Convergence of diffusions and their discretizations: from continuous to discrete processes and back

Valentin De Bortoli [*1] and Alain Durmus [†1]

[1]CMLA - École normale supérieure Paris-Saclay, CNRS, Université Paris-Saclay, 94235 Cachan, France.

April 24, 2019

Abstract

In this paper, we establish new quantitative convergence bounds for a class of functional autoregressive models in weighted total variation metrics. To derive this result, we show that under mild assumptions explicit minorization and Foster-Lyapunov drift conditions hold. Our bounds are then obtained adapting classical results from Markov chain theory. To illustrate our results we study the geometric ergodicity of Euler-Maruyama discretizations of diffusion with covariance matrix identity. Second, we provide a new approach to establish quantitative convergence of these diffusion processes by applying our conclusions in the discrete-time setting to a well-suited sequence of discretizations whose associated stepsizes decrease towards zero.

1 Introduction

The study of the convergence of Markov processes in general state space is a very attractive and active field of research motivated by applications in mathematics, physics and statistics. Among the many works on the subject, we can mention the pioneering results from [53, 51, 52] using the renewal approach. Then, the work of

---

[*]Email: valentin.debortoli@cmla.ens-cachan.fr
[†]Email: alain.durmus@cmla.ens-cachan.fr

[56, 46] paved the way for the use of *Foster-Lyapunov drift conditions* [31, 5] which, in combination of an appropriate *minorization condition*, implies (f, r)-ergodicity on general state space, drawing links with control theory, see [64, 15, 37]. This approach was successively applied to the study of Markov chains in numerous papers [9, 11, 59] and was later extended and used in the case of continuous-time Markov processes in [42, 47, 48, 18, 32, 30, 14, 65]. However, most of these results establish convergence in total variation or in V-norm and are non-quantitative. Explicit convergence bounds in the same metrics for Markov chains have been established in [62, 28, 17, 63, 61, 40, 45, 49, 29]. One of the main motivation for this work was to obtain quantitative bounds for Markov Chain Monte Carlo (MCMC) algorithms to derive stopping rules for simulations. To the authors' knowledge, the technique developed in these papers have not been adapted to continuous-time Markov processes, expect in [60]. One of the main reason is that deriving quantitative minorization conditions for continuous-time process seems to be even more difficult than for their discrete counterpart [26]. Indeed, in most cases, the constants which appear in minorization conditions are either really pessimistic or hard to quantify precisely [39, 57].

Since the last decade, in order to avoid the use of minorization conditions, attention has been paid to consider other metrics than the total variation distance or *V*-norm. In particular, Wasserstein metrics have shown to be very interesting in the study of Markov processes and to derive quantitative bounds of convergence. For example, [54, 41] introduced the notion of Ricci curvature of Markov chains and its use to derive precise bounds on variance and concentration inequalities for additive functionals. Following [33], [34] generalizes the Harris' theorem for *V*-norms to handle more general Wasserstein type metrics. In addition, the use of Wasserstein distance has been successively applied to the study of diffusion processes and MCMC algorithms. In particular, [25, 26] establish explicit convergence rates for diffusions and McKean Vlasov processes. Regarding analysis of MCMC methods, [35] establishes geometric convergence of the pre-conditioned Crank-Nicolson algorithm. Besides, [22, 12, 10, 1] study the computational complexity in Wasserstein distance to sample from a log-concave density on $\mathbb{R}^d$ using appropriate discretizations of the overdamped Langevin diffusion. One key idea introduced in [34] and [25] is the construction of an appropriate metric designed specifically for the Markov process under consideration. While this approach leads to quantitative results in the case of diffusions or their discretization, we can still wonder if appropriate minorization conditions can be found to derive similar bounds using classical results cited above.

In this paper, we show that for a class of functional auto-regressive models, sharp minorization conditions hold using an iterated Markov coupling kernel. As a result new quantitative convergence bounds can be obtained combining this conclusion and

2

drift inequalities for well-suited Lyapunov functionals. We apply them to the study of the Euler-Maruyama discretization of diffusions with identity covariance matrix under various curvature assumptions on the drift. The rates of convergence we derive in weighted total variation metric in this case improve the one recently established in [27]. Note that this study is important to control the complexity of this algorithm, when it is applied to the overdamped Langevin diffusion to sample from a target density π on $\mathbb{R}^d$. Indeed, while recent papers have established precise bounds between the n-th iterate of the Euler-Maruyama scheme and π in different metrics (e.g.total variation or Wasserstein distances), the convergence of the associated Markov kernel is in general needed to obtain quantitative bounds on the mean square error or concentration inequalities for additive functionals, see [22, 41].

In a second part, we show how the results we derive for functional auto-regressive models can be used to establish explicit convergence rates for diffusion processes. First, we show that, given an appropriate sequence of discretizations, the distance in some appropriate metric between the distributions of the diffusion at time t with different starting points, can be upper bounded by the limit of the distance between the corresponding discretizations, when the discretization stepsize decreases towards zero. Second, we design appropriate discretizations satisfying the needed conditions and which belong to the class of functional autoregressive models we study. Therefore, under the same curvature conditions as in the discrete case, we get quantitative convergence rates for diffusions by taking the limit in the bounds we derived for the Euler-Maruyama discretizations. Finally, the rates we obtained scale similarly, with respect to the parameters of the problem under consideration, to the ones given in [25, 26] for the Kantorovitch-Rubinstein distance and improve them in the case of the total variation norm.

The paper is organized as follows. First in Section 2.1 we present our main convergence results for a class of functional autoregressive models. We then specialize them to the Euler-Maruyama discretization of diffusions under various regularity assumptions on the drift term in Section 2.2. We derive sufficient conditions for the convergence of diffusion processes with identity covariance matrix in Section 3.1, based on a sequence of well-suited discretizations. Finally, in Section 3.2, we apply our results to the continuous counterpart of the situations considered in Section 2.2.

Notations

Let A, B and C three sets with $\mathsf{C} \subset \mathsf{B}$ and $f : \mathsf{A} \to \mathsf{B}$, we set $f^{\leftarrow}(\mathsf{C}) = \{x \in \mathsf{A} : f(x) \in \mathsf{C}\}$. Let $d \in \mathbb{N}^\star$ and $\langle \cdot, \cdot \rangle$ be a scalar product over $\mathbb{R}^d$, and $\|\cdot\|$ be the corresponding norm. Let $\mathsf{A} \subset \mathbb{R}^d$, we denote $\Delta_{\mathsf{A}} = \{(x,x) : x \in \mathsf{A}\} \subset \mathbb{R}^{2d}$ and for

3

$R \geq 0$, $\Delta_R = \{(x,y) \in \mathsf{A} : \|x - y\| \leq R\}$. In the rest of the paper we consider that $\mathbb{R}^d$ is endowed with the topology of the norm $\|\cdot\|$ and denote by λ the associated Lebesgue measure. Let U be an open set of $\mathbb{R}^d$, $n \in \mathbb{N}^\star$ and set $\mathrm{C}^n(\mathsf{U})$ be the set of the n-differentiable functions defined over U. Let $f \in \mathrm{C}^1(\mathsf{U})$, we denote by ∇f its gradient. Furthermore, if $f \in \mathrm{C}^2(\mathsf{U})$ we denote $\nabla^2 f$ its Hessian and Δ its Laplacian. $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel σ-field of $\mathbb{R}^d$. We also denote $\mathrm{C}(\mathsf{U})$ the set of continuous functions defined over U.

Let $\mathsf{X} \in \mathcal{B}(\mathbb{R}^d)$, X is equipped with the trace of $\mathcal{B}(\mathbb{R}^d)$ over X defined by $\mathcal{X} = \{\mathsf{A} \cap \mathsf{X} : \mathsf{A} \in \mathcal{B}(\mathbb{R}^d)\}$. Let $(\mathsf{Y}, \mathcal{Y})$ be some measurable space, we denote by $\mathbb{F}(\mathsf{X}, \mathsf{Y})$ the set of the $\mathcal{X}$-measurable functions over X. For any $f \in \mathbb{F}(\mathsf{X}, \mathbb{R})$ we define its essential supremum by $\mathrm{esssup}(f) = \inf\{a \geq 0 : \lambda(|f|^{\leftarrow}(a, +\infty)) = 0\}$. Let $\mathbb{M}(\mathcal{X})$ be the set of finite signed measures over $\mathcal{X}$ and $\mu \in \mathbb{M}(\mathcal{X})$. For $f \in \mathbb{F}(\mathsf{X}, \mathbb{R})$ a μ-integrable function we denote by $\mu(f)$ the integral of f w.r.t. to μ. Let $V \in \mathbb{F}(\mathbb{R}^d, [1, +\infty))$. We define the V-norm for any $f \in \mathbb{F}(\mathsf{X}, \mathbb{R})$ and the V-total variation norm for any $\mu \in \mathbb{M}(\mathcal{X})$ as follows

$$\|f\|_V = \mathrm{esssup}(|f|/V) , \qquad \|\mu\|_V = (1/2) \sup_{f \in \mathbb{F}(\mathsf{X}, \mathbb{R}), \|f\|_V \leq 1} \left| \int_{\mathbb{R}^d} f(x) \mathrm{d}\mu(x) \right| .$$

In the case where $V = 1$ this norm is called the total variation norm of μ. Let μ, ν be two probability measures over $\mathcal{X}$, $i.e.$ two elements of $\mathbb{M}(\mathcal{X})$ such that $\mu(\mathsf{X}) = \nu(\mathsf{X}) = 1$. A probability measure ζ over $\mathcal{X}^{\otimes 2}$ is said to be a transference plan between μ and ν if for any $\mathsf{A} \in \mathcal{X}$, $\zeta(\mathsf{A} \times \mathcal{X}) = \mu(\mathsf{A})$ and $\zeta(\mathcal{X} \times \mathsf{A}) = \nu(\mathsf{A})$. We denote by $\mathbf{T}(\mu, \nu)$ the set of all transference plans between μ and ν. Let $\mathbf{c} \in \mathbb{F}(\mathsf{X} \times \mathsf{X}, [0, +\infty))$. We define the Wasserstein cost $\mathbf{W}_\mathbf{c}(\mu, \nu)$ between μ and ν by

$$\mathbf{W}_\mathbf{c}(\mu, \nu) = \inf_{\zeta \in \mathbf{T}(\mu, \nu)} \int_{\mathsf{X}^2} \mathbf{c}(x, y) \mathrm{d}\zeta(x, y) .$$

If $\mathbf{c}(x, y) = \|x - y\|^p$ for $p \geq 1$ then we define the Wasserstein distance of order p by $\mathbf{W}_p = \mathbf{W}_\mathbf{c}^{1/p}$. Assume that $\mathbf{c}(x, y) = \mathbb{1}_{\Delta_{\mathsf{X}}^c}((x, y))W(x, y)$ with $W \in \mathbb{F}(\mathsf{X} \times \mathsf{X}, [0, +\infty))$ such that W is symmetric, satisfies the triangle inequality, $i.e.$ for any $x, y, z \in \mathsf{X}$, $W(x, z) \leq W(x, y) + W(y, z)$, and for any $x, y \in \mathsf{X}$, $W(x, y) = 0$ implies $x = y$. Then $\mathbf{c}$ is a distance over X^2 and the associated Wasserstein cost, denoted by $\mathbf{d}_W$, is a metric. Note that if $W(x, y) = \{V(x) + V(y)\}/2$ then $\mathbf{d}_W(\mu, \nu) \leq \|\mu - \nu\|_V$ with equality if V is lower semi-continuous [66, Theorem 5.10].

Assume that $\mu \ll \nu$ and denote by $\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$ its Radon-Nikodyn derivative. We define the Kullback-Leibler divergence, $\mathrm{KL}(\nu|\mu)$, between μ and ν, by

$$\mathrm{KL}(\mu|\nu) = \int_{\mathsf{X}} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}(x)\right) \mathrm{d}\mu(x) .$$

Let $\mathcal{Z}$ be a σ-field. We say that $\mathrm{P} : \mathsf{X} \times \mathcal{Z} \to [0, +\infty)$ is a Markov kernel if for any $x \in \mathsf{X}$, $\mathrm{P}(x, \cdot)$ is a probability measure over $\mathcal{Z}$ and for any $\mathsf{A} \in \mathcal{Z}$, $\mathrm{P}(\cdot, \mathsf{A}) \in \mathbb{F}(\mathsf{X}, [0, +\infty))$. Let $\mathsf{Y} \in \mathcal{B}(\mathbb{R}^d)$ be equipped with $\mathcal{Y}$ the trace of $\mathcal{B}(\mathbb{R}^d)$ over $\mathcal{Y}$, $\mathrm{P} : \mathsf{X} \times \mathcal{Z}$ and $\mathrm{Q} : \mathsf{Y} \times \mathcal{Z}$ be two Markov kernels. We say that $\mathrm{K} : \mathsf{X} \times \mathsf{Y} \to \mathcal{Z}^{\otimes 2}$ is a Markov coupling kernel if for any $(x, y) \in \mathsf{X} \times \mathsf{Y}$, $\mathrm{K}((x, y), \cdot)$ is a transference plan between $\mathrm{P}(x, \cdot)$ and $\mathrm{Q}(y, \cdot)$.

2 Quantitative convergence bounds for a class of functional autoregressive models

2.1 Main results

Let $\mathsf{X} \in \mathcal{B}(\mathbb{R}^d)$ endowed with the trace of $\mathcal{B}(\mathbb{R}^d)$ on X denoted by $\mathcal{X} = \{\mathsf{A} \cap \mathsf{X} : \mathsf{A} \in \mathcal{B}(\mathbb{R}^d)\}$. In this section we consider the Markov chain $(X_k)_{k \in \mathbb{N}}$ defined by $X_0 \in \mathsf{X}$ and the following recursion for any $k \in \mathbb{N}$

$$X_{k+1} = \Pi\left(\mathcal{T}_\gamma(X_k) + \sqrt{\gamma}\, Z_{k+1}\right) , \tag{1}$$

where $\{\mathcal{T}_{\tilde{\gamma}} : \tilde{\gamma} \in (0, \bar{\gamma}]\}$ is a family of measurable functions from X to $\mathbb{R}^d$ with $\bar{\gamma} > 0$, $\gamma \in (0, \bar{\gamma}]$ is a stepsize, $(Z_k)_{k \in \mathbb{N}^\star}$ is a family of i.i.d d-dimensional zero mean Gaussian random variables with covariance identity and $\Pi : \mathbb{R}^d \to \mathsf{X}$ is a measurable function. The Markov chain $(X_k)_{k \in \mathbb{N}}$ defined by (1) is associated with the Markov kernel R_γ defined on $\mathsf{X} \times \mathcal{B}(\mathbb{R}^d)$ for any $\gamma \in (0, \bar{\gamma}]$, $x \in \mathbb{R}^d$ and $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$ by

$$\mathrm{R}_\gamma(x, \mathsf{A}) = (2\pi\gamma)^{-d/2} \int_{\Pi^{\leftarrow}(\mathsf{A})} \exp\left[-(2\gamma)^{-1}\|y - \mathcal{T}_\gamma(x)\|^2\right] \mathrm{d}y . \tag{2}$$

Note that for any $x \in \mathsf{X}$, $\mathrm{R}_\gamma(x, \mathsf{X}) = 1$ and therefore, R_γ given in (2) is also a Markov kernel over $\mathsf{X} \times \mathcal{X}$.

In this section we state explicit convergence results for R_γ for some Wasserstein distances and discuss the rates we obtain. These results rely on appropriate minorization and Foster-Lyapunov drift conditions. We first derive the minorization condition for the n-th iterate of R_γ. To do so, we consider a Markov coupling kernel K_γ for R_γ for any $\gamma \in (0, \bar{\gamma}]$, *i.e.* for any $x, y \in \mathbb{R}^d$, $\mathrm{K}_\gamma((x, y), \cdot)$ is a transference plan between $\mathrm{R}_\gamma(x, \cdot)$ and $\mathrm{R}_\gamma(y, \cdot)$. Indeed, in that case, by [16, Theorem 19.3.7], we have for any $x, y \in \mathsf{X}$, $\gamma \in (0, \bar{\gamma}]$ and $n \in \mathbb{N}^\star$,

$$\|\delta_x \mathrm{R}_\gamma^n - \delta_y \mathrm{R}_\gamma^n\|_{\mathrm{TV}} \leq \mathrm{K}_\gamma^n((x, y), \Delta_{\mathsf{X}}^{\mathrm{c}}) . \tag{3}$$

For any $x, y, z \in \mathbb{R}^d$, $\gamma \in (0, \bar{\gamma}]$, define

$$
\mathrm{e}(x,y) = \begin{cases} \mathrm{E}(x,y)/\|\mathrm{E}(x,y)\| & \text{if } x \neq y \\ 0 & \text{otherwise} \end{cases} , \qquad \mathrm{E}(x,y) = \mathcal{T}_\gamma(y) - \mathcal{T}_\gamma(x) ,
$$

and

$$
\mathcal{S}_\gamma(x,y,z) = \mathcal{T}_\gamma(y) + \left[(\mathrm{Id} - 2\mathrm{e}(x,y)\mathrm{e}(x,y)^\top) z \right] ,
$$

$$
p_\gamma(x,y,z) = 1 \wedge \frac{\boldsymbol{\varphi}_\gamma(\|\mathrm{E}(x,y)\| - \langle \mathrm{e}(x,y), z \rangle)}{\boldsymbol{\varphi}_\gamma(\langle \mathrm{e}(x,y), z \rangle)} , \tag{4}
$$

where $\boldsymbol{\varphi}_\gamma$ is the one dimensional zero mean Gaussian distribution function with variance γ. Let $(U_k)_{k \in \mathbb{N}^\star}$ be a sequence of i.i.d. uniform random variables on $[0,1]$ independent of $(Z_k)_{k \in \mathbb{N}^\star}$. Define the Markov chain $(X_k, Y_k)_{k \in \mathbb{N}}$ starting from $(X_0, Y_0) \in \mathsf{X}^2$ by the recursion: for any $k \in \mathbb{N}$

$$
\tilde{X}_{k+1} = \mathcal{T}_\gamma(X_k) + \sqrt{\gamma} Z_{k+1} ,
$$

$$
\tilde{Y}_{k+1} = \begin{cases} \tilde{X}_{k+1} & \text{if } \mathcal{T}_\gamma(X_k) = \mathcal{T}_\gamma(Y_k) , \\ W_{k+1} \tilde{X}_{k+1} + (1 - W_{k+1}) \mathcal{S}_\gamma(X_k, Y_k, \sqrt{\gamma} Z_{k+1}) & \text{otherwise} , \end{cases}
$$

where $W_{k+1} = \mathbb{1}_{(-\infty, 0]}(U_{k+1} - p(X_k, Y_k, \sqrt{\gamma} Z_{k+1}))$ and finally set

$$
(X_{k+1}, Y_{k+1}) = (\Pi(\tilde{X}_{k+1}), \Pi(\tilde{Y}_{k+1})) . \tag{5}
$$

The Markov chain $(X_k, Y_k)_{k \in \mathbb{N}}$ is associated with the Markov kernel K_γ on $\mathsf{X}^2 \times \mathcal{X}^{\otimes 2}$ given for all $\gamma \in (0, \bar{\gamma}]$, $x, y \in \mathsf{X}$ and $\mathsf{A} \in \mathcal{X}^{\otimes 2}$ by

$$
\mathrm{K}_\gamma((x,y), \mathsf{A}) = \frac{\mathbb{1}_{\Delta_{\mathbb{R}^d}}(\mathcal{T}_\gamma(x), \mathcal{T}_\gamma(y))}{(2\pi\gamma)^{d/2}} \int_{\mathbb{R}^d} \mathbb{1}_{(\Pi^{\otimes 2}) \leftarrow (\mathsf{A})}(\tilde{x}, \tilde{x}) \mathrm{e}^{-\frac{\|\tilde{x} - \mathcal{T}_\gamma(x)\|^2}{2\gamma}} \mathrm{d}\tilde{x}
$$

$$
+ \frac{\mathbb{1}_{\Delta_{\mathbb{R}^d}^c}(\mathcal{T}_\gamma(x), \mathcal{T}_\gamma(y))}{(2\pi\gamma)^{d/2}} \left[\int_{\mathbb{R}^d} \mathbb{1}_{(\Pi^{\otimes 2}) \leftarrow (\mathsf{A})}(\tilde{x}, \tilde{x}) p_\gamma(x, y, \tilde{x} - \mathcal{T}_\gamma(x)) \mathrm{e}^{-\frac{\|\tilde{x} - \mathcal{T}_\gamma(x)\|^2}{2\gamma}} \mathrm{d}\tilde{x} \right.
$$

$$
\left. + \int_{\mathbb{R}^d} \mathbb{1}_{(\Pi^{\otimes 2}) \leftarrow (\mathsf{A})} (\tilde{x}, \mathcal{S}_\gamma(x, y, \tilde{x} - \mathcal{T}_\gamma(x))) \{1 - p_\gamma(x, y, \tilde{x} - \mathcal{T}_\gamma(x))\} \mathrm{e}^{-\frac{\|\tilde{x} - \mathcal{T}_\gamma(x)\|^2}{2\gamma}} \mathrm{d}\tilde{x} \right] , \tag{6}
$$

where $\Pi^{\otimes 2} = (\Pi, \Pi)$ and $\Delta_{\mathbb{R}^d} = \left\{ (\tilde{x}, \tilde{y}) \in \mathbb{R}^d \times \mathbb{R}^d : \tilde{x} = \tilde{y} \right\}$. Note that marginally, by definition, the distribution of X_{k+1} given X_k is $\mathrm{R}_\gamma(X_k, \cdot)$. It is well-know (see e.g. [8, Section 3.3]) that $\tilde{Y}_{k+1}$ and $\mathcal{T}_\gamma(Y_k) + \sqrt{\gamma} Z_{k+1}$ have the same distribution given

Y_k, and therefore the distribution of Y_{k+1} given Y_k is $R_\gamma(Y_k, \cdot)$. As a result for any $x, y \in X$, $K_\gamma((x, y), \cdot)$ is a transference plan between $R_\gamma(x, \cdot)$ and $R_\gamma(y, \cdot)$.

As emphasized previously, based on (3), to study convergence of R_γ for $\gamma \in (0, \bar\gamma]$ (in total variation or V-norm), we first give upper bounds for $K_\gamma^n((x, y), \Delta_X^c)$ for any $x, y \in X$ and $n \in \mathbb{N}^\star$ under appropriate conditions on $\mathcal{T}_\gamma$ and Π.

A1. *The function $\Pi : \mathbb{R}^d \to X$ is non expansive: i.e. for any $x, y \in \mathbb{R}^d$, $\|\Pi(x) - \Pi(y)\| \le \|x - y\|$.*

Note that **A**3 is satisfied if Π is the proximal operator [2, Proposition 12.27] associated with a convex lower semi-continuous function $f : \mathbb{R}^d \to (-\infty, +\infty]$. For example, if $f(x) = \sum_{i=1}^d |x_i|$, the associated proximal operator is the soft thresholding operator [55, Section 6.5.2]. For f the convex indicator of a convex closed set $C \subset \mathbb{R}^d$, defined by $f(x) = 0$ for $x \in C$, $f(x) = +\infty$ otherwise, the proximal operator is simply the projection onto C by [2, Example 12.21]:

$$\Pi_C(x) = \arg\min_{y \in C} \|y - x\| \ .$$

For the applications that we have in mind, the use of Markov chains defined by (1) have been proposed based on optimization literature to sample non-smooth log-concave densities [24], [7], [20], [3]. We now consider the following assumption on $\{\mathcal{T}_{\tilde\gamma} : \tilde\gamma \in (0, \bar\gamma]\}$ where $A \in \mathcal{B}(\mathbb{R}^{2d})$.

A 2 (A)**.** *There exists $\kappa : (0, \bar\gamma] \to \mathbb{R}$ such that for any $\gamma \in (0, \bar\gamma]$, $\gamma\kappa(\gamma) \in (-1, +\infty)$ and for any $(x, y) \in A$*

$$\|\mathcal{T}_\gamma(x) - \mathcal{T}_\gamma(y)\|^2 \le (1 + \gamma\kappa(\gamma))\|x - y\|^2 \ . \tag{7}$$

In addition, one of the following conditions holds for any $\gamma \in (0, \bar\gamma]$: (i) $\kappa(\gamma) < 0$; (ii) $\kappa(\gamma) \le 0$; (iii) $\kappa(\gamma) > 0$.

Note that **A**2(X^2)-(i) or **A**2(X^2)-(ii) imply that for any $\gamma \in (0, \bar\gamma]$, $\mathcal{T}_\gamma$ is non-expansive itself (see **A**3). For $\kappa : (0, \bar\gamma] \to [0, +\infty)$ and $\ell \in \mathbb{N}^\star$, $\gamma \in (0, \bar\gamma]$ such that $\gamma\kappa(\gamma) \in (-1, +\infty)$, define

$$\Xi_n(\kappa) = \gamma \sum_{k=1}^n (1 + \gamma\kappa(\gamma))^{-k} \ .$$

The following theorem gives a generalization of a minorization condition on autoregressive models [22, Section 6].

Theorem 1. *Let* $\mathsf{A} \in \mathcal{B}(\mathbb{R}^{2d})$ *and assume* **A**3 *and* **A**2(A). *Let* $(X_k, Y_k)_{k\in\mathbb{N}}$ *be defined by* (5) *with* $(X_0, Y_0) = (x, y) \in \mathsf{A} \cap \mathsf{X}^2$. *Then for any* $n \in \mathbb{N}^\star$

$$\mathbb{P}\left(X_n \neq Y_n \text{ and for any } k \in \{1, \ldots, n-1\}, (X_k, Y_k) \in \mathsf{A}\right)$$
$$\leq \mathbb{1}_{\Delta_{\mathsf{X}}^{\mathsf{c}}}((x, y)) \left\{1 - 2\boldsymbol{\Phi}\left(-\frac{\|x - y\|}{2\Xi_n^{1/2}(\kappa)}\right)\right\} ,$$

where $\boldsymbol{\Phi}$ *is the cumulative distribution function of the standard normal distribution on* $\mathbb{R}$.

Proof. The proof is a simple application of Theorem 25 in Appendix A. □

Based on Theorem 1, since $\mathbb{P}(X_n \neq Y_n) = \mathrm{K}^n((x, y), \Delta_{\mathsf{X}}^{\mathsf{c}})$ where $(X_k, Y_k)_{k\in\mathbb{N}}$ is defined by (5) with $(X_0, Y_0) = (x, y) \in \mathsf{X}^2$, we can derive minorization conditions for the Markov kernel R_γ^n and $n \in \mathbb{N}^\star$ for any $\gamma \in (0, \bar{\gamma}]$ depending on the assumption we make on κ in **A**2(X^2). More precisely, these minorization conditions are derived using $\mathrm{K}_\gamma^{\ell\lceil 1/\gamma\rceil}$ with $\ell \in \mathbb{N}^\star$. This is a requirement to obtain sharp bounds. Let $x, y \in \mathsf{X}$, based only on the results of Theorem 1, we obtain only that for any $\ell \in \mathbb{N}^\star$, $\lim_{\gamma\to 0}\|\delta_x \mathrm{R}_\gamma^\ell - \delta_y \mathrm{R}_\gamma^\ell\|_{\mathrm{TV}} \leq 1$, whereas the following lemma implies that for any $\ell \in \mathbb{N}^\star$, $\lim_{\gamma\to 0}\|\delta_x \mathrm{R}_\gamma^{\ell\lceil 1/\gamma\rceil} - \delta_y \mathrm{R}_\gamma^{\ell\lceil 1/\gamma\rceil}\|_{\mathrm{TV}} < 1$.

Lemma 2. *Let* $\bar{\gamma} > 0$ *and* $\kappa : (0, \bar{\gamma}] \to (0, +\infty)$, *with* $\kappa(\gamma)\gamma \in (-1, +\infty)$ *for any* $\gamma \in (0, \bar{\gamma}]$. *We have for any* $\gamma \in (0, \bar{\gamma}]$ *such that* $\kappa(\gamma) \neq 0$ *and* $\ell \in \mathbb{N}^\star$

$$\Xi_{\ell\lceil 1/\gamma\rceil}(\kappa) = -\kappa^{-1}(\gamma)\left\{\exp\left[-\ell\lceil 1/\gamma\rceil \log\{1 + \gamma\kappa(\gamma)\}\right] - 1\right\} . \tag{8}$$

In addition, for any $\ell \in \mathbb{N}^\star$, $\gamma \in (0, \bar{\gamma}]$

(a) $\Xi_{\ell\lceil 1/\gamma\rceil}(\kappa) \geq \alpha_-(\kappa, \gamma, \ell) = -\kappa^{-1}(\gamma)[\exp(-\ell\kappa(\gamma)) - 1]$ *if for any* $\gamma \in (0, \bar{\gamma}]$, $\kappa(\gamma) < 0$;

(b) $\Xi_{\ell\lceil 1/\gamma\rceil}(\kappa) \geq \alpha_0(\kappa, \gamma, \ell) = \ell$ *if for any* $\gamma \in (0, \bar{\gamma}]$, $\kappa(\gamma) \leq 0$;

(c) $\Xi_{\ell\lceil 1/\gamma\rceil}(\kappa) \geq \alpha_+(\kappa, \gamma, \ell) = \kappa^{-1}(\gamma)\left[1 - \exp\left\{-\frac{\ell\kappa(\gamma)}{1 + \gamma\kappa(\gamma)}\right\}\right]$ *if for any* $\gamma \in (0, \bar{\gamma}]$, $\kappa(\gamma) > 0$.

