde{Z}_0^i, Z_0^i)$. We get by (S19) for $k \geq 0$,

$$\tilde{\mathrm{Y}}_{k+1}^{(0)} - \mathrm{Y}_{k+1}^{(0)} = (\mathbf{I}_p - \mathbf{C}_k^{(0)})(\tilde{\mathrm{Y}}_k^{(0)} - \mathrm{Y}_k^{(0)}) + \mathbf{J}(k)\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0(\tilde{\theta}_0 - \theta_0) \;.$$

Hence, for $k \geq 0$, we obtain

$$\mathbf{M}_{k+1}^{(0)}\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\tilde{\mathrm{Y}}_{k+1}^{(0)} - \mathrm{Y}_{k+1}^{(0)}) = \mathbf{M}_k^{(0)}\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\tilde{\mathrm{Y}}_k^{(0)} - \mathrm{Y}_k^{(0)}) + \mathbf{M}_{k+1}^{(0)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{B}_0(\tilde{\theta}_0 - \theta_0) \;.$$

Summing the previous equality gives

$$\mathbf{M}_\infty^{(0)}\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\tilde{\mathrm{Y}}_{\boldsymbol{N}}^{(0)} - \mathrm{Y}_{\boldsymbol{N}}^{(0)}) = \mathbf{M}_0^{(0)}\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\tilde{\mathrm{Y}}_0^{(0)} - \mathrm{Y}_0^{(0)}) + \sum_{k=0}^{\infty}\mathbf{M}_{k+1}^{(0)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{B}_0(\tilde{\theta}_0 - \theta_0) \;.$$

Multiplying by $[\mathbf{M}_\infty^{(0)}]^{-1}$ and using the fact that $(\theta_0, \mathrm{Y}_0^{(0)}) = (\boldsymbol{\theta}, \mathbf{z})$, $(\tilde{\theta}_0, \tilde{\mathrm{Y}}_0^{(0)}) = (\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{z}})$, we get

$$\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\tilde{Z}_1 - Z_1) = [\mathbf{M}_\infty^{(0)}]^{-1}\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\tilde{\mathbf{z}} - \mathbf{z}) + \sum_{k=0}^{\infty}[\mathbf{M}_\infty^{(0)}]^{-1}\mathbf{M}_{k+1}^{(0)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{B}_0(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) \;.$$

Plugging the result in (S52) implies for any $n \geq 1$,

$$\begin{aligned} \|\tilde{Z}_n - Z_n\|_{\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1}} \leq (1 - \min\nolimits_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^{n-1}\Bigg[\|[\mathbf{M}_\infty^{(0)}]^{-1}\|\|\tilde{\mathbf{z}} - \mathbf{z}\|_{\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1}} \\ + \|\sum_{k=0}^{\infty}[\mathbf{M}_\infty^{(0)}]^{-1}\mathbf{M}_{k+1}^{(0)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{B}_0\|\|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}\|\Bigg] \;. \end{aligned} \tag{S45}$$

By $\mathbf{H}$2-(ii) and the definitions of $\mathbf{C}_l^{(0)}, \mathbf{M}_k^{(0)}$ given in (S22), (S23), we have $\|\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,0)}\| \leq 1 - \gamma_i \tilde{m}_i$. As a result and since $([\mathbf{M}_\infty^{(0)}]^{-1}\mathbf{M}_k^{(0)})^i = \prod_{l=0}^{k-1}(\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,0)})$, the triangle inequality implies

$$
\begin{aligned}
\left\|\sum_{k=0}^{\infty}[\mathbf{M}_\infty^{(0)}]^{-1}\mathbf{M}_{k+1}^{(0)}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{B}_0\right\| &\leq \sum_{i\in[b]}\sqrt{\gamma_i/N_i}(\|\mathbf{A}_i\|/\rho_i)\sum_{k=1}^{N_i}\|([\mathbf{M}_\infty^{(0)}]^{-1}\mathbf{M}_k^{(0)})^i\| \\
&\leq \sum_{i\in[b]}\sqrt{\gamma_i/N_i}(\|\mathbf{A}_i\|/\rho_i)\sum_{k=0}^{N_i-1}(1-\gamma_i\tilde{m}_i)^k \\
&\leq \sum_{i\in[b]}\|\mathbf{A}_i\|\sqrt{N_i\gamma_i}/\rho_i \;.
\end{aligned}
$$

Plugging this result in (S45), we get

$$
\begin{aligned}
\|\tilde{Z}_n - Z_n\|_{\mathbf{D}_{\boldsymbol{N\gamma}}^{-1}} \leq (1-\min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^{n-1}\Big[&\|[\mathbf{M}_\infty^{(0)}]^{-1}\|\|\tilde{\mathbf{z}}-\mathbf{z}\|_{\mathbf{D}_{\boldsymbol{N\gamma}}^{-1}} \\
&+\Big(\sum_{i\in[b]}\|\mathbf{A}_i\|\sqrt{N_i\gamma_i}/\rho_i\Big)\|\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta}\|\Big] \;.
\end{aligned}
$$

Finally, Lemma S6 gives

$$
\begin{aligned}
\|\tilde{X}_n - X_n\|^2 \leq (1-\min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^{2n-2}\cdot(1+\|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2)\frac{\max_{i\in[b]}\{N_i\gamma_i\}}{\min_{i\in[b]}\{N_i\gamma_i\}}& \\
\times\left[\|[\mathbf{M}_\infty^{(0)}]^{-1}\|\|\tilde{\mathbf{z}}-\mathbf{z}\| + \Big(\sum_{i\in[b]}\|\mathbf{A}_i\|/\rho_i\Big)\|\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta}\|\right]^2& \;.
\end{aligned}
$$

Plugging this result into (S24) concludes the proof. □

The following result gives a condition on $\max_{i\in[b]}\{N_i\gamma_i\}$ to simplify the contrating term in Lemma S11 to $1 - \min_{i\in[b]}\{N_i\gamma_i m_i\}/2$. To this end, define

$$
\begin{aligned}
A_0 &= \max_{i\in[b]}\{\tilde{M}_i\}\max_{i\in[b]}\{1/\rho_i\}/2 + 4\max_{i\in[b]}\{\tilde{M}_i\}^2 \;, \\
A_1 &= \max_{i\in[b]}\{\tilde{M}_i\}^2\max_{i\in[b]}\{1/\rho_i\} \;.
\end{aligned}
$$

**Lemma S12.** *Assume **H**1-**H**2 and let* $c \in \mathbb{R}_+^*, \boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ *such that*

$$
\begin{aligned}
&\min_{i\in[b]}\{N_i\gamma_i\}/\max_{i\in[b]}\{N_i\gamma_i\} \geq c \;, \\
&\max_{i\in[b]}\{N_i\gamma_i\} \leq \frac{c\min_{i\in[b]}\{m_i\}}{2A_0 + \sqrt{2A_1 c\min_{i\in[b]}\{m_i\}}} \wedge \frac{2}{\max_{i\in[b]}\{m_i + M_i + 1/\rho_i\}} \;.
\end{aligned}
\tag{S46}
$$

*Then,* $1 - \min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} < 1 - \min_{i\in[b]}\{N_i\gamma_i m_i\}/2 < 1$*, where* $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ *is defined in* (S33)*.*

*Proof.* The proof is straightforward solving a second order polynomial inequality and using for any $a, b \in \mathbb{R}_+^*$, $a + \frac{b^2}{2a+b} \leq \sqrt{a^2+b^2}$. □

### S2.2.2. Proof of Proposition 2

The next proposition quantifies the convergence of $\delta_{\mathbf{x}}P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}^n$ towards $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$, where $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ is the stationnary distribution derived in Proposition S5. In addition, it generalises and gives a more formal statement than Proposition 2.

**Proposition S13.** *Assume **H**1-**H**2 and let $c > 0$ and $\boldsymbol{\gamma} = \{\gamma_i\}_{i=1}^b$, $\boldsymbol{N} \in (\mathbb{N}^*)^b$ such that (S46) is satisfied, for any $i \in [b]$, $N_i\gamma_i < 2/\max_{i\in[b]}\{m_i + \tilde{M}_i\}$ and $\gamma_i < \tilde{M}_i^{-1}$. Then, $P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ defined in (S17) admits a unique invariant probability measure $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$. In addition, for any $\mathbf{x} = (\mathbf{z}^\top, \boldsymbol{\theta}^\top)^\top$ whith $(\boldsymbol{\theta}, \mathbf{z}) \in \mathbb{R}^d \times \mathbb{R}^p$, any integer $n \geq 1$, we have*

$$
\begin{aligned}
W_2^2(\delta_{\mathbf{x}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}) \leq (1 - \min_{i\in[b]}\{N_i\gamma_i m_i\}/2)^{2n-2} \cdot (1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2) \frac{\max_{i\in[b]}\{N_i\gamma_i\}}{\min_{i\in[b]}\{N_i\gamma_i\}} \\
\times \int_{\mathbb{R}^d\times\mathbb{R}^p} \left[ \|[\mathbf{M}_\infty^{(0)}]^{-1}\| \|\tilde{\mathbf{z}} - \mathbf{z}\| + \Big(\sum_{i\in[b]} \|\mathbf{A}_i\|/\rho_i\Big) \|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}\| \right]^2 \mathrm{d}\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\tilde{\mathbf{x}}) \ ,
\end{aligned}
$$

*where $\mathbf{B}_0, \bar{\mathbf{B}}_0, \mathbf{M}_\infty^{(0)}$ are defined in (S2), (S23), respectively.*

*Finally, if $\boldsymbol{N} = N(1, \ldots, 1) = N\mathbf{1}_b$ for $N \geq 1$, then $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}} = \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\mathbf{1}_b}$.*

*Proof.* Note that under the conditions on $\boldsymbol{\gamma}$ and $\boldsymbol{N}$ stated in Proposition S13, Lemma S12 ensures that $1 - \min_{i\in[b]}\{N_i\gamma_i m_i\}/2 < 1$. Then, from Lemma S11 and Douc et al. (2018, Lemma 20.3.2, Theorem 20.3.4), we deduce the existence and unicity of a stationary distribution $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ for $P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$. The proof is concluded by using the upper bound given in Lemma S11.

We now show the last statement and assume that $\boldsymbol{N} = N\mathbf{1}_b$, for $N \geq 1$. By Proposition S5, we have the existence and unicity of a stationary distribution $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\mathbf{1}_b}$ which is invariant for $P_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ defined in (S7). For ease of notation, we simply denote $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\mathbf{1}_b}$ by $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ We now show that $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ is also invariant for $P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ defined in (S17). Using the fact that $P_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ defined in (S7) leaves $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ invariant from Proposition S5 and Fubini's theorem, we get for any $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$ and $\mathsf{B} \in \mathcal{B}(\mathbb{R}^p)$,

$$
\begin{aligned}
&\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}} P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathsf{A}\times\mathsf{B}) \qquad\qquad (S47)\\
&= \int_{\mathsf{A}\times\mathsf{B}} \int_{\mathbb{R}^d\times\mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}, \mathrm{d}\tilde{\mathbf{z}}) P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}((\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{z}}), (\mathrm{d}\boldsymbol{\theta}, \mathrm{d}\mathbf{z}))\\
&= \int_{\mathsf{A}\times\mathsf{B}} \int_{\mathbb{R}^d\times\mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}, \mathrm{d}\tilde{\mathbf{z}}) Q_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\tilde{\mathbf{z}}, \mathrm{d}\mathbf{z}|\tilde{\boldsymbol{\theta}}) \Pi_{\boldsymbol{\rho}}(\mathrm{d}\boldsymbol{\theta}|\mathbf{z})\\
&= \int_{\mathsf{A}\times\mathsf{B}} \int_{\mathbb{R}^d\times\mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}, \mathrm{d}\tilde{\mathbf{z}}) \left[\prod_{i=1}^b R_{\rho_i,\gamma_i}^{N_i}(\tilde{\mathbf{z}}_i, \mathrm{d}\mathbf{z}_i|\tilde{\boldsymbol{\theta}})\right] \Pi_{\boldsymbol{\rho}}(\mathrm{d}\boldsymbol{\theta}|\mathbf{z})\\
&= \int_{\mathsf{A}\times\mathsf{B}} \int_{\mathbb{R}^d\times\mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}, \mathrm{d}\tilde{\mathbf{z}}) \int_{\mathbb{R}^p} \left[\prod_{i=1}^b R_{\rho_i,\gamma_i}(\tilde{\mathbf{z}}_i, \mathrm{d}\tilde{\mathbf{z}}_i^{(1)}|\tilde{\boldsymbol{\theta}})\right] \left[\prod_{i=1}^b R_{\rho_i,\gamma_i}^{N_i-1}(\tilde{\mathbf{z}}_i^{(1)}, \mathrm{d}\mathbf{z}_i|\tilde{\boldsymbol{\theta}})\right] \Pi_{\boldsymbol{\rho}}(\mathrm{d}\boldsymbol{\theta}|\mathbf{z})\\
&= \int_{\mathsf{A}\times\mathsf{B}} \int_{\mathbb{R}^d\times\mathbb{R}^p} \left[\int_{\mathbb{R}^d\times\mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}, \mathrm{d}\tilde{\mathbf{z}}) \left[\prod_{i=1}^b R_{\rho_i,\gamma_i}(\tilde{\mathbf{z}}_i, \mathrm{d}\tilde{\mathbf{z}}_i^{(1)}|\tilde{\boldsymbol{\theta}})\right] \Pi_{\boldsymbol{\rho}}(\mathrm{d}\tilde{\boldsymbol{\theta}}^{(1)}|\tilde{\mathbf{z}}_i^{(1)})\right]\\
&\times \left[\prod_{i=1}^b R_{\rho_i,\gamma_i}^{N_i-1}(\tilde{\mathbf{z}}_i^{(1)}, \mathrm{d}\mathbf{z}_i|\tilde{\boldsymbol{\theta}})\right] \Pi_{\boldsymbol{\rho}}(\mathrm{d}\boldsymbol{\theta}|\mathbf{z})\\
&= \int_{\mathsf{A}\times\mathsf{B}} \int_{\mathbb{R}^d\times\mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}^{(1)}, \mathrm{d}\tilde{\mathbf{z}}^{(1)}) \left[\prod_{i=1}^b R_{\rho_i,\gamma_i}^{N_i-1}(\tilde{\mathbf{z}}_i^{(1)}, \mathrm{d}\mathbf{z}_i|\tilde{\boldsymbol{\theta}}^{(1)})\right] \Pi_{\boldsymbol{\rho}}(\mathrm{d}\boldsymbol{\theta}|\mathbf{z}) \ .
\end{aligned}
$$

Using a straightforward induction, we finally get

$$
\int_{\mathsf{A}\times\mathsf{B}} \int_{\mathbb{R}^d\times\mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}, \mathrm{d}\tilde{\mathbf{z}}) P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}((\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{z}}), (\mathrm{d}\boldsymbol{\theta}, \mathrm{d}\mathbf{z})) = \int_{\mathsf{A}\times\mathsf{B}} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\boldsymbol{\theta}, \mathrm{d}\mathbf{z}) \ ,
$$

which shows that $P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ leaves $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ invariant. Since this stationary distribution is unique, we conclude that $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}} = \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$. □

We specify our result to the case where we take a specific initial distribution. To define it, consider

$$\mathbf{x}^\star = ([\boldsymbol{\theta}^\star]^\top, [\mathbf{z}^\star]^\top)^\top, \text{ where } \boldsymbol{\theta}^\star = \arg\min\{-\log\pi\} \text{ and } \mathbf{z}^\star = ([\mathbf{A}_1\boldsymbol{\theta}^\star]^\top, \cdots, [\mathbf{A}_b\boldsymbol{\theta}^\star]^\top)^\top . \tag{S48}$$

We define the probability measure

$$\mu^\star_{\boldsymbol{\rho}} = \delta_{\mathbf{z}^\star} \otimes \Pi_{\boldsymbol{\rho}}(\cdot|\mathbf{z}^\star) . \tag{S49}$$

Note that sampling from $\mu^\star_{\boldsymbol{\rho}}$ is straightforward and simply consists in setting $\mathbf{z}_0 = \mathbf{z}^\star$ and $\boldsymbol{\theta}_0 = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\mathbf{z}_0 + \bar{\mathbf{B}}_0^{-1/2}\xi$, where $\xi$ is a $d$-dimensional standard Gaussian random variable. We now specify our result when using $\mu^\star_{\boldsymbol{\rho}}$ as an initial distribution. Define the $\mathbf{z}$-marginal under $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ by

$$\pi^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}} = \int_{\mathbb{R}^d} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\boldsymbol{\theta}, \mathbf{z}) , \tag{S50}$$

and the transition kernel of the Markov chain $\{Z_n\}_{n\geq 0}$, for all $\mathbf{z} \in \mathbb{R}^p$ and $\mathsf{B} \in \mathcal{B}(\mathbb{R}^p)$, by

$$P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathbf{z}, \mathsf{B}) = \int_{\mathbb{R}^d} Q_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathbf{z}, \mathsf{B}|\boldsymbol{\theta})\Pi_{\rho}(\mathrm{d}\boldsymbol{\theta}|\mathbf{z}) , \tag{S51}$$

where $\Pi_\rho(\cdot|\cdot)$ and $Q_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ are defined in (5) and (S16), respectively.

**Proposition S14.** *Assume **H**1-**H**2 and let $c > 0$ and $\boldsymbol{\gamma} = \{\gamma_i\}_{i=1}^b$, $\boldsymbol{N} \in (\mathbb{N}^*)^b$ such that (S46) is satisfied, for any $i \in [b]$, $N_i\gamma_i < 2/\max_{i\in[b]}\{m_i + \tilde{M}_i\}$ and $\gamma_i < \tilde{M}_i^{-1}$. Then, for any integer $n \geq 1$, we have*

$$W_2(\mu^\star_{\boldsymbol{\rho}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}) \leq 2^{1/2}(1 - \min_{i\in[b]}\{N_i\gamma_i m_i\}/2)^{n-1} \cdot (1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2)^{1/2} \max_{i\in[b]}\{N_i\gamma_i\}^{1/2}$$
$$\times \left\{\int_{\mathbb{R}^d} \|\mathbf{z}_1 - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} \pi^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\mathbf{z}_1) + \int_{\mathbb{R}^d} \|\mathbf{z}_1 - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathbf{z}^\star, \mathrm{d}\mathbf{z}_1)\right\}^{1/2} ,$$

*where $\bar{\mathbf{B}}_0, \mathbf{B}_0, \tilde{\mathbf{D}}_0$ are defined in (S2)-(S3).*

*Proof.* Consider for $n \in \mathbb{N}^*$, $X_n = (\theta_n^\top, Z_n^\top)^\top, \tilde{X}_n = (\tilde{\theta}_n^\top, \tilde{Z}_n^\top)^\top$ defined in (S18) with $X_0$ distributed according to $\mu^\star_{\boldsymbol{\rho}}$ and $\tilde{X}_0$ distributed according to $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$. Combining Lemma S7, Lemma S10 and Lemma S12, we have for $n \geq 1$,

$$\|\tilde{Z}_{n+1} - Z_{n+1}\|_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} \leq (1 - \min_{i\in[b]}\{N_i\gamma_i m_i\}/2)\|\tilde{Z}_n - Z_n\|_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} .$$

Thereby, for any $n \geq 1$, we obtain by induction

$$\|\tilde{Z}_n - Z_n\|_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} \leq (1 - \min_{i\in[b]}\{N_i\gamma_i m_i\}/2)^{n-1}\|\tilde{Z}_1 - Z_1\|_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} . \tag{S52}$$

Using $\|\tilde{Z}_1 - Z_1\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} \leq 2\|\tilde{Z}_1 - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1/2}_{\boldsymbol{N}\boldsymbol{\gamma}}} + 2\|Z_1 - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1/2}_{\boldsymbol{N}\boldsymbol{\gamma}}}$ combined with the definition of the Wasserstein distance and Lemma S6 give

$$\begin{aligned} W_2(\mu^\star_{\boldsymbol{\rho}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}) &\leq \mathbb{E}\left[\|\tilde{X}_n - X_n\|^2\right]^{1/2} \\ &\leq (1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2)^{1/2} \max_{i\in[b]}\{N_i\gamma_i\}^{1/2}\mathbb{E}\left[\|\tilde{Z}_n - Z_n\|^2_{\mathbf{D}^{-1/2}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right]^{1/2} \\ &\leq 2^{1/2}(1 - \min_{i\in[b]}\{N_i\gamma_i m_i\}/2)^{n-1}(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2)^{1/2} \max_{i\in[b]}\{N_i\gamma_i\}^{1/2} \\ &\times \mathbb{E}\left[\|\tilde{Z}_1 - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1/2}_{\boldsymbol{N}\boldsymbol{\gamma}}} + \|Z_1 - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1/2}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right]^{1/2} . \end{aligned} \tag{S53}$$

Since $\tilde{X}_0$ is distributed according to the stationnary distribution $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$, $\tilde{X}_1$ also and therefore $\tilde{Z}_1$ is distributed according to $\pi^{\mathbf{z}}_{\boldsymbol{\rho},\bar{\gamma}}$. Finally, by definition $Z_1$ has distribution $P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathbf{z}^\star, \cdot)$, therefore (S53) completes the proof. □

## S3. Proof of Proposition 3

The proof of Proposition 3 stands for a generalization of Vono et al. (2019, Proposition 6) which only considered the specific case $\rho_i = \rho^2$ for $i \in [b]$. This section is divided into two parts, the first gathers lemmas which allow us to upper bound the $\xi^2$-divergence between $\pi_{\boldsymbol{\rho}}$ and $\pi$. Then, in the second subsection, we combine these results to control the Wasserstein distance $W_2(\pi_{\boldsymbol{\rho}}, \pi)$ by showing that it is smaller than $\chi^2(\pi_{\boldsymbol{\rho}}|\pi)$. For any $\boldsymbol{\theta} \in \mathbb{R}^d$ and $\boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$, define

$$U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta}) = -\log\left(\int_{\mathbf{z}_i \in \mathbb{R}^d} \exp\{-U_i(\mathbf{z}_i) - \|\mathbf{z}_i - \mathbf{A}_i\boldsymbol{\theta}\|^2/(2\rho_i)\}\,\mathrm{d}\mathbf{z}_i/(2\pi\rho_i)^{d_i/2}\right) , \tag{S54}$$

$$\overline{B}(\boldsymbol{\theta}) = \sum_{i=1}^b \rho_i \|\nabla U_i(\mathbf{A}_i\boldsymbol{\theta})\|^2/2 , \tag{S55}$$

$$\underline{B}(\boldsymbol{\theta}) = \sum_{i=1}^b \left\{\rho_i\|\nabla U_i(\mathbf{A}_i\boldsymbol{\theta})\|^2/[2(1+\rho_i M_i)] - d_i\log(1+\rho_i M_i)/2\right\} \tag{S56}$$

and consider

$$U(\boldsymbol{\theta}) = \sum_{i\in[b]} U_i(\mathbf{A}_i\boldsymbol{\theta}) , \qquad\qquad U^{\boldsymbol{\rho}}(\boldsymbol{\theta}) = \sum_{i\in[b]} U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta}) .$$

### S3.1. Technical lemmata

We start this subsection by Lemma S15 which allow us to bound the ratio between the integrals defined by $\int_{\mathbb{R}^d} \exp\{-\sum_{i\in[b]} U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta})\}$ and $\int_{\mathbb{R}^d} \exp\{-\sum_{i\in[b]} U_i(\mathbf{A}_i\boldsymbol{\theta})\}\,\mathrm{d}\boldsymbol{\theta}$.

**Lemma S15.** *Assume **H**1-**H**2-(i) and let $\boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$. Then, we have $\underline{B}(\boldsymbol{\theta}) \le U(\boldsymbol{\theta}) - U^{\boldsymbol{\rho}}(\boldsymbol{\theta})$, for any $\boldsymbol{\theta} \in \mathbb{R}^d$. If we assume in addition that for any $i \in [b]$, $U_i$ is convex, we have $U(\boldsymbol{\theta}) - U^{\boldsymbol{\rho}}(\boldsymbol{\theta}) \le \overline{B}(\boldsymbol{\theta})$, for any $\boldsymbol{\theta} \in \mathbb{R}^d$.*

*Proof.* The proof follows from the same lines as in Vono et al. (2019, Lemma 14). In what follows, we give it for the sake of completeness. First, note for any $\boldsymbol{\theta} \in \mathbb{R}^d$ and $i \in [b]$,

$$\exp\left\{U_i(\mathbf{A}_i\boldsymbol{\theta}) - U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta})\right\} = \int_{\mathbb{R}^{d_i}} \exp\left(U_i(\mathbf{A}_i\boldsymbol{\theta}) - U_i(\mathbf{z}_i) - \|\mathbf{z}_i - \mathbf{A}_i\boldsymbol{\theta}\|^2/(2\rho_i)\right)\frac{\mathrm{d}\mathbf{z}_i}{(2\pi\rho_i)^{d_i/2}} . \tag{S57}$$

Using **H**2-(i), and a second order Taylor expansion, for any $\boldsymbol{\theta} \in \mathbb{R}^d, i \in [b], \mathbf{z}_i \in \mathbb{R}^{d_i}$, we have

$$U_i(\mathbf{A}_i\boldsymbol{\theta}) - U_i(\mathbf{z}_i) \ge \nabla U_i(\mathbf{A}_i\boldsymbol{\theta})^\top(\mathbf{A}_i\boldsymbol{\theta} - \mathbf{z}_i) - M_i\|\mathbf{A}_i\boldsymbol{\theta} - \mathbf{z}_i\|^2/2 .$$

Hence, using (S57), we have for any $\boldsymbol{\theta} \in \mathbb{R}^d$ and $i \in [b]$,

$$\begin{aligned}\exp\left(\sum_{i=1}^b U_i(\mathbf{A}_i\boldsymbol{\theta}) - U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta})\right) &\ge \prod_{i=1}^b \exp\left(\frac{\rho_i}{2(1+\rho_i M_i)}\left\|\nabla U_i(\mathbf{A}_i\boldsymbol{\theta})\right\|^2\right)(1+\rho_i M_i)^{-d_i/2} \\ &= \exp(\underline{B}(\boldsymbol{\theta})) .\end{aligned}$$

Similarly, under the assumption that for any $i \in [b]$, $U_i$ is convex, the proof for the upper bound follows from the same lines using, for any $i \in [b]$, $\boldsymbol{\theta} \in \mathbb{R}^d$ and $\mathbf{z}_i \in \mathbb{R}^{d_i}$, that

$$U_i(\mathbf{A}_i\boldsymbol{\theta}) - U_i(\mathbf{z}_i) \le \nabla U_i(\mathbf{A}_i\boldsymbol{\theta})^\top(\mathbf{A}_i\boldsymbol{\theta} - \mathbf{z}_i) .$$

□

**Lemma S16.** *Assume **H**1-**H**2. Then, $U$ is $m_U$-strongly convex with $m_U = \lambda_{\min}(\sum_{i=1}^b m_i\mathbf{A}_i^\top\mathbf{A}_i)$.*

*Proof.* Using by **H**2-(i) that for any $i \in [b]$, $U_i$ is twice differentiable and by **H**2-(ii) the fact that for any $i \in [b]$, $U_i$ is $m_i$-strongly convex, we have for any $\boldsymbol{\theta} \in \mathbb{R}^d$

$$\nabla^2 U(\boldsymbol{\theta}) = \sum_{i=1}^b \mathbf{A}_i^\top\nabla^2 U_i(\mathbf{A}_i\boldsymbol{\theta})\mathbf{A}_i \succeq \sum_{i=1}^b m_i\mathbf{A}_i^\top\mathbf{A}_i \succeq \lambda_{\min}\left(\sum_{i=1}^b m_i\mathbf{A}_i^\top\mathbf{A}_i\right)\mathbf{I}_d = m_U\mathbf{I}_d .$$

□

For any $\boldsymbol{\theta} \in \mathbb{R}^d$, define

$$\beta(\boldsymbol{\theta}) = \left(\sum_{i=1}^{b} \rho_i \left\| \nabla U_i(\mathbf{A}_i \boldsymbol{\theta}) \right\|^2 \right)^{1/2} . \tag{S58}$$

**Lemma S17.** *Assume **H**2-(i) and let $\boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$. Then $\beta$ is a Lipschitz function w.r.t. $\|\cdot\|$, with Lipschitz constant*

$$L_\beta = \lambda_{\max}\left(\sum_{i=1}^{b} \rho_i M_i^2 \mathbf{A}_i^\top \mathbf{A}_i \right)^{1/2} . \tag{S59}$$

*Proof.* For any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathbb{R}^d$, we have using $|(\sum_{i=1}^b a_i^2)^{1/2} - (\sum_{i=1}^b b_i^2)^{1/2}| \le (\sum_{i=1}^b (a_i - b_i)^2)^{1/2}$, that

$$|\beta(\boldsymbol{\theta}_1) - \beta(\boldsymbol{\theta}_2)| \le \left(\sum_{i=1}^{b} \rho_i \|\nabla U_i(\mathbf{A}_i \boldsymbol{\theta}_1) - \nabla U_i(\mathbf{A}_i \boldsymbol{\theta}_2)\|^2 \right)^{1/2} \le \left(\sum_{i=1}^{b} \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2)\|^2 \right)^{1/2},$$

which completes the proof. □

Suppose **H**2-(ii) and for any $i \in [b]$, denote $\boldsymbol{\theta}_i^\star$ a minimiser of $\boldsymbol{\theta} \mapsto U_i(\mathbf{A}_i \boldsymbol{\theta})$.

**Lemma S18.** *Assume **H**1-**H**2 and let $\boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$. Then for any $s < m_U/(12 L_\beta^2)$, where $L_\beta$ is defined in (S59), we have*

$$\log \pi\left[\mathrm{e}^{s\{\beta^2 - \pi[\beta^2]\}}\right] \le 8 s^2 L_\beta^4 / m_U^2 + 4 s^2 \{\pi[\beta]\}^2 L_\beta^2 / m_U . \tag{S60}$$

*In addition,*

$$\pi(\beta^2) \le 2 d L_\beta^2 / m_U + 2 \sum_{i=1}^{b} \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2 . \tag{S61}$$

*Proof.* Using the decomposition

$$\beta^2(\boldsymbol{\theta}) - \{\pi[\beta]\}^2 = (\beta(\boldsymbol{\theta}) - \pi[\beta])^2 + 2\pi[\beta](\beta(\boldsymbol{\theta}) - \pi[\beta])$$

and the Cauchy-Schwarz inequality imply, for any $s > 0$,

$$\pi\left[\mathrm{e}^{s\{\beta^2 - \{\pi[\beta]\}^2\}}\right] \le \left\{\pi[\mathrm{e}^{2s\{\beta - \pi[\beta]\}^2}]\right\}^{1/2} \cdot \left\{\pi[\mathrm{e}^{4s\pi[\beta]\{\beta - \pi[\beta]\}}]\right\}^{1/2} . \tag{S62}$$

The proof consists in bounding the two terms in the right-hand sided. Since $\beta : \mathbb{R}^d \to \mathbb{R}$ is $L_\beta$-Lipschitz by Lemma S17, for any $0 \le s \le m_U/(12 L_\beta^2)$, using Vono et al. (2019, Lemma 16) and Lemma S16 gives setting $\bar{\beta} = \beta - \pi[\beta]$, that

$$\pi\left[\exp\{2s(\bar{\beta}^2 - \pi[\bar{\beta}^2])\}\right] \le \exp(16 s^2 L_\beta^4 / m_U^2) . \tag{S63}$$

In addition, using Bakry et al. (2013, Proposition 5.4.1), Lemma S17 and Lemma S16, we get for any $s \ge 0$,

$$\pi\left[\mathrm{e}^{4s\pi[\beta](\beta - \pi[\beta])}\right] \le \mathrm{e}^{8s^2\{\pi[\beta]\}^2 L_\beta^2 / m_U} .$$

Plugging this result and (S63) into (S62), we get

$$\pi\left[\mathrm{e}^{s\{\beta^2 - \{\pi[\beta]\}^2\}}\right] \le \exp(s\pi(\bar{\beta}^2) + 8 s^2 L_\beta^4 / m_U^2 + 4 s^2 \{\pi[\beta]\}^2 L_\beta^2 / m_U) .$$

The proof of (S60) follows using $\pi(\bar{\beta}^2) = \pi(\beta^2) - [\pi(\beta)]^2$ and rearranging terms.

Using the Young inequality, **H**2-(i),$\nabla U_i(\mathbf{A}_i\boldsymbol{\theta}_i^\star) = 0$, $\nabla U(\boldsymbol{\theta}^\star) = 0$, we have

$$\begin{aligned}
\pi(\beta^2) &= \int_{\mathbb{R}^d} \left(\sum_{i=1}^b \rho_i \|\nabla U_i(\mathbf{A}_i\boldsymbol{\theta})\|^2\right)\pi(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} \\
&\leq 2\int_{\mathbb{R}^d} \left(\sum_{i=1}^b \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta} - \boldsymbol{\theta}^\star)\|^2\right)\pi(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} + 2\sum_{i=1}^b \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2 \\
&\leq 2\lambda_{\max}\left(\sum_{i=1}^b \rho_i M_i^2 \mathbf{A}_i^\top \mathbf{A}_i\right)\int_{\mathbb{R}^d} \|\boldsymbol{\theta} - \boldsymbol{\theta}^\star\|^2 \pi(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} + 2\sum_{i=1}^b \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2 \\
&\leq 2dL_\beta^2/m_U + 2\sum_{i=1}^b \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2 \ ,
\end{aligned}$$

where we have used $\pi[\|\boldsymbol{\theta} - \boldsymbol{\theta}^\star\|^2] \leq d/m_U$ by Durmus & Moulines (2019, Proposition 1 (ii)) and Lemma S16. □

Proposition 1 shows that $\pi_{\boldsymbol{\rho}}(\cdot) = \int_{\mathbb{R}^p} \Pi_{\boldsymbol{\rho}}(\cdot, \mathbf{z})\mathrm{d}\mathbf{z}$ is well-defined and as such admits a finite normalising constant. These two quantities are defined by

$$\mathrm{Z}_{\pi_{\boldsymbol{\rho}}} = \int_{\mathbb{R}^d} \exp\left\{-\sum_{i\in[b]} U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta})\right\}\mathrm{d}\boldsymbol{\theta}\ , \qquad \pi_{\boldsymbol{\rho}}(\cdot) = \exp\left\{-\sum_{i\in[b]} U_i^{\rho_i}(\mathbf{A}_i\cdot)\right\}/\mathrm{Z}_{\pi_{\boldsymbol{\rho}}}\ . \tag{S64}$$

Finally, note that the following quantity $\mathrm{Z}_\pi$ is a normalising constant of $\pi$ associated with the potential $U$, *i.e.* $\pi = \mathrm{e}^{-U}/\mathrm{Z}_\pi$,

$$\mathrm{Z}_\pi = \int_{\mathbb{R}^d} \exp\left\{-\sum_{i\in[b]} U_i(\mathbf{A}_i\boldsymbol{\theta})\right\}\mathrm{d}\boldsymbol{\theta}\ . \tag{S65}$$

**Lemma S19.** *Assume **H**1-**H**2 and let $\boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$. Suppose in addition that $6L_\beta^2 \leq m_U$ where $L_\beta$ is given in (S59). Then, we have*

$$\log\left(\mathrm{Z}_{\pi_{\boldsymbol{\rho}}}/\mathrm{Z}_\pi\right) \leq \left\{dL_\beta^2/m_U + \sum_{i=1}^b \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2\right\}(1 + 2L_\beta^2/m_U) + 2L_\beta^4/m_U^2\ .$$

*Proof.* From the definitions (S64) and (S65), we have $\mathrm{Z}_{\pi_{\boldsymbol{\rho}}}/\mathrm{Z}_\pi = \int_{\mathbb{R}^d} \pi(\boldsymbol{\theta}) \exp\{\sum_{i=1}^b U_i(\mathbf{A}_i\boldsymbol{\theta}) - U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta})\}\,\mathrm{d}\boldsymbol{\theta}$. By Lemma S15, we obtain

$$\mathrm{Z}_{\pi_{\boldsymbol{\rho}}}/\mathrm{Z}_\pi \leq \int_{\mathbb{R}^d} \pi(\boldsymbol{\theta})\exp(\overline{B}(\boldsymbol{\theta}))\,\mathrm{d}\boldsymbol{\theta}\ .$$

Note that $\overline{B} = \beta^2/2$ by (S55)-(S58), hence using that $6L_\beta^2 \leq m_U$, Lemma S18 applied with $s = 1/2$ shows that

$$\log\left(\int_{\mathbb{R}^d} \pi(\boldsymbol{\theta})\exp(\overline{B}(\boldsymbol{\theta}))\,\mathrm{d}\boldsymbol{\theta}\right) \leq \pi[\beta^2]/2 + 2L_\beta^4/m_U^2 + \{\pi[\beta]\}^2 L_\beta^2/m_U\ .$$

Using Lemma S18-(S61) and $\pi[\beta] \leq \pi[\beta^2]$ concludes the proof. □

### S3.2. Proof of Proposition 3

Based on the technical lemmas derived in Section S3.1, we are now ready to bound the Wasserstein distance of order 2 between $\pi$ and $\pi_{\boldsymbol{\rho}}$.

*Proof of Proposition Proposition 3.* Let $\boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$ such that $\max_{i\in[b]} \rho_i = \bar{\rho} \leq \sigma_U^2/12$, where $\sigma_U^2 = \|\mathbf{A}^\top\mathbf{A}\| \max_{i\in[b]}\{M_i^2\}/m_U$. Then, by definition of $L_\beta$ (S59), we get

$$12L_\beta^2 \leq m_U\ . \tag{S66}$$

and Lemma S18 can be applied for $s = 1$ and Lemma S19 too. By Lemma S16, $U = -\log \pi$ is $m_U$-strongly convex therefore $\pi$ satisfies a log-Sobolev inequality with constant $m_U$ (Ledoux, 2001, Theorem 5.2). Finally, Otto & Villani (2000, Theorem 1) shows that $\pi$ satisfies for any $\nu \in \mathcal{P}_2(\mathbb{R}^d)$:

$$W_2(\nu, \pi) \le \sqrt{(2/m_U)\mathrm{KL}(\nu|\pi)} \le \sqrt{(2/m_U)\chi^2(\pi_{\boldsymbol{\rho}}|\pi)} \ , \tag{S67}$$

where $\chi^2$ is the chi-square divergence and where we have used for the last inequality that $\mathrm{KL}(\pi_{\boldsymbol{\rho}}|\pi) \le \chi^2(\pi_{\boldsymbol{\rho}}|\pi)$ since for any $t > 0$, $\log(t) \le t - 1$. We now bound $\chi^2(\pi_{\boldsymbol{\rho}}|\pi)$. By (S64) and (S65), for any $\boldsymbol{\theta} \in \mathbb{R}^d$, consider the decomposition given by

$$\pi_{\boldsymbol{\rho}}(\boldsymbol{\theta})/\pi(\boldsymbol{\theta}) - 1 = (\mathrm{Z}_\pi/\mathrm{Z}_{\pi_{\boldsymbol{\rho}}}) \exp\left(\sum_{i=1}^b \left(U_i(\mathbf{A}_i\boldsymbol{\theta}) - U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta})\right)\right) - 1 \ . \tag{S68}$$

In the sequel, we will both lower and upper bound (S68) in order to upper bound $|1 - \pi_{\boldsymbol{\rho}}(\boldsymbol{\theta})/\pi(\boldsymbol{\theta})|$. Using the fact that for all $x \in \mathbb{R}, \exp(x) - 1 \ge x$, Lemmas S15 and S19 yield

$$\begin{aligned} \pi_{\boldsymbol{\rho}}(\boldsymbol{\theta})/\pi(\boldsymbol{\theta}) - 1 &\ge \log\left(\mathrm{Z}_\pi/\mathrm{Z}_{\pi_{\boldsymbol{\rho}}}\right) + \sum_{i=1}^b \left(U_i(\mathbf{A}_i\boldsymbol{\theta}) - U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta})\right) \\ &\ge -\left\{dL_\beta^2/m_U + \sum_{i=1}^b \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2\right\}(1 + 2L_\beta^2/m_U) - 2L_\beta^4/m_U^2 + \underline{B}(\boldsymbol{\theta}) \ge -A_1 \ , \end{aligned} \tag{S69}$$

where

$$A_1 = \left\{dL_\beta^2/m_U + \sum_{i=1}^b \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2\right\}(1 + 2L_\beta^2/m_U) + 2L_\beta^4/m_U^2 + \sum_{i=1}^b (d_i/2)\log(1 + \rho_i M_i) \ ,$$

where we have used in the last inequality that $\underline{B}(\boldsymbol{\theta}) \ge -\sum_{i=1}^b (d_1/2)\log(1 + \rho_i M_i)$ by (S55). In addition, by (S64) and (S65) $\mathrm{Z}_{\pi_{\boldsymbol{\rho}}}/\mathrm{Z}_\pi = \int_{\mathbb{R}^d} \pi(\boldsymbol{\theta}) \exp\{\sum_{i=1}^b U_i(\mathbf{A}_i\boldsymbol{\theta}) - U_i^{\rho_i}(\mathbf{A}_i\boldsymbol{\theta})\} \, \mathrm{d}\boldsymbol{\theta}.$ , which implies by Lemma S15 and Jensen inequality

$$\mathrm{Z}_{\pi_{\boldsymbol{\rho}}}/\mathrm{Z}_\pi \ge \int_{\mathbb{R}^d} \pi(\boldsymbol{\theta}) \exp(\underline{B}(\boldsymbol{\theta})) \, \mathrm{d}\boldsymbol{\theta} \ge \exp(\pi[\underline{B}]) \ .$$

It follows by (S68) that $\pi_{\boldsymbol{\rho}}(\boldsymbol{\theta})/\pi(\boldsymbol{\theta}) - 1 \le \exp(\overline{B}(\boldsymbol{\theta}) - \pi\left(\underline{B}\right)) - 1$. Combining this result and (S69), it follows that the Pearson $\chi^2$-divergence between $\pi$ and $\pi_{\boldsymbol{\rho}}$ can be upper bounded as where

$$\chi^2(\pi_{\boldsymbol{\rho}}|\pi) \le \max(A_1^2, A_2) \ , \qquad A_2 = \int_{\mathbb{R}^d} \left(\exp(\overline{B}(\boldsymbol{\theta}) - \pi\left(\underline{B}\right)) - 1\right)^2 \pi(\boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{\theta} \ .$$

We now provide an explicit bound for $A_2$. First by Jensen inequality, we have $\pi(\exp(\overline{B})) \ge \exp(\pi(\overline{B}))$ which implies that $\exp\left(-\pi(\underline{B})\right)\pi\left[\exp(\overline{B})\right] \ge \prod_{i=1}^b (1 + \rho_i M_i)^{d_i/2}$ by (S55). Therefore, using that $\overline{B} = \beta^2/2$ by (S55)-(S58) and

Lemma S18 with $s = 1$ since (S66) holds, we get by (S55),

$$
\begin{aligned}
A_2 &= \int_{\mathbb{R}^d} \left(\exp\left(\overline{B}(\boldsymbol{\theta}) - \pi\left(\underline{B}\right)\right) - 1\right)^2 \pi(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta} \\
&= \exp\left(-2\pi\left(\underline{B}\right)\right) \pi\left[\exp\left(2\overline{B}\right)\right] - 2\exp\left(-\pi\left(\underline{B}\right)\right)\pi\left[\exp\left(\overline{B}\right)\right] + 1 \\
&\leq \prod_{i=1}^{b} (1+\rho_i M_i)^{d_i} \cdot \exp(-\pi\{\sum_{i=1}^{b} (\rho_i/(1+\rho_i M_i)) \|\nabla U_i(\mathbf{A}_i \cdot)\|^2\}) \pi\left[\exp(\beta^2)\right] \\
&\qquad - 2\prod_{i=1}^{b} (1+\rho_i M_i)^{d_i/2} + 1 \\
&\leq \prod_{i=1}^{b} (1+\rho_i M_i)^{d_i} \cdot \exp(\pi\{\sum_{i=1}^{b} (\rho_i^2 M_i/(1+\rho_i M_i)) \|\nabla U_i(\mathbf{A}_i \cdot)\|^2\}) \\
&\times \exp\left(8L_\beta^4/m_U^2 + 4\{2dL_\beta^2/m_U + 2\sum_{i=1}^{b} \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2\} L_\beta^2/m_U\right) \qquad \text{(S70)} \\
&\qquad - 2\prod_{i=1}^{b} (1+\rho_i M_i)^{d_i/2} + 1\,, \qquad \text{(S71)}
\end{aligned}
$$

where we have used for the last inequality that for $\boldsymbol{\theta} \in \mathbb{R}^d$, $\beta(\boldsymbol{\theta})^2 - \sum_{i=1}^{b} (\rho_i/(1+\rho_i M_i)) \|\nabla U_i(\mathbf{A}_i \boldsymbol{\theta})\|^2 = \sum_{i=1}^{b} (\rho_i^2 M_i/(1+\rho_i M_i)) \|\nabla U_i(\mathbf{A}_i \boldsymbol{\theta})\|^2$, $\pi[\beta]^2 \leq \pi[\beta^2]$ by the Cauchy-Schwartz inequality and Lemma S18-(S61). Similarly to the proof of Lemma S18-(S61), by **H**2-(i), $\nabla U_i(\mathbf{A}_i \boldsymbol{\theta}_i^\star) = 0$, $\nabla U(\boldsymbol{\theta}^\star) = 0$, Durmus & Moulines (2019, Proposition 1 (ii)) and Lemma S16, we have

$$
\begin{aligned}
\pi\left[\sum_{i=1}^{b} (\rho_i^2 M_i/(1+\rho_i M_i)) \|\nabla U_i(\mathbf{A}_i \cdot)\|^2\right] &\leq \pi\left[\sum_{i=1}^{b} \rho_i^2 M_i \|\nabla U_i(\mathbf{A}_i \cdot)\|^2\right] \\
&\leq 2d\lambda_{\max}\left(\sum_{i=1}^{b} \rho_i^2 M_i^3 \mathbf{A}_i^\top \mathbf{A}_i\right)/m_U + 2\sum_{i=1}^{b} \rho_i^2 M_i^3 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2\,.
\end{aligned}
$$