Proof. The proof is postponed to Appendix F.1. □

Proposition 3. *Let* $\mathsf{A} \in \mathcal{B}(\mathbb{R}^{2d})$ *and assume* **A**3 *and* **A**2(A)*. Let* $(X_k, Y_k)_{k\in\mathbb{N}}$ *be defined by* (5) *with* $(X_0, Y_0) = (x, y) \in \mathsf{A} \cap \mathsf{X}^2$*. Then for any* $\ell \in \mathbb{N}^*$ *and* $\gamma \in (0, \bar{\gamma}]$,

$$\mathbb{P}\left(X_{\ell\lceil\gamma\rceil} \neq Y_{\ell\lceil\gamma\rceil} \text{ and for any } k \in \{1, \dots, n-1\}, (X_k, Y_k) \in \mathsf{A}^2\right)$$
$$\leq 1 - 2\mathbf{\Phi}\left(-\alpha^{-1/2}(\kappa, \gamma, \ell)\|x - y\|/2\right) ,$$

where

(a) $\alpha = \alpha_-$ *is given in Lemma 2-(a) if* **A**2(A)-(i) *holds ;*

(b) $\alpha = \alpha_0$ *is given in Lemma 2-(b) if* **A**2(A)-(ii) *holds ;*

(c) $\alpha = \alpha_+$ *is given in Lemma 2-(c) if* **A**2(A)-(iii) *holds.*

Proof. The proof is a direct application of Theorem 1 and Lemma 2 with $\mathsf{\kappa}(\gamma) = \kappa(\gamma)$. $\qquad\square$

Depending on the conditions imposed on κ defined in **A**2(X^2), we obtain the following consequences of Proposition 3 which establish, either an explicit convergence bound in total variation for R_γ, or a quantitative minorization condition satisfied by this kernel.

Corollary 4. *Assume* **A**3 *and* **A**2(X^2):

(a) *if* **A**2(X^2)-(i) *holds, then, for any* $\gamma \in (0, \bar{\gamma}]$, R_γ *admits a unique invariant probability measure* π_γ *and we have for any* $x \in \mathbb{R}^d$ *and* $\ell \in \mathbb{N}^\star$,

$$\|\delta_x \mathrm{R}_\gamma^{\ell\lceil 1/\gamma\rceil} - \pi_\gamma\|_{\mathrm{TV}} \leq 1 - 2\int_{\mathbb{R}^d} \mathbf{\Phi}\left\{-\kappa_-^{1/2}\|x - y\|/\{2(\exp(-\ell\kappa_-) - 1)^{1/2}\}\right\} \mathrm{d}\pi_\gamma(y) ,$$

where $\kappa_- = \sup_{\gamma\in(0,\bar{\gamma}]} \kappa(\gamma)$;

(b) *if* **A**2(X^2)-(ii) *holds and for any* $\gamma \in (0, \bar{\gamma}]$, R_γ *admits an invariant probability measure* π_γ, *then we have for any* $\gamma \in (0, \bar{\gamma}]$, $x \in \mathbb{R}^d$ *and* $\ell \in \mathbb{N}^\star$,

$$\|\delta_x \mathrm{R}_\gamma^{\ell\lceil 1/\gamma\rceil} - \pi_\gamma\|_{\mathrm{TV}} \leq 1 - 2\int_{\mathbb{R}^d} \mathbf{\Phi}\left\{-\|x - y\|/(2\ell^{1/2})\right\} \mathrm{d}\pi_\gamma(y) .$$

Proof. (a) Consider $V : \mathsf{X} \to [1, +\infty]$ given for any $x \in \mathsf{X}$ by $V(x) = 1 + \|x\|$. Then since **A**2(X^2) with $\sup_{\gamma\in(0,\bar{\gamma}]} \kappa(\gamma) \leq \kappa_- < 0$ holds, using the triangle inequality and the Cauchy-Schwarz inequality, we have for any $\gamma \in (0, \bar{\gamma}]$ and $x \in \mathsf{X}$

$$\mathrm{R}_\gamma V(x) \leq \|\mathcal{T}_\gamma(x)\| + \sqrt{\gamma d} \leq (1 + \kappa_-\gamma)\|x\| + \|\mathcal{T}_\gamma(0)\| + \sqrt{\gamma d} + 1 \leq \lambda V(x) + A ,$$

with $\lambda \in (0,1)$ and $A \geq 0$. From this result, since for any $\gamma \in (0, \bar\gamma]$, R_γ is a Feller kernel and the level sets of V are compact, R_γ admits a unique invariant probability measure π_γ for any $\gamma \in (0, \bar\gamma]$ by [16, Theorem 12.3.3]. Then the last result is a straightforward consequence of Proposition 3-(a), (3) and the fact that for any $\ell \in \mathbb{N}^\star$ and $\gamma \in (0, \bar\gamma]$, $\alpha_-(\kappa, \gamma, \ell) \geq -(\exp(-\ell\kappa_-) - 1)/\kappa_-$ since $t \mapsto (\exp(\ell t) - 1)/t$ is increasing on $\mathbb{R}$.

(b) This result is a direct consequence of Proposition 3-(b), (3) and the fact that R_γ admits an invariant probability measure π_γ.

$\square$

In other words, if $\mathcal{T}_\gamma$ is a contractive mapping, see $\mathbf{A}2(\mathsf{X}^2)$-(i), then the convergence of $(\delta_x R_\gamma^{\ell\lceil 1/\gamma \rceil})_{\ell \in \mathbb{N}^\star}$ to π_γ in total variation is exponential in $\ell^{1/2}$. If $\mathcal{T}_\gamma$ is non expansive, see $\mathbf{A}2(\mathsf{X}^2)$-(ii), and R_γ admits an invariant probability measure the convergence of $(\delta_x R_\gamma^{\ell\lceil 1/\gamma \rceil})_{\ell \in \mathbb{N}^\star}$ to π_γ in total variation is linear in $\ell^{1/2}$. In the case where $\mathcal{T}_\gamma$ is non expansive, see $\mathbf{A}2(\mathsf{X}^2)$-(ii), or simply Lipschitz, see $\mathbf{A}2(\mathsf{X}^2)$-(iii) and no additional assumption is made, we do not directly obtain contraction in total variation but only minorization conditions.

Corollary 5. *Assume* $\mathbf{A}3$ *and* $\mathbf{A}2(\mathsf{X}^2)$. *Then, for any* $\gamma \in (0, \bar\gamma]$,

(a) *if* $\mathbf{A}2(\mathsf{X}^2)$-(ii) *holds, for any* $x, y \in \bar{\mathrm{B}}(0, M) \cap \mathsf{X}$ *and* $\ell \in \mathbb{N}^\star$, *with* $M \geq 0$ *and* $\ell \geq \lceil M^2 \rceil$,
$$K_\gamma^{\ell\lceil 1/\gamma \rceil}((x,y), \Delta_\mathsf{X}^\mathsf{c}) \leq 1 - 2\Phi(-1/2) \; ; \qquad (9)$$

(b) *if* $\mathbf{A}2(\mathsf{X}^2)$-(iii) *holds, for any* $x, y \in \mathsf{X}$ *and* $\ell \in \mathbb{N}^\star$,
$$K_\gamma^{\ell\lceil 1/\gamma \rceil}((x,y), \Delta_\mathsf{X}^\mathsf{c}) \leq 1 - 2\Phi\left\{-(1+\bar\gamma)^{1/2}(1+\kappa_+)^{1/2} \|x - y\|/2\right\} \; , \qquad (10)$$

where $\kappa_+ = \sup_{\gamma \in (0, \bar\gamma]} \kappa(\gamma)$.

Proof. (a) The proof is a direct application of Proposition 3-(b), the fact that $(X_k, Y_k) \in \mathsf{X}^2$ for any $k \in \mathbb{N}$ and that K_γ is the Markov kernel associated with $(X_k, Y_k)_{k \in \mathbb{N}}$.

(b) Consider the case where $\mathbf{A}2(\mathsf{X}^2)$-(iii) holds. Using that for any $t \geq 0$, $1 - \mathrm{e}^{-t} \geq t/(t+1)$ we obtain that for any $\gamma \in (0, \bar\gamma]$ and $\ell \in \mathbb{N}^\star$
$$\alpha_+(\kappa, \gamma, \ell) \geq \ell/(1 + (\ell + \bar\gamma)\kappa(\gamma)) \geq (1 + (1+\bar\gamma)\kappa_+)^{-1} \geq (1+\bar\gamma)^{-1}(1+\kappa_+)^{-1} \; ,$$

where α_+ is given in Lemma 2-(c). Then, combining this result and Proposition 3-(c) completes the proof.

$\square$

Under $\mathbf{A}2(\mathsf{X}^2)$-(ii) or $\mathbf{A}2(\mathsf{X}^2)$-(iii), to conclude that R_γ admits a unique invariant probability measure π_γ, for $\gamma \in (0, \bar{\gamma}]$, we need to impose some additional conditions.

In our application below, we are mainly interested in the case where R_γ satisfies a geometric drift condition. Let $(\mathsf{Y}, \mathcal{Y})$ be a measurable space, $\lambda \in (0, 1)$, $A \geq 0$ and $V : \mathbb{R}^d \to [1, +\infty)$ be a measurable function.

$\mathbf{D}_{\mathrm{d}}(V, \lambda, A, \mathsf{C})$. *A Markov kernel* R *on* $\mathsf{Y} \times \mathcal{Y}$ *satisfies the discrete Foster-Lyapunov drift condition if for all* $y \in \mathsf{Y}$

$$\mathrm{R}V(y) \leq \lambda V(y) + A\mathbb{1}_{\mathsf{C}}(y) \ .$$

Note that this drift condition implies the existence of an invariant probability measure if R is a Feller kernel and the level sets of V are compact, see [16, Theorem 12.3.3]. In the sequel, we implicitly assume that R_γ admits a unique invariant distribution π_γ for any $\gamma \in (0, \bar{\gamma}]$, and are interested in establishing a contraction for the Wasserstein metric $\mathbf{d}_W$ associated with the cost

$$\mathbf{w} : (x, y) \mapsto \mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}((x, y))W(x, y)$$

where $W : \mathsf{X} \times \mathsf{X} \to [0, +\infty)$ satisfies for any $x, y, z \in \mathsf{X}$, $W(x, y) = W(y, x)$, $W(x, z) \leq W(x, y) + W(y, z)$ and $W(x, y) = 0$ implies that $x = y$. Let μ, ν be two probability measures over $\mathcal{X}$, we highlight three cases

- total variation: if $W = 1$ then $\mathbf{d}_W(\mu, \nu) = \|\mu - \nu\|_{\mathrm{TV}}$;

- V-norm: if $W(x, y) = \left\{ \tilde{V}(x) + \tilde{V}(y) \right\}/2$ where $\tilde{V} : \mathbb{R}^d \to [1, +\infty)$ is measurable then $\mathbf{d}_W(\mu, \nu) \geq \|\mu - \nu\|_{\tilde{V}}$;

- total variation + Kantorovitch-Rubinstein metric: if $W(x, y) = 1 + \|x - y\|$ then $\mathbf{d}_W(\mu, \nu) \geq \|\mu - \nu\|_{\mathrm{TV}} + \mathbf{W}_1(\mu, \nu)$.

We now state convergence bounds for Markov kernels which satisfy one of the conclusions of Corollary 5. Indeed, in order to deal with the two assumptions $\mathbf{A}2(\mathsf{X}^2)$-(ii) and $\mathbf{A}2(\mathsf{X}^2)$-(iii) together, we provide a general result regarding the contraction of R_γ in the metric $\mathbf{d}_W$ for some function W on X^2. This result is based on an abstract condition on $\tilde{\mathrm{K}}_\gamma \mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}$, which is satisfied under $\mathbf{A}2(\mathsf{X}^2)$-(ii) or $\mathbf{A}2(\mathsf{X}^2)$-(iii) by Corollary 5 with $\tilde{\mathrm{K}}_\gamma \leftarrow \mathrm{K}_\gamma$, and a drift condition for $\tilde{\mathrm{K}}_\gamma$, where $\tilde{\mathrm{K}}_\gamma$ is a Markov coupling kernel for R_γ. Its proof is an application of Theorem 27 in Appendix B which essentially follows from [16, Lemma 19.4.2], see also [17]. We recall that for any $M \geq 0$,

$$\Delta_M = \{(x, y) \in \mathbb{R}^{2d} : \|x - y\| \leq M\} \ . \tag{11}$$

Theorem 6. *Assume that there exist $\bar\gamma > 0$, $\lambda \in (0,1)$, $A \geq 0$, a measurable function $W : \mathsf{X} \times \mathsf{X} \to [1, +\infty)$ such that for any $C \geq 0$,*

$$\operatorname{diam}\left\{ (x,y) \in \mathsf{X}^2 : \ W(x,y) \leq C \right\} < +\infty \,,$$

and $\tilde{\mathrm{K}}_\gamma$ a Markov coupling kernel for R_γ such that for any $\gamma \in (0, \bar\gamma]$, $\tilde{\mathrm{K}}_\gamma$ satisfies $\mathbf{D}_{\mathrm{d}}(W, \lambda^\gamma, A\gamma, \mathsf{X}^2)$. Further, assume that there exists $\boldsymbol{\Psi} : (0, \bar\gamma] \times \mathbb{N}^\star \times \mathbb{R}_+ \to [0,1]$ such that for any $\gamma \in (0, \bar\gamma]$, $\ell \in \mathbb{N}^\star$ and $x, y \in \mathsf{X}$

$$\tilde{\mathrm{K}}_\gamma^{\ell\lceil 1/\gamma\rceil}((x,y), \Delta_\mathsf{X}^\mathsf{c}) \leq 1 - \boldsymbol{\Psi}(\gamma, \ell, \|x - y\|) \,, \qquad \boldsymbol{\Psi}(\gamma, \ell, 0) = 1 \,,$$
$$\text{and for any } M \geq 0, \quad \inf_{(x,y) \in \Delta_M} \boldsymbol{\Psi}(\gamma, \ell, \|x - y\|) > 0 \,. \tag{12}$$

(a) For any $\gamma \in (0, \bar\gamma]$, $M_{\mathrm{d}} \geq \operatorname{diam}(\{(x,y) \in \mathsf{X}^2 : W(x,y) \leq K_{\mathrm{d}}\})$ with $K_{\mathrm{d}} = 2A(1 + \bar\gamma)\{1 + \log^{-1}(1/\lambda)\}$ and $\ell \in \mathbb{N}^\star$, there exist $C_\gamma \geq 0$ and $\rho_\gamma \in (0,1)$ such that for any $x, y \in \mathsf{X}$ and $k \in \mathbb{N}$

$$\mathbf{d}_W(\delta_x \mathrm{R}_\gamma^k, \delta_y \mathrm{R}_\gamma^k) \leq C_\gamma \rho_\gamma^{\left\lfloor k(\ell(\lceil 1/\gamma\rceil)^{-1}\right\rfloor} W(x,y) \,, \tag{13}$$

where

$$C_\gamma = 2[1 + A_\gamma][1 + c_\gamma/\{(1 - \varepsilon_{\mathrm{d},\gamma})(1 - \lambda_\gamma)\}] \,,$$
$$\log(\rho_\gamma) = \{\log(1 - \varepsilon_{\mathrm{d},\gamma})\log(\lambda_\gamma)\}/\{\log(1 - \varepsilon_{\mathrm{d},\gamma}) + \log(\lambda_\gamma) - \log(c_\gamma)\} \,,$$
$$A_\gamma = A\gamma(1 - \lambda^{\gamma\ell\lceil 1/\gamma\rceil})/(1 - \lambda^\gamma) \,, \quad c_\gamma = \lambda^{\gamma\ell\lceil 1/\gamma\rceil} A_\gamma + K_{\mathrm{d}} \,,$$
$$\varepsilon_{\mathrm{d},\gamma} = \inf_{(x,y) \in \Delta_{M_{\mathrm{d}}}} \boldsymbol{\Psi}(\gamma, \ell, \|x - y\|) \,, \quad \lambda_\gamma = (\lambda^{\gamma\ell\lceil 1/\gamma\rceil} + 1)/2 \,.$$

(b) For any $\gamma \in (0, \bar\gamma]$, $M_{\mathrm{d}} \geq \operatorname{diam}(\{(x,y) \in \mathsf{X}^2 : W(x,y) \leq K_{\mathrm{d}}\})$ with $K_{\mathrm{d}} = 2A(1 + \bar\gamma)\{1 + \log^{-1}(1/\lambda)\}$ and $\ell \in \mathbb{N}^\star$, it holds that

$$C_\gamma \leq \bar{C}_1 \,, \quad \log(\rho_\gamma) \leq \log(\bar\rho_1) \leq 0 \,,$$
$$\bar{C}_1 = 2[1 + \bar{A}_1][1 + \bar{c}_1 / \{(1 - \bar\varepsilon_{\mathrm{d},1})(1 - \bar\lambda_1)\}] \,,$$
$$\log(\rho_1) = \{\log(1 - \bar\varepsilon_{\mathrm{d},1})\log(\bar\lambda_1)\} / \{\log(1 - \bar\varepsilon_{\mathrm{d},1}) + \log(\bar\lambda_1) - \log(\bar{c}_1)\} \,,$$
$$\bar{A}_1 = A(1 + \bar\gamma)\min(\ell, 1 + \log^{-1}(1/\lambda)) \,, \quad \bar{c}_1 = \bar{A}_1 + K_{\mathrm{d}} \,, \quad \bar\lambda_1 = (\lambda + 1)/2 \,,$$
$$\bar\varepsilon_{\mathrm{d},1} = \inf_{\gamma \in (0, \bar\gamma], \ (x,y) \in \Delta_{M_{\mathrm{d}}}} \boldsymbol{\Psi}(\gamma, \ell, \|x - y\|) \,.$$

(c) In addition, if $\bar\gamma \leq 1$, $-\log(\lambda) \in [0, \log(2)]$, $A \geq 1$ and $0 < \bar\varepsilon_{\mathrm{d},1} \leq 1 - \mathrm{e}^{-1}$, then

$$\log^{-1}(1/\bar\rho_1) \leq 12\log(2)\log(6A(1 + \log^{-1}(1/\lambda)))\big/(\log(1/\lambda)\bar\varepsilon_{\mathrm{d},\bar\gamma}) \,. \tag{14}$$

Proof. First, note that $1 - \lambda^t = -\int_0^t \log(\lambda) e^{s \log(\lambda)} ds \geq -\log(\lambda) t e^{t \log(\lambda)}$ for any $t \in (0, \bar{t}]$, for $\bar{t} > 0$, and therefore

$$t/(1 - \lambda^t) = t + t\lambda^t/(1 - \lambda^t) \leq \bar{t} + \log^{-1}(\lambda^{-1}) . \tag{15}$$

(a) To establish (13), we apply an extension of [16, Theorem 19.4.1], given for completeness in Theorem 27. For any $x, y \in \mathsf{X}$ such that $W(x, y) \leq K_\mathrm{d}$ we have

$$\tilde{\mathrm{K}}_\gamma^{\ell\lceil 1/\gamma\rceil}((x, y), \Delta_\mathsf{X}^\mathsf{c}) \leq 1 - \varepsilon_{\mathrm{d},\gamma} .$$

Using that $\tilde{\mathrm{K}}_\gamma$ satisfies $\mathbf{D}_\mathrm{d}(W, \lambda^\gamma, A\gamma, \mathsf{X}^2)$, we can apply Theorem 27 with $M \leftarrow K_\mathrm{d} \geq 2A\gamma/(1 - \lambda^\gamma)$ by (15), which completes the proof of (a).

(b) We now provide upper bounds for C_γ and ρ_γ. These constants are non-decreasing in $c_\gamma, \varepsilon_{\mathrm{d},\gamma}$ and λ_γ. Therefore it suffices to give upper bounds on $c_\gamma, \varepsilon_{\mathrm{d},\gamma}$ and λ_γ. The result is then straightforward using that $(1 - \lambda^{\gamma\ell\lceil 1/\gamma\rceil})/(1 - \lambda^\gamma) \leq \ell\lceil 1/\gamma\rceil$, $(1 - \lambda^{\gamma\ell\lceil 1/\gamma\rceil})/(1 - \lambda^\gamma) \leq \bar{\gamma} + \log^{-1}(1/\lambda)$ and $\lambda^{\gamma\ell\lceil 1/\gamma\rceil} \leq \lambda$.

(c) By assumption on $\bar{\gamma}, \lambda$ and $\bar{\varepsilon}_{\mathrm{d},1}$ we have that $\log((1 - \bar{\varepsilon}_{\mathrm{d},1})^{-1}) \leq 1$ and

$$\log(\bar{\lambda}_1^{-1}) \leq \log(\lambda^{-1}) \leq \log(2) , \quad 1 \leq K_\mathrm{d} \leq \bar{c}_1 .$$

As a result, we obtain that $\log(\bar{\lambda}_1^{-1})/\log(\bar{c}_1) \leq 1$ and $\log((1 - \bar{\varepsilon}_{\mathrm{d},1})^{-1})/\log(\bar{c}_1) \leq 1$. Therefore we have

$$\log^{-1}(1/\bar{\rho}_1) = \log(\bar{\lambda}_1^{-1}) + \log((1 - \bar{\varepsilon}_{\mathrm{d},1})^{-1}) + \log(\bar{c}_1)/\Big[\log(\bar{\lambda}_1^{-1})\log((1 - \bar{\varepsilon}_{\mathrm{d},1})^{-1})\Big]$$
$$\leq 3\log[6A(1 + \log^{-1}(1/\lambda))]/\Big[\log(\bar{\lambda}_1^{-1})\log((1 - \bar{\varepsilon}_{\mathrm{d},1})^{-1})\Big] .$$

Using that $\log(1 - t) \leq -t$ for any $t \in (0, 1]$ and the definition of $\bar{\lambda}_1$, we obtain that

$$\log^{-1}(\bar{\rho}_1^{-1}) \leq 6\bar{\varepsilon}_{\mathrm{d},1}^{-1}(1 - \lambda)^{-1}\log[6A(1 + \log^{-1}(1/\lambda))] .$$

Finally, we get (14) using that for any $t \in [0, \log(2)]$, $1 - e^{-t} \geq (2\log(2))^{-1}t$.

$\square$

Note that Theorem 6-(c) gives an upper bound on the rate of convergence ρ_1 in the worst case scenario for which the minorization constant $\bar{\varepsilon}_{\mathrm{d},1}$ is small and the constant λ in $\mathbf{D}_\mathrm{d}(V, \lambda^\gamma, A\gamma, \mathsf{X}^2)$ is close to one.

The following theorem gives the same conclusion as Theorem 6 but the dependency in the problem constants in the convergence bounds is different. We compare these bounds at the end of this section. Note however that Theorem 7 requires a stronger drift condition than Theorem 6.

13

Theorem 7. *Assume that there exist $\bar{\gamma} > 0$, $\lambda \in (0,1)$, $A \geq 0$, $\tilde{M}_{\mathrm{d}} > 0$, a measurable function $W : \mathsf{X} \times \mathsf{X} \to [1, +\infty)$ and $\tilde{\mathrm{K}}_\gamma$ a Markov coupling kernel for R_γ such that for any $\gamma \in (0, \bar{\gamma}]$, $\tilde{\mathrm{K}}_\gamma$ satisfies $\mathbf{D}_{\mathrm{d}}(W, \lambda^\gamma, A\gamma, \Delta_{\tilde{M}_{\mathrm{d}}})$, where $\Delta_{\tilde{M}_{\mathrm{d}}}$ is defined by (11). Further, assume that there exists $\boldsymbol{\Psi} : (0,\bar{\gamma}] \times \mathbb{N}^\star \times \mathbb{R}_+ \to [0,1]$ such that for any $\gamma \in (0, \bar{\gamma}]$, $\ell \in \mathbb{N}^\star$ and $x,y \in \mathsf{X}$, (12) is satisfied.*

(a) For any $\gamma \in (0,\bar{\gamma}]$ and $\ell \in \mathbb{N}^\star$, there exist $D_\gamma, \tilde{C}_\gamma \geq 0$ and $\tilde{\rho}_\gamma \in (0,1)$ such that for any $x,y \in \mathsf{X}$ and $k \in \mathbb{N}$

$$\mathbf{d}_W(\delta_x \mathrm{R}_\gamma^k, \delta_y \mathrm{R}_\gamma^k) \leq \lambda^{k\gamma/2}[D_{\gamma,1}W(x,y) + D_{\gamma,2}] + \tilde{C}_\gamma \tilde{\rho}_\gamma^{k\gamma/2} \ , \qquad (16)$$

where

$$D_{\gamma,1} = 1 + A[\log(1/\lambda)\lambda^\gamma]^{-1} \ , \qquad D_{\gamma,2} = D_{\gamma,1}\left[A\lambda^{-\gamma\lceil 1/\gamma \rceil \ell}\gamma(\lambda^{-\gamma} + \ell \lceil 1/\gamma \rceil)\right] \ ,$$

$$\tilde{C}_\gamma = 2A\log^{-1}(1/\tilde{\rho}_\gamma)/\tilde{\rho}_\gamma \ ,$$

$$\log(\tilde{\rho}_\gamma) = \left\{\log(\lambda)\log(1 - \tilde{\varepsilon}_{\mathrm{d},\gamma})\right\}\big/\left\{-\log(\tilde{c}_\gamma) + \log(1 - \tilde{\varepsilon}_{\mathrm{d},\gamma})\right\} \ ,$$

$$\tilde{K}_{\mathrm{d}} = \sup_{(x,y) \in \Delta_{\tilde{M}_{\mathrm{d}}}} [W(x,y)] \ , \quad \tilde{c}_\gamma = \tilde{K}_{\mathrm{d}} + A\lambda^{-\gamma\lceil 1/\gamma \rceil \ell}\gamma(\lambda^{-\gamma} + \ell \lceil 1/\gamma \rceil) \ ,$$

$$\tilde{\varepsilon}_{\mathrm{d},\gamma} = \inf_{(x,y) \in \Delta_{\tilde{M}_{\mathrm{d}}}} \boldsymbol{\Psi}(\gamma, \ell, \|x - y\|) \ .$$

(b) For any $\gamma \in (0,\bar{\gamma}]$ and $\ell \in \mathbb{N}^\star$, it holds that

$$D_{\gamma,1} \leq \bar{D}_1 = 1 + A\log^{-1}(1/\lambda)/\lambda^{\bar{\gamma}} \ , \quad D_{\gamma,2} \leq \bar{D}_2 = \bar{D}_1 A\lambda^{-(1+\bar{\gamma})\ell}(\lambda^{-\bar{\gamma}} + \ell)(1 + \bar{\gamma}) \ ,$$

$$\tilde{C}_\gamma \leq \bar{C}_2 = 2A\log^{-1}(1/\bar{\rho}_2)/\bar{\rho}_2 \ ,$$

$$\log(\tilde{\rho}_\gamma) \leq \log(\bar{\rho}_2) = \left\{\log(\lambda)\log(1 - \bar{\varepsilon}_{\mathrm{d},2})\right\}\big/\left\{-\log(\bar{c}_2) + \log(1 - \bar{\varepsilon}_{\mathrm{d},2})\right\} \ ,$$

$$\bar{c}_2 = \tilde{K}_{\mathrm{d}} + A\lambda^{-(1+\bar{\gamma})\ell}(\lambda^{-\bar{\gamma}} + \ell)(1 + \bar{\gamma}) \ ,$$

$$\bar{\varepsilon}_{\mathrm{d},2} = \inf_{\gamma \in (0,\bar{\gamma}], \ (x,y) \in \Delta_{\tilde{M}_{\mathrm{d}}}} \boldsymbol{\Psi}(\gamma, \ell, \|x - y\|) \ .$$

(c) In addition, if $\bar{\gamma} \leq 1$, $-\log(\lambda) \in [0, \log(2)]$ and $\bar{\varepsilon}_{\mathrm{d},2} \leq 1 - \mathrm{e}^{-1}$, then

$$\log^{-1}(\bar{\rho}_2^{-1})$$

$$\leq \left[1 + \log\left(\tilde{K}_{\mathrm{d}}\right) + \log(1 + 2A(1 + \ell)) + (2\ell + 1)\log(\lambda^{-1})\right]\bigg/\left[\log(\lambda^{-1})\bar{\varepsilon}_{\mathrm{d},2}\right] \ .$$

Proof. (a) The proof of this theorem is an application of Theorem 30 in Appendix C. Let $\gamma \in (0,\bar{\gamma}]$. Consider $\mathbf{d}(x,y) = \mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}(x,y)$ which satisfies $\mathbf{H}1$. Then, since $\tilde{\mathrm{K}}_\gamma$ and

14

$\boldsymbol{\Psi}$ satisfy $\mathbf{D}_{\mathrm{d}}(W, \lambda^\gamma, A\gamma, \Delta_{\tilde{M}_\mathrm{d}})$ and (12) respectively, $\mathbf{H}2(\mathrm{K}_\gamma)$ and $\mathbf{H}3(\mathrm{K}_\gamma)$ are satisfied. In particular, $\mathbf{H}2(\mathrm{K}_\gamma)$-(i) is satisfied since for any $x, y \in \Delta_{\tilde{M}_\mathrm{d}}$, $\tilde{\mathrm{K}}_\gamma^{\lceil 1/\gamma \rceil \ell} \mathbb{1}_{\Delta_\mathsf{X}^\mathsf{c}}(x, y) \leq \{1 - \inf_{(x,y) \in \Delta_{\tilde{M}_\mathrm{d}}} \boldsymbol{\Psi}(\gamma, \ell, \|x - y\|)\} \mathbb{1}_{\Delta_\mathsf{X}^\mathsf{c}}(x, y)$. $\mathbf{H}2(\mathrm{K}_\gamma)$-(ii) is satisfied since for any $x, y \in \mathsf{X}$, $\mathrm{K}_\gamma \mathbb{1}_{\Delta_\mathsf{X}^\mathsf{c}}(x, y) \leq \mathbb{1}_{\Delta_\mathsf{X}^\mathsf{c}}(x, y)$. Finally, the conditions $\mathbf{H}2(\mathrm{K}_\gamma)$-(iii) and $\mathbf{H}3(\mathrm{K}_\gamma)$ hold by $\mathbf{D}_{\mathrm{d}}(W, \lambda^\gamma, A\gamma, \Delta_{\tilde{M}_\mathrm{d}})$ with $W_1 \leftarrow W$, $W_2 \leftarrow W\mathrm{d}$, $\lambda_1 = \lambda_2 \leftarrow \lambda^\gamma$, $A_1 = A_2 \leftarrow A\gamma$ and $\mathsf{C} \leftarrow \Delta_{\tilde{M}_\mathrm{d}}$. Applying Theorem 30, we obtain (16).