Therefore, we get by (S71)

$$
\begin{aligned}
A_2 \leq A_3 &= \prod_{i=1}^{b} (1+\rho_i M_i)^{d_i} \exp\left(2d\lambda_{\max}\left(\sum_{i=1}^{b} \rho_i^2 M_i^3 \mathbf{A}_i^\top \mathbf{A}_i\right)/m_U + 2\sum_{i=1}^{b} \rho_i^2 M_i^3 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2\right) \\
&\exp\left(8L_\beta^4/m_U^2 + 8\left[dL_\beta^2/m_U + \sum_{i=1}^{b} \rho_i M_i^2 \|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2\right] L_\beta^2/m_U\right) - 2\prod_{i=1}^{b} (1+\rho_i M_i)^{d_i/2} + 1\,. \qquad \text{(S72)}
\end{aligned}
$$

It follows by (S70) and (S67) that

$$
W_2(\pi_{\boldsymbol{\rho}}, \pi) \leq \sqrt{(2/m_U)\max(A_1^2, A_3)}\,, \qquad \text{(S73)}
$$

where $A_1$ and $A_3$ are given by (S69) and (S72) respectively. Using that $L_\beta^2 = \mathcal{O}(\bar{\rho})$ and an expansion of the bound as $\bar{\rho} \to 0$ completes the proof. □

## S4. Proof of Proposition 4 and Proposition 5

As in Section S2, we assume in all this section that $\boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$ is fixed. For any $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_b) \in (\mathbb{R}_+^*)^b$, we establish in this section explicit bounds on $W_2(\pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \pi_{\boldsymbol{\rho}})$ where $\pi_{\boldsymbol{\rho}}$ is given in (1) and $\pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ is the marginal distribution defined by

$$
\pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathsf{A}) = \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathsf{A} \times \mathbb{R}^p)\,, \qquad \mathsf{A} \in \mathcal{B}(\mathbb{R}^d)\,,
$$

of the stationary probability measure $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ associated with the Markov chain $(Z_n, \theta_n)_{n\geq 0}$ defined in Algorithm 1. Note that in the case $\boldsymbol{N} = N(1,\ldots,1)$, this distribution is independent of $N$, see Proposition S13. To this purpose, we define an "ideal" dynamics from which we cannot sample but which converges geometrically towards $\Pi_{\boldsymbol{\rho}}$ under appropriate conditions. The corresponding ideal process will play the same role as the Langevin dynamics for the study of the unadjusted Langevin algorithm (Durmus & Moulines, 2019). This dynamics is defined as follows. Consider first for any $\boldsymbol{\theta} \in \mathbb{R}^d$, $i \in [b]$, the stochastic differential equation (SDE) defined by

$$\mathrm{d}\tilde{Y}_t^{i,\boldsymbol{\theta}} = -\nabla V_i(\tilde{Y}_t^{i,\boldsymbol{\theta}})\,\mathrm{d}t - \rho_i^{-1}\mathbf{A}_i\boldsymbol{\theta} + \sqrt{2}\,\mathrm{d}B_t^i \;, \tag{S74}$$

where $(B_t^i)_{t\geq 0}$ is a $d_i$-dimensional Brownian motion and $V_i$ is defined in (S20). Note that under **H**2-(i), this SDE admits a unique strong solution (Revuz & Yor, 2013, Theorem (2.1) in Chapter IX). Denote for any $i \in [b]$, the Markov semi-group associated to (S74) by $(\tilde{R}^i_{\rho_i,t})_{t\geq 0}$ defined for any $\tilde{\mathbf{y}}_0^i \in \mathbb{R}^{d_i}$, $t \geq 0$ and $\mathsf{B}_i \in \mathcal{B}(\mathbb{R}^{d_i})$ by

$$\tilde{R}^i_{\rho_i,t}(\tilde{\mathbf{y}}_0^i, \mathsf{B}_i|\boldsymbol{\theta}) = \mathbb{P}(\tilde{Y}_t^{i,\boldsymbol{\theta},\tilde{\mathbf{y}}_0^i} \in \mathsf{B}_i) \;,$$

where $(\tilde{Y}_t^{i,\boldsymbol{\theta},\tilde{\mathbf{y}}_0^i})_{t\geq 0}$ is a solution of (S74) with $\tilde{Y}_0^{i,\boldsymbol{\theta},\tilde{\mathbf{y}}_0^i} = \tilde{\mathbf{y}}_0^i$. For any bounded measurable function $f_i : \mathbb{R}^{d_i} \to \mathbb{R}_+$, Lemma S20 shows the measurability of the function $(\boldsymbol{\theta}, \tilde{\mathbf{y}}_0^i) \mapsto \mathbb{E}[f_i(\tilde{Y}_t^{i,\boldsymbol{\theta},\tilde{\mathbf{y}}_0^i})]$ on $\mathbb{R}^d \times \mathbb{R}^{d_i}$ and therefore $\tilde{R}^i_{\rho_i,t}$ is a conditional Markov kernel.

**Lemma S20.** *For any bounded measurable function $f_i : \mathbb{R}^{d_i} \to \mathbb{R}_+$ and function $f_i$ satisfying **H**2-(i), the mapping $(\tilde{\boldsymbol{\theta}}_0, \tilde{\mathbf{y}}_0^i) \mapsto \mathbb{E}[f_i(\tilde{Y}_t^{i,\tilde{\boldsymbol{\theta}}_0,\tilde{\mathbf{y}}_0^i})]$ is Borel measurable.*

*Proof.* Consider the following stochastic differential equation

$$\begin{cases} \mathrm{d}\tilde{\theta}_t = \mathbf{0}_d \;, \\ \mathrm{d}\tilde{Y}_t^i = -\nabla V_i(\tilde{Y}_t^i)\,\mathrm{d}t - \rho_i^{-1}\mathbf{A}_i\tilde{\theta}_t + \sqrt{2}\,\mathrm{d}B_t^i \;. \end{cases}$$

Using Revuz & Yor (2013, Theorem (2.4) in Chapter IX), since $U_i$ satisfies **H**2-(i), there exists a unique solution $(\tilde{X}_t^{\tilde{\mathbf{x}}})_{t\geq 0} = (\tilde{\theta}_t, \tilde{Y}_t^i)_{t\geq 0}$ with initial condition $\tilde{\mathbf{x}} = (\tilde{\boldsymbol{\theta}}_0^\top, (\tilde{\mathbf{y}}_0^i)^\top)^\top \in \mathbb{R}^p$. Then, the proof follows from Revuz & Yor (2013, Theorem (1.9) in Chapter IX) and the fact that $\tilde{Y}_t^i$ is the unique solution of (S74) with $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. □

Define for any $\boldsymbol{\theta} \in \mathbb{R}^d$, $\mathbf{z} = (\mathbf{z}_1^\top, \cdots, \mathbf{z}_b^\top)^\top \in \mathbb{R}^p$, and for $i \in [b]$, $\mathsf{B}_i \in \mathcal{B}(\mathbb{R}^{d_i})$,

$$\tilde{Q}_{\boldsymbol{\rho},\boldsymbol{\gamma}}\left(\mathbf{z}, \mathsf{B}_1 \times \cdots \times \mathsf{B}_b|\boldsymbol{\theta}\right) = \prod_{i=1}^b \tilde{R}^i_{\rho_i,N_i\gamma_i}(\mathbf{z}_i, \mathsf{B}_i|\boldsymbol{\theta}) \;,$$

and consider the Markov kernel defined, for any $\mathbf{x}^\top = (\boldsymbol{\theta}^\top, \mathbf{z}^\top)$ and $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$, $\mathsf{B} \in \mathcal{B}(\mathbb{R}^p)$, by

$$\tilde{P}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathbf{x}, \mathsf{A} \times \mathsf{B}) = \int_{\mathsf{B}} \tilde{Q}_{\boldsymbol{\rho},\boldsymbol{\gamma}}\left(\mathbf{z}, \mathrm{d}\tilde{\mathbf{z}}|\boldsymbol{\theta}\right) \int_{\mathsf{A}} \Pi_{\boldsymbol{\rho}}(\mathrm{d}\tilde{\boldsymbol{\theta}}|\tilde{\mathbf{z}}) \;, \tag{S75}$$

where $\Pi_{\boldsymbol{\rho}}(\cdot|\tilde{\mathbf{z}})$ is defined in (5). Note that $P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ can be interpreted as a discretised version of $\tilde{P}_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ using the Euler-Maruyama scheme.

In the sequel, we first derive technical lemmata in Section S4.1 that are used to prove both Proposition 4 and Proposition 5. Based on these lemmata, we then prove each proposition in a dedicated section, namely Section S4.2 and Section S4.3.

### S4.1. Synchronous coupling and a first estimate

The main idea to prove Proposition 4 and Proposition 5 is to define $(X_n, \tilde{X}_n)_{n\in\mathbb{N}}$ such that for any $n \in \mathbb{N}$, $(X_n, \tilde{X}_n)$ is a coupling between $\updelta_{\mathbf{x}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ defined in (S17) and $\updelta_{\tilde{\mathbf{x}}}\tilde{P}^n_{\boldsymbol{\rho},\boldsymbol{\gamma}}$, and satisfies

$$\mathbb{E}\left[\left\|X_n - \tilde{X}_n\right\|^2\right] \leq c_1(\mathbf{x}, \tilde{\mathbf{x}})\mathrm{e}^{-c_2 \min_{i\in[b]}\{\gamma_i m_i\}} + c_3\gamma^\alpha \;,$$

where $c_2, c_3 > 0$ and $\alpha \in \{1, 2\}$ depending if **H**3 holds or not. Conditioning with respect to $(X_0, \tilde{X}_0)$ with distribution $\updelta_{\mathbf{x}} \otimes \Pi_{\boldsymbol{\rho}}$, using the definition of the Wasserstein distance of order 2 and taking $n \to \infty$, we obtain

$$W_2(\pi_{\boldsymbol{\rho}}, \pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}) \leq W_2(\Pi_{\boldsymbol{\rho}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}) \leq \tilde{c}_3\gamma^\alpha \;,$$

where $\tilde{c}_3 > 0$. We now provide the rigourous construction of $(X_n, \tilde{X}_n)_{n\in\mathbb{N}}$.

Let $\{(B_t^{(i,n)})_{t\geq 0} : i \in [b], n \in \mathbb{N}\}$ be independent random variables such that for any $i \in [b]$, the sequences $\{(B_t^{(i,n)})_{t\geq 0} : n \in \mathbb{N}\}$ are i.i.d. $d_i$-dimensional Brownian motions and let $(\xi_n)_{n\geq 0}$ be a sequence of i.i.d. standard $d$-dimensional Gaussian random variables independent of $\{(B_t^{(i,n)})_{t\geq 0} : i \in [b], n \in \mathbb{N}\}$. Consider the stochastic process $(\tilde{X}_n)_{n\geq 0}$ on $\mathbb{R}^d \times \mathbb{R}^p$ starting from $\tilde{X}_0$ distributed according to $\Pi_{\boldsymbol{\rho}}$ and defined by the recursion: for $n \in \mathbb{N}$, $i \in [b]$,

$$\tilde{X}_{n+1} = (\tilde{\theta}_{n+1}^\top, \tilde{Z}_{n+1}^\top)^\top \ , \quad \tilde{Z}_{n+1}^i = \tilde{Y}_{N_i\gamma_i}^{(i,n)} \ , \quad \tilde{\theta}_{n+1} = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\tilde{Z}_{n+1} + \bar{\mathbf{B}}_0^{-1/2}\xi_{n+1} \ , \tag{S76}$$

where $(\tilde{Y}_t^{(i,n)})_{t\geq 0}$, is a solution of (S74) starting from $\tilde{Z}_n^i$ with parameter $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta}_n$. Similarly to the process $(X_n)_{n\in\mathbb{N}}$ defined in Algorithm 1, the process $(\tilde{X}_n)_{n\in\mathbb{N}}$ defines a homogeneous Markov chain. Indeed, it is easy to show that for any $n \in \mathbb{N}$ and measurable function $f : \mathbb{R}^p \to \mathbb{R}_+$, $\mathbb{E}[f(\tilde{Z}_{n+1})|\tilde{X}_n] = \int_{\mathbb{R}^p} f(\tilde{\mathbf{z}})\tilde{Q}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\tilde{Z}_n, \mathrm{d}\mathbf{z}|\tilde{\theta}_n)$ and therefore $(\tilde{X}_n)_{n\in\mathbb{N}}$ is associated with (S75).

**Proposition S21.** *Assume **H**1-**H**2-(i), and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$. Then, the Markov kernel $\tilde{P}_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ defined in (S75) admits $\Pi_{\boldsymbol{\rho}}$ as an invariant probability measure.*

*Proof.* By property of the Langevin diffusion defined in (S74), for all $\boldsymbol{\theta}_0 \in \mathbb{R}^d$, the Markov kernel $\tilde{Q}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\cdot|\boldsymbol{\theta}_0)$ admits $\Pi_{\boldsymbol{\rho}}(\cdot|\boldsymbol{\theta}_0)$ as invariant measure, see e.g. (Roberts & Tweedie, 1996) or (Kent, 1978). Thus, for any $\boldsymbol{\theta}_0 \in \mathbb{R}^d$ and $\mathsf{B} \in \mathcal{B}(\mathbb{R}^p)$, we have

$$\int_{\mathsf{B}} \Pi_{\boldsymbol{\rho}}(\mathbf{z}_1|\boldsymbol{\theta}_0)\,\mathrm{d}\mathbf{z}_1 = \int_{\mathbf{z}_0\in\mathbb{R}^p} \tilde{Q}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathbf{z}_0, \mathsf{B}|\boldsymbol{\theta}_0)\Pi_{\boldsymbol{\rho}}(\mathbf{z}_0|\boldsymbol{\theta}_0)\,\mathrm{d}\mathbf{z}_0 \ . \tag{S77}$$

Denote by $\pi_{\boldsymbol{\rho}}^{\boldsymbol{\theta}}, \pi_{\boldsymbol{\rho}}^{\mathbf{z}}$ the marginals under $\Pi_{\boldsymbol{\rho}}$: $\pi_{\boldsymbol{\rho}}^{\boldsymbol{\theta}}(\mathsf{A}) = \Pi_{\boldsymbol{\rho}}(\mathsf{A}\times\mathbb{R}^p)$, $\pi_{\boldsymbol{\rho}}^{\mathbf{z}}(\mathsf{B}) = \Pi_{\boldsymbol{\rho}}(\mathbb{R}^d\times\mathsf{B})$, for $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$ and $\mathsf{B} \in \mathcal{B}(\mathbb{R}^p)$, and consider the Markov chain $(\tilde{X}_n)_{n\in\mathbb{N}}$ defined in (S76). For any measurable function $f : \mathbb{R}^{d+p} \to \mathbb{R}_+$, the Fubini-Tonelli theorem gives

$$\begin{aligned}
\mathbb{E}[f(\tilde{X}_1)] &= \int_{\mathbb{R}^{d+p}}\int_{\mathbb{R}^{d+p}} f(\mathbf{x}_1)\Pi_{\boldsymbol{\rho}}(\boldsymbol{\theta}_1|\mathbf{z}_1)\,\mathrm{d}\boldsymbol{\theta}_1\tilde{Q}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathbf{z}_0, \mathrm{d}\mathbf{z}_1|\boldsymbol{\theta}_0)\Pi_{\boldsymbol{\rho}}(\boldsymbol{\theta}_0,\mathbf{z}_0)\,\mathrm{d}\boldsymbol{\theta}_0\,\mathrm{d}\mathbf{z}_0 \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^p} f(\mathbf{x}_1)\Pi_{\boldsymbol{\rho}}(\boldsymbol{\theta}_1|\mathbf{z}_1)\int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^p}\tilde{Q}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathbf{z}_0, \mathrm{d}\mathbf{z}_1|\boldsymbol{\theta}_0)\Pi_{\boldsymbol{\rho}}(\mathbf{z}_0|\boldsymbol{\theta}_0)\,\mathrm{d}\mathbf{z}_0\right]\pi_{\boldsymbol{\rho}}^{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)\,\mathrm{d}\boldsymbol{\theta}_0\,\mathrm{d}\boldsymbol{\theta}_1 \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^p} f(\mathbf{x}_1)\Pi_{\boldsymbol{\rho}}(\boldsymbol{\theta}_1|\mathbf{z}_1)\left[\int_{\boldsymbol{\theta}_0\in\mathbb{R}^d}\Pi_{\boldsymbol{\rho}}(\mathbf{z}_1|\boldsymbol{\theta}_0)\pi_{\boldsymbol{\rho}}^{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)\,\mathrm{d}\boldsymbol{\theta}_0\right]\mathrm{d}\mathbf{z}_1\,\mathrm{d}\boldsymbol{\theta}_1 \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^p} f(\mathbf{x}_1)\Pi_{\boldsymbol{\rho}}(\boldsymbol{\theta}_1|\mathbf{z}_1)\pi_{\boldsymbol{\rho}}^{\mathbf{z}}(\mathbf{z}_1)\,\mathrm{d}\mathbf{z}_1\mathrm{d}\boldsymbol{\theta}_1 \\
&= \int_{\mathbb{R}^{d+p}} f(\mathbf{x}_1)\Pi_{\boldsymbol{\rho}}(\boldsymbol{\theta}_1,\mathbf{z}_1)\,\mathrm{d}\mathbf{z}_1\mathrm{d}\boldsymbol{\theta}_1 = \mathbb{E}[f(\tilde{X}_0)] \ ,
\end{aligned} \tag{S78}$$

where we have used (S77) in (S78). Therefore, $X_1$ has distribution $\Pi_{\boldsymbol{\rho}}$ and the Markov kernel $\tilde{P}_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ admits $\Pi_{\boldsymbol{\rho}}$ as a stationary distribution, which completes the proof. □

Define by induction the synchronous coupling $(X_n = (\theta_n, Z_n))_{n\geq 0}, (\tilde{X}_n = (\tilde{\theta}_n, \tilde{Z}_n))_{n\geq 0}$, starting from $(\theta_0, Z_0) = (\boldsymbol{\theta}, \mathbf{z})$, $(\tilde{\theta}_0, \tilde{Z}_0)$ distributed according to $\Pi_{\boldsymbol{\rho}}$, for any $i \in [b]$ and $n \geq 0$, as

$$\begin{aligned}
\tilde{Z}_{n+1}^i &= \tilde{Y}_{N_i\gamma_i}^{(i,n)} \ , & \tilde{\theta}_{n+1} &= \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\tilde{Z}_{n+1} + \bar{\mathbf{B}}_0^{-1/2}\xi_{n+1} \ , \\
Z_{n+1}^i &= Y_{N_i\gamma_i}^{(i,n)} \ , & \theta_{n+1} &= \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}Z_{n+1} + \bar{\mathbf{B}}_0^{-1/2}\xi_{n+1} \ ,
\end{aligned} \tag{S79}$$

where we consider for any $i \in [b], k \in \mathbb{N}$, for $t \in [k\gamma_i, (k+1)\gamma_i)$

$$\begin{aligned}
\tilde{Y}_t^{(i,n)} &= \tilde{Y}_{k\gamma_i}^{(i,n)} - \int_{k\gamma_i}^t \nabla V_i(\tilde{Y}_l^{(i,n)})\,\mathrm{d}l + (t-k\gamma_i)(\rho_i)^{-1}\mathbf{A}_i\tilde{\theta}_n + 2^{1/2}(B_t^{(i,n)} - B_{k\gamma_i}^{(i,n)}) \ , \\
Y_t^{(i,n)} &= Y_{k\gamma_i}^{(i,n)} - (t-k\gamma_i)\nabla V_i(Y_{k\gamma_i}^{(i,n)}) + (t-k\gamma_i)(\rho_i)^{-1}\mathbf{A}_i\theta_n + 2^{1/2}(B_t^{(i,n)} - B_{k\gamma_i}^{(i,n)}) \ .
\end{aligned} \tag{S80}$$

Let $\mathcal{G}_0 = \sigma(Z_0, \tilde{Z}_0, \theta_0, \tilde{\theta}_0)$, for any $n \in \mathbb{N}^*$, let

$$\mathcal{G}_n = \sigma\{(Z_0, \tilde{Z}_0, \theta_0, \tilde{\theta}_0), (B_t^{(i,k)})_{t\geq 0} : i \in [b], k \leq n\} \,, \tag{S81}$$

and for any $t \geq 0$, let $\mathcal{H}_t^{(n)} = \sigma(\{(B_s^{(i,n)})_{s\leq t} : i \in [b]\})$, and

$$\mathcal{F}_t^{(n)} \text{ the } \sigma\text{-field generated by } \mathcal{H}_t^{(n)} \text{ and } \mathcal{G}_{n-1} \,. \tag{S82}$$

Note that $X_n$ and $\tilde{X}_n$ are distributed according to $\Pi_{\boldsymbol{\rho}} \tilde{P}_{\boldsymbol{\rho}}^n$ and $\delta_{\tilde{\mathbf{x}}} P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}^n$, respectively. Hence, by definition of the Wasserstein distance of order 2, it follows since $\Pi_{\boldsymbol{\rho}} \tilde{P}_{\boldsymbol{\rho}}^n = \Pi_{\boldsymbol{\rho}}$ by Proposition S21 that

$$W_2(\Pi_{\boldsymbol{\rho}}, \delta_{\mathbf{x}} P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}^n) \leq \mathbb{E}\left[\|X_n - \tilde{X}_n\|^2\right]^{1/2} \,. \tag{S83}$$

We start this section by a first estimate on $\mathbb{E}[\|X_n - \tilde{X}_n\|^2]^{1/2}$ and some technical results needed for the proof of Proposition 4 and Proposition 5. The following result holds regarding the process $(\tilde{Y}_t^{(i,n)})_{t\in\mathbb{R}_+}$ defined, for any $i \in [b]$ and $n \in \mathbb{N}$, in (S80).