(b) The proof is straightforward using that $\lambda^\gamma \geq \lambda^{\bar{\gamma}}$.

(c) First, by assumption on $\bar{\gamma}$ and λ, we have the following inequality

$$\lambda^{-(1+\bar{\gamma})\ell}(\lambda^{-\bar{\gamma}} + \ell) \leq \lambda^{-2\ell}(\lambda^{-1} + \ell) \leq \lambda^{-(2\ell+1)}(1 + \ell) \ .$$

As a result and using the fact that $\log(1-t) \leq -t$ for any $t \in (0, 1)$, $\log((1-\bar{\varepsilon}_{\mathrm{d},2})^{-1}) \leq 1$ and $W(x, y) \geq 1$ for any $x, y \in \mathsf{X}$, we obtain that

$$\begin{aligned}
\log^{-1}(\bar{\rho}_2^{-1}) &\leq [\log(\lambda^{-1}) \log((1 - \bar{\varepsilon}_{\mathrm{d},\bar{\gamma}})^{-1})]^{-1} [1 + \log(\bar{c}_2)] \\
&\leq \left[\log(\lambda^{-1})\bar{\varepsilon}_{\mathrm{d},2}\right]^{-1} \left[1 + \log\left(\tilde{K}_\mathrm{d}\right) + \log(1 + 2A(1 + \ell)\lambda^{-(2\ell+1)})\right] \ , \\
&\leq \left[\log(\lambda^{-1})\bar{\varepsilon}_{\mathrm{d},2}\right]^{-1} \left[1 + \log\left(\tilde{K}_\mathrm{d}\right) + \log(1 + 2A(1 + \ell)) + (2\ell + 1) \log(\lambda^{-1})\right] \ ,
\end{aligned}$$

which completes the proof.

$\square$

First, note that in (16), the leading term, $\tilde{C}_\gamma \tilde{\rho}_\gamma^{k\gamma/2}$, does not depend on $x, y \in \mathsf{Y}$. The rate in front of the initial conditions $W(x, y)$ is given by $\lambda^{k/2}$ which is a better rate than $\tilde{\rho}_\gamma^{k\gamma/2}$. Some remarks are in order here concerning the bounds obtained in Theorem 6-(c) and Theorem 7-(c). Assume that $\ell = 1$, we will see in Section 2.2 that the leading term in the upper bound in Theorem 6-(c), respectively Theorem 7-(c), is given by $\log(A)/(\log(\lambda^{-1})\bar{\varepsilon}_{\mathrm{d},1})$, respectively $\log(A)/(\log(\lambda^{-1})\bar{\varepsilon}_{\mathrm{d},2})$. In addition, we will show that in some applications given in Section 2.2 below, $\bar{\varepsilon}_{\mathrm{d},1}$ is smaller than $\bar{\varepsilon}_{\mathrm{d},2}$. Therefore, the bounds provided in Theorem 7-(c) yield better rates than the one in Theorem 6-(c). The main difference between the two results is that in the proof of Theorem 6 a drift condition on the *iterated* coupling kernel $\tilde{\mathrm{K}}_\gamma^{\lceil 1/\gamma \rceil}$ is required. However, even if such drift conditions can be derived from a drift condition on $\tilde{\mathrm{K}}_\gamma$, the constants obtained by this technique are not sharp in general. On the contrary, the proof of Theorem 7 uses the iterated minorization condition and a drift condition on the *original* coupling $\tilde{\mathrm{K}}_\gamma$.

2.2 Application to the projected Euler-Maruyama discretization

Here we consider the case in which the operator $\mathcal{T}_\gamma$ in (1) is given by the discretization of a diffusion. For $b : \mathbb{R}^d \to \mathbb{R}^d$, we study the projected Euler-Maruyama discretization, *i.e.* we consider the following assumption for $\mathsf{X} \subset \mathbb{R}^d$.

B 1 (X). X *is assumed to be a convex closed subset of $\mathbb{R}^d$, $\Pi = \Pi_\mathsf{X}$ is the projection onto X and*

$$\mathcal{T}_\gamma(x) = x + \gamma b(x) \text{ for any } \gamma > 0 \text{ and } x \in \mathsf{X} . \tag{17}$$

Note that if $\mathsf{X} = \mathbb{R}^d$ and $\Pi = \mathrm{Id}$, then this scheme is the classical Euler-Maruyama discretization of a diffusion with drift b and diffusion coefficient $\sigma = \mathrm{Id}$. The application to the tamed Euler-Maruyama discretization of the results of Section 2.1 is given in Appendix D. In what follows, we show the convergence in V-norm for the projected Euler-Maruyama discretization and discuss the dependency of the constants appearing in the bounds we obtain with respect to the properties we assume on the drift b.

First, we show that some regularity/curvature conditions on the drift b imply condition **A**2(X^2) for $\mathcal{T}_\gamma$ given by (17). Let $\mathtt{m} \in \mathbb{R}$.

B2. *There exists $\mathtt{L} \geq 0$ such that b is $\mathtt{L}$-Lipschitz, i.e. for any $x, y \in \mathsf{X}$, $\|b(x) - b(y)\| \leq \mathtt{L}\|x - y\|$ and $b(0) = 0$.*

B 3 ($\mathtt{m}$). *For any $x, y \in \mathsf{X}$,*

$$\langle b(x) - b(y), x - y \rangle \leq -\mathtt{m} \|x - y\|^2 .$$

Note that if there exists $U \in \mathrm{C}^1(\mathsf{X}, \mathbb{R}^d)$ such that for any $x \in \mathsf{X}$, $b(x) = -\nabla U(x)$ and **B** 3($\mathtt{m}$) holds with $\mathtt{m} = 0$, respectively $\mathtt{m} > 0$ then U is convex, respectively strongly convex. Note that **B** 3(0) does not imply that $\mathcal{T}_\gamma$ given by (17) is non-expansive, therefore we consider the following assumption.

B4. *There exists $\mathtt{m}_b > 0$ such that for any $x, y \in \mathsf{X}$,*

$$\langle b(x) - b(y), x - y \rangle \leq -\mathtt{m}_b \|b(x) - b(y)\|^2 . \tag{18}$$

Proposition 8. *Assume* **B**1(X) *for $\mathsf{X} \subset \mathbb{R}^d$ and*

(a) **B** 2 *and* **B** 3($\mathtt{m}$) *with $\mathtt{m} \in \mathbb{R}$. Then (7) in* **A**2($\mathsf{X}^2$) *holds for any $\gamma > 0$ with $\kappa(\gamma) = -2\mathtt{m} + \mathtt{L}^2\gamma$. In particular, if $\mathtt{m} > 0$ then* **A**2(X^2)-(i) *holds for any $\bar{\gamma} < 2\mathtt{m}/\mathtt{L}^2$ and if $\mathtt{m} \leq 0$ then* **A**2(X^2)-(iii) *holds for any $\bar{\gamma} > 0$;*

(b) **B**4, *then* **A**2(X^2)-(ii) *holds with* $\kappa(\gamma) = 0$ *for any* $\bar{\gamma} \leq 2m_b$.

Proof. See Appendix F.2 □

Note that (18) implies that **B**2 with $L = m_b^{-1}$ and **B**3(0) hold. Assume $X = \mathbb{R}^d$ and that there exists $U \in C^1(\mathbb{R}^d, \mathbb{R})$ such that for any $x \in \mathbb{R}^d$, $b(x) = -\nabla U(x)$. If U is convex and ∇U is Lipschitz, *i.e.* if **B**2 and **B**3(0) hold, then (18) holds with $m_b = L^{-1}$ by [50, Theorem 2.1.5]. Based on Proposition 8 and assuming **B**1, we obtain the following results on the Markov kernel R_γ defined by (2) with $\gamma > 0$.

- If **B**2 and **B**3(m) hold with $m > 0$, then we obtain, by Proposition 8-(a) and Lemma 2-(a), that for any $\gamma \in (0, 2m/L^2)$ and $\ell \in \mathbb{N}^\star$,

$$\alpha_-(\kappa, \gamma, \ell) = -\frac{\exp(-\ell(-2m + L^2\gamma)) - 1}{-2m + L^2\gamma} .$$

  In addition, Corollary 4-(a) implies that for any $\gamma \in (0, 2m/L^2]$ and $x \in X$, $(\delta_x R_\gamma^{\lceil 1/\gamma \rceil \ell})_{\ell \in \mathbb{N}}$ converges exponentially fast to its invariant probability measure π_γ in total variation, with a rate which does not depend on γ, but only on m and L.

- In the case where (18) holds, then combining Proposition 8-(b) and Corollary 5-(a) we obtain that on any compact set $K \subset X$, $R_\gamma^{\lceil 1/\gamma \rceil \ell}$ satisfies the minorization condition (9) with universal constants for $\ell \geq \mathrm{diam}(K)^2$. In addition, if R_γ admits an invariant probability measure π_γ, then Corollary 4-(b) implies that for any $\gamma \in (0, 2m_b]$ and $x \in X$, $(\delta_x R_\gamma^{\lceil 1/\gamma \rceil \ell})_{\ell \in \mathbb{N}}$ converges linearly in $\ell^{1/2}$ to π_γ in total variation.

- In the case where **B**2 and **B**3(m) are satisfied with $m < 0$, we obtain that for any $\gamma > 0$ and $\ell \in \mathbb{N}^\star$

$$\alpha_+(\kappa, \gamma, \ell) = (-2m + L^2\gamma)^{-1} \left\{ 1 - \exp\left[-\ell(-2m + L^2\gamma)/(1 + \gamma(-2m + L^2\gamma)) \right] \right\}$$
$$\leq (-2m + L^2\gamma)^{-1} ,$$

  which implies that the bound given by Proposition 3 does not go to 0 when ℓ goes to infinity. Therefore we cannot directly conclude that the Markov chain converges in total variation. However, by Proposition 8-(a), Corollary 5-(b) shows that for any $\gamma \in (0, \bar{\gamma}]$ with $\bar{\gamma} > 0$, $R_\gamma^{\lceil 1/\gamma \rceil}$ satisfies the minorization condition (10), with constants which only depend on m and L.

17

Now we discuss several assumptions which imply drift conditions on the Markov coupling kernel K_γ for R_γ with $\gamma \in (0, \bar{\gamma}]$, with $\bar{\gamma} > 0$. This will allow us to use Theorem 6 and Theorem 7. First, we consider conditions on b which imply that R_γ for $\gamma \in (0, \bar{\gamma}]$, with $\bar{\gamma} > 0$, is geometrically convergent in a metric which dominates the total variation distance and the Wasserstein distance of order 1. This result is an application of Theorem 7 and the constants we end up with are independent of the dimension d. To do so, we establish that there exists a Lyapunov function W for which K_γ satisfies for $\gamma \in (0, \bar{\gamma}]$, $\mathbf{D}_d(W, \lambda^\gamma, A\gamma, \Delta_{M_d})$ where Δ_{M_d} is given by (11), $\bar{\gamma} > 0$ and $M_d \geq 0$ which do not depend on the dimension.

C 1. *There exist $R_1 > 0$ and $\mathtt{m}_1^+ > 0$ such that for any $x, y \in \mathsf{X}$ with $\|x - y\| \geq R_1$,*

$$\langle b(x) - b(y), x - y \rangle \leq -\mathtt{m}_1^+ \|x - y\|^2 .$$

This assumption is common in the literature [27, 25, 43] and implies strong convexity of the drift b outside of a ball. This condition is met if $b = -\nabla U$, with $U = U_1 + U_2$ with $U_1, U_2 \in \mathrm{C}^2(\mathbb{R}^d, \mathbb{R})$, $\sup_{\mathbb{R}^d} \|\nabla^2 U_1\| < +\infty$ and U_2 strongly convex. In the next proposition, we derive the announced drift for $W : \mathsf{X}^2 \to [1, +\infty)$ defined for any $x, y \in \mathsf{X}$ by

$$W(x, y) = 1 + \|x - y\| / R_1 . \tag{19}$$

Proposition 9. *Assume* $\mathbf{B}1(\mathsf{X})$ *for* $\mathsf{X} \subset \mathbb{R}^d$, $\mathbf{B}2$, $\mathbf{B}3(\mathtt{m})$ *for* $\mathtt{m} \in \mathbb{R}$ *and* $\mathbf{C}1$. *Then* K_γ *defined by* (6) *satisfies* $\mathbf{D}_d(W, \lambda^\gamma, A\gamma, \Delta_{R_1})$ *for any* $\gamma \in (0, \bar{\gamma}]$ *where* Δ_{R_1} *is given by* (11), $\bar{\gamma} \in (0, 2\mathtt{m}_1^+/\mathrm{L}^2)$ *and*

$$\lambda = \exp\left[-(\mathtt{m}_1^+/2 - \bar{\gamma}\mathrm{L}^2/4)\right] , \quad A = \mathtt{m}_1^+ - \mathtt{m} . \tag{20}$$

Proof. Let $x, y \in \mathsf{X}$, $\mathrm{E} = \mathcal{T}_\gamma(y) - \mathcal{T}_\gamma(x)$. If $\mathrm{E} = 0$ then the proposition is trivial, therefore we suppose that $E \neq 0$ and let $\mathrm{e} = \mathrm{E}/\|\mathrm{E}\|$. Consider Z_1, a d-dimensional Gaussian random variable with zero mean and covariance identity. By (6) and the fact that Π_X is non expansive, we have for any $\gamma \in (0, \bar{\gamma}]$

$$
\begin{aligned}
K_\gamma \|x - y\| &\leq \mathbb{E}\left[(1 - p_\gamma(x, y, \sqrt{\gamma}Z_1)) \left\|(\mathcal{T}_\gamma(x) + \sqrt{\gamma}Z_1) - (\mathcal{T}_\gamma(y) + \sqrt{\gamma}(\mathrm{Id} - 2\mathrm{ee}^\top)Z_1)\right\|\right] \\
&\leq \mathbb{E}\left[\left\|\mathrm{E} - 2\sqrt{\gamma}\mathrm{ee}^\top Z_1\right\| (1 - p_\gamma(x, y, \sqrt{\gamma}Z_1))\right] \\
&\leq \int_\mathbb{R} \|\mathrm{E} - 2z\mathrm{e}\| \left\{\boldsymbol{\varphi}_\gamma(z) - (\boldsymbol{\varphi}_\gamma(z) \wedge \boldsymbol{\varphi}_\gamma(\|\mathrm{E}\| - z))\right\} \mathrm{d}z \\
&\leq \int_{-\infty}^{\|\mathrm{E}\|/2} (\|E\| - 2z) \left\{\boldsymbol{\varphi}_\gamma(z) - \boldsymbol{\varphi}_\gamma(\|\mathrm{E}\| - z)\right\} \mathrm{d}z \leq \|\mathrm{E}\| , \tag{21}
\end{aligned}
$$

where we have used a simple change of variable for the last line, p_γ is defined in (4) and $\boldsymbol{\varphi}_\gamma$ is the one dimensional zero mean Gaussian distribution function with variance γ.

Consider now the case $(x, y) \in \Delta^{\mathrm{c}}_{R_1}$. By **B**2, **C**1, and since for any $t \in [-1, +\infty)$, $\sqrt{1 + t} \leq 1 + t/2$, we have that

$$\|\mathcal{T}_\gamma(x) - \mathcal{T}_\gamma(y)\| \leq (1 - \gamma \mathtt{m}_1^+ + \gamma^2 \mathtt{L}^2/2) \|x - y\| . \tag{22}$$

Combining (21) and (22), we obtain that for any $(x, y) \in \Delta^{\mathrm{c}}_{R_1}$,

$$
\begin{aligned}
\mathrm{K}_\gamma W(x, y) &\leq (1 - \gamma \mathtt{m}_1^+ + \gamma^2 \mathtt{L}^2/2) \|x - y\| / R_1 + 1 \\
&\leq (1 - \gamma \mathtt{m}_1^+/2 + \gamma^2 \mathtt{L}^2/4)(1 + \|x - y\| / R_1) \\
&\leq \lambda^\gamma W(x, y) .
\end{aligned} \tag{23}
$$

Similarly, we obtain using Proposition 8-(a) that for any $(x, y) \in \Delta_{R_1}$

$$
\begin{aligned}
\mathrm{K}_\gamma(1 + \|x - y\|) &\leq (1 - \gamma \mathtt{m} + \gamma^2 \mathtt{L}^2/2) \|x - y\| / R_1 + 1 \\
&\leq (1 - \gamma \mathtt{m}_1^+/2 + \gamma^2 \mathtt{L}^2/4) W(x, y) + \gamma \left[\mathtt{m}_1^+/2 - \mathtt{m} + \mathtt{m}_1^+/2 \right] \\
&\leq \lambda^\gamma W(x, y) + A\gamma .
\end{aligned} \tag{24}
$$

We conclude the proof upon combining (23) and (24). □

Theorem 10. *Assume* **B**1(X) *for* X $\subset \mathbb{R}^d$, **B**2, **B**3(m) *for* $\mathtt{m} \in \mathbb{R}$ *and* **C**1*. Then the conclusions of Theorem 7 hold with $\bar{\gamma}$, λ and A given by Proposition 9 and for any $\gamma \in (0, \bar{\gamma}]$, $\ell \in \mathbb{N}^\star$ and $t > 0$,*

$$\boldsymbol{\Psi}(\gamma, \ell, t) = 2\boldsymbol{\Phi}\{-t/(2\Xi^{1/2}_{\ell\lceil 1/\gamma \rceil}(\kappa))\} , \tag{25}$$

where κ is given in Proposition 8-(a) and $\Xi_{\ell\lceil 1/\gamma \rceil}$ in (8).

Proof. Let $\bar{\gamma} \in (0, 2\mathtt{m}_1^+/\mathtt{L}^2)$. Using Proposition 9 we obtain that W given by (19) satisfies $\mathbf{D}_{\mathrm{d}}(W, \lambda^\gamma, A\gamma, \Delta_{R_1})$ for any $\gamma \in (0, \bar{\gamma}]$ with λ and A given in (20). Using Theorem 1 and Lemma 2, we have for any $\gamma \in (0, \bar{\gamma}]$, $\ell \in \mathbb{N}^\star$ and $x, y \in X$

$$\mathrm{K}_\gamma((x, y), \Delta^{\mathrm{c}}_X) \leq 1 - 2\boldsymbol{\Phi}(-\Xi^{-1/2}_{\ell\lceil 1/\gamma \rceil}(\kappa) \|x - y\| / 2) ,$$

where $\kappa(\gamma) = -2\mathtt{m} + \gamma \mathtt{L}^2$, which concludes the proof. □

Let $\bar{\gamma} \in (0, \max(2\mathtt{m}_1^+/\mathtt{L}^2, 1))$, $\lambda_{\bar{\gamma},a}, \rho_{\bar{\gamma},a} \in (0, 1)$ and $D_{\bar{\gamma},1,a}, D_{\bar{\gamma},2,a}, C_{\bar{\gamma},a} \geq 0$ the constants given by Theorem 10 such that for any $k \in \mathbb{N}$, $\gamma \in (0, \bar{\gamma}]$ and $x, y \in X$

$$\mathbf{d}_W(\delta_x \mathrm{R}^k_\gamma, \delta_y \mathrm{R}^k_\gamma) \leq \lambda^{k\gamma/2}_{\bar{\gamma},a}[D_{\bar{\gamma},1,a} W(x, y) + D_{\bar{\gamma},2,a}] + C_{\bar{\gamma},a} \rho^{k\gamma/2}_{\bar{\gamma},a} .$$

If $\lim_{\bar\gamma \to 0} \ell$ exists and does not depend on L, using Theorem 7-(b), we obtain that the following limits exist and do not depend on L

$$
\begin{cases}
D_{1,a} = \lim_{\bar\gamma \to 0} D_{\bar\gamma,1,a}\ , & D_{2,a} = \lim_{\bar\gamma \to 0} D_{\bar\gamma,2,a}\ , & C_a = \lim_{\bar\gamma \to 0} C_{\bar\gamma,a}\ , \\
\lambda_a = \lim_{\bar\gamma \to 0} \lambda_{\bar\gamma,a}\ , & \rho_a = \lim_{\bar\gamma \to 0} \rho_{\bar\gamma,a}\ .
\end{cases}
\tag{26}
$$

If $\mathbf{B}3(\mathtt{m})$ is replaced by $\mathbf{B}4$ in Theorem 10 then the conclusion still holds and (25) can be replaced by the following expression for any $\gamma \in (0,\bar\gamma]$, $\ell \in \mathbb{N}^\star$ and $t > 0$,

$$
\boldsymbol{\Psi}(\gamma, \ell, t) = \begin{cases}
2\boldsymbol{\Phi}\{-1/2\} & \text{if } \ell \geq \lceil R_1 \rceil^2 \text{ and } t \leq R_1\ ; \\
2\boldsymbol{\Phi}\{-t/(2\Xi_{\ell\lceil 1/\gamma\rceil}^{1/2}(\kappa))\} & \text{otherwise .}
\end{cases}
$$

We now give upper bounds on the rate $\rho_{\bar\gamma,a}$ and ρ_a using Theorem 7-(c) depending on the assumptions in Theorem 10

(a) if $\mathbf{B}4$ holds, we set $\ell = \lceil R_1^2 \rceil$. Using that $2\boldsymbol{\Phi}(-1/2) \leq 1 - \mathrm{e}^{-1}$ and choosing $\mathtt{m}_1^+$ sufficiently small such that the conditions of Theorem 7-(c) hold, we have

$$
\begin{aligned}
\log(\rho_{\bar\gamma,a}^{-1})^{-1} \leq \Big[1 + \log(2) + \log\Big(1 + 2(2 + R_1^2)\mathtt{m}_1^+ \Big) \\
+ 2(2 + R_1^2)(\mathtt{m}_1^+ - \bar\gamma \mathtt{L}^2/2) \Big] \Big/ \Big[(\mathtt{m}_1^+ - \bar\gamma \mathtt{L}^2/2)\boldsymbol{\Phi}\{-1/2\} \Big]\ . \tag{27}
\end{aligned}
$$

Taking the limit $\bar\gamma \to 0$ in (27) and using that for any $t \geq 0$, $\log(1 + t) \leq t$, we get that

$$
\log(\rho_a^{-1})^{-1} \leq (1 + \log(2))/(\boldsymbol{\Phi}\{-1/2\}\mathtt{m}_1^+) + 4R_1^2/\boldsymbol{\Phi}\{-1/2\}\ .
$$

The leading term in (43) is of order $\max(R_1^2, 1/\mathtt{m}_1^+)$, which corresponds to the one identified in [27, Theorem 2.8] and is optimal, see [25, Remark 2.10].

(b) if $\mathbf{B}3(\mathtt{m})$ holds with $\mathtt{m} < 0$, we set $\ell = 1$ and choosing $\mathtt{m}_1^+ > 0$ sufficiently small and $R_1, |\mathtt{m}|$ sufficiently large such that the conditions of Theorem 7-(c) hold, we have

$$
\begin{aligned}
\log(\rho_{\bar\gamma,a}^{-1})^{-1} \leq \Big[1 + \log(2) + \log\Big(1 + 4\{\mathtt{m}_1^+ - \mathtt{m}\} \Big) \\
+ 3(\mathtt{m}_1^+ - \bar\gamma \mathtt{L}^2/2) \Big] \Big/ \Big[(\mathtt{m}_1^+ - \bar\gamma \mathtt{L}^2/2)\boldsymbol{\Phi}\{-\Xi_{\lceil 1/\bar\gamma\rceil}^{-1/2}(\kappa)R_1/2\} \Big]\ . \tag{28}
\end{aligned}
$$

Taking the limit $\bar\gamma \to 0$ in (28), we get that

$$
\begin{aligned}
\log(\rho_a^{-1})^{-1} \leq \Big[1 + \log(2) + \log(1 + 4\{\mathtt{m}_1^+ - \mathtt{m}\}) + 3\mathtt{m}_1^+ \Big] \\
\Big/ \Big[\mathtt{m}_1^+ \boldsymbol{\Phi}\{-(-\mathtt{m})^{1/2}R_1/(2 - 2\mathrm{e}^{2\mathtt{m}})^{1/2}\} \Big]\ . \tag{29}
\end{aligned}
$$

20

The leading term in $\log(\log(\rho_a)^{-1})$ is $\log(\mathbf{\Phi}\{-(-\mathtt{m})^{1/2}R_1/(2-2\mathrm{e}^{2\mathtt{m}})\})$ which is dominated by $\exp(-\mathtt{m}R_1^2/4(1-\mathrm{e}^{2\mathtt{m}}))$. This quantity is asymptotically in $\mathtt{m}$ equal to the one identified by [25] for the Kantorovitch-Rubinstein distance. Denoting ρ_E the rate obtained taking the limit $h \to 0$ in [27, Theorem 2.10, Equation (2.64)], we get that

$$\log(\log(\rho_E^{-1})^{-1}) = \psi(\mathtt{m}_1^+, R_1) \exp(-\tilde{c}_0^{-1}\mathtt{m}R_1^2) \, ,$$

with ψ a rational function in both variables. Since $\tilde{c}_0 \leq 0.00051$, the leading term in $\log(\log(\rho_E^{-1})^{-1})$ is lower bounded by $\exp(-1960\mathtt{m}R_1^2)$ and if $\mathtt{m} \geq 1$, the leading term in $\log(\log(\rho_a))$ is upper bounded by $\exp(-0.29\mathtt{m}R_1^2)$.

We now derive uniform ergodic convergence in V-norm under weaker conditions. The following assumption enforces that the radial part of b decreases faster than a linear function with slope $-\mathtt{m}_2^+ < 0$.

C2. *There exist $R_2 \geq 0$ and $\mathtt{m}_2^+ > 0$ such that for any $x \in \bar{\mathrm{B}}(0, R_2)^{\mathrm{c}} \cap \mathsf{X}$,*

$$\langle b(x), x \rangle \leq -\mathtt{m}_2^+ \|x\|^2 \, .$$

In the next proposition we derive a Foster-Lyapunov drift condition for $W : \mathsf{X}^2 \to [1, +\infty)$ defined for any $x, y \in \mathsf{X}$ by

$$W(x, y) = 1 + \|x\|^2 + \|y\|^2 \, .$$

Note that for any $x, y \in \mathsf{X}$, $W(x, y) = \{V(x) + V(y)\}/2$ with $V(x) = 1 + \|x\|^2$.

Proposition 11. *Assume* $\mathbf{B}1(\mathsf{X})$ *for* $\mathsf{X} \subset \mathbb{R}^d$, $\mathbf{B}2$, $\mathbf{B}3(\mathtt{m})$ *for* $\mathtt{m} \in \mathbb{R}$ *and* $\mathbf{C}2$. *Then* K_γ *defined by* (6) *satisfies* $\mathbf{D}_{\mathrm{d}}(W, \lambda^\gamma, A\gamma, \mathsf{X}^2)$ *for any* $\gamma \in (0, \bar{\gamma}]$ *where* $\bar{\gamma} \in (0, 2\mathtt{m}_2^+/\mathtt{L}^2)$ *and*

$$\lambda = \exp[-(2\mathtt{m}_2^+ - \bar{\gamma}\mathtt{L}^2)] \, , \quad A = d + 2(\mathtt{m}_2^+ - \mathtt{m})(1 + R_2^2) \, .$$

Proof. See Appendix F.3. $\qquad\qquad\square$

Theorem 12. *Assume* $\mathbf{B}1(\mathsf{X})$ *for* $\mathsf{X} \subset \mathbb{R}^d$, $\mathbf{B}2$, $\mathbf{B}3(\mathtt{m})$ *for* $\mathtt{m} \in \mathbb{R}$ *and* $\mathbf{C}2$. *Then the conclusions of Theorem 6 hold with* $\bar{\gamma}$, λ *and* A *given by Proposition 11,*

$$M_{\mathrm{d}} = 2\sqrt{2K_{\mathrm{d}}} \quad \text{with } K_{\mathrm{d}} = 2A(1 + \bar{\gamma})(1 + \log^{-1}(1/\lambda)) \, ,$$

and $\mathbf{\Psi}$ *is given in Theorem 10.*

Proof. The proof is postponed to Appendix F.4. $\qquad\qquad\square$

Let $\bar{\gamma} \in (0, 2\mathtt{m}_2^+/\mathtt{L}^2)$, $\tilde{\rho}_{\bar{\gamma},b} \in (0,1)$ and $\tilde{C}_{\bar{\gamma},b} \geq 0$ the constants given by Theorem 12 such that for any $\ell \in \mathbb{N}^\star$, $\gamma \in (0, \bar{\gamma}]$, $k \in \mathbb{N}$ and $x, y \in \mathsf{X}$

$$\mathbf{d}_W(\delta_x \mathrm{R}_\gamma^k, \delta_y \mathrm{R}_\gamma^k) \leq \tilde{C}_{\bar{\gamma},b} \tilde{\rho}_{\bar{\gamma},b}^{\left\lfloor k(\ell \lceil 1/\gamma \rceil)^{-1} \right\rfloor} W(x,y) \;.$$

Using that $W(x,y) = \{V(x) + V(y)\}/2$, the fact that $\lfloor k/(\ell \lceil 1/\gamma \rceil) \rfloor \geq k\gamma/(\ell(1 + \bar{\gamma})) - 1$, and letting $C_{\bar{\gamma},b} = \tilde{C}_{\bar{\gamma},b} \tilde{\rho}_{\bar{\gamma},b}^{-1}/2$ and $\rho_{\bar{\gamma},b} = \tilde{\rho}_{\bar{\gamma},b}^{1/(\ell(1+\bar{\gamma}))}$, we obtain that for any $\gamma \in (0, \bar{\gamma}]$, $k \in \mathbb{N}$ and $x, y \in \mathsf{X}$

$$\|\delta_x \mathrm{R}_\gamma^k - \delta_y \mathrm{R}_\gamma^k\|_V \leq C_{\bar{\gamma},b} \rho_{\bar{\gamma},b}^{k\gamma}\{V(x) + V(y)\} \;.$$

If $\lim_{\bar{\gamma} \to 0} \ell$ exists and does not depend on $\mathtt{L}$, using Theorem 6-(b), we obtain that the following limits exist and do not depend on $\mathtt{L}$

$$\rho_b = \lim_{\bar{\gamma} \to 0} \rho_{\bar{\gamma},b} \;, \qquad C_b = \lim_{\bar{\gamma} \to 0} C_{\bar{\gamma},b} \;. \tag{30}$$

We now discuss the dependency of ρ_b with respect to the problem constants, depending on the sign of $\mathtt{m}$, based on Theorem 6-(c). Consider $\mathtt{m}_2^+$ sufficiently small and $\mathtt{m}$ and R_2 sufficiently large such that the conditions of Theorem 6-(c) hold. Then,

(a) if $\mathbf{B}4$ holds, we set $\ell = \lceil M_\mathtt{d}^2 \rceil$ and we have

$$\log(\rho_b) \leq 3\log(2) \left(1 + 16\{d + 2\mathtt{m}_2^+(1 + R_2^2)\}\{1 + 1/(2\mathtt{m}_2^+)\}\right)$$
$$\times \log\left(6\{d + 2\mathtt{m}_2^+(1 + R_2^2)\}\{1 + 1/(2\mathtt{m}_2^+)\}\right) \Big/ \left[\mathtt{m}_2^+ \Phi(-1/2)\right] \;; \quad (31)$$

(b) if $\mathbf{B}3(\mathtt{m})$ with $\mathtt{m} < 0$, we set $\ell = 1$ and we have

$$\log(\rho_b) \leq 3\log(2)\log\left(6\{d + 2(\mathtt{m}_2^+ - \mathtt{m})(1 + R_2^2)\}\{1 + 1/(2\mathtt{m}_2^+)\}\right)$$
$$\Big/ \left[\mathtt{m}_2^+ \Phi\left\{-2\left[-\mathtt{m}\{d + 2(\mathtt{m}_2^+ - \mathtt{m})(1 + R_2^2)\}\{1 + 1/(2\mathtt{m}_2^+)\}/(1 - \mathrm{e}^{2\mathtt{m}})\right]^{1/2}\right\}\right] \;.$$
$$(32)$$

If $\mathbf{B}4$ holds, we obtain that the leading term in (31) is dominated by $\max(d, R_2^2/\mathtt{m}_2^+)$. If $\mathbf{B}3(\mathtt{m})$ holds with $\mathtt{m} < 0$, the leading term in $\log(\log(\rho_b))$ in (32) is exponential in $\mathtt{m}d$, $\mathtt{m}R_2^2$ and $\mathtt{m}/\mathtt{m}_2^+$. Therefore, $\log^{-1}(\rho_b^{-1})$ is linear in the dimension, respectively exponential in the dimension, if $\mathbf{B}4$ holds, respectively $\mathbf{B}3(\mathtt{m})$ holds with $\mathtt{m} < 0$. We now consider a condition which enforces weak curvature outside of a compact set.

C 3. *There exist $R_3, c \geq 0$, $\mathtt{m}_1, \mathtt{m}_2 > 0$, such that for any $x \in \mathbb{R}^d$*

$$\langle b(x), x \rangle \leq -\mathtt{m}_1 \|x\| \mathbb{1}_{\bar{\mathrm{B}}(0,R_3)^c}(x) - \mathtt{m}_2 \|b(x)\|^2 + c/2 \ .$$

Assume $\mathsf{X} = \mathbb{R}^d$ and $\Pi_{\mathsf{X}} = \mathrm{Id}$, if there exists $U \in \mathrm{C}^1(\mathbb{R}^d, \mathbb{R})$ such that **B** 2 and **B** 3(0) hold with $b = -\nabla U$ and $\int_{\mathbb{R}^d} \mathrm{e}^{-U(x)} \mathrm{d}x < +\infty$, then there exist $R_3, c \geq 0$ and $\mathtt{m}_1 > 0$ such that **C** 3 holds with $\mathtt{m}_2 = (4\mathrm{L})^{-1}$, see [13, Proposition 5]. Assuming **C** 3 we define $V : \mathsf{X} \to [1, +\infty)$ for any $x \in \mathsf{X}$ by

$$V(x) = \exp(\mathtt{m}_3 \phi(x)) \ , \qquad \phi(x) = \sqrt{1 + \|x\|^2} \ , \qquad \mathtt{m}_3 \in (0, \mathtt{m}_1/2] \ .$$

We also define for any $x, y \in \mathsf{X}$, $W(x, y) = \{V(x) + V(y)\}/2$.

Proposition 13. *Assume* **B** 1(X) *for* $\mathsf{X} \subset \mathbb{R}^d$ *and* **C** 3. *Then for any* $\gamma \in (0, \bar{\gamma}]$, K_γ *defined by* (6) *satisfies* $\mathbf{D}_{\mathrm{d}}(W, \lambda^\gamma, A\gamma, \mathsf{X}^2)$ *where* $\bar{\gamma} \in (0, 2\mathtt{m}_2)$, $R_4 = \max(1, R_3, (d + c)/\mathtt{m}_1)$ *and*

$$\lambda = e^{-\mathtt{m}_3^2/2} \ , \quad A = \exp\left[\bar{\gamma}(\mathtt{m}_3(d+c) + \mathtt{m}_3^2)/2 + \mathtt{m}_3(1 + R_4^2)^{1/2}\right](\mathtt{m}_3(d+c)/2 + \mathtt{m}_3^2) \ .$$

Proof. The proof is postponed to Appendix F.5. $\qquad\square$

Theorem 14. *Assume* **B** 1(X) *for* $\mathsf{X} \subset \mathbb{R}^d$, **B** 2, **B** 3($\mathtt{m}$) *for* $\mathtt{m} \in \mathbb{R}$ *and* **C** 3. *Then the conclusions of Theorem 6 hold with* $\bar{\gamma}$, λ *and* A *given by Proposition 13,*

$$K_{\mathrm{d}} = 2A(1 + \bar{\gamma})(1 + \log^{-1}(1/\lambda)) \quad \text{and } M_{\mathrm{d}} = 2\log(2K_{\mathrm{d}})/\mathtt{m}_3 \ ,$$

and $\mathbf{\Psi}$ *is given in Theorem 10.*

Proof. The proof is postponed to Appendix F.6. $\qquad\square$

The discussion on the dependency of the rate with respect to the problem constants is given in Appendix E.

3 Quantitative convergence bounds for diffusions

3.1 Main results

In this section, we aim at deriving quantitative convergence bounds with respect to some Wasserstein metrics for diffusion processes under regularity and curvature assumptions on the drift b. Consider the following Stochastic Differential Equation

$$\mathrm{d}\mathbf{X}_t = b(\mathbf{X}_t)\mathrm{d}t + \mathrm{d}\mathbf{B}_t \ , \tag{33}$$

where $(\mathbf{B}_s)_{s \geq 0}$ is a d-dimensional Brownian motion and $b : \mathbb{R}^d \to \mathbb{R}^d$ is a continuous drift. In the sequel we will always assume the following assumption.

L 1. *There exists a unique strong solution of* (33) *for any starting point* $\mathbf{X}_0 = x$.

drift b, which ensures us that the solution to (33) is unique. Consider the extended infinitesimal generator $\mathcal{A}$ defined for any $f \in \mathrm{C}^2(\mathbb{R}^d, \mathbb{R})$ by

$$\mathcal{A}f = (1/2)\Delta f + \langle \nabla f, b \rangle \ .$$

Let $V \in \mathrm{C}^2(\mathbb{R}^d, [1, +\infty))$, $\zeta \in \mathbb{R}$ and $B \geq 0$

$\mathbf{D}_\mathrm{c}(V, \zeta, B)$. *The extended infinitesimal generator $\mathcal{A}$ satisfies the continuous Foster-Lyapunov drift condition if for all $x \in \mathbb{R}^d$*

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

This assumption is the continuous counterpart of $\mathbf{D}_\mathrm{d}(V, \lambda, A, \mathbb{R}^d)$. We define the semi-group $(\mathrm{P}_t)_{t \geq 0}$ such that for any $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ and $t \geq 0$ by $\mathrm{P}_t(x, \mathsf{A}) = \mathbb{P}(\mathbf{X}_t \in \mathsf{A})$ where $(\mathbf{X}_t)_{t \geq 0}$ is the solution of (33) starting from $\mathbf{X}_0 = x \in \mathbb{R}^d$.

We now turn to establishing that $(\mathrm{P}_t)_{t \geq 0}$ converges for some Wasserstein metrics. In order to prove these results we will rely on discretizations of the Stochastic Differential Equation (33). If the hypotheses of Theorem 6 or Theorem 7 are satisfied, these discretized processes are uniformly geometrically ergodic.

First, we draw a link between the continuous drift condition $\mathbf{D}_\mathrm{c}(V, \zeta, B)$ and the discrete drift condition $\mathbf{D}_\mathrm{d}(V, \lambda, A, \mathbb{R}^d)$. The result and its proof are standard [48, Theorem 2.1] but are given here for completeness. Denote by $(\mathcal{F}_t)_{t \geq 0}$ the filtration associated to $(B_t)_{t \geq 0}$ satisfying the usual conditions.

Lemma 15. *Let $\zeta \in \mathbb{R}$, $B \geq 0$ and $V \in \mathrm{C}^2(\mathbb{R}^d, [1, +\infty))$ such that $\lim_{\|x\| \to +\infty} V(x) = +\infty$. Assume $\mathbf{L}$1 and $\mathbf{D}_\mathrm{c}(V, \zeta, B)$.*

> *(a) If $B = 0$, then for any $x \in \mathbb{R}^d$, $(V(\mathbf{X}_t)\mathrm{e}^{\zeta t})_{t \geq 0}$ is a $(\mathcal{F}_t)_{t \geq 0}$-supermartingale where $(\mathbf{X}_t)_{t \geq 0}$ is the solution of (33) starting from $\mathbf{X}_0 = x$.*

> *(b) For any $t_0 > 0$, P_{t_0} satisfies $\mathbf{D}_\mathrm{d}(V, \exp(-\zeta t_0), B(1 - \exp(-\zeta t_0))/\zeta, \mathbb{R}^f)$.*

Proof. (a) Let $x \in \mathbb{R}^d$ and let $(\mathbf{X}_t)_{t \geq 0}$ a solution of (33) starting from x. Define for any $k \in \mathbb{N}^\star$, $\tau_k = \inf\{t \geq 0 : \|\mathbf{X}_t\| \geq k\}$ and for any $t \geq 0$, $\mathbf{M}_t = \int_0^t \langle \nabla V(\mathbf{X}_s), \mathrm{d}B_s \rangle$. Using the Itô formula we obtain that for every $t \geq 0$ and $k \in \mathbb{N}^\star$

$$V(\mathbf{X}_{t \wedge \tau_k})\mathrm{e}^{\zeta(t \wedge \tau_k)} = \int_0^{t \wedge \tau_k} \left[\mathrm{e}^{\zeta(t \wedge \tau_k)}\mathcal{A}V(\mathbf{X}_u) + \zeta \mathrm{e}^{\zeta u}V(\mathbf{X}_u) \right] \mathrm{d}u + \mathbf{M}_{t \wedge \tau_k} + V(x)$$

$$= V(\mathbf{X}_{s \wedge \tau_k})\mathrm{e}^{\zeta(s \wedge \tau_k)} + \mathbf{M}_{t \wedge \tau_k} - \mathbf{M}_{s \wedge \tau_k} + \int_{s \wedge \tau_k}^{t \wedge \tau_k} \left[\mathrm{e}^{\zeta(t \wedge \tau_k)}\mathcal{A}V(\mathbf{X}_u) + \zeta \mathrm{e}^{\zeta u}V(\mathbf{X}_u) \right] \mathrm{d}u$$

$$\leq V(\mathbf{X}_{s \wedge \tau_k})\mathrm{e}^{\zeta(s \wedge \tau_k)} + \mathbf{M}_{t \wedge \tau_k} - \mathbf{M}_{s \wedge \tau_k} \ .$$

Therefore since for any $k \in \mathbb{N}^\star$, $(\mathbf{M}_{t \wedge \tau_k})_{t \geq 0}$ is a $(\mathcal{F}_t)_{t \geq 0}$-martingale, we get for every $t \geq s \geq 0$ and $k \in \mathbb{N}^\star$

$$\mathbb{E}\left[V(\mathbf{X}_{t \wedge \tau_k})\mathrm{e}^{\zeta(t \wedge \tau_k)} \mid \mathcal{F}_s\right] \leq V(\mathbf{X}_{s \wedge \tau_k})\mathrm{e}^{\zeta(s \wedge \tau_k)} \;,$$

which concludes the proof of (a) taking $k \to +\infty$ and using Fatou's lemma.

(b) Similarly as in (a) we have that $(V(\mathbf{X}_t)\mathrm{e}^{\zeta t} - B(1 - \exp(-\zeta t))/\zeta)_{t \geq 0}$ is a $(\mathcal{F}_t)_{t \geq 0}$-supermartingale which concludes the proof of (a) upon taking the expectation of $V(\mathbf{X}_t)\mathrm{e}^{\zeta t} - B(1 - \exp(-\zeta t))/\zeta$.

$\square$

We give in the following conditions under which a family of discretizations approximates the underlying continuous process. More specifically, we introduce a family of drifts $\left\{b_{\gamma,n} : \mathbb{R}^d \to \mathbb{R}^d : \gamma \in (0, \bar{\gamma}], n \in \mathbb{N}\right\}$ for some $\bar{\gamma} > 0$. For all $\gamma \in (0, \bar{\gamma}]$ and $n \in \mathbb{N}$, we denote by $\tilde{\mathrm{R}}_{\gamma,n}$ the Markov kernel associated with (1) where $\mathcal{T}_\gamma(x) = x + \gamma b_{\gamma,n}(x)$, $\mathsf{X} = \mathbb{R}^d$ and $\Pi = \mathrm{Id}$. We will show that under the following assumptions the family $\left\{\mathrm{R}_{\gamma,n}^{\lceil T/\gamma \rceil} : \mathbb{R}^d \to \mathbb{R}^d : \gamma \in (0, \bar{\gamma}], n \in \mathbb{N}\right\}$ approximates P_T for $T \geq 0$ as $\gamma \to 0$ and $n \to +\infty$.

L 2. *There exist $\beta > 0$ and $C_1 \geq 0$ such that for any $\gamma \in (0, \bar{\gamma}]$, $n \in \mathbb{N}$, $b_{\gamma,n} \in \mathrm{C}(\mathbb{R}^d, \mathbb{R}^d)$ and for any $x \in \mathbb{R}^d$,*

$$\|b(x) - b_{\gamma,n}(x)\|^2 \leq C_1 \gamma^\beta \|b(x)\|^2 \;.$$

The following assumption is mainly technical and is satisfied in our applications.

L 3. *There exists $\varepsilon_b > 0$ such that $\sup_{s \in [0,T]}\left\{\mathbb{E}\left[\|b(\mathbf{X}_s)\|^{2(1+\varepsilon_b)}\right]\right\} < +\infty$, for any $x \in \mathbb{R}^d$ and $T \geq 0$, where $(\mathbf{X}_t)_{t \geq 0}$ is the solution of (33) starting from $\mathbf{X}_0 = x$.*

By Lemma 15-(a), if $\mathbf{D}_\mathrm{c}(V, \zeta, 0)$ is satisfied with $\zeta \in \mathbb{R}$, for any $x \in \mathbb{R}^d$, it holds that $\sup_{t \in [0,T]} \mathbb{E}[V(\mathbf{X}_t)] \leq \mathrm{e}^{-\zeta T}V(x)$, where $(\mathbf{X}_t)_{t \geq 0}$ is solution of (33) starting from x. Therefore, if $\|b(x)\|^{2(1+\varepsilon_b)} \leq V(x)$ for any $x \in \mathbb{R}^d$, **L** 3 is satisfied.

Proposition 16. *Let $V : \mathbb{R}^d \to [1, +\infty)$. Assume **L** 1, **L** 2 and **L** 3. In addition, assume that for any $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$*

$$\mathrm{P}_T V^2(x) < +\infty \;, \qquad \limsup_{m \to +\infty} \tilde{\mathrm{R}}_{T/m,n}^m V^2(x) < +\infty \;. \tag{34}$$

Then for any $n \in \mathbb{N}$, $T \geq 0$ and $x \in \mathbb{R}^d$

$$\lim_{m \to +\infty} \|\delta_x \mathrm{P}_T - \delta_x \tilde{\mathrm{R}}_{T/m,n}^m\|_V = 0 \;,$$

where $(\mathrm{P}_t)_{t \geq 0}$ is the semigroup associated with (33) and for any $\gamma \in (0, \bar{\gamma}]$ and $n \in \mathbb{N}$, $\tilde{\mathrm{R}}_{\gamma,n}$ is the Markov kernel associated with (1) where $\mathcal{T}_\gamma(x) = x + \gamma b_{\gamma,n}(x)$ and $\Pi = \mathrm{Id}$.

Proof. Let $T \geq 0$, $x \in \mathbb{R}^d$, $n \in \mathbb{N}$ and $m \in \mathbb{N}^\star$ with $T/m \leq \bar{\gamma}$. Using [23, Lemma 24], we obtain

$$\|\delta_x \mathrm{P}_T - \delta_x \tilde{\mathrm{R}}^m_{T/m,n}\|_V \leq \sqrt{2} \left(\delta_x \mathrm{P}_T V^2(x) + \delta_x \tilde{\mathrm{R}}^m_{T/m,n} V^2(x) \right)^{1/2} \mathrm{KL} \left(\delta_x \mathrm{P}_T | \delta_x \tilde{\mathrm{R}}^m_{T/m,n} \right)^{1/2} .$$

Let $M \geq 0$, $n \in \mathbb{N}^\star$ with $n^{-1} < \bar{\gamma}$, $x \in \mathbb{R}^d$ and $k \in \mathbb{N}$. Therefore by (34), we only need to show that $\lim_{m \to +\infty} \mathrm{KL}(\delta_x \mathrm{P}_T | \delta_x \tilde{\mathrm{R}}^m_{T/m,n}) = 0$. Consider the two processes $(\mathbf{X}_t)_{t \in [0,T]}$ and $(\tilde{\mathbf{X}}_t)_{t \in [0,T]}$ defined by (33) with $\mathbf{X}_0 = \tilde{\mathbf{X}}_0 = x$ and

$$\mathrm{d}\tilde{\mathbf{X}}_t = \tilde{b}_{T/m,n}(t, (\tilde{\mathbf{X}}_s)_{s \in [0,T]}) \mathrm{d}t + \mathrm{d}\mathbf{B}_t , \qquad \tilde{\mathbf{X}}_0 = x ,$$

where for any $(\mathrm{w}_s)_{s \in [0,T]} \in \mathrm{C}([0,T], \mathbb{R}^d)$, $t \in [0,T]$,

$$\tilde{b}_{T/m,n}(t, (\mathrm{w}_s)_{s \in [0,T]}) = \sum_{i=0}^{m-1} b_{T/m,n}(\mathrm{w}_{iT/n}) \mathbb{1}_{[iT/m,(i+1)T/m)}(t) . \tag{35}$$

Note for any $i \in \{0, \ldots, m\}$, the distribution of $\tilde{\mathbf{X}}_{iT/m}$ is $\delta_x \tilde{\mathrm{R}}^i_{T/m,n}$. Using that b and $b_{T/m,n}$ are continuous and that $(\mathbf{X}_t)_{t \in [0,T]}$ and $(\tilde{\mathbf{X}}_t)_{t \in [0,T]}$ take their values in $\mathrm{C}([0,T], \mathbb{R}^d)$, we obtain that

$$\mathbb{P}\left(\int_0^T \|b(\mathbf{X}_t)\|^2 \mathrm{d}t < +\infty \right) = 1 , \qquad \mathbb{P}\left(\int_0^T \|\tilde{b}_{T/m,n}(t, (\tilde{\mathbf{X}}_s)_{s \in [0,T]})\|^2 \mathrm{d}t < +\infty \right) = 1 ,$$

and

$$\mathbb{P}\left(\int_0^T \|b(\mathbf{B}_t)\|^2 \mathrm{d}t < +\infty \right) = 1 , \qquad \mathbb{P}\left(\int_0^T \|\tilde{b}_{T/m,n}(t, (\mathbf{B}_s)_{s \in [0,T]})\|^2 \mathrm{d}t < +\infty \right) = 1 ,$$

where $(\mathbf{B}_t)_{t \in [0,T]}$ is the d-dimensional Brownian motion associated with (33). Therefore by [44, Theorem 7.7] the distributions of $(\mathbf{X}_t)_{t \in [0,T]}$ and $(\tilde{\mathbf{X}}_t)_{t \in [0,T]}$, denoted by μ^x and $\tilde{\mu}^x$ respectively, are equivalent to the distribution of the Brownian motion μ^x_B starting at x. In addition, μ^x admits a Radon-Nikodyn density w.r.t. to μ^x_B and μ^x_B admits a Radon-Nikodyn density w.r.t. to $\tilde{\mu}^x$, given μ^x_B-almost surely for any $(\mathrm{w}_t)_{t \in [0,T]} \in \mathrm{C}([0,T], \mathbb{R}^d)$ by

$$\frac{\mathrm{d}\mu^x}{\mathrm{d}\mu^x_B}((\mathrm{w}_t)_{t \in [0,T]}) = \exp\left((1/2) \int_0^T \langle b(\mathrm{w}_s), \mathrm{d}\mathrm{w}_s \rangle - (1/4) \int_0^T \|b(\mathrm{w}_s)\|^2 \mathrm{d}s \right) ,$$

$$\frac{\mathrm{d}\mu^x_B}{\mathrm{d}\tilde{\mu}^x}((\mathrm{w}_t)_{t \in [0,T]}) = \exp\left(-(1/2) \int_0^T \langle \tilde{b}_{T/m,n}(s, (\mathrm{w}_u)_{u \in [0,T]}), \mathrm{d}\mathrm{w}_s \rangle \right.$$

$$\left. + (1/4) \int_0^T \|\tilde{b}_{T/m,n}(s, (\mathrm{w}_u)_{u \in [0,T]})\|^2 \mathrm{d}s \right) .$$

Finally we obtain that μ_B^x-almost surely for any $(\mathrm{w}_s)_{s\in[0,T]} \in \mathrm{C}([0,T],\mathbb{R}^d)$

$$\frac{\mathrm{d}\mu^x}{\mathrm{d}\tilde{\mu}^x}((\mathrm{w}_t)_{t\in[0,T]}) = \exp\left((1/2)\int_0^T \langle b(\mathrm{w}_s) - \tilde{b}_{T/m,n}(s,(\mathrm{w}_u)_{u\in[0,T]}),\mathrm{d}\mathrm{w}_s\rangle\right.$$
$$\left. +(1/4)\int_0^T \|\tilde{b}_{T/m,n}(s,(\mathrm{w}_u)_{u\in[0,T]})\|^2 - \|b(\mathrm{w}_s)\|^2\mathrm{d}s\right) . \qquad (36)$$

Now define for any $(\mathrm{w}_s)_{s\in[0,T]} \in \mathrm{C}([0,T],\mathbb{R}^d)$ and $t \in [0,T]$

$$b_{T/m}(t,(\mathrm{w}_s)_{s\in[0,T]}) = \sum_{i=0}^{m-1} b(\mathrm{w}_{iT/m})\mathbb{1}_{[iT/m,(i+1)T/m)}(t) . \qquad (37)$$

Using (33), (35), (36), **L**2, and for any $a_1, a_1 \in \mathbb{R}^d$, $\|a_1 - a_2\|^2 \leq 2(\|a_1\|^2 + \|a_2\|^2)$, we obtain that

$$2\mathrm{KL}\left(\delta_x\mathrm{P}_T|\delta_x\tilde{\mathrm{R}}_{T/m,n}^m\right) \leq 2^{-1}\mathbb{E}\left[\int_0^T \|b(\mathbf{X}_s) - \tilde{b}_{T/m,n}(s,(\mathbf{X}_u)_{u\in[0,T]})\|^2\mathrm{d}s\right]$$
$$\leq \mathbb{E}\left[\int_0^T \|b(\mathbf{X}_s) - b_{T/m}(s,(\mathbf{X}_u)_{u\in[0,T]})\|^2\mathrm{d}s\right]$$
$$+ \sum_{i=0}^{m-1}\mathbb{E}\left[\int_{iT/m}^{(i+1)T/m} \|b(\mathbf{X}_{iT/n}) - b_{T/m,n}(\mathbf{X}_{iT/n})\|^2\mathrm{d}s\right]$$
$$\leq \mathbb{E}\left[\int_0^T \|b(\mathbf{X}_s) - b_{T/m}(s,(\mathbf{X}_u)_{u\in[0,T]})\|^2\mathrm{d}s\right] + C_1 T^{1+\beta}m^{-\beta}\sup_{s\in[0,T]}\mathbb{E}\left[\|b(\mathbf{X}_s)\|^2\right] . \qquad (38)$$

It only remains to show that the first term goes to 0 as $m \to +\infty$. Note that since $(\mathbf{X}_s)_{s\in[0,T]}$ is almost surely continuous, b is continuous on $\mathbb{R}^d$, $\lim_{m\to+\infty}\|b(\mathbf{X}_s) - b_{T/m}(s,(\mathbf{X}_u)_{u\in[0,T]})\|^2 = 0$ for any $s \in [0,T]$ almost surely. Then, using the Lebesgue dominated convergence theorem and the continuity of b, we obtain that for any $M \geq 0$,

$$\lim_{m\to+\infty}\mathbb{E}\left[\mathbb{1}_{[0,M]}\left(\sup_{s\in[0,T]}\|\mathbf{X}_s\|\right)\int_0^T \|b(\mathbf{X}_s) - b_{T/m}(s,(\mathbf{X}_u)_{u\in[0,T]})\|^2\mathrm{d}s\right] = 0 . \qquad (39)$$