**Lemma S22.** *Assume **H**1-**H**2. For $i \in [b], n \in \mathbb{N}$, denote by $\mathbf{z}_{n,\star}^i$ the unique minimiser of $\mathbf{z}_i \in \mathbb{R}^{d_i} \mapsto U_i(\mathbf{z}_i) + \|\mathbf{z}_i - \mathbf{A}_i\tilde{\theta}_n\|/(2\rho_i)$. Then, for any $i \in [b], k \in \mathbb{N}$ and $n \in \mathbb{N}$,*

$$\mathbb{E}^{\mathcal{G}_n}\left[\|\tilde{Y}_{k\gamma_i}^{(i,n)} - \mathbf{z}_{n,\star}^i\|^2\right] \leq d_i/\tilde{m}_i \,. \tag{S84}$$

*where $\tilde{m}_i$ is defined in (S25).*

*Proof.* Let $n \in \mathbb{N}$. By Durmus & Moulines (2019, Proposition 1), for $i \in [b]$ and $k \in \mathbb{N}$, we have

$$\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\|\tilde{Y}_{k\gamma_i}^{(i,n)} - \mathbf{z}_{n,\star}^i\|^2 \leq \|\tilde{Z}_n^i - \mathbf{z}_{n,\star}^i\|^2 \mathrm{e}^{-2k\gamma_i\tilde{m}_i} + (d_i/\tilde{m}_i)(1 - \mathrm{e}^{-2k\gamma_i\tilde{m}_i}) \,. \tag{S85}$$

By (S80), using Proposition S21 we get that $\tilde{X}_n$ has distribution $\Pi_{\boldsymbol{\rho}}$, therefore given $\tilde{\theta}_n$, $\tilde{Z}_n$ has distribution $\Pi_{\boldsymbol{\rho}}(\cdot|\tilde{\theta}_n)$. Then, using (S85), Durmus & Moulines (2019, Proposition 1(ii)) combined with **H**2, and since $(\tilde{Z}_n^1, \ldots, \tilde{Z}_n^b)$ are independent given $\tilde{\theta}_n$, we get the stated result. □

**Lemma S23.** *Assume **H**1 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$. Then, for any $n \in \mathbb{N}$, the random variables $X_n = (\theta_n^\top, Z_n^\top)^\top, \tilde{X}_n = (\tilde{\theta}_n^\top, \tilde{Z}_n^\top)^\top$ defined in (S79) sastify*

$$\|\tilde{X}_{n+1} - X_{n+1}\|^2 \leq (1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2)\|\tilde{Z}_{n+1} - Z_{n+1}\|^2 \,,$$

*where $\bar{\mathbf{B}}_0, \mathbf{B}_0, \tilde{\mathbf{D}}_0$ are defined in (S2)-(S3).*

*Proof.* The proof is similar to the proof of Lemma S3 and is omitted. □

For any $k, n \in \mathbb{N}, s \in \mathbb{R}_+$ consider the $p \times p$ matrices defined by

$$\mathbf{J}(k,s) = \mathrm{diag}\left(\mathbb{1}_{[N_1]}(k+1)\mathbb{1}_{[0,\gamma_1]}(s)\cdot\mathbf{I}_{d_1}, \cdots, \mathbb{1}_{[N_b]}(k+1)\mathbb{1}_{[0,\gamma_b]}(s)\cdot\mathbf{I}_{d_b}\right) \,, \tag{S86}$$

$$\mathbf{H}_{U,k}^{(n)} = \mathrm{diag}\Big(\gamma_1\int_0^1 \nabla^2 U_1((1-s)Y_{k\gamma_1}^{(1,n)} + s\tilde{Y}_{k\gamma_1}^{(1,n)})\,\mathrm{d}s, \ldots, \gamma_b\int_0^1 \nabla^2 U_b((1-s)Y_{k\gamma_b}^{(b,n)} + s\tilde{Y}_{k\gamma_b}^{(b,n)})\,\mathrm{d}s\Big) \,, \tag{S87}$$

$$\mathbf{C}_k^{(n)} = \mathbf{J}(k,0)(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}} + \mathbf{H}_{U,k}^{(n)}) \,, \tag{S88}$$

$$\mathbf{M}_{k+1}^{(n)} = (\mathbf{I}_p - \mathbf{C}_0^{(n)})^{-1}\ldots(\mathbf{I}_p - \mathbf{C}_k^{(n)})^{-1} \,, \qquad \text{with } \mathbf{M}_0^{(n)} = \mathbf{I}_p \,. \tag{S89}$$

Similarly to (S19), for $n, k \in \mathbb{N}$ and $i \in [b]$, consider $\mathbf{C}_k^{(i,n)}$ corresponding to the $i$-th diagonal block of $\mathbf{C}_k^{(n)}$ defined in (S88), *i.e.*

$$\mathbf{C}_k^{(i,n)} = \mathbb{1}_{[N_i]}(k+1)\gamma_i \left\{ \rho_i^{-1}\mathbf{I}_{d_i} + \int_0^1 \nabla^2 U_i((1-s)Y_{k\gamma_i}^{(i,n)} + s\tilde{Y}_{k\gamma_i}^{(i,n)})\,\mathrm{d}s \right\} \in \mathbb{R}^{d_i \times d_i} \;, \tag{S90}$$

where, for any $n \in \mathbb{N}$ and $i \in [b]$, $(Y_{k\gamma_i}^{(i,n)}, \tilde{Y}_{k\gamma_i}^{(i,n)})_{k\in\mathbb{N}}$ is defined in (S80).

**Lemma S24.** *Assume **H**1-**H**2 and let $\boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $\gamma_i < 1/\tilde{M}_i$. Then, for any $n, k \in \mathbb{N}$, the matrix $(\mathbf{I}_p - \mathbf{C}_k^{(n)})$ is invertible and in addition, for any $i \in [b]$, we have*

$$\|\mathbf{I}_{d_i} - \mathbf{C}_k^{(i,n)}\| \leq 1 - \gamma_i \tilde{m}_i \;,$$

*where $\mathbf{C}_k^{(i,n)}$ is defined in (S90).*

*Proof.* Let $i \in [b], n, k \in \mathbb{N}$. By **H**2, we have $\|\nabla^2 U_i\| \leq M_i$ which implies by (S90) that $\|\mathbf{C}_k^{(i,n)}\| \leq \gamma_i \tilde{M}_i$. Since $\gamma_i < 1/\tilde{M}_i$, the matrix $\mathbf{I}_p - \mathbf{C}_k^{(i,n)}$ is invertible and so is $\mathbf{I}_p - \mathbf{C}_k^{(n)}$. In addition, following the same lines as the proof of Lemma S9 implies $\|\mathbf{I}_{d_i} - \mathbf{C}_k^{(i,n)}\| \leq \max\{|1-\gamma_i\tilde{m}_i|, |1-\gamma_i\tilde{M}_i|\} = 1 - \gamma_i\tilde{m}_i$. □

For any $n, k \in \mathbb{N}, i \in [b]$, if $\gamma_i \in (0, 1/\tilde{M}_i)$, Lemma S24 shows the invertibility of the matrices $\mathbf{I}_p - \mathbf{C}_k^{(n)}$. Therefore, $\mathbf{M}_\infty^{(n)}$ is invertible and we can define

$$\mathrm{T}_1^{(n)} = [\mathbf{M}_\infty^{(n)}]^{-1} + \sum_{k=0}^{\infty} [\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_{k+1}^{(n)}\mathbf{J}(k,0)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2} \;, \tag{S91}$$

$$\mathrm{T}_2^{(n)} = \sum_{k=0}^{\infty} \left\{ [\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_{k+1}^{(n)}\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2} \int_0^{+\infty} \mathbf{J}(k,l)[\nabla V(\tilde{Y}_{k\boldsymbol{\gamma}+l}^{(n)}) - \nabla V(\tilde{Y}_{k\boldsymbol{\gamma}}^{(n)})]\,\mathrm{d}l \right\} . \tag{S92}$$

Using these matrices, we have the following result.

**Lemma S25.** *Assume **H**1-**H**2 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $\gamma_i < 1/\tilde{M}_i$. Then, for any $n \geq 1$,*

$$\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\tilde{Z}_{n+1} - Z_{n+1}) = \mathrm{T}_1^{(n)}(\tilde{Z}_n - Z_n) - \mathrm{T}_2^{(n)} \;, \tag{S93}$$

*where $(Z_n, \tilde{Z}_n)_{n\in\mathbb{N}}$ is defined in (S79) and $\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}} = \operatorname{diag}(N_1\gamma_1\mathbf{I}_{d_1}, \ldots, N_b\gamma_b\mathbf{I}_{d_b}) \in \mathbb{R}^{p\times p}$.*

*Proof.* Let $i \in [b]$ and $n \geq 1$. Recall that $V_i$ is defined in (S20) and for $\mathbf{z} \in \mathbb{R}^p$, denote $V(\mathbf{z}) = \sum_{i=1}^b V_i(\mathbf{z}_i)$. For any $k \in \mathbb{N}$, we have

$$\nabla V_i(\tilde{Y}_{k\gamma_i}^{(i,n)}) - \nabla V_i(Y_{k\gamma_i}^{(i,n)}) = \left[\int_0^1 \nabla^2 V_i((1-s)Y_{k\gamma_i}^{(i,n)} + s\tilde{Y}_{k\gamma_i}^{(i,n)})\,\mathrm{d}s\right](\tilde{Y}_{k\gamma_i}^{(i,n)} - Y_{k\gamma_i}^{(i,n)}) \;.$$

For $k \geq 0$, it follows from (S80) that

$$\begin{aligned} \tilde{Y}_{(k+1)\gamma_i}^{(i,n)} - Y_{(k+1)\gamma_i}^{(i,n)} &= \left(\mathbf{I}_{d_i} - \gamma_i \int_0^1 \nabla^2 V_i((1-s)Y_{k\gamma_i}^{(i,n)} + s\tilde{Y}_{k\gamma_i}^{(i,n)})\,\mathrm{d}s\right)(\tilde{Y}_{k\gamma_i}^{(i,n)} - Y_{k\gamma_i}^{(i,n)}) \\ &\quad - \int_0^{\gamma_i} \left[\nabla V_i(\tilde{Y}_{k\gamma_i+l}^{(i,n)}) - \nabla V_i(\tilde{Y}_{k\gamma_i}^{(i,n)})\right]\mathrm{d}l + (\gamma_i/\rho_i)\mathbf{A}_i(\tilde{\theta}_n - \theta_n) \;. \end{aligned} \tag{S94}$$

Consider the process $(\tilde{\mathrm{Y}}_t^{(n)}, \mathrm{Y}_t^{(n)})_{t\in\mathbb{R}_+}$ valued in $\mathbb{R}^p \times \mathbb{R}^p$ and defined for any $t \geq 0$ by

$$\tilde{\mathrm{Y}}_t^{(n)} = \tilde{Y}_{\min(t, N_i\gamma_i)}^{(n)} \;, \qquad \mathrm{Y}_t^{(n)} = Y_{\min(t, N_i\gamma_i)}^{(n)} \;. \tag{S95}$$

The process (S95) is continuous with respect to $t$ and defined so that its component $(\tilde{\mathrm{Y}}_t^{(i,n)}, \mathrm{Y}_t^{(i,n)})$ equals $(\tilde{Y}_t^i, Y_t^i)$ for $t \le N_i\gamma_i$ and is constant for $t > N_i\gamma_i$. For $l \ge 0$, we write $(\tilde{\mathrm{Y}}^{(n)}_{k\boldsymbol{\gamma}+l}, \mathrm{Y}^{(n)}_{k\boldsymbol{\gamma}+l}) = (\tilde{\mathrm{Y}}^{(i,n)}_{k\gamma_i+l}, \mathrm{Y}^{(i,n)}_{k\gamma_i+l})_{i\in[b]} \in \mathbb{R}^p \times \mathbb{R}^p$. Using the matrices defined in (S89), for $k \in \mathbb{N}$, we obtain

$$\tilde{\mathrm{Y}}^{(n)}_{(k+1)\boldsymbol{\gamma}} - \mathrm{Y}^{(n)}_{(k+1)\boldsymbol{\gamma}} = (\mathbf{I}_p - \mathbf{C}_k^{(n)})(\tilde{\mathrm{Y}}^{(n)}_{k\boldsymbol{\gamma}} - \mathrm{Y}^{(n)}_{k\boldsymbol{\gamma}}) - \int_0^\infty \mathbf{J}(k,l)\big[\nabla V(\tilde{\mathrm{Y}}^{(n)}_{k\boldsymbol{\gamma}+l}) - \nabla V(\tilde{\mathrm{Y}}^{(n)}_{k\boldsymbol{\gamma}})\big]\,\mathrm{d}l + \mathbf{J}(k,0)\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{P}_0\tilde{\mathbf{D}}_0^{1/2}(\tilde{\mathrm{Y}}_0^{(n)} - \mathrm{Y}_0^{(n)}) \,, \quad \text{(S96)}$$

where $\mathbf{P}_0$ is defined in (S3). Recall the matrix $\mathbf{M}_k^{(n)}$ defined in (S89) with $\mathbf{M}_0^{(n)} = \mathbf{I}_p$ and for $k \ge 1$, $\mathbf{M}_k^{(n)} = (\mathbf{I}_p - \mathbf{C}_0^{(n)})^{-1} \dots (\mathbf{I}_p - \mathbf{C}_{k-1}^{(n)})^{-1}$. By multiplying (S96) by $\mathbf{M}_{k+1}^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}$, we have

$$\begin{aligned}\mathbf{M}_{k+1}^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}(\tilde{\mathrm{Y}}^{(n)}_{(k+1)\boldsymbol{\gamma}} - \mathrm{Y}^{(n)}_{(k+1)\boldsymbol{\gamma}}) &= \mathbf{M}_k^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}(\tilde{\mathrm{Y}}^{(n)}_{k\boldsymbol{\gamma}} - \mathrm{Y}^{(n)}_{k\boldsymbol{\gamma}})\\ &\quad - \mathbf{M}_{k+1}^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}\int_0^\infty \mathbf{J}(k,l)\big[\nabla V(\tilde{Y}^{(n)}_{k\boldsymbol{\gamma}+l}) - \nabla V(\tilde{Y}^{(n)}_{k\boldsymbol{\gamma}})\big]\,\mathrm{d}l\\ &\quad + \mathbf{M}_{k+1}^{(n)}\mathbf{J}(k,0)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\tilde{\mathbf{D}}_0^{1/2}(\tilde{\mathrm{Y}}_0^{(n)} - \mathrm{Y}_0^{(n)}) \,.\end{aligned}$$

By (S95) and (S79), we have for $t \ge \max_{i\in[b]}\{\gamma_i N_i\}$, $(\tilde{Z}_{n+1}, Z_{n+1}) = (\tilde{\mathrm{Y}}_t, \mathrm{Y}_t)$. Therefore, summing the previous expression over $k$, we get

$$\begin{aligned}\mathbf{M}_\infty^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}(\tilde{Z}_{n+1} - Z_{n+1}) &= -\sum_{k=0}^\infty \mathbf{M}_{k+1}^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}\int_0^\infty \mathbf{J}(k,l)[\nabla V(\tilde{Y}^{(n)}_{k\boldsymbol{\gamma}+l}) - \nabla V(\tilde{Y}^{(n)}_{k\boldsymbol{\gamma}})]\,\mathrm{d}l\\ &+ \left[\mathbf{M}_0^{(n)} + \sum_{k=0}^\infty \mathbf{M}_{k+1}^{(n)}\mathbf{J}(k,0)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{P}_0\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{D}_{\boldsymbol{N}}^{1/2}\right]\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}\cdot(\tilde{Z}_n - Z_n) \,.\end{aligned}$$

By Lemma S24, $\mathbf{M}_\infty^{(n)}$ is invertible and the proof is concluded by multiplying the previous equality by $[\mathbf{M}_\infty^{(n)}]^{-1}$. □

Based on Lemma S25, we have the following relation between $\|\tilde{Z}_{n+1} - Z_{n+1}\|^2$ and $\|\tilde{Z}_n - Z_n\|^2$.

**Lemma S26.** *Assume **H**1-**H**2 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b], \gamma_i < 1/\tilde{M}_i$. Then, for any $\epsilon > 0$ and $n \ge 1$,*

$$\|\tilde{Z}_{n+1} - Z_{n+1}\|^2_{\mathbf{D}_{\boldsymbol{N\gamma}}^{-1}} \le (1+2\epsilon)\|\mathrm{T}_1^{(n)}\|^2\|\tilde{Z}_n - Z_n\|^2_{\mathbf{D}_{\boldsymbol{N\gamma}}^{-1}} + (1 + 1/\{2\epsilon\})\|\mathrm{T}_2^{(n)}\|^2 \,.$$

*where $(Z_n, \tilde{Z}_n)_{n\in\mathbb{N}}$ is defined in (S79) and $\mathbf{D}_{\boldsymbol{N\gamma}} = \mathrm{diag}(N_1\gamma_1\mathbf{I}_{d_1}, \dots, N_b\gamma_b\mathbf{I}_{d_b}) \in \mathbb{R}^{p\times p}$.*

*Proof.* The proof follows from Lemma S25 and by using the fact that for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^p, \epsilon > 0$ we have $2\langle\mathbf{a}, \mathbf{b}\rangle \le 2\epsilon\|\mathbf{a}\|^2 + (1/\{2\epsilon\})\|\mathbf{b}\|^2$. □

Similarly to Lemma S10, we have the following result regarding the contracting term.

**Lemma S27.** *Assume **H**1-**H**2 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $\gamma_i < 1/\tilde{M}_i$ and $N_i\gamma_i \le 2/(m_i + \tilde{M}_i)$. Then, for any $n \ge 0$, we have*

$$\|\mathrm{T}_1^{(n)}\| \le 1 - \min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \,,$$

*where $\mathrm{T}_1^{(n)}$ and $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ are defined in (S91) and (S33), respectively.*

*Proof.* The proof is similar to the proof of Lemma S10 and therefore is omitted. □

In the next lemma, we upper bound the coefficient $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ defined in (S33). For this, we explicit a choice of $\boldsymbol{N}$ that we denote $\boldsymbol{N}^\star = (N_1^\star(\gamma_1), \dots, N_b^\star(\gamma_b)) \in (\mathbb{N}^*)^b$ defined for any $i \in [b]$, any $\gamma_i > 0$, by

$$N_i^\star(\gamma_i) = \left\lfloor m_i \min_{i\in[b]}\{m_i/\tilde{M}_i\}^2 / \big(20\gamma_i\tilde{M}_i^2 \max_{i\in[b]}\{m_i/\tilde{M}_i\}^2\big)\right\rfloor \,, \quad \text{(S97)}$$

where $\tilde{M}_i = M_i + 1/\rho_i$.

**Lemma S28.** *Assume **H**1-**H**2 and let $\boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$,*

$$\gamma_i \leq \frac{m_i}{40\tilde{M}_i^2}\left(\frac{\min_{i\in[b]}\{m_i/\tilde{M}_i\}}{\max_{i\in[b]}\{m_i/\tilde{M}_i\}}\right)^2 .$$

*Then, for any $i \in [b]$, we have $N_i^\star(\gamma_i) \in \mathbb{N}^*$ and*

$$r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}^\star} < \min_{i\in[b]}\{N_i^\star(\gamma_i)\gamma_i m_i\}/2 ,$$

*where $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}^\star}$ is defined in (S33).*

*Proof.* The assumption on $\gamma_i$ combined with the definition (S97) of $N_i^\star(\gamma_i)$ imply $N_i^\star(\gamma_i) \geq 2$, using in addition $m_i \leq M_i$, $\max_{i\in[b]}\{N_i^\star(\gamma_i)\gamma_i\tilde{M}_i\mathbb{1}_{N_i^\star(\gamma_i)>1}\} \leq 1/20$ and

$$\begin{aligned}\frac{1}{20}\left(\frac{\min_{i\in[b]}\{m_i/\tilde{M}_i\}}{\max_{i\in[b]}\{m_i/\tilde{M}_i\}}\right)^2 \geq \frac{N_i^\star(\gamma_i)\gamma_i\tilde{M}_i^2}{m_i} &> \frac{1}{20}\left(\frac{\min_{i\in[b]}\{m_i/\tilde{M}_i\}}{\max_{i\in[b]}\{m_i/\tilde{M}_i\}}\right)^2 - \frac{\gamma_i\tilde{M}_i^2}{m_i} \\ &\geq \frac{1}{40}\left(\frac{\min_{i\in[b]}\{m_i/\tilde{M}_i\}}{\max_{i\in[b]}\{m_i/\tilde{M}_i\}}\right)^2 . \end{aligned} \tag{S98}$$

Using the definition (S33) of $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$, we have $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} < 5\max_{i\in[b]}\{N_i^\star(\gamma_i)\gamma_i\tilde{M}_i\mathbb{1}_{N_i^\star(\gamma_i)>1}\}^2$. Thus, plugging (S98) in the previous inequality gives

$$r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \leq \max_{i\in[b]}\{m_i/\tilde{M}_i\}^2 \max_{i\in[b]}\left\{\frac{N_i^\star(\gamma_i)\gamma_i\tilde{M}_i^2}{m_i}\right\} < \frac{\min_{i\in[b]}\{m_i/\tilde{M}_i\}^4}{80\max_{i\in[b]}\{m_i/\tilde{M}_i\}^2} . \tag{S99}$$

In addition, (S98) also shows that

$$\frac{1}{40}\left(\frac{\min_{i\in[b]}\{m_i/\tilde{M}_i\}}{\max_{i\in[b]}\{m_i/\tilde{M}_i\}}\right)^2\left(\frac{m_i}{\tilde{M}_i}\right)^2 \leq N_i^\star(\gamma_i)\gamma_i m_i . \tag{S100}$$

Therefore, combining (S99) and (S100) completes the proof. ☐

### S4.2. Proof of Proposition 4

We first give the formal statement of Proposition 4.

**Proposition S29.** *Assume **H** 1-**H** 2 and let $\boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$, $\boldsymbol{N} \in (\mathbb{N}^*)^b$ such that for any $i \in [b]$, $\gamma_i \leq m_i/40\tilde{M}_i^2(\min_{i\in[b]}\{m_i/\tilde{M}_i\}/\max_{i\in[b]}\{m_i/\tilde{M}_i\})^2$ and $N_i = \lfloor m_i \min_{i\in[b]}\{m_i/\tilde{M}_i\}^2/(20\gamma_i\tilde{M}_i^2\max_{i\in[b]}\{m_i/\tilde{M}_i^2\})\rfloor$. Then, we have*

$$W_2^2(\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}},\Pi_{\boldsymbol{\rho}}) \leq \frac{4(1+\|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2)\max_{i\in[b]}\{m_i/\tilde{M}_i^2\}}{5\min_{i\in[b]}\{m_i/\tilde{M}_i\}^2\max_{i\in[b]}\{m_i/\tilde{M}_i\}^2} \times \sum_{i=1}^b d_i\gamma_i m_i(1+\gamma_i^2\tilde{M}_i^2/12+\gamma_i\tilde{M}_i^2/(2\tilde{m}_i)) ,$$

*where $\bar{\mathbf{B}}_0, \mathbf{B}_0, \tilde{\mathbf{D}}_0$ are defined in (S2)-(S3), and for any $i \in [b]$, $\tilde{m}_i$, $\tilde{M}_i$ are defined in (S25).*

By Lemma S23 and Lemma S26, we can note that the proof of Proposition S29 boils down to derive an upper bound on $\|\mathrm{T}_2^{(n)}\|^2$ defined in (S92) for $n \in \mathbb{N}$. The following lemma provides such a bound.

**Lemma S30.** *Assume **H**1-**H**2 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $\gamma_i < 1/\tilde{M}_i$. Then, for any $n \in \mathbb{N}$, we have*

$$\mathbb{E}\left[\|\mathrm{T}_2^{(n)}\|^2\right] \leq \sum_{i=1}^b d_i N_i\gamma_i^2\tilde{M}_i^2\left[1+\gamma_i^2\tilde{M}_i^2/12+\gamma_i\tilde{M}_i^2/(2\tilde{m}_i)\right] ,$$

*where $\tilde{m}_i, \tilde{M}_i, \mathrm{T}_2^{(n)}$ are defined in (S25) and (S92), respectively.*

*Proof.* Let $n \in \mathbb{N}$. Using (S86), we can write, for any $l \in \mathbb{R}_+$ and $k \in \mathbb{N}$, $\mathbf{J}(k,l)$ as a block-diagonal matrix $\mathrm{diag}(\mathbf{J}^1(k,l), \ldots, \mathbf{J}^b(k,l))$ with $\mathbf{J}^i(k,l) = \mathbb{1}_{[N_i]}(k+1)\mathbb{1}_{[0,\gamma_i]}(s) \cdot \mathbf{I}_{d_i}$ for any $i \in [b]$. By (S89) and using for any $k \in \mathbb{N}$, that $[\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_{k+1}^{(n)} = \prod_{l=k+1}^\infty (\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)})$ is finite by (S88), we have

$$\begin{aligned}\|\mathrm{T}_2^{(n)}\|^2 &= \Big\|\sum_{k=0}^\infty [\mathbf{M}_\infty^{(n)}]^{-1}\mathbf{M}_{k+1}^{(n)}\mathbf{D}_{\boldsymbol{N\gamma}}^{-1/2}\int_0^\infty \mathbf{J}(k,l)\big[\nabla V(\tilde{Y}_{k\boldsymbol{\gamma}+l}^{(n)}) - \nabla V(\tilde{Y}_{k\boldsymbol{\gamma}}^{(n)})\big]\,\mathrm{d}l\Big\|^2 \\ &= \sum_{i=1}^b \frac{1}{N_i\gamma_i}\Big\|\sum_{k=0}^\infty \prod_{l=k+1}^\infty (\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)})\int_0^{\gamma_i} \mathbf{J}^i(k,0)\big[\nabla V_i(\tilde{Y}_{k\gamma_i+l}^{(i,n)}) - \nabla V_i(\tilde{Y}_{k\gamma_i}^{(i,n)})\big]\,\mathrm{d}l\Big\|^2 . \end{aligned} \tag{S101}$$

Since for any $i \in [b]$, $k \geq N_i$ we have $\mathbf{J}^i(k,0) = \mathbf{C}_l^{(i,n)} = \mathbf{0}_{d_i \times d_i}$, (S101) can be rewritten as

$$\|\mathrm{T}_2^{(n)}\|^2 = \sum_{i=1}^b \frac{1}{N_i\gamma_i}\Big\|\sum_{k=0}^{N_i-1} \prod_{l=k+1}^{N_i-1} (\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)})\int_0^{\gamma_i} \mathbf{J}^i(k,0)\big[\nabla V_i(\tilde{Y}_{k\gamma_i+l}^{(i,n)}) - \nabla V_i(\tilde{Y}_{k\gamma_i}^{(i,n)})\big]\,\mathrm{d}l\Big\|^2 ,$$

and the Cauchy-Schwarz inequality gives

$$\|\mathrm{T}_2^{(n)}\|^2 \leq \sum_{i=1}^b \frac{1}{\gamma_i}\left(\sum_{k=0}^{N_i-1}\Big\|\prod_{l=k+1}^{N_i-1}(\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)})\Big\|^2\Big\|\int_0^{\gamma_i}\Big[\nabla V_i(\tilde{Y}_{k\gamma_i+l}^{(i,n)}) - \nabla V_i(\tilde{Y}_{k\gamma_i}^{(i,n)})\Big]\,\mathrm{d}l\Big\|^2\right) . \tag{S102}$$

Since, for any $i \in [b]$, $\gamma_i\tilde{M}_i < 1$, we get using Lemma S24,

$$\Big\|\prod_{l=k+1}^{N_i-1}(\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)})\Big\|^2 \leq \{1 - \gamma_i\tilde{m}_i\}^{2(N_i-k-1)} .$$

By combining (S102) with the previous result and the Jensen inequality, we have

$$\|\mathrm{T}_2^{(n)}\|^2 \leq \sum_{i=1}^b\sum_{k=0}^{N_i-1}\{1-\gamma_i\tilde{m}_i\}^{2(N_i-k-1)}\int_0^{\gamma_i}\Big\|\nabla V_i(\tilde{Y}_{k\gamma_i+l}^{(i,n)}) - \nabla V_i(\tilde{Y}_{k\gamma_i}^{(i,n)})\Big\|^2\,\mathrm{d}l . \tag{S103}$$

For $i \in [b]$, using Durmus & Moulines (2019, Lemma 21) applied to the potential $V_i^\theta : \mathbf{y}^i \mapsto U_i(\mathbf{y}^i) + \|\mathbf{y}^i - \mathbf{A}_i\theta\|^2/(2\rho_i)$ yields

$$\begin{aligned}\int_0^{\gamma_i}\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\|\nabla V_i(\tilde{Y}_{k\gamma_i+l}^{(i,n)}) - \nabla V_i(\tilde{Y}_{k\gamma_i}^{(i,n)})\|^2\,\mathrm{d}l &= \int_0^{\gamma_i}\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\|\nabla V_i^{\tilde{\theta}_n}(\tilde{Y}_{k\gamma_i+l}^{(i,n)}) - \nabla V_i^{\tilde{\theta}_n}(\tilde{Y}_{k\gamma_i}^{(i,n)})\|^2\,\mathrm{d}l \\ &\leq \gamma_i^2\tilde{M}_i^2\left[d_i + d_i\gamma_i^2\tilde{M}_i^2/12 + (\gamma_i\tilde{M}_i^2/2)\|\tilde{Y}_{k\gamma_i}^{(i,n)} - \mathbf{z}_{n,\star}^i\|^2\right] ,\end{aligned} \tag{S104}$$

where $\mathbf{z}_{n,\star}^i = \arg\min_{\mathbf{z}_i \in \mathbb{R}^{d_i}} V_i^{\tilde{\theta}_n}(\mathbf{z}_i)$.

By (S104), (S84), Lemma S22 and since $\max_{i\in[b]}\gamma_i\tilde{m}_i < 1$, we get

$$\begin{aligned}&\sum_{i=1}^b\sum_{k=0}^{N_i-1}\{1-\gamma_i\tilde{m}_i\}^{2(N_i-k-1)}\int_0^{\gamma_i}\mathbb{E}\|\nabla V_i(\tilde{Y}_{k\gamma_i+l}^{(i,n)}) - \nabla V_i(\tilde{Y}_{k\gamma_i}^{(i,n)})\|^2\,\mathrm{d}l \\ &\qquad\leq \sum_{i=1}^b d_iN_i\gamma_i^2\tilde{M}_i^2[1 + \gamma_i^2\tilde{M}_i^2/12 + \gamma_i\tilde{M}_i^2/(2\tilde{m}_i)] .\end{aligned}$$

Combining this result with (S103) completes the proof. $\square$

We can now combine Lemma S30 and Lemma S27 with Lemma S26 to get the following bound.

**Lemma S31.** *Assume **H**1-**H**2 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $\gamma_i < 1/\tilde{M}_i$, $N_i\gamma_i \leq 2/(m_i + \tilde{M}_i)$. Suppose in addition $\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} = \min_{i\in[b]}\{N_i\gamma_i m_i\} - r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \in (0,1)$, where $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ is defined in (S33). Then, for $n \geq 1$, we have*

$$\mathbb{E}\Big[\|\tilde{Z}_n - Z_n\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\Big] \leq (1 - \kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} + \kappa^2_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}/2)^{2(n-1)}\mathbb{E}\Big[\|\tilde{Z}_1 - Z_1\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\Big] + 2\kappa^{-2}_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}\sum_{i=1}^b d_i N_i \gamma_i^2 \tilde{M}_i^2\left(1 + \frac{\gamma_i^2\tilde{M}_i^2}{12} + \frac{\gamma_i\tilde{M}_i^2}{2\tilde{m}_i}\right) ,$$

*where, for any $i \in [b]$, $\tilde{M}_i$ and $\tilde{m}_i$ are defined in (S25).*

*Proof.* Taking expectation in Lemma S26, we get for any $n \in \mathbb{N}, \epsilon > 0$ that

$$\mathbb{E}\left[\left\|\tilde{Z}_{n+1} - Z_{n+1}\right\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] \leq (1+2\epsilon)\mathbb{E}\left[\|\mathrm{T}_1^{(n)}\|^2\|\tilde{Z}_n - Z_n\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] + (1 + 1/\{2\epsilon\})\mathbb{E}\left[\|\mathrm{T}_2^{(n)}\|^2\right] ,$$

where $\mathrm{T}_1^{(n)}$ and $\mathrm{T}_2^{(n)}$ are defined in (S91) and (S92), respectively. To ease notation, denote $\mathrm{B} = \sum_{i=1}^b d_i N_i\gamma_i^2\tilde{M}_i^2(1 + \gamma_i^2\tilde{M}_i^2/12 + \gamma_i\tilde{M}_i^2/(2\tilde{m}_i))$. Using Lemma S30, we obtain for any $n \in \mathbb{N}, \epsilon > 0$

$$\mathbb{E}\left[\|\tilde{Z}_{n+1} - Z_{n+1}\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] \leq (1+2\epsilon)\mathbb{E}\left[\|\mathrm{T}_1^{(n)}\|^2\|\tilde{Z}_n - Z_n\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] + (1 + 1/\{2\epsilon\})\mathrm{B} . \tag{S105}$$

In addition, Lemma S27 implies that $\|\mathrm{T}_1^{(n)}\|^2 \leq (1 - \kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2$ almost surely. Therefore, taking $\epsilon = (1 - [1 - \kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}]^2)/(4[1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}]^2)$, (S105) yields for any $n \geq 0$,

$$\mathbb{E}\left[\left\|\tilde{Z}_{n+1} - Z_{n+1}\right\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] \leq \frac{1 + (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2}{2}\mathbb{E}\left[\left\|\tilde{Z}_n - Z_n\right\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] + \frac{1 + (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2}{1 - (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2}\mathrm{B} .$$

An easy induction implies for any $n \geq 1$,

$$\mathbb{E}\big[\|\tilde{Z}_n - Z_n\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\big] \leq \left(\frac{1 + (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2}{2}\right)^{n-1}\mathbb{E}\big[\|\tilde{Z}_1 - Z_1\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\big] + 2\frac{1 + (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2}{(1 - (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2)^2}\mathrm{B} . \tag{S106}$$

Since $\kappa^2_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} = (\min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2$ and using $\kappa^2_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \leq 1$, we obtain

$$\begin{aligned}
&(1 + (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2)/2 = 1 - \kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} + \kappa^2_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}/2 , \\
&(1 + (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2)/(1 - (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^2)^2 \leq \kappa^{-2}_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} .
\end{aligned}$$

Combining these inequalities with (S106) and (S105) completes the proof. □

**Lemma S32.** *Assume **H**1-**H**2 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $\gamma_i < 1/\tilde{M}_i, N_i\gamma_i \leq 2/(m_i + \tilde{M}_i)$ and $\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} = \min_{i\in[b]}\{N_i\gamma_i m_i\} - r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \in (0,1)$, where $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ is defined in (S33). Then, for any $\mathbf{x} \in \mathbb{R}^{d+p}$ and $n \geq 1$, we have*

$$\begin{aligned}
&W_2^2(\delta_{\mathbf{x}}P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \Pi_{\boldsymbol{\rho}}) \\
&\leq (1 - \kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} + \kappa^2_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}/2)^{2(n-1)}(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2)\max_{i\in[b]}\{N_i\gamma_i\}\mathbb{E}\big[\|\tilde{Z}_1 - Z_1\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\big] \\
&\quad + \frac{2(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\|^2)\max_{i\in[b]}\{N_i\gamma_i\}}{\kappa^2_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}}\sum_{i=1}^b d_i N_i\gamma_i^2\tilde{M}_i^2[1 + \gamma_i^2\tilde{M}_i^2/12 + \gamma_i\tilde{M}_i^2/(2\tilde{m}_i)] ,
\end{aligned}$$

*where $\bar{\mathbf{B}}_0, \mathbf{B}_0, \tilde{\mathbf{D}}_0$ are defined in (S2)-(S3), $P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ is defined in (S17), $(\tilde{Z}_n, Z_n)_{n\in\mathbb{N}}$ is defined in (S79) and for any $i \in [b]$, $\tilde{M}_i$, $\tilde{m}_i$ are defined in (S25).*

*Proof.* By Lemma S31, we have the following upper bound for $n \geq 1$,

$$\mathbb{E}\left[\|\tilde{Z}_n - Z_n\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] \leq (1 - \kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} + \kappa^2_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}/2)^{2(n-1)}\mathbb{E}\left[\|\tilde{Z}_1 - Z_1\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] + 2\kappa^{-2}_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \sum_{i=1}^b d_i N_i \gamma_i^2 \tilde{M}_i^2 \left(1 + \frac{\gamma_i^2 \tilde{M}_i^2}{12} + \frac{\gamma_i \tilde{M}_i^2}{2\tilde{m}_i}\right) .$$

Using (S79), Lemma S23, combined with the previous inequality, we get for any $n \geq 1, \mathbf{x} \in \mathbb{R}^{d+p}$,

$$\begin{aligned}
&W_2^2(\Pi_{\boldsymbol{\rho}}, \delta_{\mathbf{x}} P^n_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}) \\
&\leq (1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2)\mathbb{E}\big[\|\tilde{Z}_n - Z_n\|^2\big] \\
&\leq (1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2) \max_{i\in[b]}\{N_i\gamma_i\}\mathbb{E}\left[\|\tilde{Z}_n - Z_n\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] \\
&\leq (1 - \kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} + \kappa^2_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}/2)^{2(n-1)}(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2) \max_{i\in[b]}\{N_i\gamma_i\}\mathbb{E}\left[\|\tilde{Z}_1 - Z_1\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right] \\
&+ \frac{2(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2) \max_{i\in[b]}\{N_i\gamma_i\}}{\kappa^2_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}} \sum_{i=1}^b d_i N_i \gamma_i^2 \tilde{M}_i^2 \left(1 + \frac{\gamma_i^2 \tilde{M}_i^2}{12} + \frac{\gamma_i \tilde{M}_i^2}{2\tilde{m}_i}\right) .
\end{aligned}$$

Hence the stated result. □

**Proof of Proposition 4/Proposition S29.**

*Proof.* Since for any $i \in [b]$, $\gamma_i \leq m_i/40\tilde{M}_i^2(\min_{i\in[b]}\{m_i/\tilde{M}_i\}/\max_{i\in[b]}\{m_i/\tilde{M}_i\})^2$, setting

$$N_i^\star(\gamma_i) = \big\lfloor m_i \min_{i\in[b]}\{m_i/\tilde{M}_i\}^2 / \big(20\gamma_i \tilde{M}_i^2 \max_{i\in[b]}\{m_i/\tilde{M}_i\}^2\big)\big\rfloor$$

implies $\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}^\star} \in (0, 1)$ by Lemma S28. Thereby, letting $n$ tend towards infinity in Lemma S32 and using Proposition S13 conclude the proof. □

### S4.3. Proof of Proposition 5

We first give the formal statement of Proposition 5.

**Proposition S33.** *Assume **H** 1-**H** 2-**H** 3 and let $\boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$, $\boldsymbol{N} \in (\mathbb{N}^*)^b$ such that for any $i \in [b]$, $\gamma_i \leq m_i/40\tilde{M}_i^2(\min_{i\in[b]}\{m_i/\tilde{M}_i\}/\max_{i\in[b]}\{m_i/\tilde{M}_i\})^2$ and $N_i = \lfloor m_i \min_{i\in[b]}\{m_i/\tilde{M}_i\}^2/(20\gamma_i\tilde{M}_i^2 \max_{i\in[b]}\{m_i/\tilde{M}_i^2\})\rfloor$. Then, we have*

$$W_2^2(\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \Pi_{\boldsymbol{\rho}}) \leq 4(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2)\frac{\max_{i\in[b]}\{m_i/\tilde{M}_i^2\}}{\min_{i\in[b]}\{m_i/\tilde{M}_i\}^2}\mathscr{R}^\star(\boldsymbol{\gamma}) ,$$

*where setting $\mathfrak{f}_i = m_i/(20\tilde{M}_i)$,*

$$\mathscr{R}^\star(\boldsymbol{\gamma}) = \sum_{i=1}^b \left\{ d_i\gamma_i^2\tilde{M}_i^2 + \frac{d_i\gamma_i^2\mathfrak{f}_i}{\tilde{M}_i}\left(d_i L_i^2 + \frac{\tilde{M}_i^4}{\tilde{m}_i}\right) + d_i\gamma_i\tilde{M}_i\mathfrak{f}_i^3(1 + \mathfrak{f}_i + \mathfrak{f}_i^2)\right\} , \tag{S107}$$

*$\bar{\mathbf{B}}_0, \mathbf{B}_0, \tilde{\mathbf{D}}_0$ are defined in (S2)-(S3), and for any $i \in [b]$, $\tilde{m}_i$, $\tilde{M}_i$ are defined in (S25).*

We provide the proof of Proposition 5 in what follows. Similarly to Lemma S26 for the proof of Proposition 4, we derive an explicit relation between $\|\tilde{Z}_{n+1} - Z_{n+1}\|$ and $\|\tilde{Z}_n - Z_n\|$.

**Lemma S34.** *Assume **H**1-**H**2-**H**3 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that for any $i \in [b]$, $N_i\gamma_i \leq 2/(m_i + \tilde{M}_i)$ and $\gamma_i < 1/\tilde{M}_i$. Then, for $n \geq 1$, we have*

$$\mathbb{E}\big[\|\tilde{Z}_{n+1} - Z_{n+1}\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\big]^{1/2} \leq \big(1 - \min_{i\in[b]}\{N_i\gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}\big)\mathbb{E}\big[\|\tilde{Z}_n - Z_n\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\big]^{1/2} + \mathscr{R}(\boldsymbol{\gamma}, \boldsymbol{N})^{1/2} ,$$

*where*

$$\mathscr{R}(\boldsymbol{\gamma}, \boldsymbol{N}) = \sum_{i=1}^{b} d_i N_i \gamma_i^3 (d_i L_i^2 + \tilde{M}_i^4/\tilde{m}_i) + \sum_{i=1}^{b} \left( d_i \gamma_i^2 \tilde{M}_i^2 + d_i N_i^3 \gamma_i^4 \tilde{M}_i^4 \right) + \sum_{i=1}^{b} d_i N_i^4 \gamma_i^5 \tilde{M}_i^5 (1 + N_i \gamma_i \tilde{M}_i) \;, \tag{S108}$$

$(\tilde{Z}_n, Z_n)_{n\in\mathbb{N}}$ *is defined in* (S79), $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ *in* (S33) *and for any* $i \in [b]$, $\tilde{m}_i$, $\tilde{M}_i$ *are defined in* (S25).

*Proof.* Let $n \in \mathbb{N}$. For any $k \in \mathbb{N}$, recall that $\mathbf{M}_k^{(n)}$ is defined in (S89) and invertible by Lemma S24. Define

$$w_n = \mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2} (\tilde{Z}_n - Z_n) \;.$$

Under this notation, the result given in Lemma S25 can be rewritten as

$$w_{n+1} = \mathrm{T}_1^{(n)} w_n - \mathrm{T}_2^{(n)} \;,$$

where $\mathrm{T}_1^{(n)}$ and $\mathrm{T}_2^{(n)}$ are defined in (S91) and (S92), respectively. By the Minkowsky inequality and using (S81), we have

$$\mathbb{E}^{\mathcal{G}_n}\left[\|w_{n+1}\|^2\right]^{1/2} \leq \mathbb{E}^{\mathcal{G}_n}\left[\|\mathrm{T}_1^{(n)} w_n\|^2\right]^{1/2} + \mathbb{E}^{\mathcal{G}_n}\left[\|\mathrm{T}_2^{(n)}\|^2\right]^{1/2} . \tag{S109}$$

Since by Lemma S27,

$$\|\mathrm{T}_1^{(n)}\| \leq 1 - \min_{i\in[b]}\{N_i \gamma_i m_i\} + r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \;, \tag{S110}$$

it remains to bound $\mathbb{E}^{\mathcal{G}_n}[\|\mathrm{T}_2^{(n)}\|^2]$ to complete the proof.