On the other hand, using Hölder's inequality and the definition of $b_{T/m}$ (37), we

obtain that for any $M \geq 0$,

$$\mathbb{E}\left[\mathbb{1}_{(M,+\infty)}\left(\sup_{s\in[0,T]}\|\mathbf{X}_s\|\right)\int_0^T\|b(\mathbf{X}_s) - b_{T/m}(s,(\mathbf{X}_s)_{s\in[0,T]})\|^2\mathrm{d}s\right]$$

$$\leq 2\left(\mathbb{P}\left(\sup_{s\in[0,T]}\|\mathbf{X}_s\| > M\right)\right)^{\varepsilon_b/(1+\varepsilon_b)}$$

$$\int_0^T\left(\mathbb{E}^{1/(1+\varepsilon_b)}\left[\|b(\mathbf{X}_s)\|^{2(1+\varepsilon_b)}\right] + \mathbb{E}^{1/(1+\varepsilon_b)}\left[\left\|b_{T/m}(s,(\mathbf{X}_u)_{u\in[0,T]})\right\|^{2(1+\varepsilon_b)}\right]\right)\mathrm{d}s$$

$$\leq 4T\left(\mathbb{P}\left(\sup_{s\in[0,T]}\|\mathbf{X}_s\| > M\right)\right)^{\varepsilon_b/(1+\varepsilon_b)}\left(\sup_{s\in[0,T]}\mathbb{E}\left[\|b(\mathbf{X}_s)\|^{2(1+\varepsilon_b)}\right]\right)^{1/(1+\varepsilon_b)}.$$

Combining this result, $\mathbf{L}3$, and (39) in (38), we obtain that for any $M \geq 0$,

$$\limsup_{m\to+\infty}\mathrm{KL}\left(\delta_x\mathrm{P}_T|\delta_x\tilde{\mathrm{R}}_{T/m,n}^m\right)$$

$$\leq 2T\left(\mathbb{P}\left(\sup_{s\in[0,T]}\|\mathbf{X}_s\| > M\right)\right)^{\varepsilon_b/(1+\varepsilon_b)}\left(\sup_{s\in[0,T]}\mathbb{E}\left[\|b(\mathbf{X}_s)\|^{2(1+\varepsilon_b)}\right]\right)^{1/(1+\varepsilon_b)}. \quad (40)$$

Since $(\mathbf{X}_s)_{s\in[0,T]}$ is a.s. continuous, we get by the monotone convergence theorem and $\mathbf{L}3$, taking $M \to +\infty$, that $\lim_{m\to+\infty}\mathrm{KL}(\delta_x\mathrm{P}_T|\delta_x\tilde{\mathrm{R}}_{T/m,n}^m) = 0$, which concludes the proof. $\qquad\square$

Note that if $V = 1$, Proposition 16 implies that $\lim_{m\to+\infty}\|\delta_x\mathrm{P}_T - \delta_x\tilde{\mathrm{R}}_{T/m,n}^m\|_{\mathrm{TV}} = 0$. Let $V : \mathbb{R}^d \to [1,+\infty)$ and $W : \mathbb{R}^d \times \mathbb{R}^d \to [1,+\infty)$ such that for any $x, y \in \mathbb{R}^d$, $W(x,y) \leq \{V(x) + V(y)\}/2$. Then, under the conditions of Proposition 16, we obtain that for any $T \geq 0$, $n \in \mathbb{N}$ and $x, y \in \mathbb{R}^d$

$$\mathbf{d}_W(\delta_x\mathrm{P}_T, \delta_y\mathrm{P}_T) \leq \lim_{m\to+\infty}\mathbf{d}_W(\delta_x\tilde{\mathrm{R}}_{T/m,n}^m, \delta_y\tilde{\mathrm{R}}_{T/m,n}^m),$$

Therefore, if for any $T \geq 0$, $\mathbf{d}_W(\delta_x\tilde{\mathrm{R}}_{T/m,n}^m, \delta_y\tilde{\mathrm{R}}_{T/m,n}^m)$ can be bounded using Theorem 6 or Theorem 7 uniformly in m, we obtain an explicit bound for $\mathbf{d}_W(\delta_x\mathrm{P}_T, \delta_y\mathrm{P}_T)$ for any $T \geq 0$. Then, this result easily implies non-asymptotic convergence bounds of $(\mathrm{P}_t)_{t\geq 0}$ to its invariant measure if it exists. However, in our applications, global Lipschitz regularity on $b_{T/m,n} : \mathbb{R}^d \to \mathbb{R}^d$ is needed in order to apply Theorem 6 or Theorem 7 to $\tilde{\mathrm{R}}_{T/m,n}$ for $T \geq 0$, $m \in \mathbb{N}^\star$ and $n \in \mathbb{N}$. To be able to deal with the fact that $b_{T/m,n}$ is non necessarily globally Lipschitz, we consider an appropriate sequence of projected Euler Maruyama schemes associated to a sequence of subset of $\mathbb{R}^d$, $(\mathsf{K}_n)_{n\in\mathbb{N}}$ satisfying the following assumption.

L 4. *For any $n \in \mathbb{N}$, K_n is convex and closed, and $\bar{\mathrm{B}}(0,n) \subset \mathsf{K}_n$.*

Consider for any $\gamma \in (0, \bar\gamma]$ and $n \in \mathbb{N}$ the Markov chain associated (1), where for any $x \in \mathbb{R}^d$, $\mathcal{T}_\gamma(x) = x + \gamma b_{\gamma,n}(x)$, $\mathsf{X} = \mathsf{K}_n$ and $\Pi = \Pi_{\mathsf{K}_n}$, the projection on K_n. The Markov kernel associated with this chain is denoted $\mathrm{R}_{\gamma,n}$ for any $\gamma \in (0, \bar\gamma]$ and $n \in \mathbb{N}$. Assuming only local Lipschitz regularity we can apply these theorems to the projected version of the Markov chain associated with $\mathrm{R}_{T/m,n}$. Therefore we want to replace $\tilde{\mathrm{R}}_{T/m,n}$ by $\mathrm{R}_{T/m,n}$ in (3.1). In order to do so we consider the following assumption on the family of drifts $\{b_{\gamma,n} \; ; \; \gamma \in (0, \bar\gamma], n \in \mathbb{N}\}$.

L 5. *There exist $\tilde{A} > 0$ and $\tilde{V} : \mathbb{R}^d \to [1, +\infty)$ such that for any $n \in \mathbb{N}$ there exist $E_n \geq 0$, $\varepsilon_n > 0$ and $\bar\gamma_n \in (0, \bar\gamma]$ satisfying for any $\gamma \in (0, \bar\gamma_n]$ and $x \in \mathbb{R}^d$,*

$$\tilde{\mathrm{R}}_{\gamma,n}\tilde{V}(x) \leq \exp\left[\log(\tilde{A})\gamma(1 + E_n\gamma^{\varepsilon_n})\right]\tilde{V}(x) \;, \qquad \sup_{x \in \mathbb{R}^d} \tilde{V}(x)^{-1}\|x\| \leq 1 \;,$$

where for any $\gamma \in (0, \bar\gamma]$ and $n \in \mathbb{N}$, $\tilde{\mathrm{R}}_{\gamma,n}$ is the Markov kernel associated with (1) where $\mathcal{T}_\gamma(x) = x + \gamma b_{\gamma,n}(x)$ and $\Pi = \mathrm{Id}$.

Proposition 17. *Let $V : \mathbb{R}^d \to [1, +\infty)$. Assume **L 1**, **L 4**, **L 5** and that for any $T \geq 0$, $x \in \mathbb{R}^d$*

$$\limsup_{n \to +\infty} \limsup_{m \to +\infty} \left(\mathrm{R}_{T/m,n}^m + \tilde{\mathrm{R}}_{T/m,n}^m\right) V^2(x) < +\infty \;.$$

Then for any $T \geq 0$ and $x \in \mathbb{R}^d$

$$\lim_{n \to +\infty} \limsup_{m \to +\infty} \|\delta_x \mathrm{R}_{T/m,n}^{mk} - \delta_x \tilde{\mathrm{R}}_{T/m,n}^{mk}\|_V = 0 \;,$$

Proof. For any $n \in \mathbb{N}$ and $\gamma \in (0, \bar\gamma]$, we consider the synchronous Markov coupling $\mathrm{Q}_{\gamma,n}$ for $\mathrm{R}_{\gamma,n}$ and $\tilde{\mathrm{R}}_{\gamma,n}$ defined for any $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ and $\mathsf{A} \in \mathcal{B}(\mathbb{R}^d)$ by

$$\mathrm{Q}_{\gamma,n}((x,y),\mathsf{A}) = (2\pi\gamma)^{-d/2}\int_{\mathbb{R}^d} \mathbb{1}_{(\mathrm{Id},\Pi_{\mathsf{K}_n})\leftarrow(\mathsf{A})}(\mathcal{T}_\gamma(x) + \sqrt{\gamma}z, \mathcal{T}_\gamma(y) + \sqrt{\gamma}z)\mathrm{e}^{-\|z\|^2/2}\mathrm{d}\tilde{x} \;. \tag{41}$$

Let $T \geq 0$, $n \in \mathbb{N}$, $m \in \mathbb{N}^\star$ such that $T/m \leq \bar\gamma$ and define $W : \mathbb{R}^{2d} \to \mathbb{R}$ by $W(x,y) = \mathbb{1}_{\Delta_{\mathbb{R}^d}^c}(y,z)\{V(y) + V(z)\}/2$ for any $y,z \in \mathbb{R}^d$. Consider $(X_j, \tilde{X}_j)_{j \in \mathbb{N}}$ a Markov chain with Markov kernel $\mathrm{Q}_{T/m,n}$ and started from $X_0 = \tilde{X}_0 = x$ for a fixed $x \in \mathbb{R}^d$. Note that by definition and **L 4**, we have that for $k < \tau$, $X_k = \tilde{X}_k$ where $\tau = \inf\{j \in \mathbb{N} : \tilde{X}_j \notin \bar{\mathrm{B}}(0,n)\}$. and that using **L 5**, we have that $\left(\tilde{V}(\tilde{X}_j)\exp\left[-j\log(\tilde{A})(T/m)(1 + E_n(T/m)^{\varepsilon_n})\right]\right)_{j \in \mathbb{N}}$ is a supermartingale. Using (41),

the Cauchy-Schwarz inequality, **L**5 and the Doob maximal inequality, we get for any $x \in \mathbb{R}^d$

$$\|\delta_x \mathrm{R}^m_{T/m,n} - \delta_x \tilde{\mathrm{R}}^m_{T/m,n}\|_V \leq \mathbb{E}\left[\mathbb{1}_{\Delta^{\mathrm{c}}_{\mathbb{R}^d}}((X_m, \tilde{X}_m))(V(X_m) + V(\tilde{X}_m))/2\right]$$

$$\leq 2^{-1}\mathbb{P}\left(\sup\left\{\left\|\tilde{X}_i\right\| : j \in \{0,\ldots,m\}\right\} \geq n\right)\left(\mathbb{E}\left[V^2(X_m)\right]^{1/2} + \mathbb{E}\left[V^2(\tilde{X}_m)\right]^{1/2}\right)$$

$$\leq 2^{-1}\mathbb{P}\left(\sup\left\{\tilde{V}(\tilde{X}_j) : j \in \{0,\ldots,m\}\right\} \geq n\right)\left(\mathbb{E}\left[V^2(X_m)\right]^{1/2} + \mathbb{E}\left[V^2(\tilde{X}_m)\right]^{1/2}\right)$$

$$\leq (2n)^{-1}\exp\left[\log(\tilde{A})T(1 + E_n T^{\varepsilon_n} m^{-\varepsilon_n})\right]\tilde{V}(x)\left(\mathrm{R}^m_{T/m,n}V^2(x) + \tilde{\mathrm{R}}^m_{T/m,n}V^2(x)\right),$$

which concludes the proof upon taking $m \to +\infty$ then $n \to +\infty$. $\qquad \square$

Based on Proposition 16 and Proposition 17, we have the following result which establishes a clear link between the convergence of the family of the projected Euler-Maruyama scheme $\{\mathrm{R}_{\gamma,n} : \gamma \in (0, \bar{\gamma}], n \in \mathbb{N}\}$ and the semigroup $(\mathrm{P}_t)_{t \geq 0}$ associated with (33).

Theorem 18. *Let $W : \mathbb{R}^d \times \mathbb{R}^d \to [1, +\infty)$ and $V : \mathbb{R}^d \to [1, +\infty)$ satisfying for any $x, y \in \mathbb{R}^d$, $W(x,y) \leq \{V(x) + V(y)\}/2$. Assume* **L**1*,* **L**2*,* **L**3*,* **L**4 *and* **L**5*. In addition, assume that for any $T \geq 0$ and $x \in \mathbb{R}^d$*

$$\mathrm{P}_T V^2(x) < +\infty, \quad \limsup_{n \to +\infty} \limsup_{m \to +\infty} \left(\mathrm{R}^m_{T/m,n} + \tilde{\mathrm{R}}^m_{T/m,n}\right) V^2(x) < +\infty.$$

Then,
$$\mathbf{d}_W(\delta_x \mathrm{P}_T, \delta_y \mathrm{P}_T) \leq \limsup_{n \to +\infty} \limsup_{m \to +\infty} \mathbf{d}_W(\delta_x \mathrm{R}^m_{T/m,n}, \delta_y \mathrm{R}^m_{T/m,n}),$$

where $(\mathrm{P}_t)_{t \geq 0}$ is the semigroup associated with (33) and for any $\gamma \in (0, \bar{\gamma}]$, $n \in \mathbb{N}$, $\mathrm{R}_{\gamma,n}$ is the Markov kernel associated with (1) where $\mathcal{T}_\gamma(x) = x + \gamma b_{\gamma,n}(x)$, $\mathsf{X} = \mathsf{K}_n$ and $\Pi = \Pi_{\mathsf{K}_n}$, $\tilde{\mathrm{R}}_{\gamma,n}$ is the Markov kernel associated with (1) where $\mathcal{T}_\gamma(x) = x + \gamma b_{\gamma,n}(x)$, $\mathsf{X} = \mathbb{R}^d$ and $\Pi = \mathrm{Id}$.

Proof. Let $T \geq 0$ and $x, y \in \mathbb{R}^d$. We have the following inequality for any $n \in \mathbb{N}$ and $m \in \mathbb{N}^\star$ such that $x, y \in \mathsf{K}_n$ and $T/m \leq \bar{\gamma}$

$$\mathbf{d}_W(\delta_x \mathrm{P}_T, \delta_y \mathrm{P}_T) \leq \|\delta_x \mathrm{P}_T - \delta_x \tilde{\mathrm{R}}^m_{T/m,n}\|_V + \|\delta_x \mathrm{R}^m_{T/m,n} - \delta_x \tilde{\mathrm{R}}^m_{T/m,n}\|_V$$
$$+ \mathbf{d}_W(\delta_x \mathrm{R}^m_{T/m,n}, \delta_y \mathrm{R}^m_{T/m,n}) + \|\delta_y \mathrm{P}_T - \delta_y \tilde{\mathrm{R}}^m_{T/m,n}\|_V + \|\delta_y \mathrm{R}^m_{T/m,n} - \delta_y \tilde{\mathrm{R}}^m_{T/m,n}\|_V,$$

which concludes the proof upon using Proposition 16 and Proposition 17. $\qquad \square$

3.2 Applications

In this section, we combine the results of Theorem 18 with the convergence bounds for discrete processes derived in Section 2.2, in order to obtain convergence bounds for continuous processes. Assuming the following condition we are ensured that the solution of (33) exists locally.

B 5. *b is locally Lipschitz, i.e. for any $M \geq 0$, there exists $\mathtt{L}_M \geq 0$ such that for any $x, y \in \bar{\mathrm{B}}(0, M)$, $\|b(x) - b(y)\| \leq \mathtt{L}_M \|x - y\|$ and $b(0) = 0$.*

If **B** 5 holds, by [36, Chapter 4, Theorem 2.3], (33) admits a unique solution $(\mathbf{X}_t)_{t \geq 0}$ with $\mathbf{X}_0 = x \in \mathbb{R}^d$ such that for any $t < e = \inf \{t \geq 0 : \|\mathbf{X}_t\| = +\infty\}$ with $\|\mathbf{X}_t\| < +\infty$. In particular, the condition $e = +\infty$ is met a.s. if we assume that b is sub-linear [36, Chapter 4, Theorem 2.3] or the condition $\mathbf{D}_c(V, \zeta, 0)$ holds with $\zeta \in \mathbb{R}$ and $\lim_{\|x\| \to +\infty} V(x) = +\infty$ [42, Theorem 3.5].

Proposition 19. *Assume **B**5 and **B**3($\mathtt{m}$) for $\mathtt{m} \in \mathbb{R}$, then **L**1 holds. In addition:*

*(a) if there exists $\varepsilon_b > 0$ and $p \in \mathbb{N}^\star$ such that $\sup_{x \in \mathbb{R}^d} \|b(x)\|^{2(1+\varepsilon)} (1 + \|x\|^{2p})^{-1} < +\infty$ then **L**3 holds ;*

*(b) if **C**2 holds and there exists $\varepsilon_b > 0$ such that $\sup_{x \in \mathbb{R}^d} \|b(x)\|^{2(1+\varepsilon)} \mathrm{e}^{-\mathtt{m}_2^+ \|x\|^2} < +\infty$ then **L**3 holds.*

Proof. The proof is postponed to Appendix F.7. $\qquad \square$

We now apply Theorem 18 to the diffusion process defined by (33). We prove in Section 2.2 that some associated Markov chain: the projected Euler-Maruyama discretization converges in V-norm and in Wasserstein distance $\mathbf{d}_W$ with explicit bounds. Therefore, it only remains to 1) find a family of drifts which satisfies **L**2, **L**3, **L**4 and **L**5, 2) take the limit when $m \to +\infty$, *i.e.* when the discretization step goes to zero, in the bounds found in Section 2.2. In what follows, we consider the following families of drift $b_{\gamma,n}$ defined for any $\gamma > 0$, $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$ by

$$b_{\gamma,n}(x) = \varphi_n(x)b(x) + (1 - \varphi_n(x))\frac{b(x)}{1 + \gamma^\alpha \|b(x)\|} , \tag{42}$$

with $\alpha < 1/2$ and $\varphi_n \in \mathrm{C}(\mathbb{R}^d, \mathbb{R})$ such that for any $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$, $\varphi_n(x) \in [0, 1]$, for any $x \in \bar{\mathrm{B}}(0, n)$, $\varphi_n(x) = 1$ and for any $x \in \bar{\mathrm{B}}(0, n+1)^c$, $\varphi_n(x) = 0$. An example of such a family is displayed in Figure 1. We now give some properties on $\{b_{\gamma,n} \in \mathrm{C}(\mathbb{R}^d, \mathbb{R}^d) : \gamma > 0, n \in \mathbb{N}\}$.
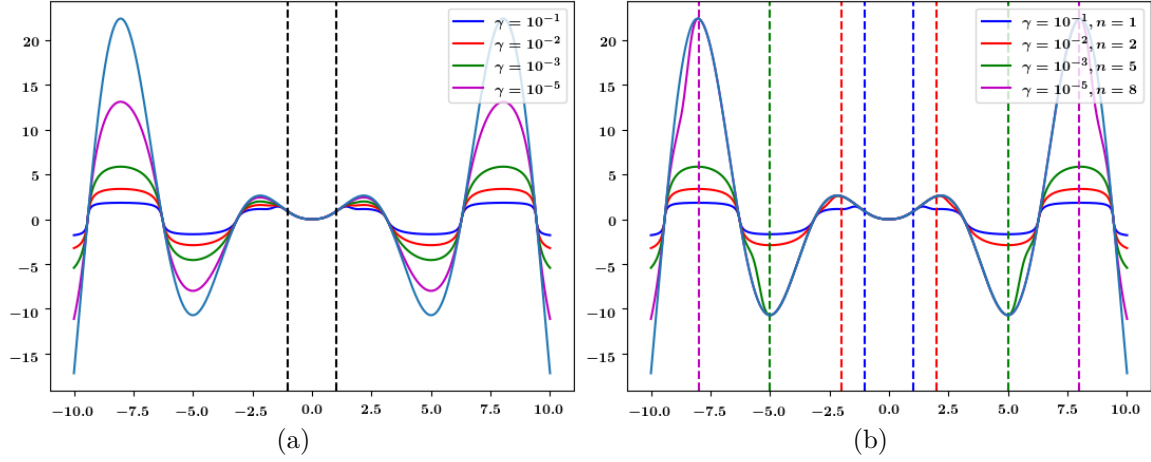
Figure 1: In this figure we illustrate the approximation properties of the family of drifts defined by (42). Let $b(x) = |x|^{1.5}\sin(x)$ and for any $n \in \mathbb{N}$, $\varphi_n(x) = d(x, \bar{\mathrm{B}}(0, +1)^c)^2/(d(x, \bar{\mathrm{B}}(0, n))^2 + d(x, \bar{\mathrm{B}}(0, n+1)^c)^2)$. In both figures the original drift is displayed in cyan and we fix $\alpha = 0.3$. In (a), we fix $n = 1$, represented by the black dashed lines, and observe the behavior of the drifts for different values of $\gamma > 0$. In (b), we plot the drift for different $\gamma > 0$ and $n \in \mathbb{N}$.

Proposition 20. *Assume* **B**5 *and* **B**3(m) *for* $\mathtt{m} \in \mathbb{R}$, *then*

(a) **L**2, **L**4 *and* **L**5 *hold for the family* $\{b_{\gamma,n} \in \mathrm{C}(\mathbb{R}^d, \mathbb{R}^d) : \gamma \in (0, \bar{\gamma}], n \in \mathbb{N}\}$ *defined by* (42) ;

(b) *if in addition* **C**2 *holds, then for any* $T \geq 0$ *and* $x \in \mathbb{R}^d$

$$\limsup_{n \to +\infty} \limsup_{m \to +\infty} \left(\mathrm{R}_{T/m,n}^m + \tilde{\mathrm{R}}_{T/m,n}^m\right) V(x) < +\infty ,$$

with $V(x) = \exp(\phi(x))$ *and* $\phi(x) = \sqrt{1 + \|x\|^2}$.

Proof. The proof is postponed to Appendix F.8. $\qquad\square$

We also show that under general conditions on b, $\mathrm{P}_T V(x) < +\infty$ which is a necessary conditions in Theorem 18

Proposition 21. *Assume* **L**1 *and that* $\sup_{x \in \mathbb{R}^d}\langle b(x), x\rangle < +\infty$. *Then for any* $M \geq 0$, *there exists* $\zeta \in \mathbb{R}$ *such that* $\mathbf{D}_c(V, \zeta, 0)$ *holds with* $V(x) = \exp[M\phi(x)]$ *and* $\phi(x) = (1 + \|x\|^2)^{1/2}$. *In particular, for any* $T \geq 0$, $\mathrm{P}_T V(x) < +\infty$.

32

Proof. The proof is postponed to Appendix F.9. □

We now have all the tools to apply Theorem 18. As in Section 2.2, we consider three different curvature assumptions on the drift b and derive convergence of the associated continuous process.

Proposition 22. *Assume* **B**5, **B**3($\mathtt{m}$) *for* $\mathtt{m} \in \mathbb{R}$, **C**1 *and that there exists* $\varepsilon_b > 0$ *such that* $\sup_{x \in \mathbb{R}^d} \|b(x)\|^{2(1+\varepsilon_b)} \mathrm{e}^{-\mathtt{m}_1^+ \|x\|^2} < +\infty$. *Then, for any* $T \geq 0$, *and* $x, y \in \mathbb{R}^d$

$$\mathbf{d}_W(\delta_x \mathrm{P}_T, \delta_y \mathrm{P}_T) \leq D_{1,a} \lambda_a^{T/2}(W(x,y) + D_{2,a}) + C_a \rho^{T/2} \,,$$

with $D_{1,a}, D_{2,a}, C_a \geq 0$ *and* $\lambda_a, \rho_a \in (0,1)$ *given by* (26) *and for any* $x, y \in \mathbb{R}^d$, $W(x,y) = 1 + \|x - y\|$.

Proof. Let $T \geq 0$. Using **B**5, that **C**1 implies **C**2 and $\sup_{x \in \mathbb{R}^d} \|b(x)\| \mathrm{e}^{-\mathtt{m}^+ \|x\|} < +\infty$, we obtain that **L**1 and **L**3 hold using Proposition 19. Applying Proposition 20 we get that **L**2, **L**4, **L**5 hold and

$$\limsup_{n \to +\infty} \limsup_{m \to +\infty} \left(\mathrm{R}_{T/m,n}^m + \tilde{\mathrm{R}}_{T/m,n}^m \right) V^2(x) < +\infty \,,$$

with $V(x) = C_V \exp(\phi(x))$, $\phi(x) = (1 + \|x\|^2)^{1/2}$ and $C_V = \sup_{x \in \mathbb{R}^d}(1 + \|x\|) \exp(-\phi(x))$. Using Proposition 21 we obtain that $\mathrm{P}_T V^2(x) < +\infty$. It is clear that for any $x, y \in \mathbb{R}^d$, $W(x,y) \leq V(x) + V(y)$, therefore we can apply Theorem 18 and we obtain that for any $x, y \in \mathbb{R}^d$

$$\mathbf{d}_W(\delta_x \mathrm{P}_T, \delta_y \mathrm{P}_T) \leq \limsup_{n \to +\infty} \limsup_{m \to +\infty} \mathbf{d}_W(\delta_x \mathrm{R}_{T/m,n}^m, \delta_y \mathrm{R}_{T/m,n}^m) \,.$$

Let $n \in \mathbb{N}$ and $m \in \mathbb{N}^\star$ such that $x, y \in \bar{\mathrm{B}}(0,n)$ and $T/\mathtt{m} \leq 2\mathtt{m}_1^+ / \mathrm{L}_n^2$. Since **B**1($\bar{\mathrm{B}}(0,n)$) holds and **B**5 implies **B**2 on $\bar{\mathrm{B}}(0,n)$, we can apply Theorem 10 and we get

$$\mathbf{d}_W(\delta_x \mathrm{R}_{T/m,n}^m, \delta_y \mathrm{R}_{T/m,n}^m) \leq D_{T/m,1} \lambda_{T/m,n}^{T/2}(W(x,y) + D_{T/m,2}) + C_{T/m} \rho_{T/m}^{T/2} \,,$$

where $D_{T/m,1}, D_{T/m,2}, C_{T/m}, \lambda_{T/m,n}$ and $\rho_{T/m,n}$ are given in Theorem 6. In addition, these quantities admit limits $D_{1,a}, D_{2,a}, C_a \geq 0$ and $\lambda_a, \rho_a \in (0,1)$ when $m \to +\infty$ which do not depend on L_n, hence on n. □

The discussion on the dependency of ρ_a with respect to the parameters of the problem conducted in Section 2.2 still holds. We distinguish the following cases, assuming that the conditions of Theorem 7-(c) hold:

33

(a) if **B**4 holds, we have

$$\log(\rho_a^{-1})^{-1} \leq (1 + \log(2))/(\Phi\{-1/2\}\mathsf{m}_1^+) + 4R_1^2/\Phi\{-1/2\} . \qquad (43)$$

The leading term in (43) is of order $\max(R_1^2, 1/\mathsf{m}_1^+)$, which corresponds to the one identified in [27, Theorem 2.8] and is optimal, see [25, Remark 2.10].

(b) if **B**3(m) holds with $\mathsf{m} < 0$, we have

$$\log(\rho_a^{-1})^{-1} \leq \left[1 + \log(2) + \log(1 + 4\{\mathsf{m}_1^+ - \mathsf{m}\}) + 3\mathsf{m}_1^+ \right]$$
$$\Big/ \left[\mathsf{m}_1^+ \Phi\{-(-\mathsf{m})^{1/2} R_1/(2 - 2\mathrm{e}^{2\mathsf{m}})^{1/2}\} \right] .$$

Since for any $t \geq C$ with $C \geq 0$ we have

$$\Phi(-t)^{-1} \leq \sqrt{2\pi}(1 + C^{-2})t\mathrm{e}^{t^2/2} . \qquad (44)$$

As a consequence if we also have $R_1 \geq 2$, $\mathsf{m}_1^+ \leq 1 \leq -\mathsf{m}$ and using that for any $t \in (0,1)$, $-\log(1 - t) \leq t$ as well as (44) we get that $\log(\rho_a^{-1})^{-1} \leq \log(\rho_{\max}^{-1})^{-1}$

$$\log^{-1}(\rho_{\max}^{-1}) = \left[1 + \log(1 + 4\{\mathsf{m}_1^+ - \mathsf{m}\}) + 3\mathsf{m}_1^+ \right] R_1(-\mathsf{m})^{1/2}$$
$$\times \exp\left[-\mathsf{m}R_1/(1 - \mathrm{e}^{2\mathsf{m}}) \right] \Big/ \left[\mathsf{m}_1^+ (1 - \mathrm{e}^{2\mathsf{m}})^{1/2} \right] ,$$

with $C = 2(1 + \log(2))\sqrt{\pi} \approx 6.00$. [25, Equation (2.18)] identifies a rate which we denote ρ_E. We get that

$$\lim_{\mathsf{m} \to -\infty} \frac{\log(\log(\rho_{\max}^{-1}))}{\log(\log(\rho_E^{-1}))} = 1 .$$

Note that the rate identified in [25, Lemma 2.9, Equation (2.18)] ensures a *contraction* of P_t for the Wasserstein distance of order 1 whereas our rate ensures the *convergence* of P_t for the Wasserstein distance of order 1 as well as in total variation.