For any $i \in [b]$, recall the function $V_i^{\theta_n} : \mathbb{R}^{d_i} \to \mathbb{R}$ defined for any $\mathbf{y}^i \in \mathbb{R}^{d_i}$ by $V_i^{\theta_n}(\mathbf{y}^i) = U_i(\mathbf{y}^i) + \|\mathbf{y}^i - \mathbf{A}_i\theta_n\|^2/(2\rho_i)$. For any $i \in [b], k \in \mathbb{N}$, using the Itô formula, we have for $l \in [k\gamma_i, (k+1)\gamma_i)$,

$$\nabla V_i(\tilde{Y}_{k\gamma_i+l}^{(i,n)}) - \nabla V_i(\tilde{Y}_{k\gamma_i}^{(i,n)}) = \int_{k\gamma_i}^{k\gamma_i+l} \left\{\nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)}) \nabla V_i^{\theta_n}(\tilde{Y}_u) + \vec{\Delta}(\nabla V_i^{\theta_n})(\tilde{Y}_u^{(i,n)})\right\} \mathrm{d}u + \sqrt{2}\int_{k\gamma_i}^{k\gamma_i+l} \nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\, \mathrm{d}B_u^i \;. \tag{S111}$$

For any $i \in [b], k \in \mathbb{N}$, define

$$\begin{aligned}
a_{1,k}^{(i,n)} &= \mathbb{1}_{[N_i]}(k+1)[\mathbf{M}_\infty^{(i,n)}]^{-1}\mathbf{M}_{k+1}^{(i,n)} \int_0^{\gamma_i}\int_{k\gamma_i}^{k\gamma_i+l} \nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)}) \nabla V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\, \mathrm{d}u\, \mathrm{d}l \;, \\
a_{2,k}^{(i,n)} &= \mathbb{1}_{[N_i]}(k+1)[\mathbf{M}_\infty^{(i,n)}]^{-1}\mathbf{M}_{k+1}^{(i,n)} \int_0^{\gamma_i}\int_{k\gamma_i}^{k\gamma_i+l} \vec{\Delta}(\nabla V_i^{\theta_n})(\tilde{Y}_u^{(i,n)})\, \mathrm{d}u\, \mathrm{d}l \;, \\
a_{3,k}^{(i,n)} &= \sqrt{2}\mathbb{1}_{[N_i]}(k+1)[\mathbf{M}_\infty^{(i,n)}]^{-1}\mathbf{M}_{k+1}^{(i,n)} \int_0^{\gamma_i}\int_{k\gamma_i}^{k\gamma_i+l} \nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\, \mathrm{d}B_u^i\, \mathrm{d}l \;.
\end{aligned}$$

With these notation and by (S111), we have

$$\begin{aligned}
\|T_2^{(n)}\|^2 &= \sum_{i\in[b]} \frac{1}{N_i\gamma_i} \Big\| \sum_{k\in\mathbb{N}} \{a_{1,k}^{(i,n)} + a_{2,k}^{(i,n)} + a_{3,k}^{(i,n)}\} \Big\|^2 \\
&\leq E_1 + E_2 + E_3 \;,
\end{aligned} \tag{S112}$$

where for any $j \in [3]$, $E_j = 3\sum_{i\in[b]} \|\sum_{k=0}^{N_i-1} a_{j,k}^{(i,n)}\|^2/(N_i\gamma_i)$. We now bound $\{E_j\}_{j\in[3]}$.

**Upper bound on $E_1$.** For any $i \in [b], k \in \mathbb{N}$, recall that we have $\left[\mathbf{M}_\infty^{(i,n)}\right]^{-1}\mathbf{M}_{k+1}^{(i,n)} = \prod_{l=k+1}^{\infty}(\mathbf{I}_{d_i} + \mathbf{C}_l^{(i,n)})$ where $\mathbf{C}_l^{(i,n)}$ is defined in (S88). In addition, since we suppose for any $i \in [b]$, that $\gamma_i \tilde{M}_i < 1$, Lemma S24 implies

$$\left\|\prod_{l=k+1}^{N_i-1}(\mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)})\right\|^2 \leq \{1 - \gamma_i \tilde{m}_i\}^{2(N_i-k-1)} \ .$$

Combining this result with the Cauchy-Schwarz inequality, we obtain

$$\frac{1}{N_i}\left\|\sum_{k=0}^{N_i-1} a_{1,k}^{(i,n)}\right\|^2 \leq \sum_{k=0}^{N_i-1}\left\|\int_0^{\gamma_i}\int_{k\gamma_i}^{k\gamma_i+l} \nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\nabla V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\,\mathrm{d}u\,\mathrm{d}l\right\|^2 . \tag{S113}$$

For $i \in [b]$, using the definition of $\mathbf{z}_{n,\star}^i = \arg\min_{\mathbf{y}^i \in \mathbb{R}^{d_i}} V_i^{\theta_n}(\mathbf{y}^i) \in \mathbb{R}^{d_i}$, we have $\nabla V_i^{\theta_n}(\mathbf{z}_{n,\star}^i) = \mathbf{0}_{d_i}$. Therefore, for $i \in [b], k \in \mathbb{N}$, conditioning with respect to $\mathcal{F}_{k\gamma_i}^{(n)}$ defined in (S82) and using the $\tilde{M}_i$-Lipschitz property of $V_i^{\theta_n}$ by **H**2 gives

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\left[\|\nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\nabla V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\|^2\right] &\leq \tilde{M}_i^2\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\left[\|\nabla V_i^{\theta_n}(\tilde{Y}_u^{(i,n)}) - \nabla V_i^{\theta_n}(\mathbf{z}_{n,\star}^i)\|^2\right] \\ &\leq \tilde{M}_i^4\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\left[\|\tilde{Y}_u^{(i,n)} - \mathbf{z}_{n,\star}^i\|^2\right] .\end{aligned}$$

For any $i \in [b], k \in \mathbb{N}$, combining this result with the Jensen inequality yields

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\left[\left\|\int_0^{\gamma_i}\int_{k\gamma_i}^{k\gamma_i+l}\nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\nabla V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\,\mathrm{d}u\,\mathrm{d}l\right\|\right]^2 \\ &\leq \gamma_i\int_0^{\gamma_i} l\int_{k\gamma_i}^{k\gamma_i+l}\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\left[\|\nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\nabla V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\|^2\right]\mathrm{d}u\,\mathrm{d}l \\ &\leq \gamma_i\tilde{M}_i^4\int_0^{\gamma_i} l\int_{k\gamma_i}^{k\gamma_i+l}\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\left[\|\tilde{Y}_u^{(i,n)} - \mathbf{z}_{n,\star}^i\|^2\right]\mathrm{d}u\,\mathrm{d}l \ .\end{aligned} \tag{S114}$$

By Lemma S22, we have for any $i \in [b]$, $u \in \mathbb{R}_+$,

$$\mathbb{E}^{\mathcal{G}_n}\left[\|\tilde{Y}_u^{(i,n)} - \mathbf{z}_{n,\star}^i\|^2\right] \leq d_i/\tilde{m}_i \ . \tag{S115}$$

Injecting this result in (S114) yields

$$\mathbb{E}\left[\int_0^{\gamma_i} l\int_{k\gamma_i}^{k\gamma_i+l}\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}}\left[\|\tilde{Y}_u^{(i,n)} - \mathbf{z}_{n,\star}^i\|^2\right]\mathrm{d}u\,\mathrm{d}l\right] \leq d_i\gamma_i^3/(3\tilde{m}_i) \ .$$

Finally, this inequality, (S114) and (S113), we get

$$\mathbb{E}\left[E_1\right] \leq \sum_{i=1}^{b} d_i N_i \gamma_i^3 \tilde{M}_i^4/\tilde{m}_i \ . \tag{S116}$$

**Upper bound on $E_2$.** Using the Cauchy-Schwarz inequality, we have

$$\frac{1}{N_i}\left\|\sum_{k=0}^{N_i-1} a_{2,k}^{(i,n)}\right\|^2 \leq \sum_{k=0}^{N_i-1}\left\|\int_0^{\gamma_i}\int_{k\gamma_i}^{k\gamma_i+l}\vec{\Delta}(\nabla V_i^{\theta_n})(\tilde{Y}_u^{(i,n)})\,\mathrm{d}u\,\mathrm{d}l\right\|^2 .$$

By **H**3, we have for any $\mathbf{z}_i \in \mathbb{R}^{d_i}$, $\|\vec{\Delta}(\nabla V_i^{\theta_n})(\mathbf{z}_i)\|^2 \leq d_i^2 L_i^2$. Therefore, we obtain

$$\begin{aligned}\left\|\int_0^{\gamma_i}\int_{k\gamma_i}^{k\gamma_i+l}\vec{\Delta}(\nabla V_i^{\theta_n})(\tilde{Y}_u^{(i,n)})\,\mathrm{d}u\,\mathrm{d}l\right\|^2 &\leq \gamma_i\int_0^{\gamma_i} l\int_{k\gamma_i}^{k\gamma_i+l}\|\vec{\Delta}(\nabla V_i^{\theta_n})(\tilde{Y}_u^{(i,n)})\|^2\,\mathrm{d}u\,\mathrm{d}l \\ &\leq d_i^2\gamma_i^4 L_i^2/3 \ .\end{aligned}$$

Thus, we get

$$\mathbb{E}\left[E_2\right] \leq \sum_{i=1}^{b} d_i^2 N_i \gamma_i^3 L_i^2 \ . \tag{S117}$$

**Upper bound on $E_3$.** For any $i \in [b], k \in \mathbb{N}$, define

$$\Delta_{3,k}^{(i,n)} = \int_0^{\gamma_i} \int_{k\gamma_i}^{k\gamma_i+l} \nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\, \mathrm{d}B_u^i\, \mathrm{d}l \; .$$

Using for any $i \in [b], k \in \mathbb{N}$, $[\mathbf{M}_\infty^{(i,n)}]^{-1}\mathbf{M}_{k+1}^{(i,n)} = \mathbf{I}_{d_i} - \sum_{l=k+1}^{\infty} \mathbf{C}_l^{(i,n)} + \mathbf{R}_k^{(i,n)}$ where $\mathbf{R}_k^{(i,n)}$ is defined in (S34), we have, for any $i \in [b], k \in \mathbb{N}$,

$$\begin{aligned}
\left\| \sum_{k=0}^{N_i-1} a_{3,k}^{(i,n)} \right\|^2 &= \left\| \sqrt{2} \sum_{k=0}^{N_i-1} \prod_{l=k+1}^{N_i} \left[ \mathbf{I}_{d_i} - \mathbf{C}_l^{(i,n)} \right] \Delta_{3,k}^{(i,n)} \right\|^2 \\
&= 2 \sum_{k_1,k_2=0}^{N_i-1} \langle \mathbf{R}_{k_1}^{(i,n)} \Delta_{3,k_1}^{(i,n)}, \mathbf{R}_{k_2}^{(i,n)} \Delta_{3,k_2}^{(i,n)} \rangle + 2 \sum_{k_1,k_2=0}^{N_i-1} \langle \Delta_{3,k_1}^{(i,n)}, \Delta_{3,k_2}^{(i,n)} \rangle \\
&+ 2 \sum_{k_1,k_2=0}^{N_i-1} \langle \sum_{l=k_1+1}^{N_i} \mathbf{C}_l^{(i,n)} \Delta_{3,k_1}^{(i,n)}, \sum_{l=k_2+1}^{N_i} \mathbf{C}_l^{(i,n)} \Delta_{3,k_2}^{(i,n)} \rangle \\
&- 4 \sum_{k_1,k_2=0}^{N_i-1} \langle \sum_{l=k_1+1}^{N_i} \mathbf{C}_l^{(i,n)} \Delta_{3,k_1}^{(i,n)}, \Delta_{3,k_2}^{(i,n)} \rangle + 4 \sum_{k_1,k_2=0}^{N_i-1} \langle \mathbf{R}_{k_1}^{(i,n)} \Delta_{3,k_1}^{(i,n)}, \Delta_{3,k_2}^{(i,n)} \rangle \\
&- 4 \sum_{k_1,k_2=0}^{N_i-1} \langle \mathbf{R}_{k_1}^{(i,n)} \Delta_{3,k_1}^{(i,n)}, \sum_{l=k_2+1}^{N_i} \mathbf{C}_l^{(i,n)} \Delta_{3,k_2}^{(i,n)} \rangle \; .
\end{aligned} \tag{S118}$$

We now control the quantities which appear in (S118). First, by **H**2, for any $i \in [b], \mathbf{x}^i, \mathbf{y}^i \in \mathbb{R}^{d_i}$, note that we have

$$\|\nabla^2 V_i^{\theta_n}(\mathbf{x}^i)\mathbf{y}^i\| \leq \tilde{M}_i \|\mathbf{y}^i\| \; .$$

By the Jensen inequality and the Itô isometry, for any $k \in \mathbb{N}$, we get

$$\begin{aligned}
\mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}} \left[ \|\Delta_{3,k}^{(i,n)}\|^2 \right] &= \mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}} \left[ \left\| \int_0^{\gamma_i} \int_{k\gamma_i}^{k\gamma_i+l} \nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\, \mathrm{d}B_u^i\, \mathrm{d}l \right\|^2 \right] \\
&\leq \gamma_i \tilde{M}_i^2 \int_0^{\gamma_i} \mathbb{E}^{\mathcal{F}_{k\gamma_i}^{(n)}} \left[ \left\| \int_{k\gamma_i}^{k\gamma_i+l} \mathrm{d}B_u^i \right\|^2 \right] \mathrm{d}l = d_i \gamma_i^3 \tilde{M}_i^2 / 2 \; .
\end{aligned} \tag{S119}$$

In addition, since for $i \in [b]$, $(\int_0^t \nabla^2 V_i^{\theta_n}(\tilde{Y}_u^{(i,n)})\, \mathrm{d}B_u^i)_{t \geq 0}$ is a $(\mathcal{F}_t^{(n)})_{t \geq 0}$-martingale, for $(k_1,k_2) \in \{0,\dots,N_i-1\}^2$ such that $k_1 < k_2$, we obtain

$$\mathbb{E}^{\mathcal{G}_n}\left[ \left[\Delta_{3,k_1}^{(i,n)}\right]^\top \Delta_{3,k_2}^{(i,n)} \right] = \mathbb{E}^{\mathcal{G}_n}\left[ \mathbb{E}^{\mathcal{F}_{k_2\gamma_i}^{(n)}} \left[ \Delta_{3,k_1}^{(i,n)\top} \Delta_{3,k_2}^{(i,n)} \right] \right] = 0 \; .$$

Therefore,

$$\sum_{k_1,k_2=0}^{N_i-1} \mathbb{E}^{\mathcal{G}_n} \left[ \langle \Delta_{3,k_1}^{(i,n)}, \Delta_{3,k_2}^{(i,n)} \rangle \right] = d_i N_i \gamma_i^3 \tilde{M}_i^2 / 2 \; .$$

Second, since for any $i \in [b], l \in \mathbb{N}$, $\mathbf{C}_l^{(i,n)} \in \mathbb{R}^{d_i \times d_i}$ is symmetric positive semi-definite, we have

$$\sum_{k_1,k_2=0}^{N_i-1} \langle \sum_{l=k_1+1}^{N_i} \mathbf{C}_l^{(i,n)} \Delta_{3,k_1}^{(i,n)}, \Delta_{3,k_2}^{(i,n)} \rangle = \left\langle \left\{ \sum_{l=1}^{N} C_l \right\} \sum_{k_1=0}^{l-1} \Delta_{3,k_1}, \sum_{k_1=0}^{l-1} \Delta_{3,k_2} \right\rangle \geq 0 \; .$$

Third, using for any $i \in [b], l \in \mathbb{N}$, using $\|\mathbf{C}_l^{(i,n)}\| \leq \gamma_i \tilde{M}_i$ by definition (S88) and **H**2 and combining the Cauchy-Schwarz inequality with (S119), for any $i \in [b], (k_1,k_2) \in \{0,\dots,N_i-1\}^2$, we get

$$\sum_{k_1,k_2=0}^{N_i-1} \mathbb{E}^{\mathcal{G}_n} \left[ \langle \sum_{l=k_1+1}^{N_i} \mathbf{C}_l^{(i,n)} \Delta_{3,k_1}^{(i,n)}, \sum_{l=k_2+1}^{N_i} \mathbf{C}_l^{(i,n)} \Delta_{3,k_2}^{(i,n)} \rangle \right] \leq d_i N_i^4 \gamma_i^5 \tilde{M}_i^4 / 8 \; .$$

Using (S119) again and Lemma S8, for $i \in [b]$, we obtain

$$\sum_{k_1,k_2=0}^{N_i-1} \mathbb{E}^{\mathcal{G}_n}\left[\langle \mathbf{R}_{k_1}^{(i,n)}\Delta_{3,k_1}^{(i,n)}, \mathbf{R}_{k_2}^{(i,n)}\Delta_{3,k_2}^{(i,n)}\rangle\right] \le (d_i\gamma_i^3\tilde{M}_i^2/2)\sum_{k_1,k_2=0}^{N_i-1}\mathbb{E}\left[\|\mathbf{R}_{k_1}^{(i,n)}\|\|\mathbf{R}_{k_2}^{(i,n)}\|\right]$$
$$\le (d_i\gamma_i^3\tilde{M}_i^2/2)\left\{\sum_{k=0}^{N_i-1}(\exp[(N_i-k)\gamma_i\tilde{M}_i]-1-[(N_i-k)\gamma_i\tilde{M}_i])\right\}^2$$
$$\le (d_i\gamma_i^3\tilde{M}_i^2/2)\left\{(\tilde{M}_i\gamma_i)^{-1}\int_0^{N_i\gamma_i\tilde{M}_i}\{\mathrm{e}^t-1-t\}\mathrm{d}t\right\}^2$$
$$\le \frac{(\mathrm{e}^{N_i\gamma_i\tilde{M}_i}+1)^2}{288}d_iN_i^6\gamma_i^7\tilde{M}_i^6\ .$$

Similarly, we get Moreover, using the Cauchy-Schwarz inequality, for any $i \in [b]$ we get

$$\sum_{k_1,k_2=0}^{N_i-1}\mathbb{E}[\langle\Delta_{k_1}^{(i,n)}, \mathbf{R}_{k_2}^{(i,n)}\Delta_{k_2}^{(i,n)}\rangle] \le \sum_{k_1,k_2=0}^{N_i-1}\mathbb{E}\left[\|\Delta_{k_1}^{(i,n)}\|\|\Delta_{k_2}^{(i,n)}\|\|\mathbf{R}_{k_2}^{(i,n)}\|\right]$$
$$\le \frac{d_iN_i\gamma_i^3\tilde{M}_i^2}{24}(\mathrm{e}^{N_i\gamma_i\tilde{M}_i}+1)N_i^3\gamma_i^2\tilde{M}_i^2$$
$$\le d_iN_i^4\gamma_i^5\tilde{M}_i^4\frac{\mathrm{e}^{N_i\gamma_i\tilde{M}_i}+1}{24}\ .$$

In addition, for any $i \in [b]$, we have also

$$\sum_{k_1,k_2=0}^{N_i-1}\mathbb{E}\left[\langle\mathbf{R}_{k_1}^{(i,n)}\Delta_{3,k_1}^{(i,n)}, \sum_{l=k_2+1}^{N_i}\mathbf{C}_l^{(i,n)}\Delta_{3,k_2}^{(i,n)}\rangle\right] \le d_iN_i^5\gamma_i^6\tilde{M}_i^5\frac{\mathrm{e}^{N_i\gamma_i\tilde{M}_i}+1}{24}\ . \tag{S120}$$

For any $i \in [b], k \in \mathbb{N}$, regrouping the previous results and using that $N_i\gamma_i\tilde{M}_i \le 2$ give

$$\mathbb{E}[E_3] \le \sum_{i=1}^b\{d_iN_i\gamma_i^2\tilde{M}_i^2 + d_iN_i^3\gamma_i^4\tilde{M}_i^4\} + \sum_{i=1}^b d_iN_i^4\gamma_i^5\tilde{M}_i^5(1+N_i\gamma_i\tilde{M}_i)\ . \tag{S121}$$

**Combination of our previous results.** Injecting the three upper bounds (S116), (S117), (S121) in (S112), we get

$$\mathbb{E}\left[\|T_2^{(n)}\|^2\right] \le \sum_{i=1}^b d_iN_i\gamma_i^3(d_iL_i^2+\tilde{M}_i^4/\tilde{m}_i) + \sum_{i=1}^b\{d_i\gamma_i^2\tilde{M}_i^2 + d_iN_i^3\gamma_i^4\tilde{M}_i^4\}$$
$$+\sum_{i=1}^b d_iN_i^4\gamma_i^5\tilde{M}_i^5(1+N_i\gamma_i\tilde{M}_i)\ . \tag{S122}$$

Using the recursion defined in (S109), and combining the upper bounds derived in (S110) and (S122) completes the proof. □

**Lemma S35.** *Assume **H**1-**H**2-**H**3 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma} \in (\mathbb{R}_+^*)^b$ such that for any $i \in [b]$, $N_i\gamma_i \le 2/(m_i+\tilde{M}_i)$, $\gamma_i < 1/\tilde{M}_i$ and $\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} = \min_{i\in[b]}\{N_i\gamma_im_i\} - r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}} \in (0,1)$, where $r_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}$ is defined in (S33). Then, for $n \ge 1$, we have*

$$\mathbb{E}\left[\|\tilde{Z}_{n+1}-Z_{n+1}\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right]^{1/2} \le (1-\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}})^{n-1}\mathbb{E}\left[\|\tilde{Z}_1-Z_1\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}}\right]^{1/2} + \{\kappa_{\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{N}}\}^{-1}\mathscr{R}(\boldsymbol{\gamma},\boldsymbol{N})\ ,$$

*where $\mathscr{R}(\boldsymbol{\gamma},\boldsymbol{N})$ is given in (S108).*

*Proof.* The proof follows from Lemma S34 combined with a straightforward induction. □

**Proof of Proposition 5/Proposition S33.**

*Proof of Proposition 5/Proposition S33.* For any $i \in [b]$, consider

$$N_i^\star(\gamma_i) = \left\lfloor m_i \min_{i\in[b]}\{m_i/\tilde{M}_i\}^2 / \left(20\gamma_i \tilde{M}_i^2 \max_{i\in[b]}\{m_i/\tilde{M}_i\}^2\right)\right\rfloor .$$

By Proposition S13 and Lemma S28, $P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ converges in $W_2$ to $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$. Therefore, using (S83), Lemma S23 and Lemma S35 and taking $n \to +\infty$, we obtain

$$W_2^2(\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \Pi_{\boldsymbol{\rho}}) \leq 4(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2)\frac{\max_{i\in[b]}\{N_i^\star(\gamma_i)\gamma_i\}}{\min_{i\in[b]}\{N_i^\star(\gamma_i)\gamma_i m_i\}} \mathscr{R}(\boldsymbol{\gamma}, \boldsymbol{N}^\star(\boldsymbol{\gamma})) . \tag{S123}$$

By definition of $N_i^\star(\gamma_i)$, we have $\gamma_i \tilde{M}_i N_i^\star(\gamma_i) \leq \mathfrak{f}_i = m_i/(20\tilde{M}_i)$ which completes the proof upon using it in (S123). $\square$

## S5. Explicit mixing times

This section aims at providing mixing times for DG-LMC with explicit dependencies w.r.t. the dimension $d$ and the prescribed precision $\varepsilon$. We specify our result to the case where for any $i \in [b]$, $m_i = m$, $M_i = M$, $L_i = L$, $\rho_i = \rho$, $\gamma_i = \gamma$, $N_i = N$ and for the specific initial distribution

$$\mu_{\boldsymbol{\rho}}^\star = \updelta_{\mathbf{z}^\star} \otimes \Pi_{\boldsymbol{\rho}}(\cdot|\mathbf{z}^\star) , \tag{S124}$$

where

$$\mathbf{x}^\star = ([\boldsymbol{\theta}^\star]^\top, [\mathbf{z}^\star]^\top)^\top, \text{ where } \boldsymbol{\theta}^\star = \arg\min\{-\log\pi\} \text{ and } \mathbf{z}^\star = ([\mathbf{A}_1\boldsymbol{\theta}^\star]^\top, \cdots, [\mathbf{A}_b\boldsymbol{\theta}^\star]^\top)^\top . \tag{S125}$$

Note that sampling from $\mu_{\boldsymbol{\rho}}^\star$ is straightforward and simply consists in setting $\mathbf{z}_0 = \mathbf{z}^\star$ and $\boldsymbol{\theta}_0 = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\mathbf{z}_0 + \bar{\mathbf{B}}_0^{-1/2}\xi$, where $\xi$ is a $d$-dimensional standard Gaussian random variable. Starting from this initialisation, we consider the marginal law of $\theta_n$ for $n \geq 1$ and denote it $\Gamma_{\mathbf{x}^\star}^n$. By Proposition S13, since for any $i \in [b]$, $N_i = N$, the stationary distribution associated to $P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ is $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}} = \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma},\mathbf{1}_b}$. We build upon the natural decomposition of the bias:

$$W_2(\Gamma_{\mathbf{x}^\star}^n, \pi) \leq W_2(\mu_{\boldsymbol{\rho}}^\star P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}^n, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}) + W_2(\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}, \Pi_{\boldsymbol{\rho}}) + W_2(\pi_{\boldsymbol{\rho}}, \pi) ,$$

where $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$, $\Pi_{\boldsymbol{\rho}}$ and $\pi_{\boldsymbol{\rho}}$ are defined in Proposition 2, (2) and (3), respectively. The following subsections focus on deriving conditions on $n_\varepsilon$, $\gamma_\varepsilon$, $N_\varepsilon$ and $\rho_\varepsilon$ to satisfy $W_2(\Gamma_{\mathbf{x}^\star}^{n_\varepsilon}, \pi) \leq \varepsilon$, where $\varepsilon > 0$.

### S5.1. Lower bound on the number of iterations $n_\varepsilon$

In this section, we derive a lower bound on $n_\varepsilon$ such that $W_2(\mu_{\boldsymbol{\rho}}^\star P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}^{n_\varepsilon}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}) \leq \varepsilon/3$ following the result provided in Proposition S14. Recall that we define the $\mathbf{z}$-marginal under $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ by

$$\pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}^{\mathbf{z}} = \int_{\mathbb{R}^d} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\boldsymbol{\theta}, \mathbf{z})\, \mathrm{d}\boldsymbol{\theta} , \tag{S126}$$

and the transition kernel of the Markov chain $\{Z_n\}_{n\geq 0}$, for all $\mathbf{z} \in \mathbb{R}^p$ and $\mathsf{B} \in \mathcal{B}(\mathbb{R}^p)$, by

$$P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}^{\mathbf{z}}(\mathbf{z}, \mathsf{B}) = \int_{\mathbb{R}^d} Q_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathbf{z}, \mathsf{B}|\boldsymbol{\theta})\Pi_\rho(\boldsymbol{\theta}|\mathbf{z})\, \mathrm{d}\boldsymbol{\theta} , \tag{S127}$$

where $\Pi_\rho(\cdot|\mathbf{z})$ and $Q_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ are defined in (5) and (S16), respectively. In the case $\boldsymbol{N} = \mathbf{1}_b$, we simply denote $P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}^{\mathbf{z}}$ by $P_{\boldsymbol{\rho},\boldsymbol{\gamma}}^{\mathbf{z}}$. We need to bound in Proposition S14 the factor

$$\left\{\int_{\mathbb{R}^d} \|\mathbf{z}_1 - \mathbf{z}^\star\|_{\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1}}^2 \pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}^{\mathbf{z}}(\mathrm{d}\mathbf{z}_1) + \int_{\mathbb{R}^d} \|\mathbf{z}_1 - \mathbf{z}^\star\|_{\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1}}^2 P_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}^{\mathbf{z}}(\mathbf{z}^\star, \mathrm{d}\mathbf{z}_1)\right\}^{1/2} . \tag{S128}$$

Our next results provide such bounds.

**Lemma S36.** *Assume **H**1. Then, the transition kernel $P_{\boldsymbol{\rho},\boldsymbol{\gamma}}^{\mathbf{z}}$ leaves $\pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}^{\mathbf{z}}$ invariant, that is $\pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}^{\mathbf{z}} P_{\boldsymbol{\rho},\boldsymbol{\gamma}}^{\mathbf{z}} = \pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}^{\mathbf{z}}$, where $\pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}^{\mathbf{z}}$ is defined by* (S50).

*Proof.* We have for any $\mathsf{B} \in \mathcal{B}(\mathbb{R}^p)$

$$\int_{\mathsf{B}} \pi^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\mathbf{z}) = \int_{\mathsf{B}} \int_{\mathbb{R}^d} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\boldsymbol{\theta}, \mathrm{d}\mathbf{z}) = \int_{\mathsf{B}} \pi^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\mathbf{z}) \int_{\mathbb{R}^d} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\boldsymbol{\theta}|\mathbf{z}) \ .$$

Therefore, using the fact that $P_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ leaves $\Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ invariant from Proposition S5 and Fubini's theorem, we get

$$\begin{aligned}
\int_{\mathsf{B}} \pi^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\mathbf{z}) = \int_{\mathsf{B}} \int_{\mathbb{R}^d} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\boldsymbol{\theta}, \mathrm{d}\mathbf{z}) &= \int_{\mathsf{B}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d \times \mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}, \mathrm{d}\tilde{\mathbf{z}}) P_{\boldsymbol{\rho},\boldsymbol{\gamma}}((\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{z}}), (\mathrm{d}\boldsymbol{\theta}, \mathrm{d}\mathbf{z})) \\
&= \int_{\mathsf{B}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d \times \mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}, \mathrm{d}\tilde{\mathbf{z}}) Q_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\tilde{\mathbf{z}}, \mathrm{d}\mathbf{z}|\tilde{\boldsymbol{\theta}}) \Pi_{\boldsymbol{\rho}}(\boldsymbol{\theta}|\mathbf{z})\, \mathrm{d}\boldsymbol{\theta} \\
&= \int_{\mathsf{B}} \int_{\mathbb{R}^d \times \mathbb{R}^p} \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\boldsymbol{\theta}}, \mathrm{d}\tilde{\mathbf{z}}) Q_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\tilde{\mathbf{z}}, \mathrm{d}\mathbf{z}|\tilde{\boldsymbol{\theta}}) \int_{\mathbb{R}^d} \Pi_{\boldsymbol{\rho}}(\boldsymbol{\theta}|\mathbf{z})\, \mathrm{d}\boldsymbol{\theta} \qquad \text{(S129)} \\
&= \int_{\mathbb{R}^d} \pi^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\mathbf{z}}) P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\tilde{\mathbf{z}}, \mathsf{B}) \ .
\end{aligned}$$

□

For any $i \in [b]$, let $\boldsymbol{\theta}_i^\star$ a minimiser of $\boldsymbol{\theta} \mapsto U_i(\mathbf{A}_i\boldsymbol{\theta})$, and define

$$\mathbf{u}^\star = ([\mathbf{A}_1(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_1^\star)]^\top, \cdots, [\mathbf{A}_b(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_b^\star)]^\top)^\top \qquad \text{(S130)}$$

**Lemma S37.** *Assume **H**1-**H**2 and let* $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma}, \boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$ *such that, for any* $i \in [b]$, $\gamma_i \leq 2/(m_i + M_i + 1/\rho_i)$ *and denote* $\mathbf{z}^\star = ([\mathbf{A}_1\boldsymbol{\theta}^\star]^\top, \cdots, [\mathbf{A}_b\boldsymbol{\theta}^\star]^\top)^\top$. *Then, for any* $\mathbf{z} \in \mathbb{R}^p$ *and* $\varepsilon > 0$,

$$\begin{aligned}
\int_{\mathbb{R}^p} \|\tilde{\mathbf{z}} - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathbf{z}, \mathrm{d}\tilde{\mathbf{z}}) &\leq \min_{i\in[b]}\{N_i\}^{-1} \Big[\kappa^2_{\boldsymbol{\gamma}}(1 + 2\varepsilon)\|\mathbf{z} - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1}_{\boldsymbol{\gamma}}} \\
&\quad + (1 + 1/(2\varepsilon)) \max_{i\in[b]}\{\gamma_i M_i^2\}\|\mathbf{u}^\star\|^2 + \mathrm{Tr}(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}\mathbf{P}_0) + 2\sum_{i=1}^b d_i\Big] \ ,
\end{aligned}$$

*where the transition kernel* $P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}$ *is defined in* (S51) *with* $\boldsymbol{N} = \mathbf{1}_b$.

*Proof.* Let $\gamma_i \leq 2/(m_i + M_i + 1/\rho_i)$ for any $i \in [b]$. Let $\xi$ be a $d$-dimensional Gaussian random variable independent of $\{\eta^i : i \in [b]\}$ where for any $i \in [b]$, $\eta^i$ is a $d_i$-dimensional Gaussian random variable. Let $\mathbf{z} \in \mathbb{R}^p$ and $Z$ be the random variable distributed according to $\updelta_{\mathbf{z}} P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}$, and defined by

$$\theta = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\mathbf{z} + \bar{\mathbf{B}}_0^{-1/2}\xi \ ,$$

and for any $i \in [b]$,

$$\begin{aligned}
Z^i &= \left(1 - \gamma_i/\rho_i\right)\mathbf{z}_i - \gamma_i\nabla U_i(\mathbf{z}_i) + \frac{\gamma_i}{\rho_i}\mathbf{A}_i\theta + \sqrt{2\gamma_i}\eta^i \\
&= \left(1 - \gamma_i/\rho_i\right)\mathbf{z}_i - \gamma_i\nabla U_i(\mathbf{z}_i) + \frac{\gamma_i}{\rho_i}\mathbf{A}_i\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\mathbf{z} + \frac{\gamma_i}{\rho_i}\mathbf{A}_i\bar{\mathbf{B}}_0^{-1/2}\xi + \sqrt{2\gamma_i}\eta^i \\
&= \left(1 - \gamma_i/\rho_i\right)\mathbf{z}_i - \gamma_i[\nabla U_i(\mathbf{z}_i) - \nabla U_i(\mathbf{A}_i\boldsymbol{\theta}^\star)] - \gamma_i[\nabla U_i(\mathbf{A}_i\boldsymbol{\theta}^\star) - \nabla U_i(\mathbf{A}_i\boldsymbol{\theta}_i^\star)] \\
&+ \frac{\gamma_i}{\rho_i}\mathbf{A}_i\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top\tilde{\mathbf{D}}_0^{1/2}\mathbf{z} + \frac{\gamma_i}{\rho_i}\mathbf{A}_i\bar{\mathbf{B}}_0^{-1/2}\xi + \sqrt{2\gamma_i}\eta^i \ .
\end{aligned}$$

Let

$$\begin{aligned}
\mathbf{D}_U^\star &= \mathrm{diag}\Big(\gamma_1\int_0^1 \nabla^2 U_1(\mathbf{z}_1 + t(\mathbf{A}_1\boldsymbol{\theta}^\star - \mathbf{z}_1))\,\mathrm{d}t, \cdots, \gamma_b\int_0^1 \nabla^2 U_b(\mathbf{z}_b + t(\mathbf{A}_b\boldsymbol{\theta}^\star - \mathbf{z}_b))\,\mathrm{d}t\Big) \ , \\
\tilde{\mathbf{D}}_U^\star &= \mathrm{diag}\Big(\gamma_1\int_0^1 \nabla^2 U_1(\mathbf{A}_1\boldsymbol{\theta}^\star + t(\mathbf{A}_1\boldsymbol{\theta}_1^\star - \mathbf{A}_1\boldsymbol{\theta}^\star))\,\mathrm{d}t, \cdots, \gamma_b\int_0^1 \nabla^2 U_b(\mathbf{A}_b\boldsymbol{\theta}^\star + t(\mathbf{A}_b\boldsymbol{\theta}_b^\star - \mathbf{A}_b\boldsymbol{\theta}^\star))\,\mathrm{d}t\Big) \ . \qquad \text{(S131)}
\end{aligned}$$

Since $\mathbf{P}_0\mathbf{D}_{\boldsymbol{\rho}}^{-1/2}\mathbf{z}^\star = \mathbf{D}_{\boldsymbol{\rho}}^{-1/2}\mathbf{z}^\star$, it follows that

$$Z - \mathbf{z}^\star = \left[\mathbf{I}_p - \mathbf{D}_U^\star - \mathbf{D}_{\boldsymbol{\gamma}}^{1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}(\mathbf{I}_p - \mathbf{P}_0)\mathbf{D}_{\boldsymbol{\rho}}^{-1/2}\right](\mathbf{z} - \mathbf{z}^\star) - \tilde{\mathbf{D}}_U^\star\mathbf{u}^\star + \mathbf{D}_{\boldsymbol{\gamma}}^{1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{B}_0\bar{\mathbf{B}}_0^{-1/2}\xi + \mathbf{D}_{2\boldsymbol{\gamma}}^{1/2}\eta \ .$$

With the notation $\mathbf{H} = \mathbf{I}_p - \mathbf{D}_U^\star - \mathbf{D}_{\boldsymbol{\gamma}}^{1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}(\mathbf{I}_p - \mathbf{P}_0)\mathbf{D}_{\boldsymbol{\rho}}^{-1/2}$, (S14), and using the fact that for any $\varepsilon > 0$, $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$, $|\langle \mathbf{a}, \mathbf{b}\rangle| \leq \varepsilon\|\mathbf{a}\|^2 + (4\varepsilon)^{-1}\|\mathbf{b}\|^2$, it follows, for any $\mathbf{z} \in \mathbb{R}^p$, that

$$\begin{aligned}
&\int_{\mathbb{R}^p} \left\|\tilde{\mathbf{z}} - \mathbf{z}^\star\right\|^2_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathbf{z}, \mathrm{d}\tilde{\mathbf{z}}) \\
&= \int_{\mathbb{R}^p}\int_{\mathbb{R}^d} \left\|\mathbf{H}(\mathbf{z} - \mathbf{z}^\star) - \tilde{\mathbf{D}}_U^\star\mathbf{u}^\star + \mathbf{D}_{\boldsymbol{\gamma}}^{1/2}\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}^{1/2}\mathbf{B}_0\bar{\mathbf{B}}_0^{-1/2}\boldsymbol{\xi} + \mathbf{D}_{2\boldsymbol{\gamma}}^{1/2}\boldsymbol{\eta}\right\|^2_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} \phi_d(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\,\phi_p(\boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\eta} \\
&= \left\|\mathbf{H}(\mathbf{z} - \mathbf{z}^\star) - \tilde{\mathbf{D}}_U^\star\mathbf{u}^\star\right\|^2_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} + \mathrm{Tr}(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}\mathbf{P}_0) + 2\sum_{i=1}^b d_i \\
&\leq \kappa_{\boldsymbol{\gamma}}^2\|\mathbf{z} - \mathbf{z}^\star\|^2_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} - 2\langle \mathbf{H}(\mathbf{z} - \mathbf{z}^\star), \tilde{\mathbf{D}}_U^\star\mathbf{u}^\star\rangle_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} + \left\|\tilde{\mathbf{D}}_U^\star\mathbf{u}^\star\right\|^2_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} + \mathrm{Tr}(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}\mathbf{P}_0) + 2\sum_{i=1}^b d_i \\
&\leq \kappa_{\boldsymbol{\gamma}}^2(1 + 2\varepsilon)\|\mathbf{z} - \mathbf{z}^\star\|^2_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} + \left(1 + \frac{1}{2\varepsilon}\right)\max_{i\in[b]}\{\gamma_i M_i^2\}\|\mathbf{u}^\star\|^2 + \mathrm{Tr}(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}\mathbf{P}_0) + 2\sum_{i=1}^b d_i \ .
\end{aligned}$$

□

**Proposition S38.** *Assume **H**1-**H**2 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma}, \boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $\gamma_i \leq 2/(m_i + M_i + 1/\rho_i)$. Then, we have*

$$\int_{\mathbb{R}^d} \|\mathbf{z}_1 - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} \pi^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\mathbf{z}_1) \leq \min_{i\in[b]}\{N_i\}^{-1}\frac{2}{1-\kappa_{\boldsymbol{\gamma}}^2}\left(\frac{1+\kappa_{\boldsymbol{\gamma}}^2}{1-\kappa_{\boldsymbol{\gamma}}^2}\max_{i\in[b]}\{\gamma_i M_i^2\}\|\mathbf{u}^\star\|^2 + \mathrm{Tr}(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}\mathbf{P}_0) + 2\sum_{i=1}^b d_i\right) ,$$

*with $\kappa_{\boldsymbol{\gamma}}$ defined in (S12).*

*Proof.* With the choice $\varepsilon = (1-\kappa_{\boldsymbol{\gamma}}^2)/(4\kappa_{\boldsymbol{\gamma}}^2)$ in Lemma S37 and using Lemma S36, we have

$$\begin{aligned}
\int_{\mathbb{R}^p}\|\tilde{\mathbf{z}} - \mathbf{z}^\star\|^2_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} \pi^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\tilde{\mathbf{z}}) \leq &\frac{\kappa_{\boldsymbol{\gamma}}^2+1}{2}\int_{\mathbb{R}^p}\|\mathbf{z} - \mathbf{z}^\star\|^2_{\mathbf{D}_{\boldsymbol{\gamma}}^{-1}} \pi^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma}}(\mathrm{d}\mathbf{z}) + \frac{1+\kappa_{\boldsymbol{\gamma}}^2}{1-\kappa_{\boldsymbol{\gamma}}^2}\max_{i\in[b]}\{\gamma_i M_i^2\}\|\mathbf{u}^\star\|^2 \\
&+\mathrm{Tr}(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}}\mathbf{P}_0) + 2\sum_{i=1}^b d_i \ .
\end{aligned}$$

Rearranging terms concludes the proof. □

**Lemma S39.** *Assume **H**1-**H**2 and let $\boldsymbol{N} \in (\mathbb{N}^*)^b, \boldsymbol{\gamma}, \boldsymbol{\rho} \in (\mathbb{R}_+^*)^b$ such that, for any $i \in [b]$, $N_i\gamma_i \leq 2/(m_i + M_i + 1/\rho_i), \gamma_i\tilde{M}_i < 1$ and denote $\mathbf{z}^\star = ([\mathbf{A}_1\boldsymbol{\theta}^\star]^\top, \cdots, [\mathbf{A}_b\boldsymbol{\theta}^\star]^\top)^\top$. Then, we have*

$$\int_{\mathbb{R}^p}\|\tilde{\mathbf{z}} - \mathbf{z}^\star\|^2_{\mathbf{D}^{-1}_{\boldsymbol{N}\boldsymbol{\gamma}}} P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathbf{z}^\star, \mathrm{d}\tilde{\mathbf{z}}) \leq 2\sum_{i=1}^b \gamma_i N_i\left(1 + \mathrm{Tr}(\mathbf{P}_0)/\rho_i\right) + 4\sum_{i=1}^b d_i \ .$$

*where the transition kernel $P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}$ is defined in (S51).*

*Proof.* Let $\{(\eta_k^i)_{k\geq 1} : i \in [b]\}$ be independent random variables such that for any $i \in [b]$, the sequences $\{(\eta_k^i)_{k\geq 1}\}$ are i.i.d. $d_i$-dimensional Brownian motions and let $\xi$ a $d$-dimensional standard Gaussian random variable independent of

$\{(\eta_k^i)_{k\geq 1} : i \in [b]\}$. Consider the stochastic process $(Y_k)_{k\in\mathbb{N}}$ initialised for any $i \in [b]$ at $Y_0^i = \mathbf{A}_i\boldsymbol{\theta}^\star$ and defined, for any $i \in [b]$, $k \in \mathbb{N}$, by

$$Y_{k+1}^i = Y_k^i - \gamma_i \nabla V_i(Y_k^i) + (\gamma_i/\rho_i)\mathbf{A}_i\theta + \sqrt{2\gamma_i}\eta_{k+1}^i \;, \tag{S132}$$

where the potential $V_i = \mathbf{y}^i \mapsto U_i(\mathbf{y}^i) + \|\mathbf{y}^i\|^2/(2\rho_i)$ and

$$\theta = \bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\mathbf{z}^\star + \bar{\mathbf{B}}_0^{-1/2}\xi \;. \tag{S133}$$

In addition, we define the random variable $Z = (Z^1, \ldots, Z^b)$, for any $i \in [b]$, as

$$Z^i = Y_{N_i}^i \;.$$

By definition, note that $Z$ is distributed according to $P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathbf{z}^\star, \cdot)$. Define the process $(\mathrm{Y}_k = \{\mathrm{Y}_k^i\}_{i=1}^b)_{k\in\mathbb{N}}$ valued in $\mathbb{R}^p \times \mathbb{R}^p$ defined for any $i \in [b]$, $k \geq 0$ by

$$\mathrm{Y}_k^i = Y^i_{\min(k,N_i)} \;.$$

and consider the following matrices defined, for any $k \in \mathbb{N}$, by

$$\begin{aligned}
\mathbf{H}_{U,k} &= \mathrm{diag}\bigg(\gamma_1 \int_0^1 \nabla^2 U_1((1-s)Y_k^1 + s\mathbf{z}^\star)\,\mathrm{d}s, \\
&\qquad\qquad \ldots, \gamma_b \int_0^1 \nabla^2 U_b((1-s)Y_k^b + s\mathbf{z}^\star)\,\mathrm{d}s\bigg) \;, \\
\mathbf{J}(k) &= \mathrm{diag}\left(\mathbb{1}_{[N_1]}(k+1)\cdot \mathbf{I}_{d_1}, \cdots, \mathbb{1}_{[N_b]}(k+1)\cdot\mathbf{I}_{d_b}\right) \;, && \text{(S134)}\\
\mathbf{C}_k &= \mathbf{J}(k)(\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{\rho}} + \mathbf{H}_{U,k}) \;, && \text{(S135)}\\
\mathbf{M}_{k+1} &= (\mathbf{I}_p - \mathbf{C}_0)^{-1}\ldots(\mathbf{I}_p - \mathbf{C}_k)^{-1} \;, \qquad \text{with } \mathbf{M}_0 = \mathbf{I}_p \;. && \text{(S136)}
\end{aligned}$$