Proposition 23. *Assume* **B**5, **B**3(m) *for* $\mathsf{m} \in \mathbb{R}$, **C**2 *and that there exists* $\varepsilon_b > 0$ *such that* $\sup_{x \in \mathbb{R}^d} \|b(x)\|^{2(1+\varepsilon_b)} \mathrm{e}^{-\mathsf{m}^+\|x\|^2} < +\infty$. *Then for any* $T \geq 0$ *and* $x, y \in \mathbb{R}^d$

$$\|\delta_x \mathrm{P}_T - \delta_y \mathrm{P}_T\|_V \leq C_b \rho_b^T \{V(x) + V(y)\} ,$$

with $C_b \geq 0$ *and* $\rho_b \in (0, 1)$ *given by* (61) *and for any* $x \in \mathbb{R}^d$, $V(x) = 1 + \|x\|^2$.

Proof. The proof is identical to the one of Proposition 22 upon replacing Theorem 10 by Theorem 12. □

The rates we obtain are identical to the ones derived taking the limit $\bar{\gamma} \to 0$ in Theorem 12. An upper bound on ρ_b depending on the curvature condition on b is provided in (31) and (32).

Proposition 24. *Assume* **B** 5*,* **B** 3(m)*,* **C** 3 *and that there exists $\varepsilon_b > 0$ such that* $\sup_{x \in \mathbb{R}^d} \|b(x)\|^{2(1+\varepsilon_b)} e^{-m_1(1+\|x\|)^{1/2}/2} < +\infty$. *Then*

$$\|\delta_x P_T - \delta_y P_T\|_V \leq C_c \rho_c^T \{V(x) + V(y)\} \ ,$$

with $C_c \geq 0$ and $\rho_c \in (0,1)$ given by Appendix E and for any $x \in \mathbb{R}^d$, $V(x) = \exp(m_3(1+\|x\|^2)^{1/2})$ and $m_3 \in (0, m_1/4)$.

Proof. Let $T \geq 0$, $V(x) = \exp(m_3 \phi(x))$, $V_1(x) = \exp(m_1 \phi(x))$ with $\phi(x) = (1 + \|x\|^2)^{1/2}$. Note that for any $x \in \mathbb{R}^d$, $V(x) \leq V_1(x)$. Since **B** 3(m) holds we obtain using Proposition 19 that **L** 1 holds. In addition, since $\sup_{x \in \mathbb{R}^d} \langle x, b(x) \rangle < +\infty$ by **C** 3 there exists $\zeta \in \mathbb{R}$ such that $\mathbf{D}_c(V_1, \zeta, 0)$ holds. Therefore, using Lemma 15 and that there exists $\varepsilon_b > 0$ such that $\sup_{x \in \mathbb{R}^d} \|b(x)\|^{2(1+\varepsilon_b)} e^{-m_1 \phi(x)} < +\infty$, we get that **L** 3 holds. Using Proposition 20-(a) we get that **L** 2, **L** 4 and **L** 5 hold. In the proof of Proposition 13, see Appendix F.5 we show that there exists $\lambda > 0$ and $A \geq 0$ such that $\mathbf{D}_d(V^2, \lambda^{1/m}, A/m, \bar{B}(0,n))$ for $R_{T/m,n}$ and $\tilde{R}_{T/m,n}$ with $m, n \in \mathbb{N}$ and m large enough. Since λ and A admit limits when $m \to +\infty$ which do not depend on n we get that for any $x \in \mathbb{R}^d$

$$\limsup_{n \to +\infty} \limsup_{m \to +\infty} \mathbf{d}_W(\delta_x R_{T/m,n}^m, \delta_y R_{T/m,n}^m) < +\infty \ .$$

In addition since $\mathbf{D}_c(V_1, \zeta, 0)$ holds, we get that $P_T V_1(x) < +\infty$, hence $P_T V(x) < +\infty$. We can apply Theorem 18 and we obtain that

$$\|\delta_x P_T - \delta_y P_T\|_V \leq \limsup_{n \to +\infty} \limsup_{m \to +\infty} \|\delta_x R_{T/m,n}^m - \delta_y R_{T/m,n}^m\|_V \ .$$

Since **B** 1($\bar{B}(0,n)$) holds and **B** 5 implies **B** 2 on $\bar{B}(0,n)$, we can apply Theorem 14 and we get that for any $m, n \in \mathbb{N}$ with $x, y \in \bar{B}(0,n)$ and $1/m \in (0, 2m_2)$

$$\|\delta_x R_{T/m,n}^m - \delta_y R_{T/m,n}^m\|_V \leq C_{1/m,c} \rho_{1/m,c}^T \{V(x) + V(y)\} \ ,$$

where $C_{1/m,c} \geq 0$ and $\rho_{1/m,c} \in (0,1)$, see Appendix E. We conclude upon remarking that $C_{1/m,c}$ and $\rho_{1/m,c}$ admit limits C_c and ρ_c when $m \to +\infty$ which do not depend on n. □

The rates we obtain are identical to the ones derived taking the limit $\bar{\gamma} \to 0$ in Theorem 14. An upper bound on ρ_c depending on the curvature condition on b is given in Appendix F.5.

References

[1] J. Baker, P. Fearnhead, E. B. Fox, and C. Nemeth. Control variates for stochastic gradient mcmc. *Statistics and Computing*, pages 1–17.

[2] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer Publishing Company, Incorporated, 1st edition, 2011.

[3] E. Bernton. Langevin monte carlo and jko splitting. *arXiv preprint arXiv:1802.08671*, 2018.

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

[5] P. Brémaud. *Markov chains*, volume 31 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999. Gibbs fields, Monte Carlo simulation, and queues.

[6] N. Brosse, A. Durmus, Eric Moulines, and S. Sabanis. The tamed unadjusted langevin algorithm. *Stochastic Processes and their Applications*, 2018.

[7] S. Bubeck, R. Eldan, and J. Lehec. Finite-time analysis of projected langevin monte carlo. In *Proceedings of the 28th International Conference on Neural Information Processing Systems*, NIPS'15, pages 1243–1251, Cambridge, MA, USA, 2015. MIT Press.

[8] R. Bubley, M. Dyer, and M. Jerrum. An elementary analysis of a procedure for sampling points in a convex body. *Random Structures Algorithms*, 12(3):213–235, 1998.

[9] K. S. Chan. Asymptotic behavior of the Gibbs sampler. *J. Amer. Statist. Assoc.*, 88(421):320–326, 1993.

[10] N. Chatterji, N. Flammarion, Y. Ma, P. Bartlett, and M. Jordan. On the theory of variance reduction for stochastic gradient monte carlo. In *International Conference on Machine Learning*, pages 763–772, 2018.

[11] R. Chen and R. S. Tsay. On the ergodicity of TAR(1) processes. *Ann. Appl. Probab.*, 1(4):613–634, 1991.

[12] A. S. Dalalyan and A. Karagulyan. User-friendly guarantees for the langevin monte carlo with inaccurate gradient. *Stochastic Processes and their Applications*, 2019.

[13] V. De Bortoli, A. Durmus, M. Pereyra, and A. Fernandez Vidal. Stochastic Optimization with Unadjusted Kernel: the SOUK Algorithm. working paper or preprint, Jan. 2019.

[14] R. Douc, G. Fort, and A. Guillin. Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stochastic Process. Appl.*, 119(3):897–923, 2009.

[15] R. Douc, G. Fort, E. Moulines, and P. Soulier. Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.*, 14(3):1353–1377, 2004.

[16] R. Douc, E. Moulines, P. Priouret, and P. Soulier. *Markov Chains*. Springer, 2019.

[17] R. Douc, E. Moulines, and J. S. Rosenthal. Quantitative bounds on convergence of time-inhomogeneous Markov chains. *Ann. Appl. Probab.*, 14(4):1643–1665, 2004.

[18] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23(4):1671–1691, 1995.

[19] A. Durmus, G. Fort, and E. Moulines. Subgeometric rates of convergence in Wasserstein distance for Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1799–1822, 2016.

[20] A. Durmus, S. Majewski, and B. Miasojedow. Analysis of langevin monte carlo via convex optimization. *arXiv preprint arXiv:1802.09188*, 2018.

[21] A. Durmus and É. Moulines. Quantitative bounds of convergence for geometrically ergodic markov chain in the wasserstein distance with application to the metropolis adjusted langevin algorithm. *Statistics and Computing*, 25(1):5–19, 2015.

[22] A. Durmus and E. Moulines. High-dimensional Bayesian inference via the Unadjusted Langevin Algorithm. *ArXiv e-prints*, May 2016.

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

[24] A. Durmus, E. Moulines, and M. Pereyra. Efficient bayesian computation by proximal markov chain monte carlo: when langevin meets moreau. *SIAM Journal on Imaging Sciences*, 11(1):473–506, 2018.

[25] A. Eberle. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields*, 166(3-4):851–886, 2016.

[26] A. Eberle, A. Guillin, and R. Zimmer. Quantitative harris-type theorems for diffusions and mckean–vlasov processes. *Transactions of the American Mathematical Society*, 2018.

[27] A. Eberle and M. B. Majka. Quantitative contraction rates for markov chains on general state spaces. *arXiv preprint arXiv:1808.07033*, 2018.

[28] G. Fort. *Controle explicite d'ergodicitÃ© de chaÃ®nes de Markov : Applications a l'analyse de convergence de l'algorithme Monte-Carlo EM*. PhD thesis, UniversitÃ© Pierre et Marie Curie, Paris, Paris, 2001.

[29] G. Fort. Computable bounds for V-geometric ergodicity of Markov transition kernels. *Rapport de Recherche, Univ. J. Fourier, RR 1047-M.,* `https: // www. math. univ-toulouse. fr/ %7Egfort/ Preprints/ fort: 2002. pdf` , 2002.

[30] G. Fort and G. O. Roberts. Subgeometric ergodicity of strong Markov processes. *Ann. Appl. Probab.*, 15(2):1565–1589, 2005.

[31] F. G. Foster. On the stochastic matrices associated with certain queuing processes. *Ann. Math. Statistics*, 24:355–360, 1953.

[32] B. Goldys and B. Maslowski. Lower estimates of transition densities and bounds on exponential ergodicity for stochastic PDE's. *Ann. Probab.*, 34(4):1451–1496, 2006.

[33] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011.

[34] M. Hairer, J. C. Mattingly, and M. Scheutzow. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields*, 149(1-2):223–259, 2011.

[35] M. Hairer, A. M. Stuart, and S. J. Vollmer. Spectral gaps for a Metropolis-Hastings algorithm in infinite dimensions. *Ann. Appl. Probab.*, 24(6):2455–2490, 2014.

[36] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, second edition, 1989.

[37] S. F. Jarner and G. O. Roberts. Polynomial convergence rates of Markov chains. *Ann. Appl. Probab.*, 12(1):224–247, 2002.

[38] S. F. Jarner and R. L. Tweedie. Locally contracting iterated functions and stability of Markov chains. *J. Appl. Probab.*, 38(2):494–507, 2001.

[39] G. L. Jones and J. P. Hobert. Honest exploration of intractable probability distributions via Markov chain Monte Carlo. *Statist. Sci.*, 16(4):312–334, 2001.

[40] G. L. Jones and J. P. Hobert. Sufficient burn-in for Gibbs samplers for a hierarchical random effects model. *Ann. Statist.*, 32(2):784–817, 2004.

[41] A. Joulin and Y. Ollivier. Curvature, concentration and error estimates for markov chain monte carlo. *The Annals of Probability*, 38(6):2418–2442, 2010.

[42] R. Khasminskii. *Stochastic stability of differential equations*, volume 66. Springer Science & Business Media, 2011.

[43] T. Lindvall and L. C. G. Rogers. Coupling of multidimensional diffusions by reflection. *Ann. Probab.*, 14(3):860–872, 1986.

[44] R. Liptser and A. N. Shiryaev. *Statistics of random Processes: I. general Theory*, volume 5. Springer Science & Business Media, 2013.

[45] R. B. Lund and R. L. Tweedie. Geometric convergence rates for stochastically ordered Markov chains. *Math. Oper. Res.*, 21(1):182–194, 1996.

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

[47] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab.*, 25(3):487–517, 1993.

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

[49] S. P. Meyn and R. L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Probab.*, 4(4):981–1011, 1994.

[50] Y. Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course.* Applied Optimization. Springer, 2004.

[51] E. Nummelin and P. Tuominen. Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. *Stochastic Process. Appl.*, 12(2):187–202, 1982.

[52] E. Nummelin and P. Tuominen. The rate of convergence in Orey's theorem for Harris recurrent Markov chains with applications to renewal theory. *Stochastic Process. Appl.*, 15(3):295–311, 1983.

[53] E. Nummelin and R. L. Tweedie. Geometric ergodicity and R-positivity for general Markov chains. *Ann. Probability*, 6(3):404–420, 1978.

[54] Y. Ollivier. Ricci curvature of markov chains on metric spaces. *Journal of Functional Analysis*, 256(3):810–864, 2009.

[55] N. Parikh and S. Boyd. *Proximal Algorithms.* Foundations and Trends(r) in Optimization. Now Publishers, 2013.

[56] N. N. Popov. Geometric ergodicity conditions for countable Markov chains. *Dokl. Akad. Nauk SSSR*, 234(2):316–319, 1977.

[57] Q. Qin and J. P. Hobert. Wasserstein-based methods for convergence complexity analysis of mcmc with application to albert and chib's algorithm. *arXiv preprint arXiv:1810.08826*, 2018.

[58] Q. Qin and J. P. Hobert. Geometric convergence bounds for markov chains in wasserstein distance based on generalized drift and contraction conditions. *arXiv preprint arXiv:1902.02964*, 2019.

[59] G. O. Roberts and N. G. Polson. On the geometric convergence of the gibbs sampler. *Journal of the Royal Statistical Society, Series B*, 56:377–384, 1994.

[60] G. O. Roberts and J. S. Rosenthal. Quantitative bounds for convergence rates of continuous time Markov processes. *Electron. J. Probab.*, 1:no. 9, approx. 21 pp. 1996.

[61] G. O. Roberts and R. L. Tweedie. Bounds on regeneration times and convergence rates for Markov chains. *Stochastic Process. Appl.*, 80(2):211–229, 1999.

[62] J. S. Rosenthal. Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.*, 90(430):558–566, 1995.

[63] J. S. Rosenthal. Quantitative convergence rates of Markov chains: a simple account. *Electron. Comm. Probab.*, 7:123–128, 2002.

[64] P. Tuominen and R. L. Tweedie. Subgeometric rates of convergence of f-ergodic Markov chains. *Adv. in Appl. Probab.*, 26(3):775–798, 1994.

[65] A. Y. Veretennikov. On polynomial mixing bounds for stochastic differential equations. *Stochastic Process. Appl.*, 70(1):115–127, 1997.

[66] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

A Minorization conditions for functional autoregressive models

In this section, we extend and complete the results of [22, Section 6] on functional autoregressive models. Let $\mathsf{X} \in \mathcal{B}(\mathbb{R}^d)$ equipped with its trace σ-field $\mathcal{X} = \{\mathsf{A} \cap \mathsf{X} : \mathsf{A} \in \mathcal{B}(\mathbb{R}^d)\}$. In fact, we consider a slightly more general class of models than [22] which is associated with non-homogeneous Markov chains $(X_k^{(\mathrm{a})})_{k \in \mathbb{N}}$ with state space $(\mathsf{X}, \mathcal{X})$ defined for $k \geq 0$ by

$$X_{k+1}^{(\mathrm{a})} = \Pi \left(\mathcal{T}_{k+1}(X_k^{(\mathrm{a})}) + \sigma_{k+1} Z_{k+1} \right) \ ,$$

where Π is a measurable function from $\mathbb{R}^d$ to X, $(\mathcal{T}_k)_{k \geq 1}$ is a sequence of measurable functions from $\mathbb{R}^d$ to $\mathbb{R}^d$, $(\sigma_k)_{k \geq 1}$ is a sequence of positive real numbers and $(Z_k)_{k \geq 1}$ is a sequence of i.i.d. d dimensional standard Gaussian random variables. We assume that Π satisfies the following assumption

A3. *The function* $\Pi : \mathbb{R}^d \to \mathsf{X}$ *is non-expansive: i.e. for any* $x, y \in \mathbb{R}^d$, $\|\Pi(x) - \Pi(y)\| \leq \|x - y\|$.

We also assume some Lipschitz regularity on the operator $\mathcal{T}_k$ for any $k \in \mathbb{N}^\star$

AR 1 (A). *For all* $k \geq 1$ *there exists* $\varpi_k \in \mathbb{R}$ *such that for all* $(x, y) \in \mathsf{A}$,

$$\|\mathcal{T}_k(x) - \mathcal{T}_k(y)\|^2 \leq (1 + \varpi_k) \|x - y\|^2 .$$

The sequence $\{X_k^{(\mathrm{a})}, \ k \in \mathbb{N}\}$ is an inhomogeneous Markov chain associated with the family of Markov kernels $(P_k^{(\mathrm{a})})_{k \geq 1}$ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ given for all $x \in \mathbb{R}^d$ and $\mathsf{A} \in \mathbb{R}^d$ by

$$P_k^{(\mathrm{a})}(x, \mathsf{A}) = \frac{1}{(2\pi\sigma_k^2)^{d/2}} \int_{\Pi^{-1}(\mathsf{A})} \exp\left(-\|y - \mathcal{T}_k(x)\|^2 / (2\sigma_k^2)\right) \mathrm{d}y .$$

We denote for all $n \geq 1$ by $Q_n^{(\mathrm{a})}$ the marginal distribution of $X_n^{(\mathrm{a})}$ given by $Q_n^{(\mathrm{a})} = P_1^{(\mathrm{a})} \cdots P_n^{(\mathrm{a})}$. To obtain an upper bound of $\|\delta_x Q_n^{(\mathrm{a})} - \delta_y Q_n^{(\mathrm{a})}\|_{\mathrm{TV}}$ for any $x, y \in \mathbb{R}^d$, $n \in \mathbb{N}^\star$, we introduce a Markov coupling $(X_k^{(\mathrm{a})}, Y_k^{(\mathrm{a})})_{k \in \mathbb{N}}$ such that for any $n \in \mathbb{N}^\star$, the distribution of $X_n^{(\mathrm{a})}$ and $Y_n^{(\mathrm{a})}$ are $\delta_x Q_n^{(\mathrm{a})}$ and $\delta_x Q_n^{(\mathrm{a})}$ respectively, exactly as we have introduced in the homogeneous setting the Markov coupling with kernel K_γ defined by (6) for R_γ defined in (2). For completeness and readability, we have decided to recall the construction of $(X_k^{(\mathrm{a})}, Y_k^{(\mathrm{a})})_{k \in \mathbb{N}}$. For all $k \in \mathbb{N}^\star$ and $x, y, z \in \mathbb{R}^d$, define

$$\mathrm{e}_k(x, y) = \begin{cases} \mathrm{E}_k(x, y)/\|\mathrm{E}_k(x, y)\| & \text{if } x \neq y \\ 0 & \text{otherwise} \end{cases}, \qquad \mathrm{E}_k(x, y) = \mathcal{T}_k(y) - \mathcal{T}_k(x) , \quad (45)$$

$$\mathcal{S}_k(x, y, z) = \mathcal{T}_k(y) + \sigma_{k+1}\left[(\mathrm{Id} - 2\mathrm{e}_k(x, y)\mathrm{e}_k(x, y)^\top)z\right] ,$$
$$p_k(x, y, z) = 1 \wedge \frac{\boldsymbol{\varphi}_{\sigma_k^2}(\|\mathrm{E}_k(x, y)\| - \langle \mathrm{e}_k(x, y), z\rangle)}{\boldsymbol{\varphi}_{\sigma_k^2}(\langle \mathrm{e}_k(x, y), z\rangle)} , \qquad (46)$$

where $\boldsymbol{\varphi}_{\sigma_k^2}$ is the one-dimensional zero mean Gaussian distribution function with variance σ_k^2. Let $(U_k)_{k \in \mathbb{N}^\star}$ be a sequence of i.i.d. uniform random variables on $[0, 1]$ and define the Markov chain $(X_k^{(\mathrm{a})}, Y_k^{(\mathrm{a})})_{k \in \mathbb{N}}$ starting from $(X_0^{(\mathrm{a})}, Y_0^{(\mathrm{a})}) \in \mathsf{X}^2$ by the recursion: for any $k \geq 0$

$$\tilde{X}_{k+1}^{(\mathrm{a})} = \mathcal{T}_{k+1}(X_k^{(\mathrm{a})}) + \sigma_{k+1}Z_{k+1} ,$$
$$\tilde{Y}_{k+1}^{(\mathrm{a})} = \begin{cases} \tilde{X}_{k+1}^{(\mathrm{a})} & \text{if } \mathcal{T}_{k+1}(X_k^{(\mathrm{a})}) = \mathcal{T}_{k+1}(Y_k^{(\mathrm{a})}) ; \\ W_{k+1}^{(\mathrm{a})}\tilde{X}_{k+1}^{(\mathrm{a})} + (1 - W_{k+1}^{(\mathrm{a})})\mathcal{S}_{k+1}(X_k^{(\mathrm{a})}, Y_k^{(\mathrm{a})}, Z_{k+1}) & \text{otherwise} , \end{cases}$$

$$(47)$$

where $W_{k+1}^{(a)} = \mathbb{1}_{(-\infty,0]}(U_{k+1} - p_{k+1}(X_k^{(a)}, Y_k^{(a)}, Z_{k+1}))$ and finally set

$$(X_{k+1}^{(a)}, Y_{k+1}^{(a)}) = (\Pi(\tilde{X}_{k+1}^{(a)}), \Pi(\tilde{Y}_{k+1}^{(a)})) . \qquad (48)$$

For any $k \in \mathbb{N}^\star$, marginally, the distribution of $X_{k+1}^{(a)}$ given $X_k^{(a)}$ is $P_{k+1}^{(a)}(X_k^{(a)}, \cdot)$, and it is well-know (see e.g. [8, Section 3.3]) that $\tilde{Y}_{k+1}^{(a)}$ and $\mathcal{T}_\gamma(Y_k^{(a)}) + \sigma_{k+1} Z_{k+1}$ have the same distribution given Y_k, and therefore the distribution of Y_{k+1} given Y_k is $P_{k+1}^{(a)}(Y_k, \cdot)$. As a result for any $(x,y) \in \mathsf{X}^2$ and $n \in \mathbb{N}^\star$, $(X_n^{(a)}, Y_n^{(a)})$ with $(X_0^{(a)}, Y_0^{(a)}) = (x, y)$ is coupling between $\delta_x Q_n^{(a)}$ and $\delta_y Q_n^{(a)}$. Therefore, we obtain that $\|\delta_x Q_n^{(a)} - \delta_y Q_n^{(a)}\|_{\mathrm{TV}} \leq \mathbb{P}(X_n^{(a)} \neq Y_n^{(a)})$. Therefore to get an upper bound on $\|\delta_x Q_n^{(a)} - \delta_y Q_n^{(a)}\|_{\mathrm{TV}}$, it is sufficient to obtain a bound on $\mathbb{P}(X_n^{(a)} \neq Y_n^{(a)})$ which is a simple consequence of the following more general result.

Theorem 25. *Let* $\mathsf{A} \in \mathcal{B}(\mathbb{R}^{2d})$ *and assume* $\mathbf{A}3$ *and* $\mathbf{AR}1(\mathsf{A})$. *Let* $(X_k^{(a)}, Y_k^{(a)})_{k \in \mathbb{N}}$ *be defined by* (48), *with* $(X_0^{(a)}, Y_0^{(a)}) = (x, y) \in \mathsf{A}$. *Then for any* $n \in \mathbb{N}^\star$,

$$\mathbb{P}\left[X_n^{(a)} \neq Y_n^{(a)} \text{ and for any } k \in \{1, \ldots, n-1\}, (X_k^{(a)}, Y_k^{(a)}) \in \mathsf{A}^2 \right]$$
$$\leq \mathbb{1}_{\Delta_{\mathsf{X}}^c}((x,y)) \left\{ 1 - 2\mathbf{\Phi}\left(-\frac{\|x-y\|}{2(\Xi_n^{(a)})^{1/2}} \right) \right\} ,$$

where $\mathbf{\Phi}$ *is the cumulative distribution function of standard normal distribution on* $\mathbb{R}$ *and* $(\Xi_i^{(a)})_{i \geq 1}$ *is defined for all* $k \geq 1$ *by* $\Xi_k^{(a)} = \sum_{i=1}^k \{\sigma_i^2 / \prod_{j=1}^i (1 + \varpi_j)\}$.