Using these notation and (S132), for any $k \in \mathbb{N}$, we get

$$\mathrm{Y}_{k+1} - \mathbf{z}^\star = (\mathbf{I}_p - \mathbf{C}_k)(\mathrm{Y}_k - \mathbf{z}^\star) + \mathbf{J}(k)\left(\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\theta - \mathbf{D}_{\boldsymbol{\gamma}}\nabla V(\mathbf{z}^\star) + \mathbf{D}_{2\boldsymbol{\gamma}}^{1/2}\eta_{k+1}\right) \;.$$

Multiplying the previous equality by $\mathbf{M}_{k+1}\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}$, we obtain, for $k \geq 0$,

$$\begin{aligned}
\mathbf{M}_{k+1}\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\mathrm{Y}_{k+1} - \mathbf{z}^\star) &= \mathbf{M}_k\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\mathrm{Y}_k - \mathbf{z}^\star) \\
&\quad + \mathbf{M}_{k+1}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}\left(\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\theta - \mathbf{D}_{\boldsymbol{\gamma}}\nabla V(\mathbf{z}^\star) + \mathbf{D}_{2\boldsymbol{\gamma}}^{1/2}\eta_{k+1}\right) \;.
\end{aligned}$$

Summing the previous equality over $k \in \mathbb{N}$ gives

$$\begin{aligned}
\mathbf{M}_\infty\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\mathrm{Y}_{\boldsymbol{N}} - \mathbf{z}^\star) &= \mathbf{M}_0\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(\mathrm{Y}_0 - \mathbf{z}^\star) \\
&\quad + \sum_{k=0}^{\infty}\mathbf{M}_{k+1}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}\left(\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\theta - \mathbf{D}_{\boldsymbol{\gamma}}\nabla V(\mathbf{z}^\star) + \mathbf{D}_{2\boldsymbol{\gamma}}^{1/2}\eta_{k+1}\right) \;.
\end{aligned}$$

Multiplying the last equality by $[\mathbf{M}_\infty]^{-1}$ and using the fact that $\mathrm{Y}_0 = \mathbf{z}^\star$, we get

$$\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(Z - \mathbf{z}_\star) = \sum_{k=0}^{\infty}[\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}\left(\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\theta - \mathbf{D}_{\boldsymbol{\gamma}}\nabla V(\mathbf{z}^\star) + \mathbf{D}_{2\boldsymbol{\gamma}}^{1/2}\eta_{k+1}\right) \;. \tag{S137}$$

Recall that $\mathbf{P}_0 = \mathbf{B}_0\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top$. Hence, by (S133) and using $\mathbf{P}_0\mathbf{D}_{\boldsymbol{\rho}}^{-1/2}\mathbf{z}^\star = \mathbf{D}_{\boldsymbol{\rho}}^{-1/2}\mathbf{z}^\star$, we get

$$\mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\theta - \mathbf{D}_{\boldsymbol{\gamma}}\nabla V(\mathbf{z}^\star) = \mathbf{D}_{\boldsymbol{\gamma}/\sqrt{\boldsymbol{\rho}}}\mathbf{B}_0\bar{\mathbf{B}}_0^{-1/2}\xi - \mathbf{D}_{\boldsymbol{\gamma}}\nabla U(\mathbf{z}^\star) \;.$$

Plugging this equality into (S137) yields

$$\begin{aligned}\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1/2}(Z-\mathbf{z}^\star) &= -\sum_{k=0}^{\infty}[\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{N}}^{1/2}\nabla U(\mathbf{z}^\star)\\ &\quad+\sum_{k=0}^{\infty}[\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{\gamma}/(\boldsymbol{N}\boldsymbol{\rho})}^{1/2}\mathbf{B}_0\bar{\mathbf{B}}_0^{-1/2}\xi\\ &\quad+\sqrt{2}\sum_{k=0}^{\infty}[\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\eta_{k+1}\ .\end{aligned}\tag{S138}$$

Recall that $[\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1}=(([\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1})^1,\ldots,([\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1})^b)$ is a block-diagonal matrix where, for any $i\in[b]$, $([\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1})^i=\prod_{l=k+1}^{\infty}(\mathbf{I}_{d_i}-\mathbf{C}_l^i)$ where $\mathbf{C}_l^i$ is defined in (S135). In addition, since we suppose for any $i\in[b]$, that $\gamma_i\tilde{M}_i<1$, Lemma S24 implies

$$\left\|\prod_{l=k+1}^{N_i-1}(\mathbf{I}_{d_i}-\mathbf{C}_l^i)\right\|^2\le(1-\gamma_i\tilde{m}_i)^{2(N_i-k-1)}\ .$$

We now upper bound separately each term on the right-hand side of (S138). First, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\left\|\sum_{k=0}^{\infty}[\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{\gamma}/\boldsymbol{N}}^{1/2}\right\|^2 &\le\sum_{i=1}^{b}(\gamma_i/N_i)\left\|\sum_{k=0}^{\infty}([\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1})^i\mathbf{J}^i(k)\right\|^2\\ &\le\sum_{i=1}^{b}(\gamma_i/N_i)\left\|\sum_{k=0}^{N_i-1}\prod_{l=k+1}^{N_i-1}\left(\mathbf{I}_{d_i}-\mathbf{C}_l^i\right)\right\|^2\\ &\le\sum_{i=1}^{b}\gamma_i\sum_{k=0}^{N_i-1}\left\|\prod_{l=k+1}^{N_i-1}\left(\mathbf{I}_{d_i}-\mathbf{C}_l^i\right)\right\|^2\\ &\le\sum_{i=1}^{b}\gamma_i\sum_{k=0}^{N_i-1}(1-\gamma_i\tilde{m}_i)^{2(N_i-k-1)}\\ &\le\sum_{i=1}^{b}N_i\gamma_i\ .\end{aligned}\tag{S139}$$

Second, using the same techniques as for the above inequality, we obtain

$$\left\|\sum_{k=0}^{\infty}[\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{\gamma}/(\boldsymbol{N}\boldsymbol{\rho})}^{1/2}\mathbf{B}_0\bar{\mathbf{B}}_0^{-1/2}\xi\right\|^2\le\sum_{i=1}^{b}\frac{N_i\gamma_i}{\rho_i}\left\|\mathbf{B}_0\bar{\mathbf{B}}_0^{-1/2}\xi\right\|^2\tag{S140}$$

Finally, the third term can be upper-bounded as

$$\mathbb{E}\left[\left\|\sqrt{2}\sum_{k=0}^{\infty}[\mathbf{M}_\infty]^{-1}\mathbf{M}_{k+1}\mathbf{J}(k)\mathbf{D}_{\boldsymbol{N}}^{-1/2}\eta_{k+1}\right\|^2\right]\le 2\sum_{i=1}^{b}d_i\ .\tag{S141}$$

Combining (S138), (S139), (S140) and (S141), we get

$$\int_{\mathbb{R}^p}\|\tilde{\mathbf{z}}-\mathbf{z}^\star\|^2_{\mathbf{D}_{\boldsymbol{N}\boldsymbol{\gamma}}^{-1}}P^{\mathbf{z}}_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}(\mathbf{z}^\star,\mathrm{d}\tilde{\mathbf{z}})\le\sum_{i=1}^{b}\gamma_iN_i\left(1+\mathrm{Tr}(\mathbf{P}_0)/\rho_i\right)+2\sum_{i=1}^{b}d_i\ .$$

□

Given $\varepsilon > 0$, we are now ready to provide a condition on the number of iterations $n_\varepsilon$ to achieve $W_2(\mu^\star_{\boldsymbol{\rho}} P^{n_\varepsilon}_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}) \leq \varepsilon/3$ in the case where for any $i \in [b]$, $m_i = m$, $M_i = M$, $\rho_i = \rho$, $\gamma_i = \gamma$ and $N_i = N$. Define

$$\mathsf{E}_0^2 = 9(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|)2N\gamma\Bigg[\frac{2}{N(1-\kappa_{\boldsymbol{\gamma}}^2)}\Bigg(\frac{1+\kappa_{\boldsymbol{\gamma}}^2}{1-\kappa_{\boldsymbol{\gamma}}^2}\cdot\gamma M^2\|\mathbf{u}^\star\|^2 + (\gamma/\rho)\mathrm{Tr}(\mathbf{P}_0) + 2\sum_{i=1}^b d_i\Bigg) + 2b\gamma N\left(1 + \mathrm{Tr}(\mathbf{P}_0)/\rho\right) + 4\sum_{i=1}^b d_i\Bigg] .$$

**Theorem S40.** *Assume **H**1-**H**2 and assume that for any $i \in [b]$, $m_i = m$ and $M_i = M$. In addition, let $\boldsymbol{N} = N\mathbf{1}_b, \boldsymbol{\gamma} = \gamma\mathbf{1}_b, \boldsymbol{\rho} = \rho\mathbf{1}_b$, $\rho > 0, \gamma > 0, N \geq 1$, such that $\gamma < 1/\tilde{M}$, $N\gamma < 2/(m+\tilde{M})$, and (S46) is satisfied. Then, for any $\varepsilon > 0$, any*

$$n_\varepsilon \geq 2\log\left(\mathsf{E}_0/\varepsilon\right)/(N\gamma m),$$

*we have, $W_2(\mu^\star_{\boldsymbol{\rho}} P^{n_\varepsilon}_{\boldsymbol{\rho},\boldsymbol{\gamma},\boldsymbol{N}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}}) \leq \varepsilon/3$.*

*Proof.* By some algebra and using $1/\log(1/(1-x)) \leq 1/x$ for $0 < x < 1$, the proof directly follows from Proposition S14 combined with Proposition S38 and Lemma S39. □

### S5.2. Upper bound on the tolerance parameter $\boldsymbol{\rho}_\varepsilon$

Define

$$R_0 = 2\sigma_U^2\left(d\sigma_U^2 + \sum_{i=1}^b M_i^2\|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2\right) + 2\sigma_U^4 ,$$

$$R_1 = d\sigma_U^2 + \sum_{i=1}^b M_i^2\|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2 + \sum_{i=1}^b d_i M_i/2$$

$$R_2 = 2d\max_{i\in[b]}\{M_i\}\sigma_U^2 + 2\sum_{i=1}^b M_i^3\|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2 + 8\sigma_U^4 + 8\sigma_U^2\Big[2d\sigma_U^2 + 2\sum_{i=1}^b M_i^2\|\mathbf{A}_i(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i^\star)\|^2\Big] .$$

Recall that $\bar{\rho} = \max_{i\in[b]}\{\rho_i\}$. Then, the following result holds.

**Lemma S41.** *Assume **H**1-**H**2. For any $\varepsilon > 0$, let $\boldsymbol{\rho}_\varepsilon \in (\mathbb{R}_+^*)^b$ such that*

$$\bar{\rho}_\varepsilon \leq \frac{-R_1 + \sqrt{R_1^2 + 4R_0\varepsilon m_U^{1/2}/(3\sqrt{2})}}{2R_0} \wedge \frac{\varepsilon\sqrt{m_U}}{3\sqrt{2}\sqrt{R_2 + [R_2/(12\sigma_U^2) + \sum_{i=1}^b d_i M_i]^2}} \wedge \frac{1}{12\sigma_U^2} \wedge \frac{-\sum_{i=1}^b d_i M_i + \sqrt{(\sum_{i=1}^b d_i M_i)^2 + 6R_2}}{2R_2} .$$

*Then, $W_2(\pi_{\boldsymbol{\rho}_\varepsilon}, \pi) \leq \varepsilon/3$.*

*Proof.* Let $\varepsilon > 0$. From (S73), for any $\bar{\rho} \leq 1/(12\sigma_U^2)$, $W_2(\pi_{\boldsymbol{\rho}}, \pi) \leq \sqrt{\frac{2}{m_U}}\max(A_1, A_3^{1/2})$ , where $A_1, A_3$ are defined in (S69) and (S72) respectively. This implies that $W_2(\pi_{\boldsymbol{\rho}}, \pi) \leq \varepsilon/3$ is verified if $\max(A_1, A_3^{1/2}) \leq \varepsilon\sqrt{m_U}/(3\sqrt{2})$. First, $A_1 \leq \varepsilon\sqrt{m_U}/(3\sqrt{2})$ holds if

$$\bar{\rho} \leq \frac{-R_1 + \sqrt{R_1^2 + 4R_0\varepsilon m_U^{1/2}/(3\sqrt{2})}}{2R_0} \wedge \frac{1}{12\sigma_U^2} . \tag{S142}$$

We now focus on $A_3$. Using the fact that for any $x \in \mathbb{R}, \mathrm{e}^x \geq x+1$, we have $2\prod_{i=1}^b (1+\rho_i M_i)^{d_i} \geq 2 + \sum_{i=1}^b d_i \log(1+\rho_i M_i)$ and therefore

$$A_3 \leq \exp\left(\bar{\rho}^2 R_2 + \sum_{i=1}^b d_i \log(1+\rho_i M_i)\right) - 1 - \sum_{i=1}^b d_i \log(1+\rho_i M_i) \; .$$

Since $\sum_{i=1}^b d_i \log(1+\rho_i M_i) \leq \bar{\rho}\sum_{i=1}^b d_i M_i$, $\bar{\rho}^2 R_2 + \sum_{i=1}^b d_i \log(1+\rho_i M_i) \leq 3/2$ holds for

$$\bar{\rho} \leq \frac{-\sum_{i=1}^b d_i M_i + \sqrt{(\sum_{i=1}^b d_i M_i)^2 + 6R_2}}{2R_2} \; . \tag{S143}$$

Since for any $x \leq 3/2, \mathrm{e}^x \leq 1 + x + x^2$ and using the fact that $\bar{\rho} \leq 1/(12\sigma_U^2)$, it follows that

$$A_3 \leq \bar{\rho}^2 R_2 + \left(\bar{\rho}^2 R_2 + \bar{\rho}\sum_{i=1}^b d_i M_i\right)^2 \leq \bar{\rho}^2 \left[B_1 + \left(\frac{R_2}{12\sigma_U^2} + \sum_{i=1}^b d_i M_i\right)^2\right] \; .$$

Hence $A_3^{1/2} \leq \varepsilon\sqrt{m_U}/(3\sqrt{2})$ holds under (S143) and

$$\bar{\rho} \leq \frac{\varepsilon\sqrt{m_U}}{3\sqrt{2}\sqrt{R_2 + \left(\frac{R_2}{12\sigma_U^2} + \sum_{i=1}^b d_i M_i\right)^2}} \; . \tag{S144}$$

The proof is concluded by combining (S142), (S143) and (S144). □

### S5.3. Upper bound on the step-size $\boldsymbol{\gamma}_\varepsilon$ and number of local iteration $\boldsymbol{N}_\varepsilon$

Based on Proposition S29 or Proposition S33, we now determine an upper bound on $\boldsymbol{\gamma}_\varepsilon$ to ensure $W_2(\Pi_{\boldsymbol{\rho}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}_\varepsilon}) \leq \varepsilon/3$ in the case $\boldsymbol{N} = N\mathbf{1}_b, \boldsymbol{\gamma} = \gamma\mathbf{1}_b, \boldsymbol{\rho} = \rho\mathbf{1}_b$ where $\rho > 0, \gamma > 0, N \geq 1$. The following results hold depending if **H**3 is considered. Define

$$\begin{aligned} C_\rho &= \frac{4\tilde{M}^2(1 + \|\bar{\mathbf{B}}_0^{-1}\mathbf{B}_0^\top \tilde{\mathbf{D}}_0^{1/2}\|^2)}{5m} \; , \\ C_0 &= (\tilde{M}^2/2)\left[\tilde{M}/\tilde{m} + 1/6\right]\sum_{i=1}^b d_i \; , \quad C_1 = \sum_{i=1}^b d_i \; , \quad C_2 = \varepsilon^2/(9C_\rho) \; . \end{aligned} \tag{S145}$$

**Lemma S42.** *Assume **H**1-**H**2 and assume for any $i \in [b]$, $m_i = m$ and $M_i = M$. In addition, let $\rho, \gamma_\varepsilon > 0$ and $N_\varepsilon \geq 1$ such that $\boldsymbol{\rho} = \rho\mathbf{1}_b$, $\boldsymbol{\gamma}_\varepsilon = \gamma_\varepsilon\mathbf{1}_b$, $\boldsymbol{N}_\varepsilon = N_\varepsilon\mathbf{1}_b$ and $\varepsilon > 0$ satisfying*

$$\gamma_\varepsilon \leq \frac{-C_1 + \sqrt{C_1^2 + 4C_0C_2}}{2C_0} \wedge \frac{m}{40\tilde{M}^2} \; . \tag{S146}$$

*Then $W_2(\Pi_{\boldsymbol{\rho}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}_\varepsilon}) \leq \varepsilon/3$.*

*Proof.* Let $\varepsilon > 0$. By Proposition S29, note that $W_2^2(\Pi_{\boldsymbol{\rho}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}_\varepsilon}) \leq \varepsilon^2/9$ is satisfied if

$$C_0\gamma_\varepsilon^2 + C_1\gamma_\varepsilon \leq C_2 \; .$$

This inequality is satisfied under the choice (S147). □

We now provide a condition on $\boldsymbol{N}$ and $\boldsymbol{\gamma}$ when **H**3 is considered.

**Lemma S43.** *Assume* **H**1-**H**2 *and assume for any* $i \in [b]$, $m_i = m$, $M_i = M$ *and* $L_i = L$. *In addition, let* $\rho, \gamma_\varepsilon > 0$ *and* $N_\varepsilon \geq 1$ *such that* $\boldsymbol{\rho} = \rho \mathbf{1}_b$, $\boldsymbol{\gamma}_\varepsilon = \gamma_\varepsilon \mathbf{1}_b$, $\boldsymbol{N}_\varepsilon = N_\varepsilon \mathbf{1}_b$ *and* $\varepsilon > 0$ *satisfying*

$$\gamma_\varepsilon \leq \frac{\varepsilon}{6b\sqrt{5\max_{i\in[b]}\{d_i\}C_\rho \tilde{M}^2[4 + (\max_{i\in[b]}\{d_i\}L^2 m)/(20\tilde{M}^4)]}} \wedge \frac{m}{40\tilde{M}^2} \tag{S147}$$

$$\wedge \frac{\varepsilon}{6b(5C_\rho \max_{i\in[b]}\{d_i\}m^3/\tilde{M}^2)}\,, \tag{S148}$$

*where* $C_\rho$ *is defined in* (S145). *Then* $W_2(\Pi_{\boldsymbol{\rho}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}_\varepsilon}) \leq \varepsilon/3$.

*Proof.* In Proposition S33, we dissociate $R^\star(\boldsymbol{\gamma})$ into two contributions and the conditions we impose on $\gamma_\varepsilon$ ensure $W_2(\Pi_{\boldsymbol{\rho}}, \Pi_{\boldsymbol{\rho},\boldsymbol{\gamma}_\varepsilon}) \leq \varepsilon/3$. More precisely, we have $\sum_{i=1}^b d_i\gamma_i^2\tilde{M}_i^2 + \frac{d_i\gamma_i^2\mathfrak{f}_i}{\tilde{M}_i}(d_i L_i^2 + \frac{\tilde{M}_i^4}{\tilde{m}_i}) \leq 2\varepsilon^2/9$ and $\sum_{i=1}^b d_i\gamma_i\tilde{M}_i\mathfrak{f}_i^3(1 + \mathfrak{f}_i + \mathfrak{f}_i^2) \leq 2\varepsilon^2/9$ where $\mathfrak{f}_i < 1$ for any $i \in [b]$. □

### S5.4. Discussion

Let $\boldsymbol{\rho}_\varepsilon = \rho_\varepsilon \mathbf{1}_b$ such that $W_2(\pi_{\boldsymbol{\rho}_\varepsilon}, \pi) \leq \varepsilon/3$. From Lemma S41, $\rho_\varepsilon = \mathcal{O}(\varepsilon/d)$ when $\varepsilon \to 0$ and $d \to \infty$. Similarly, let $\boldsymbol{\gamma}_\varepsilon = \gamma_\varepsilon \mathbf{1}_b$ such that $W_2(\Pi_{\boldsymbol{\rho}_\varepsilon}, \Pi_{\boldsymbol{\rho}_\varepsilon,\boldsymbol{\gamma}_\varepsilon}) < \varepsilon/3$. Under **H**1-**H**2, we obtain by Lemma S42 $\gamma_\varepsilon = \mathcal{O}(\varepsilon^4/d^3)$. On the other hand, when **H**3 is additionally assumed, we get by Lemma S43 $\gamma_\varepsilon = \mathcal{O}(\varepsilon^2/d^2)$. Finally, to apply Theorem S40 for the previous choices $\gamma_\varepsilon$ and $\rho_\varepsilon$, we obtain for $\boldsymbol{N}_\varepsilon = N_\varepsilon \mathbf{1}_b$ the conditions $N_\varepsilon = \mathcal{O}(d/\varepsilon^2)$ and $N_\varepsilon = \mathcal{O}(1)$ under **H**1-**H**2 and **H**1-**H**2-**H**3, respectively. In both scenarios, Theorem S40 implies $n_\varepsilon = \mathcal{O}(d^2\log(d)/(\varepsilon^2|\log(\varepsilon)|)$. This concludes the results depicted in Table 1 in the main paper.

## References


Bakry, D., Gentil, I., and Ledoux, M. *Analysis and geometry of Markov diffusion operators*, volume 348. Springer Science & Business Media, 2013.

Douc, R., Moulines, E., Priouret, P., and Soulier, P. *Markov chains*. Springer, 2018.

Durmus, A. and Moulines, E. High-dimensional Bayesian inference via the unadjusted Langevin algorithm. *Bernoulli*, 25 (4A):2854–2882, 2019. doi: 10.3150/18-BEJ1073.

Kent, J. Time-reversible diffusions. *Advances in Applied Probability*, 10:819–835, 12 1978. ISSN 1475-6064. doi: 10.1017/S0001867800031396.

Ledoux, M. *The Concentration of Measure Phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

Otto, F. and Villani, C. Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality. *Journal of Functional Analysis*, 173(2):361 – 400, 2000. ISSN 0022-1236.

Revuz, D. and Yor, M. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

Roberts, G. O. and Tweedie, R. L. Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 12 1996.

Villani, C. *Optimal Transport: Old and New*. Springer Berlin Heidelberg, 2008.

Vono, M., Paulin, D., and Doucet, A. Efficient MCMC sampling with dimension-free convergence rate using ADMM-type splitting. *arXiv preprint arXiv:1905.11937*, 2019.