Proof. Let $(\mathcal{F}_k^{(a)})_{k \in \mathbb{N}}$ be the filtration associated to $(X_k^{(a)}, Y_k^{(a)})_{k \in \mathbb{N}}$. Denote for any $k \in \mathbb{N}$,

$$\mathscr{A}_k = \bigcap_{i=0}^k \{(X_i^{(a)}, Y_i^{(a)}) \in \mathsf{A}\} , \qquad \mathscr{A}_{-1} = \mathscr{A}_0 ,$$

and for all $k_1, k_2 \in \mathbb{N}^*$, $k_1 \leq k_2$, $\Xi_{k_1,k_2}^{(a)} = \sum_{i=k_1}^{k_2} \{\sigma_i^2 / \prod_{j=k_1}^i (1 + \varpi_j)\}$. Let $n \geq 1$ and $(x, y) \in \mathsf{A}^2$. We show by backward induction that for all $k \in \{0, \cdots, n-1\}$,

$$\mathbb{P}(\{X_n^{(a)} \neq Y_n^{(a)}\} \cap \mathscr{A}_{n-1}) \leq \mathbb{E}\left[\mathbb{1}_{\Delta_{\mathsf{X}}^c}(X_k^{(a)}, Y_k^{(a)}) \mathbb{1}_{\mathscr{A}_{k-1}} \left[1 - 2\mathbf{\Phi}\left\{ -\frac{\left\| X_k^{(a)} - Y_k^{(a)} \right\|}{2 \left(\Xi_{k+1,n}^{(a)}\right)^{1/2}} \right\} \right] \right] ,$$
$$(49)$$

Note that the inequality for $k = 0$ will conclude the proof. Using by (47) that $\tilde{X}_n^{(a)} = \tilde{Y}_n^{(a)}$ if $X_{n-1}^{(a)} = Y_{n-1}^{(a)}$ or $W_n = \mathbb{1}_{(-\infty,0]}(U_n - p_n(X_{n-1}^{(a)}, Y_{n-1}^{(a)}, Z_n)) = 1$, where

p_n is defined by (46), and (U_n, Z_n) are independent random variables independent of $\mathcal{F}_{n-1}^{(\mathrm{a})}$, we obtain

$$\mathbb{E}\left[\mathbb{1}_{\Delta_{\mathsf{X}}}(\tilde{X}_n^{(\mathrm{a})}, \tilde{Y}_n^{(\mathrm{a})}) \Big| \mathcal{F}_{n-1}^{(\mathrm{a})}\right] = \mathbb{E}\left[p_n(X_{n-1}^{(\mathrm{a})}, Y_{n-1}^{(\mathrm{a})}, Z_n) \Big| \mathcal{F}_{n-1}^{(\mathrm{a})}\right] = 2\Phi\left\{-\left\|(2\sigma_n)^{-1}\mathrm{E}_n(X_{n-1}^{(\mathrm{a})}, Y_{n-1}^{(\mathrm{a})})\right\|\right\} .$$

Since $\{X_n^{(\mathrm{a})} \neq Y_n^{(\mathrm{a})}\} \subset \{\tilde{X}_n^{(\mathrm{a})} \neq \tilde{Y}_n^{(\mathrm{a})}\} \subset \{X_{n-1}^{(\mathrm{a})} \neq Y_{n-1}^{(\mathrm{a})}\}$ by (48) and (47), we get

$$\mathbb{P}\left[\{X_n^{(\mathrm{a})} \neq Y_n^{(\mathrm{a})}\} \cap \mathscr{A}_{n-1}\right] = \mathbb{E}\left[\mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}(X_{n-1}^{(\mathrm{a})}, Y_{n-1}^{(\mathrm{a})})\mathbb{1}_{\mathscr{A}_{n-1}}\mathbb{E}\left[\mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}(\tilde{X}_n^{(\mathrm{a})}, \tilde{Y}_n^{(\mathrm{a})}) \Big| \mathcal{F}_{n-1}^{(\mathrm{a})}\right]\right]$$
$$= \mathbb{E}\left[\mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}(X_{n-1}^{(\mathrm{a})}, Y_{n-1}^{(\mathrm{a})})\mathbb{1}_{\mathscr{A}_{n-1}}\left[1 - 2\Phi\left\{-\left\|(2\sigma_n)^{-1}\mathrm{E}_n(X_{n-1}^{(\mathrm{a})}, Y_{n-1}^{(\mathrm{a})})\right\|\right\}\right]\right] ,$$

Using that $(X_{n-1}^{(\mathrm{a})}, Y_{n-1}^{(\mathrm{a})}) \in \mathsf{A}^2$ on $\mathscr{A}_{n-1}$, $\mathbf{AR}1(\mathsf{A})$ and (45), we have $\|\mathrm{E}_n(X_{n-1}^{(\mathrm{a})}, Y_{n-1}^{(\mathrm{a})})\|^2 \leq (1+\varpi_n)\|X_{n-1}^{(\mathrm{a})} - Y_{n-1}^{(\mathrm{a})}\|^2$, showing (49) holds for $k = n-1$ since $\mathscr{A}_{n-2} \subset \mathscr{A}_{n-1}$. Assume that (49) holds for $k \in \{1, \ldots, n-1\}$. On $\{\tilde{X}_k^{(\mathrm{a})} \neq \tilde{Y}_k^{(\mathrm{a})}\}$, we have

$$\left\|\tilde{X}_k^{(\mathrm{a})} - \tilde{Y}_k^{(\mathrm{a})}\right\| = \left|-\left\|\mathrm{E}_k(X_{k-1}^{(\mathrm{a})}, Y_{k-1}^{(\mathrm{a})})\right\| + 2\sigma_k\mathrm{e}_k(X_{k-1}^{(\mathrm{a})}, Y_{k-1}^{(\mathrm{a})})^{\mathrm{T}}Z_k\right| ,$$

which implies by (48) and since Π is non expansive by $\mathbf{A}3$

$$\mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}(X_k^{(\mathrm{a})}, Y_k^{(\mathrm{a})})\left[1 - 2\Phi\left\{-\frac{\left\|X_k^{(\mathrm{a})} - Y_k^{(\mathrm{a})}\right\|}{2(\Xi_{k+1,n}^{(\mathrm{a})})^{1/2}}\right\}\right]$$
$$\leq \mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}(X_k^{(\mathrm{a})}, Y_k^{(\mathrm{a})})\left[1 - 2\Phi\left\{-\frac{\left\|\tilde{X}_k^{(\mathrm{a})} - \tilde{Y}_k^{(\mathrm{a})}\right\|}{2(\Xi_{k+1,n}^{(\mathrm{a})})^{1/2}}\right\}\right]$$
$$\leq \mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}(X_k^{(\mathrm{a})}, Y_k^{(\mathrm{a})})\left[1 - 2\Phi\left\{-\frac{\left|2\sigma_k\mathrm{e}_k(X_{k-1}^{(\mathrm{a})}, Y_{k-1}^{(\mathrm{a})})^{\mathrm{T}}Z_k - \left\|\mathrm{E}_k(X_{k-1}^{(\mathrm{a})}, Y_{k-1}^{(\mathrm{a})})\right\|\right|}{2(\Xi_{k+1,n}^{(\mathrm{a})})^{1/2}}\right\}\right] .$$

Since Z_k is independent of $\mathcal{F}_k^{(\mathrm{a})}$, $\sigma_k\mathrm{e}_k(X_{k-1}^{(\mathrm{a})}, X_{k-1}^{(\mathrm{a})})^{\mathrm{T}}Z_k$ is a real Gaussian random variable with zero mean and variance σ_k^2, therefore by [22, Lemma 20] and since $\mathscr{A}_{k-1}$ is $\mathcal{F}_{k-1}^{(\mathrm{a})}$-measurable, we get

$$\mathbb{E}\left[\mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}(X_k^{(\mathrm{a})}, Y_k^{(\mathrm{a})})\mathbb{1}_{\mathscr{A}_{k-1}}\left[1 - 2\Phi\left\{-\frac{\left\|X_k^{(\mathrm{a})} - Y_k^{(\mathrm{a})}\right\|}{2(\Xi_{k+1,n}^{(\mathrm{a})})^{1/2}}\right\}\right] \Big| \mathcal{F}_{k-1}^{(\mathrm{a})}\right]$$
$$\leq \mathbb{1}_{\mathscr{A}_{k-1}}\mathbb{1}_{\Delta_{\mathsf{X}}^{\mathrm{c}}}(X_{k-1}^{(\mathrm{a})}, Y_{k-1}^{(\mathrm{a})})\left[1 - 2\Phi\left\{-\frac{\left\|\mathrm{E}_k(X_{k-1}^{(\mathrm{a})}, Y_{k-1}^{(\mathrm{a})})\right\|}{2\left(\sigma_k + \Xi_{k+1,n}^{(\mathrm{a})}\right)^{1/2}}\right\}\right] .$$

Using by $\mathbf{A}2(\mathsf{A})$ that $\|\mathrm{E}_k(X_{k-1}^{(\mathrm{a})}, Y_{k-1}^{(\mathrm{a})})\|^2 \leq (1 + \varpi_{k-1})\left\|X_{k-1}^{(\mathrm{a})} - Y_{k-1}^{(\mathrm{a})}\right\|^2$ on $\mathscr{A}_{k-1}$ and $\mathscr{A}_{k-2} \subset \mathscr{A}_{k-1}$ concludes the induction of (49). $\qquad\square$

B Quantitative convergence results based on [16, 17]

We start by recalling the following lemma from [16] which is inspired from the results of [17].

Lemma 26 ([16, Lemma 19.4.2]). *Let* $(\mathsf{Y}, \mathcal{Y})$ *be a measurable space and* R *be a Markov kernel over* $(\mathsf{Y}, \mathcal{Y})$. *Let* Q *be a Markov coupling kernel for* R. *Assume there exist* $\mathsf{C} \in \mathcal{Y}^{\otimes 2}$, $M \geq 0$, *a measurable function* $W : \mathsf{Y} \times \mathsf{Y} \to [1, +\infty)$, $\lambda \in (0, 1)$ *and* $A \geq 0$ *such that for any* $x, y \in \mathsf{Y}$,

$$\mathrm{Q}W(x, y) \leq \lambda W(x, y) \mathbb{1}_{\mathsf{C}^c}((x, y)) + A \mathbb{1}_{\mathsf{C}}((x, y)) \ .$$

In addition, assume that there exists $\varepsilon > 0$ *such that for any* $(x, y) \in \mathsf{C}$,

$$\mathrm{Q}((x, y), \Delta_{\mathsf{Y}}^c) \leq 1 - \varepsilon \ ,$$

where $\Delta_{\mathsf{Y}} = \{(x, y) \in \mathsf{Y}^2 : x = y\}$. *Then there exist* $\rho \in (0, 1)$ *and* $C \geq 0$ *such that for any* $x, y \in \mathcal{Y}$ *and* $n \in \mathbb{N}^\star$

$$\mathbf{d}_W(\delta_x \mathrm{R}^n, \delta_y \mathrm{R}^n) \leq C \rho^n W(x, y) \ ,$$

where $\mathbf{d}_W$ *is the Wasserstein metric associated with the cost function* $\mathbf{w} : (x, y) \mapsto \mathbb{1}_{\Delta_{\mathsf{Y}}^c}((x, y)) W(x, y)$ *and*

$$C = 2(1 + A/\{(1 - \varepsilon)(1 - \lambda)\}) \ ,$$
$$\log(\rho) = \{\log(1 - \varepsilon) \log(\lambda)\}/\{\log(1 - \varepsilon) + \log(\lambda) - \log(A)\} \ .$$

Theorem 27. *Let* $(\mathsf{Y}, \mathcal{Y})$ *be a measurable space and* R *be a Markov kernel over* $(\mathsf{Y}, \mathcal{Y})$. *Let* Q *be a Markov coupling kernel of* R. *Assume that there exist* $\lambda \in (0, 1)$, $A \geq 0$ *and a measurable function* $W : \mathsf{Y} \times \mathsf{Y} \to [1, +\infty)$, *such that* Q *satisfies* $\mathbf{D}_{\mathrm{d}}(W, \lambda, A, \mathsf{Y})$. *In addition, assume that there exist* $\ell \in \mathbb{N}^\star$, $\varepsilon > 0$ *and* $M \geq 1$ *such that for any* $(x, y) \in \mathsf{C}_M = \{(x, y) \in \mathsf{Y} \times \mathsf{Y}, \ W(x, y) \leq M\}$,

$$\mathrm{Q}^\ell((x, y), \Delta_{\mathsf{Y}}^c) \leq 1 - \varepsilon \ , \tag{50}$$

with $\Delta_{\mathsf{Y}} = \{(x, y) \in \mathsf{Y}^2 : x = y\}$ *and* $M \geq 2A/(1 - \lambda)$. *Then, there exist* $\rho \in (0, 1)$ *and* $C \geq 0$ *such that for any* $n \in \mathbb{N}$ *and* $x, y \in \mathsf{Y}$

$$\mathbf{d}_W(\delta_x \mathrm{R}^n, \delta_y \mathrm{R}^n) \leq C \rho^{\lfloor n/\ell \rfloor} W(x, y) \ ,$$

with

$$C = 2(1 + A_\ell)(1 + c_\ell/\{(1 - \varepsilon)(1 - \lambda_\ell)\}) \ ,$$
$$\lambda_\ell = (\lambda^\ell + 1)/2 \ , \quad c_\ell = \lambda^\ell M + A_\ell \ , \quad A\ell = A(1 - \lambda^\ell/(1 - \lambda) \ ,$$
$$\log(\rho_\ell) = \{\log(1 - \varepsilon)\log(\lambda_\ell)\}/\{\log(1 - \varepsilon) + \log(\lambda_\ell) - \log(c_\ell)\} \ .$$

Proof. We first show that for any $(x, y) \in \mathsf{C}_M$,

$$Q^\ell(x, y) \leq \lambda_\ell W(x, y) \mathbb{1}_{\mathsf{C}_M^c}((x, y)) + c_\ell \mathbb{1}_{\mathsf{C}_M}((x, y)) \ , \tag{51}$$

to apply Lemma 26 with coupling kernel Q^ℓ. By a straightforward induction, for any $x, y \in \mathsf{Y}$ we have

$$Q^\ell W(x, y) \leq \lambda^\ell W(x, y) + A(1 - \lambda^\ell)/(1 - \lambda) \ . \tag{52}$$

We distinguish two cases. If $(x, y) \notin \mathsf{C}_M$, using that $A/M \geq (1 - \lambda)/2$ we have that

$$Q^\ell W(x, y) \leq \lambda^\ell W(x, y) + A(1 - \lambda^\ell)W(x, y)/(M(1 - \lambda)) \leq 2^{-1}(\lambda^\ell + 1)W(x, y) \ .$$

If $(x, y) \in \mathsf{C}_M$, we have

$$Q^\ell W(x, y) \leq \lambda^\ell M + A(1 - \lambda^\ell)/(1 - \lambda) \ .$$

Therefore (51) holds. As a result and since by assumption we have (50), we can apply Lemma 26 to R^ℓ. Then, we obtain that for any $n \in \mathbb{N}$ and $x, y \in \mathsf{Y}$

$$\mathbf{d}_W(\delta_x R^{n\ell}, \delta_y R^{n\ell}) \leq 2\left\{1 + c_\ell\left[(1 - \lambda_\ell)(1 - \varepsilon)\right]^{-1}\right\} \rho_\ell^n W(x, y) \ ,$$

with
$$\log(\rho_\ell) = \{\log(1 - \varepsilon)\log(\lambda_\ell)\}/\{\log(1 - \varepsilon) + \log(\lambda_\ell) - \log(c_\ell)\} \ .$$

In addition, using (52), for any $r \in \{0, \ldots, \ell-1\}$, $\mathbf{d}_W(\delta_x R^r, \delta_y R^r) \leq (1 + A\ell)W(x, y)$. Therefore, for any $n \in \mathbb{N}$, since $n = \lfloor n/\ell \rfloor \ell + r_n$ with $r_n \in \{0, \ldots, \ell - 1\}$, we obtain that

$$\mathbf{d}_W(\delta_x R^n, \delta_y R^n) \leq 2(1 + A\ell)\left\{1 + c_\ell\left[(1 - \lambda_\ell)(1 - \varepsilon)\right]^{-1}\right\} \rho_\ell^{\lfloor n/\ell \rfloor} W(x, y) \ ,$$

which concludes the proof. $\qquad\square$

In the following lemma we draw a link between drift conditions on K_γ and drift conditions on R_γ in the case where W is separable.

Lemma 28. *Let $\bar{\gamma} > 0$, such that for any $\gamma \in (0, \bar{\gamma}]$, P_γ is a Markov kernel and Q_γ is a Markov coupling kernel for P_γ. In addition, let $V : X \to [1, +\infty)$, $\lambda \in (0,1)$ and $A \geq 0$ such that for any $\gamma \in (0, \bar{\gamma}]$, P_γ satisfies $\mathbf{D_d}(V, \lambda^\gamma, A\gamma, X)$ then Q_γ satisfies $\mathbf{D_d}(W, \lambda^\gamma, A\gamma, X^2)$, where for all $x, y \in X$, $W(x,y) = \{V(x) + V(y)\}/2$.*

Proof. Let $\gamma \in (0, \bar{\gamma}]$ and $x, y \in X$. Since $\delta_{(x,y)}Q_\gamma$ is a transference plan between $\delta_x P_\gamma$ and $\delta_y P_\gamma$ we have

$$Q_\gamma W(x,y) = Q_\gamma \{V(x) + V(y)\}/2 = P_\gamma V(x)/2 + P_\gamma V(y)/2 \leq \lambda^\gamma W(x,y) + A\gamma .$$

$\square$

C   Quantitative bounds for geometric convergence of Markov chains in Wasserstein distance

In this section, we establish new quantitative bounds for Markov chains in Wasserstein distance. We consider a Markov kernel P on the measurable space $(Y, \mathcal{Y})$ equipped with the bounded semi-metric $\mathbf{d} : Y \times Y \to \mathbb{R}_+$, *i.e.* which satisfies the following condition.

H 1. *For any $x, y \in Y$, $\mathbf{d}(x,y) \leq 1$, $\mathbf{d}(x,y) = \mathbf{d}(y,x)$ and $\mathbf{d}(x,y) = 0$ if and only if $x = y$.*

Let K be a Markov coupling kernel for P. We assume in this subsection the following condition on K.

H 2 (K)**.** *There exists $C \in \mathcal{Y}^{\otimes 2}$ such that*

(i) there exist $\mathbf{n_0} \in \mathbb{N}^\star$ and $\varepsilon > 0$ such that for any $x, y \in C$, $K^{\mathbf{n_0}}\mathbf{d}(x,y) \leq (1 - \varepsilon)\mathbf{d}(x,y)$;

(ii) for any $x, y \in Y$, $K\mathbf{d}(x,y) \leq \mathbf{d}(x,y)$;

(iii) there exist $W_1 : Y^2 \to [1, +\infty)$, $\lambda_1 \in (0,1)$ and $A_1 \geq 0$ such that K satisfies $\mathbf{D_d}(W_1, \lambda_1, A_1, C)$.

We consider the Markov chain $(X_n, Y_n)_{n \in \mathbb{N}}$ associated with the Markov kernel K defined on the canonical space $((Y \times Y)^\mathbb{N}, (\mathcal{Y}^{\otimes 2})^\mathbb{N})$ and denote by $\mathbb{P}_{(x,y)}$ and $\mathbb{E}_{(x,y)}$ the corresponding probability distribution and expectation respectively when $(X_0, Y_0) = (x,y)$. Denote by $(\mathcal{G}_n)_{n \in \mathbb{N}}$ the canonical filtration associated with $(X_n, Y_n)_{n \in \mathbb{N}}$. Note that for any $n \in \mathbb{N}$ and $x, y \in Y$, under $\mathbb{P}_{(x,y)}$, (X_n, Y_n) is by definition a coupling of $\delta_x P^n$ and $\delta_y P^n$. The main result of this section is the following which by the previous observation implies quantitative bounds on $\mathbf{W_d}(\delta_x P^n, \delta_y P^n)$.

Theorem 29. *Let* K *be a Markov coupling kernel for* P *and assume* **H**1 *and* **H**2(K). *Then for any* $n \in \mathbb{N}$ *and* $x, y \in \mathsf{Y}$,

$$\mathbb{E}_{(x,y)} \left[\mathbf{d}(\mathsf{X}_n, \mathsf{Y}_n) \right]$$
$$\leq \min \left[\rho^n (M_{\mathsf{C},\mathbf{n}_0} \mathbf{\Xi}(x, y, \mathbf{n}_0) + \mathbf{d}(x, y)), \rho^{n/2}(1 + \mathbf{d}(x, y)) + \lambda_1^{n/2} \mathbf{\Xi}(x, y, \mathbf{n}_0) \right] ,$$

where

$$\mathbf{\Xi}(x, y, \mathbf{n}_0) = W_1(x, y) + A_1 \lambda_1^{-\mathbf{n}_0} \{ \mathbf{n}_0 + \lambda_1^{-1} \}$$
$$\log(\rho) = \log(1 - \varepsilon) \log(\lambda_1) / [-\log(M_{\mathsf{C},\mathbf{n}_0}) + \log(1 - \varepsilon)] ,$$
$$M_{\mathsf{C},\mathbf{n}_0} = \sup_{(x,y) \in \mathsf{C}} \mathbf{\Xi}(x, y, \mathbf{n}_0) = \sup_{(x,y) \in \mathsf{C}} [W_1(x, y)] + A_1 \lambda_1^{-\mathbf{n}_0} \{ \mathbf{n}_0 + \lambda_1^{-1} \} . \tag{53}$$

In Theorem 29, we obtain geometric contraction for P in bounded Wasserstein metric $\mathbf{W_d}$ since $\mathbf{d}$ is assumed to be bounded. To obtain convergence associated with unbounded Wasserstein metric associated with $W_2 : \mathsf{Y}^2 \to [0, +\infty)$, we consider the next assumption which is a generalized drift condition linking W_2 and the bounded semi-metric $\mathbf{d}$.

H 3 (K). *There exist* $W_2 : \mathsf{Y}^2 \to [0, +\infty)$, $\lambda_2 \in (0, 1)$ *and* $A_2 \geq 0$ *such that for any* $x, y \in \mathsf{Y}$,
$$\mathrm{K}W_2(x, y) \leq \lambda_2 W_2(x, y) + A_2 \mathbf{d}(x, y) .$$

In the special case where $\mathbf{d}(x, y) = \mathbb{1}_{\Delta_\mathsf{Y}^c}((x, y))$, $W_2(x, y) = \mathbb{1}_{\Delta_\mathsf{Y}^c}((x, y)) W_1(x, y)$ and for any $x \in \mathsf{Y}$, $\mathrm{K}((x, x), \Delta_\mathsf{Y}) = 1$, we obtain that $\mathbf{D}_\mathbf{d}(W_1, \lambda_1, A_1, \mathsf{Y})$ implies **H** 3(K). The following result implies quantitative bounds on the Wasserstein distance $\mathbf{d}_{W_2}(\delta_x \mathrm{P}^n, \delta_y \mathrm{P}^n)$ for any $x, y \in \mathsf{Y}$ and $n \in \mathbb{N}^\star$.

Theorem 30. *Let* K *be a Markov coupling kernel for* P *and assume* **H**1, **H** 2(K) *and* **H**3(K). *Then for any* $n \in \mathbb{N}$ *and* $x, y \in \mathsf{Y}$,

$$\mathbb{E}_{(x,y)} \left[W_2(\mathsf{X}_n, \mathsf{Y}_n) \right] \leq \lambda_2^n W_2(x, y)$$
$$+ A_2 \min \left[\tilde{\rho}^n r_\rho (M_{\mathsf{C},\mathbf{n}_0} \mathbf{\Xi}(x, y, \mathbf{n}_0) + \mathbf{d}(x, y)), \tilde{\rho}^{n/2} r_\rho (1 + \mathbf{d}(x, y)) + \tilde{\lambda}^{n/2} r_\lambda \mathbf{\Xi}(x, y, \mathbf{n}_0) \right] ,$$

where
$$\begin{cases} \tilde{\rho} = \max(\lambda_2, \rho) \in (0, 1) , & \tilde{\lambda} = \max(\lambda_1, \lambda_2) \in (0, 1) , \\ r_\rho = \log^{-1}(1/\tilde{\rho})/\tilde{\rho} , & r_\lambda = \log^{-1}(1/\tilde{\lambda})/\tilde{\lambda} , \end{cases}$$

where $\mathbf{\Xi}(x, y, \mathbf{n}_0)$, $M_{\mathsf{C},\mathbf{n}_0}$ *and* ρ *are given in* (53).

Theorem 29 and Theorem 30 share some connections with [62, Theorem 5], [34] and [21] but hold under milder assumptions than the ones considered in these works. Compared to [34] and [21], the main difference is that we allow here only a contraction for the $\mathtt{n}_0$-th iterate of the Markov chain (condition $\mathbf{H}2$-(i)) which is necessary if we want to use Theorem 1 to obtain sharp quantitative convergence bounds for (1). Finally, [62, Theorem 5] also considers minorization condition for the the $\mathtt{n}_0$-th iterate, however our results compared favourably for large $\mathtt{n}_0$. Indeed, Theorem 29 implies that the rate of convergence $\min(\rho, \lambda_1)$ is of the form $C\mathtt{n}_0^{-1}$ for $C \geq 0$ independent of $\mathtt{n}_0$. Applying [62, Theorem 5], we found a rate of convergence of the form $C\mathtt{n}_0^{-2}$. Finally, a recent work [58] has established new results based on the technique used in [34]. However, we were not able to apply them since they assume as in [34], a contraction for $\mathtt{n}_0 = 1$ which does not imply sharp bounds on the situations we consider.

The rest of this section is devoted to the proof of Theorem 29 and Theorem 30. Denote by $\theta : (\mathsf{Y} \times \mathsf{Y})^{\mathbb{N}} \to (\mathsf{Y} \times \mathsf{Y})^{\mathbb{N}}$ the shift operator defined for any $(x_n, y_n)_{n\in\mathbb{N}} \in (\mathsf{Y} \times \mathsf{Y})^{\mathbb{N}}$ by $\theta((x_n, y_n)_{n\in\mathbb{N}}) = (x_{n+1}, y_{n+1})_{n\in\mathbb{N}}$. Define by recursion, for any $m \in \mathbb{N}$, the sequence of $(\mathcal{G}_n)_{n\in\mathbb{N}}$-stopping times $(T_{\mathsf{C},\mathtt{n}_0}^{(m)})_{m\in\mathbb{N}}$, with $T_{\mathsf{C},\mathtt{n}_0}^{(0)} = 0$ and for any $m \in \mathbb{N}^\star$

$$
\begin{aligned}
T_{\mathsf{C},\mathtt{n}_0}^{(m)} &= \inf\left\{k \geq T_{\mathsf{C},\mathtt{n}_0}^{(m-1)} + \mathtt{n}_0 \, : \, (\mathrm{X}_k, \mathrm{Y}_k) \in \mathsf{C}\right\} = T_{\mathsf{C},\mathtt{n}_0}^{(m-1)} + \mathtt{n}_0 + \tilde{T}_{\mathsf{C}} \circ \theta^{T_{\mathsf{C}}^{(m-1)}+\mathtt{n}_0} \\
&= T_{\mathsf{C},\mathtt{n}_0}^{(1)} + \sum_{i=1}^{(m-1)} T_{\mathsf{C},\mathtt{n}_0}^{(1)} \circ \theta^{T_{\mathsf{C},\mathtt{n}_0}^{(i)}} \, , \\
\tilde{T}_{\mathsf{C}} &= \inf\left\{k \geq 0 \, : \, (\mathrm{X}_k, \mathrm{Y}_k) \in \mathsf{C}\right\} \, .
\end{aligned}
\tag{54}
$$

Note that $(T_{\mathsf{C},\mathtt{n}_0}^{(m)})_{m\in\mathbb{N}^\star}$ are the successive return times to C delayed by $\mathtt{n}_0 - 1$ and $\tilde{T}_{\mathsf{C}}$ is the first hitting time to C. We will use the following lemma which borrows from [19] and [38, Lemma 3.1].

Lemma 31 ([19, Proposition 14]). *Let* K *be a Markov coupling kernel for* P *and assume* $\mathbf{H}2(\mathrm{K})$-(i)-(ii). *Then for any* $n, m \in \mathbb{N}$, $x, y \in \mathsf{Y}$,

$$
\mathbb{E}_{(x,y)}\left[\mathbf{d}(\mathrm{X}_n, \mathrm{Y}_n)\right] \leq (1-\varepsilon)^m \mathbf{d}(x,y) + \mathbb{E}_{(x,y)}\left[\mathbf{d}(\mathrm{X}_n, \mathrm{Y}_n)\mathbb{1}_{[n,+\infty]}(T_{\mathsf{C},\mathtt{n}_0}^{(m)})\right] \, .
$$

Proof. Using $\mathbf{H}2(\mathrm{K})$-(ii), we have that $(\mathbf{d}(\mathrm{X}_n, \mathrm{Y}_n))_{n\in\mathbb{N}}$ is a $(\mathcal{G}_n)_{n\in\mathbb{N}}$-supermartingale and therefore using the strong Markov property and $\mathbf{H}2(\mathrm{K})$-(i) we obtain for any $m \in \mathbb{N}$ and $x, y \in \mathsf{Y}$ that

$$
\mathbb{E}_{(x,y)}\left[\mathbf{d}(\mathrm{X}_{T_{\mathsf{C},\mathtt{n}_0}^{(m+1)}}, \mathrm{Y}_{T_{\mathsf{C},\mathtt{n}_0}^{(m+1)}})\right] \leq \mathbb{E}_{(x,y)}\left[\mathbb{E}\left[\mathbf{d}(\mathrm{X}_{T_{\mathsf{C},\mathtt{n}_0}^{(m)}+\mathtt{n}_0}, \mathrm{Y}_{T_{\mathsf{C},\mathtt{n}_0}^{(m)}+\mathtt{n}_0}) \Big| \mathcal{G}_{T_{\mathsf{C},\mathtt{n}_0}^{(m)}}\right]\right]
$$
$$
\leq (1-\varepsilon)\mathbb{E}_{(x,y)}\left[\mathbf{d}(\mathrm{X}_{T_{\mathsf{C},\mathtt{n}_0}^{(m)}}, \mathrm{Y}_{T_{\mathsf{C},\mathtt{n}_0}^{(m)}})\right] \, . \tag{55}
$$

Therefore by recursion and using (55) we obtain that for any $m \in \mathbb{N}$ and $x, y \in \mathsf{Y}$

$$\mathbb{E}_{(x,y)} \left[\mathbf{d}(\mathrm{X}_{T_{\mathsf{C},\mathbf{n}_0}^{(m)}}, \mathrm{Y}_{T_{\mathsf{C},\mathbf{n}_0}^{(m)}}) \right] \leq (1 - \varepsilon)^m \mathbf{d}(x, y) . \tag{56}$$

For any $n, m \in \mathbb{N}$ we have using (56) and that $(\mathbf{d}(X_n, Y_n))_{n \in \mathbb{N}}$ is a supermartingale,

$$\mathbb{E}_{(x,y)} \left[\mathbf{d}(\mathrm{X}_n, \mathrm{Y}_n) \right] \leq \mathbb{E}_{(x,y)} \left[\mathbf{d}(\mathrm{X}_{n \wedge T_{\mathsf{C},\mathbf{n}_0}^{(m)}}, \mathrm{Y}_{n \wedge T_{\mathsf{C},\mathbf{n}_0}^{(m)}}) \right]$$
$$\leq \mathbb{E}_{(x,y)} \left[\mathbf{d}(\mathrm{X}_{T_{\mathsf{C},\mathbf{n}_0}^{(m)}}, \mathrm{Y}_{T_{\mathsf{C},\mathbf{n}_0}^{(m)}}) \mathbb{1}_{[0,n]}(T_{\mathsf{C},\mathbf{n}_0}^{(m)}) \right] + \mathbb{E}_{(x,y)} \left[\mathbf{d}(\mathrm{X}_n, \mathrm{Y}_n) \mathbb{1}_{[n,+\infty]}(T_{\mathsf{C},\mathbf{n}_0}^{(m)}) \right]$$
$$\leq (1 - \varepsilon)^m \mathbf{d}(x, y) + \mathbb{E}_{(x,y)} \left[\mathbf{d}(\mathrm{X}_n, \mathrm{Y}_n) \mathbb{1}_{[n,+\infty]}(T_{\mathsf{C},\mathbf{n}_0}^{(m)}) \right] .$$

$\square$

By Lemma 31 and since $\mathbf{d}$ is bounded by 1, we need to obtain a bound on $\mathbb{P}_{(x,y)}(T_{\mathsf{C},\mathbf{n}_0}^{(m)} \geq n)$ for $x, y \in \mathsf{Y}$ and $m, n \in \mathbb{N}^\star$. For this, we will use the following proposition which gives an upper bound on exponential moment of the hitting times $(T_{\mathsf{C},\mathbf{n}_0}^{(m)})_{m \in \mathbb{N}^\star}$.

Lemma 32. *Let* K *be a Markov coupling kernel for* P *and assume* **H**2(K)-(iii)*. Then for any* $x, y \in \mathsf{Y}$ *and* $m \in \mathbb{N}^\star$,

$$\mathbb{E}_{(x,y)} \left[\lambda_1^{-T_{\mathsf{C},\mathbf{n}_0}^{(1)}} \right] \leq W_1(x, y) + A_1 \lambda_1^{-\mathbf{n}_0} \{ \mathbf{n}_0 + \lambda_1^{-1} \} , \qquad \mathbb{E}_{(x,y)} \left[\lambda_1^{-T_{\mathsf{C},\mathbf{n}_0}^{(m)} + T_{\mathsf{C},\mathbf{n}_0}^{(1)}} \right] \leq M_{\mathsf{C},\mathbf{n}_0}^{m-1} ,$$

$$\mathbb{E}_{(x,y)} \left[\lambda_1^{-T_{\mathsf{C},\mathbf{n}_0}^{(m)}} \right] \leq \left[W_1(x, y) + A_1 \lambda_1^{-\mathbf{n}_0} \{ \mathbf{n}_0 + \lambda_1^{-1} \} \right] M_{\mathsf{C},\mathbf{n}_0}^{m-1} ,$$

where $M_{\mathsf{C},\mathbf{n}_0}$ *is defined by* (53).

Proof. Let $x, y \in \mathsf{Y}$ and $(\mathrm{S}_n)_{n \in \mathbb{N}}$ be defined for any $n \in \mathbb{N}$ by, $\mathrm{S}_n = \lambda_1^{-n} W_1(\mathrm{X}_n, \mathrm{Y}_n)$. For any $n \in \mathbb{N}$ we have using **H**2(K)-(iii) and the weak Markov property

$$\mathbb{E}\left[\mathrm{S}_{n+1} \mid \mathcal{G}_n \right] \leq \lambda_1^{-n} W_1(\mathrm{X}_n, \mathrm{Y}_n) + A_1 \lambda_1^{-(n+1)} \mathbb{1}_{\mathsf{C}}((\mathrm{X}_n, \mathrm{Y}_n))$$
$$\leq \mathrm{S}_n + A_1 \lambda_1^{-(n+1)} \mathbb{1}_{\mathsf{C}}((\mathrm{X}_n, \mathrm{Y}_n)) . \tag{57}$$

Using the weak Markov property, the comparison theorem [16, Theorem 4.3.1], (57)

and the definition of $T_{\mathsf{C},\mathbf{n}_0}^{(1)}$ given in (54), $\mathbf{H}2(\mathrm{K})$-(iii) we obtain that

$$
\mathbb{E}_{(x,y)}\left[\mathrm{S}_{T_{\mathsf{C},\mathbf{n}_0}^{(1)}}\right] \leq \mathbb{E}_{(x,y)}\left[\mathbb{E}_{(x,y)}\left[\mathrm{S}_{T_{\mathsf{C},\mathbf{n}_0}^{(1)}}\,|\mathcal{G}_{\mathbf{n}_0}\right]\right] \leq \mathbb{E}_{(x,y)}\left[\mathbb{E}_{(x,y)}\left[\mathrm{S}_{\mathbf{n}_0+\tilde{T}_{\mathsf{C}}\circ\theta^{\mathbf{n}_0}}\,|\mathcal{G}_{\mathbf{n}_0}\right]\right]
$$

$$
\leq \mathbb{E}_{(x,y)}\left[\lambda_1^{-\mathbf{n}_0}\,\mathbb{E}_{(x,y)}\left[W_1(X_{\mathbf{n}_0+T_{\mathsf{C},\mathbf{n}_0}^{(1)}}, Y_{\mathbf{n}_0+T_{\mathsf{C},\mathbf{n}_0}^{(1)}})\lambda_1^{-T_{\mathsf{C},\mathbf{n}_0}^{(1)}}\,|\mathcal{G}_{\mathbf{n}_0}\right]\right] \leq \mathbb{E}_{(x,y)}\left[\lambda_1^{-\mathbf{n}_0}\,\mathbb{E}_{(X_{\mathbf{n}_0},Y_{\mathbf{n}_0})}\left[\mathrm{S}_{\tilde{T}_{\mathsf{C}}}\right]\right]
$$

$$
\leq \mathbb{E}_{(x,y)}\left[\lambda_1^{-\mathbf{n}_0}\,\mathbb{E}_{(X_{\mathbf{n}_0},Y_{\mathbf{n}_0})}\left[\mathrm{S}_0 + A_1\sum_{k=0}^{\tilde{T}_{\mathsf{C}}-1}\lambda_1^{-(k+1)}\mathbb{1}_{\mathsf{C}}((X_k,Y_k))\right]\right]
$$

$$
\leq \mathbb{E}_{(x,y)}\left[\lambda_1^{-\mathbf{n}_0}W_1(X_{\mathbf{n}_0},Y_{\mathbf{n}_0}) + A_1\lambda_1^{-(\mathbf{n}_0+1)}\right] \leq W_1(x,y) + A_1\lambda_1^{-\mathbf{n}_0}\{\mathbf{n}_0 + \lambda_1^{-1}\}\,. \qquad (58)
$$

Combining (58) and the fact that for any $x,y \in \mathsf{Y}$, $W_1(x,y) \geq 1$, we obtain that

$$
\mathbb{E}_{(x,y)}\left[\lambda_1^{-T_{\mathsf{C},\mathbf{n}_0}^{(1)}}\right] \leq W_1(x,y) + A_1\lambda_1^{-\mathbf{n}_0}\{\mathbf{n}_0 + \lambda_1^{-1}\}\,. \qquad (59)
$$

We conclude by a straightforward recursion and using (59), the definition of $T_{\mathsf{C},\mathbf{n}_0}^{(m)}$ (54) for $m \geq 1$, the strong Markov property and the fact that for any $m \in \mathbb{N}^{\star}$, $(X_{T_{\mathsf{C},\mathbf{n}_0}^{(m)}}, Y_{T_{\mathsf{C},\mathbf{n}_0}^{(m)}}) \in \mathsf{C}$. $\qquad\square$

Proof of Theorem 29. Let $x,y \in \mathsf{Y}$ and $n \in \mathbb{N}$. By Lemma 31, Lemma 32, $\mathbf{H}1$, the fact that $M_{\mathsf{C},\mathbf{n}_0} \geq 1$ and the Markov inequality, we have for any $m \in \mathbb{N}$,

$$
\mathbb{E}_{(x,y)}\left[\mathbf{d}(X_n,Y_n)\right] \leq (1-\varepsilon)^m\mathbf{d}(x,y) + \mathbb{P}_{(x,y)}\left[T_{\mathsf{C},\mathbf{n}_0}^{(m)} \geq n\right]
$$

$$
\leq (1-\varepsilon)^m\mathbf{d}(x,y) + \lambda_1^n\,\mathbb{E}_{(x,y)}\left[\lambda_1^{-T_{\mathsf{C},\mathbf{n}_0}^{(m)}}\right]
$$

$$
\leq (1-\varepsilon)^m\mathbf{d}(x,y) + \lambda_1^n M_{\mathsf{C},\mathbf{n}_0}^m\Xi(x,y,\mathbf{n}_0)\,,
$$

where $\Xi(x,y,\mathbf{n}_0)$ is given in Theorem 29. Combining this result and Lemma 32, we can conclude that $\mathbb{E}_{(x,y)}\left[\mathbf{d}(X_n,Y_n)\right] \leq \rho^n(M_{\mathsf{C},\mathbf{n}_0}\Xi(x,y,\mathbf{n}_0) + \mathbf{d}(x,y))$ setting

$$
m = \lceil n\log(\lambda_1)/\{\log(1-\varepsilon) - \log(M_{\mathsf{C},\mathbf{n}_0})\}\rceil\,.
$$

To show that $\mathbb{E}_{(x,y)}\left[\mathbf{d}(X_n,Y_n)\right] \leq \rho^{n/2}(1+\mathbf{d}(x,y)) + \lambda_1^{n/2}\Xi(x,y,\mathbf{n}_0)$, first note that Lemma 31 and $\mathbf{H}1$ imply that for any $m \in \mathbb{N}$,

$$
\mathbb{E}_{(x,y)}\left[\mathbf{d}(X_n,Y_n)\right] \leq (1-\varepsilon)^m\mathbf{d}(x,y) + \mathbb{P}_{(x,y)}\left[T_{\mathsf{C},\mathbf{n}_0}^{(m)} - T_{\mathsf{C},\mathbf{n}_0}^{(1)} \geq n/2\right] + \mathbb{P}_{(x,y)}\left[T_{\mathsf{C},\mathbf{n}_0}^{(1)} \geq n/2\right]
$$

$$
\leq (1-\varepsilon)^m\mathbf{d}(x,y) + \lambda_1^{n/2}\mathbb{E}_{(x,y)}\left[\lambda_1^{-T_{\mathsf{C},\mathbf{n}_0}^{(m)}+T_{\mathsf{C},\mathbf{n}_0}^{(1)}}\right] + \lambda_1^{n/2}\mathbb{E}_{(x,y)}\left[\lambda_1^{-T_{\mathsf{C},\mathbf{n}_0}^{(1)}}\right]\,,
$$

where we have used the Markov inequality in the last line. Combining this result and Lemma 32, we can conclude that $\mathbb{E}_{(x,y)}\left[\mathbf{d}(\mathrm{X}_n, \mathrm{Y}_n)\right] \le \rho^{n/2}(1 + \mathbf{d}(x,y)) + \lambda_1^{n/2}\boldsymbol{\Xi}(x,y,\mathtt{n}_0)$ setting

$$m = \lceil n\log(\lambda_1)/\{2\log(1-\varepsilon) - 2\log(M_{\mathsf{C},\mathtt{n}_0})\}\rceil \ .$$

$\square$

Proof of Theorem 30. Let $x, y \in \mathsf{Y}$ and $n \in \mathbb{N}$. Using **H**3(K), we obtain by recursion

$$\mathbb{E}_{(x,y)}\left[W_2(X_n, Y_n)\right] \le \lambda_2^n W_2(x,y) + A_2 \sum_{k=0}^{n-1} \lambda_2^{n-1-k} \, \mathbb{E}_{(x,y)}\left[\mathbf{d}(X_k, Y_k)\right] \ . \qquad (60)$$

Applying Theorem 29 we obtain

$$\sum_{k=0}^{n-1} \lambda_2^{n-1-k} \, \mathbb{E}_{(x,y)}\left[\mathbf{d}(X_k, Y_k)\right]$$

$$\le \sum_{k=0}^{n-1} \lambda_2^{n-1-k} \min\left[\rho^k(M_{\mathsf{C},\mathtt{n}_0}\boldsymbol{\Xi}(x,y,\mathtt{n}_0) + \mathbf{d}(x,y)), \rho^{k/2}(1 + \mathbf{d}(x,y)) + \lambda^{k/2}\boldsymbol{\Xi}(x,y,\mathtt{n}_0)\right]$$

$$\le \min\left[n\tilde{\rho}^{n-1}(M_{\mathsf{C},\mathtt{n}_0}\boldsymbol{\Xi}(x,y,\mathtt{n}_0) + \mathbf{d}(x,y)), n\tilde{\rho}^{(n-1)/2}(1 + \mathbf{d}(x,y)) + n\tilde{\lambda}^{(n-1)/2}\boldsymbol{\Xi}(x,y,\mathtt{n}_0)\right] \ .$$

We conclude plugging this result in (60) and using that for any $n \in \mathbb{N}$ and $t \in (0,1)$, $nt^n \le \log^{-1}(1/t)$.

$\square$

D   Tamed Euler-Maruyama discretization

In this subsection we consider the following assumption.

T 1. $\mathsf{X} = \mathbb{R}^d$ *and* $\Pi = \mathrm{Id}$ *and*

$$\mathcal{T}_\gamma(x) = x + \gamma b(x)/(1 + \gamma\,\|b(x)\|) \text{ for any } \gamma > 0 \text{ and } x \in \mathbb{R}^d \ .$$

Here, we focus on drift b which is no longer assumed to be Lipschitz. Therefore, the ergodicity results obtained in Section 2.2 no longer hold since the minorization condition we derived relied heavily on one-sided Lipschitz condition or Lipschitz regularity for b. Here we focus on the following assumption on b.

T 2. *There exists* $\tilde{\mathsf{L}}, \tilde{\ell} \ge 0$ *such that for any* $x, y \in \mathbb{R}^d$

$$\|b(x) - b(y)\| \le \tilde{\mathsf{L}}(1 + \|x\|^{\tilde{\ell}} + \|y\|^{\tilde{\ell}})\,\|x - y\| \ .$$

In addition, assume that $b(0) = 0$ *and* $M_{\tilde{\ell}} = \sup_{x \in \mathbb{R}^d} (1 + \|x\|^{\tilde{\ell}})(1 + \|b(x)\|)^{-1} < +\infty.$

Proposition 33. *Assume* **T**1 *and* **T**2 *then* $\mathbf{A}2(\mathbb{R}^{2d})$-(iii) *holds with* $\bar{\gamma} > 0$ *and for any* $\gamma \in (0, \bar{\gamma}]$, $\kappa(\gamma) = 2\tilde{\mathsf{L}}_\gamma + \gamma\tilde{\mathsf{L}}_\gamma^2$ *where*

$$\tilde{\mathsf{L}}_\gamma = 2\gamma^{-1}M_{\tilde{\ell}}(1 + M_{\tilde{\ell}})\tilde{\mathsf{L}} .$$

Proof. Let $x, y \in \mathbb{R}^d$ and assume that $\|x\| \geq \|y\|$. We have the following inequalities

$$\left\| \frac{b(x)}{1 + \gamma \|b(x)\|} - \frac{b(y)}{1 + \gamma \|b(y)\|} \right\| \leq \frac{\|b(x) - b(y)\|}{1 + \gamma \|b(x)\|} + \left| \frac{\|b(y)\|}{1 + \gamma \|b(x)\|} - \frac{\|b(y)\|}{1 + \gamma \|b(y)\|} \right|$$

$$\leq \gamma^{-1}2M_{\tilde{\ell}}\tilde{\mathsf{L}} \|x - y\| + \gamma \frac{\|b(y)\| \|b(x) - b(y)\|}{(1 + \gamma \|b(x)\|)(1 + \gamma \|b(y)\|)}$$

$$\leq \gamma^{-1}M_{\tilde{\ell}}(1 + M_{\tilde{\ell}})\tilde{\mathsf{L}} \|x - y\| .$$

The same inequality holds with $\|y\| \geq \|x\|$. Therefore, we have

$$\|\mathcal{T}_\gamma(x) - \mathcal{T}_\gamma(y)\|^2 \leq \left(1 + 2\gamma\tilde{\mathsf{L}}_\gamma + \gamma^2\tilde{\mathsf{L}}_\gamma^2\right) \|x - y\|^2 ,$$

which concludes the proof. $\qquad\square$

Proposition 33 implies that the conclusions of Proposition 3-(c) hold. Note that contrary to the conclusion of Proposition 8, we do not get that $\sup_{\gamma \in (0,\bar{\gamma}]} \kappa(\gamma) < +\infty$. Hence we have for any $\tilde{\ell} \in \mathbb{N}^\star$, $\inf_{\gamma \in (0,\bar{\gamma}]} \alpha_+(\kappa, \gamma, \tilde{\ell}) = 0$.

T 3. *There exist* $\tilde{R}$ *and* $\tilde{\mathfrak{m}}^+$ *For any* $x \in \bar{\mathrm{B}}(0, \tilde{R})^{\mathrm{c}}$,

$$\langle b(x), x \rangle \leq -\tilde{\mathfrak{m}}^+ \|x\| \|b(x)\| .$$

Under **T** 2 and **T** 3 it is shown in [6] that there exists $\bar{\gamma} > 0$, $\lambda \in (0, 1)$ and $A \geq 0$ such that for any $\gamma \in (0, \bar{\gamma}]$, R_γ satisfies $\mathbf{D}_{\mathrm{d}}(V, \lambda^\gamma, A\gamma, \mathsf{X})$ with $V(x) = \exp(a(1 + \|x\|^2)^{1/2})$ for some fixed a.

Theorem 34. *Assume* **T**2 *and* **T**3 *then there exists* $\bar{\gamma} > 0$ *such that for any* $\gamma \in (0, \bar{\gamma}]$ *there exists* $C_\gamma \geq 0$ *and* $\rho_\gamma \in (0, 1)$ *with for any* $\gamma \in (0, \bar{\gamma}]$, $x, y \in \mathbb{R}^d$ *and* $k \in \mathbb{N}$

$$\|\delta_x \mathrm{R}_\gamma^k - \delta_y \mathrm{R}_\gamma^k\|_V \leq C_\gamma \rho_\gamma^{k\gamma} \{V(x) + V(y)\} .$$

Proof. The proof is a direct application of Theorem 6-(a). $\qquad\square$

It is shown in [6, Theorem 4] that the following result holds: there exists $V : \mathbb{R}^d \to [1, +\infty)$, $\bar{\gamma} > 0$, $C, D \geq 0$ and $\rho \in (0, 1)$ such that for any $k \in \mathbb{N}$, $\gamma \in (0, \bar{\gamma}]$ and $x \in \mathbb{R}^d$

$$\|\delta_x \mathrm{R}_\gamma^k - \pi\|_V \leq C\rho^{k\gamma}V(x) + D\sqrt{\gamma} ,$$

where π is an invariant measure for the diffusion with drift b.

E Explicit rates and asymptotics in Theorem 14

We recall that b satisfies

$$\langle b(x), x \rangle \leq -\mathbf{m}_1 \|x\| \mathbb{1}_{\bar{B}(0,R_3)^c}(x) - \mathbf{m}_2 \|b(x)\|^2 + c/2 \;,$$

with $\mathbf{m}_1, \mathbf{m}_2 > 0$ and $R_3, c \geq 0$. Therefore, by Proposition 13, K_γ satisfies for any $\gamma \in (0, \bar{\gamma}]$, $\mathbf{D}_\mathrm{d}(W, \lambda^\gamma, A\gamma, \mathsf{X}^2)$ where $\bar{\gamma} \in (0, 2\mathbf{m}_2)$, $R_4 = \max(1, R_3, (d+c)/\mathbf{m}_1)$, $\mathbf{m}_3 \in (0, \mathbf{m}_1/2]$ and

$$\lambda = e^{-\mathbf{m}_3^2/2}\;, \quad A = \exp\left[\bar{\gamma}(\mathbf{m}_3(d+c) + \mathbf{m}_3^2)/2 + \mathbf{m}_3(1 + R_4^2)^{1/2}\right](\mathbf{m}_3(d+c)/2 + \mathbf{m}_3^2)\;.$$

We now discuss the rates given by Theorem 14. Let $\tilde{\rho}_{\bar{\gamma},c} \in (0,1)$ and $\tilde{C}_{\bar{\gamma},c} \geq 0$ the constants given by Theorem 14 such that for any $\gamma \in (0, \bar{\gamma}]$, $k \in \mathbb{N}$ and $x, y \in \mathsf{X}$

$$\mathbf{d}_W(\delta_x \mathrm{R}_\gamma^k, \delta_y \mathrm{R}_\gamma^k) \leq \tilde{C}_{\bar{\gamma},c} \tilde{\rho}_{\bar{\gamma},c}^{\left\lfloor k(\ell \lceil 1/\gamma \rceil)^{-1}\right\rfloor} W(x,y)\;.$$

Using that $W(x,y) = \{V(x) + V(y)\}/2$, the fact that $\lfloor k/(\ell \lceil 1/\gamma \rceil)\rfloor \geq k\gamma/(\ell(1 + \bar{\gamma})) - 1$, and letting $C_{\bar{\gamma},c} = \tilde{C}_{\bar{\gamma}} \tilde{\rho}_{\bar{\gamma},c}^{-1}/2$ and $\rho_{\bar{\gamma},c} = \tilde{\rho}_{\bar{\gamma},c}^{1/(\ell(1+\bar{\gamma}))}$, we obtain that for any $\gamma \in (0,\bar{\gamma}]$, $k \in \mathbb{N}$ and $x, y \in \mathsf{X}$

$$\|\delta_x \mathrm{R}_\gamma^k - \delta_y \mathrm{R}_\gamma^k\|_V \leq C_{\bar{\gamma},c} \rho_{\bar{\gamma},c}^{k\gamma}\{V(x) + V(y)\}\;.$$

If $\lim_{\bar{\gamma} \to 0} \ell$ exists and does not depend on $\mathtt{L}$, using Theorem 6-(c), we obtain that the following limits exist and do not depend on $\mathtt{L}$

$$\rho_c = \lim_{\bar{\gamma} \to 0} \rho_{\bar{\gamma},c}\;, \qquad C_c = \lim_{\bar{\gamma} \to 0} C_{\bar{\gamma},c}\;. \tag{61}$$

We now discuss the dependency of ρ_c with respect to the problem constants, depending on the sign of $\mathbf{m}$, based on Theorem 6-(c). Let $\vartheta = \mathbf{m}_3 \exp[\mathbf{m}_3(1 + R_4^2)]\{(d+c)/2 + \mathbf{m}_3\}\{1 + 2/\mathbf{m}_3^2\}$ and consider $\mathbf{m}_1 \leq \sqrt{8\log(2)}$ and $A \geq 1$:

(a) if and for any $\gamma \in (0, \bar{\gamma}]$, $x, y \in \mathsf{X}$, $\|x + \gamma b(x) - y - \gamma b(y)\| \leq \|x - y\|$, we set $\ell = \lceil M_\mathrm{d}^2 \rceil$ and we have

$$\log(\rho_c) \leq 12 \log(2) \log(6\vartheta)(1 + 4\log(4\vartheta)/\mathbf{m}_3)/\left[\mathbf{m}_3^2 \Phi\{-1/2\}\right]\;.$$

(b) if $\mathbf{m} < 0$, we set $\ell = 1$ and we have

$$\log(\rho_c) \leq 12 \log(2) \log(6\vartheta)/\left[\mathbf{m}_3^2 \Phi\{-(-\mathbf{m})^{1/2}\log(4\vartheta)/(2\mathbf{m}_3^2(1 - e^{2\mathbf{m}})^{1/2})\}\right]\;.$$

Note that a similar result was already proven in [23, Theorem 10] but the scheme of the proof was different as the authors compared the discretization to the associated diffusion process and used the contraction of the continuous process.

F Postponed Proofs

F.1 Proof of Lemma 2

Let $\ell \in \mathbb{N}^\star$ and $\gamma \in (0, \bar{\gamma}]$. First note that the following equalities hold if $\kappa(\gamma) \neq 0$

$$
\begin{aligned}
\Xi_{\ell\lceil 1/\gamma\rceil}(\kappa) &= \gamma \sum_{i=1}^{\ell\lceil 1/\gamma\rceil} (1 + \gamma\kappa(\gamma))^{-i} \\
&= \gamma(1 + \gamma\kappa(\gamma))^{-1} \frac{1 - (1 + \gamma\kappa(\gamma))^{-\ell\lceil 1/\gamma\rceil}}{1 - (1 + \gamma\kappa(\gamma))^{-1}} \\
&= -\kappa^{-1}(\gamma) \left\{ [1 + \gamma\kappa(\gamma)]^{-\ell\lceil 1/\gamma\rceil} - 1 \right\} \\
&= -\kappa^{-1}(\gamma) \left\{ \exp\left[-\ell\lceil 1/\gamma\rceil \log\left\{1 + \gamma\kappa(\gamma)\right\}\right] - 1 \right\} .
\end{aligned}
\tag{62}
$$

We now give a lower-bound on $\Xi_{\ell\lceil 1/\gamma\rceil}(\kappa)$ depending on the condition satisfied by $\gamma \mapsto \kappa(\gamma)$.

(a) Assume that for any $\tilde{\gamma} \in (0, \bar{\gamma}]$, $\kappa(\tilde{\gamma}) < 0$. Using that $\log(1 - t) \leq -t$ for $t \in (0, 1)$, we obtain that

$$
\exp\left[-\ell\lceil 1/\gamma\rceil \log\left\{1 + \gamma\kappa(\gamma)\right\}\right] \geq \exp(-\ell\lceil 1/\gamma\rceil \gamma\kappa(\gamma)) \geq \exp(-\ell\kappa(\gamma)) ,
$$

which together with (62) concludes the proof for (a).

(b) Assume that for any $\tilde{\gamma} \in (0, \bar{\gamma}]$, $\kappa(\tilde{\gamma}) \leq 0$. Then,

$$
\Xi_{\ell\lceil 1/\gamma\rceil}(\kappa) = \gamma \sum_{i=1}^{\ell\lceil 1/\gamma\rceil} (1 + \gamma\kappa(\gamma))^{-i} = \gamma \lceil 1/\gamma\rceil \ell \geq \ell .
$$

(c) Assume that for any $\tilde{\gamma} \in (0, \bar{\gamma}]$, $\kappa(\tilde{\gamma}) > 0$. Using that $\log(1 + at)/t \geq a/(1 + at)$ for $t, a > 0$, we obtain that

$$
\begin{aligned}
\exp\left[-\ell\lceil 1/\gamma\rceil \log\left\{1 + \gamma\kappa(\gamma)\right\}\right] &\leq \exp\left[-(\ell/\gamma) \log\left\{1 + \gamma\kappa(\gamma)\right\}\right] \\
&\leq \exp\left[-\ell\kappa(\gamma)/(1 + \gamma\kappa(\gamma))\right] ,
\end{aligned}
$$

which concludes the proof for (c).

F.2 Proof of Proposition 8

(a) By **B**2 and **B**3(m) we have for any $\gamma > 0$ and $x, y \in \mathsf{X}$, $\|\mathcal{T}_\gamma(x) - \mathcal{T}_\gamma(y)\|^2 \leq (1 - 2\gamma\mathtt{m} + \gamma^2\mathtt{L}^2) \|x - y\|^2 \leq (1 + \gamma\kappa(\gamma)) \|x - y\|^2$, which concludes the proof.

(b) By (18), we have for any $\gamma > 0$ and $x, y \in \mathsf{X}$, $\|\mathcal{T}_\gamma(x) - \mathcal{T}_\gamma(y)\|^2 \leq \|x - y\|^2 + \gamma(-2\mathtt{m}_b + \gamma) \|b(x) - b(y)\|^2$. Then if $\gamma \leq 2\mathtt{m}_b$, $\|\mathcal{T}_\gamma(x) - \mathcal{T}_\gamma(y)\|^2 \leq \|x - y\|^2$, which concludes the proof.

F.3 Proof of Proposition 11

Let $\gamma \in (0, \bar{\gamma}]$ and $x \in \mathsf{X}$. Using (1), **B**2, **B**3(m), **C**2 and that the projection Π_X is non expansive we obtain the following inequalities

$$
\begin{aligned}
\mathrm{R}_\gamma V(x) &\leq 1 + \|x + \gamma b(x)\|^2 + \gamma d \\
&\leq 1 + \|x\|^2 + 2\gamma\langle x, b(x)\rangle + \gamma^2 \|b(x)\|^2 + \gamma d \\
&\leq (1 + \|x\|^2) \left[1 - \gamma(2\mathtt{m}_2^+ - \bar{\gamma}\mathtt{L}^2)\right] + \gamma \left(d + 2R_2^2(\mathtt{m}_2^+ - \mathtt{m})_+ + 2\mathtt{m}_2^+\right) ,
\end{aligned}
$$

which concludes the proof using Lemma 28.

F.4 Proof of Theorem 12

Let $\bar{\gamma} \in (0, 2\mathtt{m}^+/\mathtt{L}^2)$. Using Proposition 11 we obtain that for any $\gamma \in (0, \bar{\gamma}]$, K_γ satisfies $\mathbf{D}_\mathrm{d}(W, \lambda^\gamma, A\gamma, \mathsf{X})$, with λ and A explicitly given by in Proposition 11. For any $x, y \in \mathsf{X}$ such that $\|x - y\| \geq M_\mathrm{d}$, either $V(x) \geq 2K_\mathrm{d}$ or $V(y) \geq 2K_\mathrm{d}$, and therefore $W(x, y) > K_\mathrm{d}$. Note also that (12) is satisfied for any $\gamma \in (0, \bar{\gamma}]$ and $x, y \in \mathsf{X}$ and Theorem 1 and Lemma 2. Therefore we can apply Theorem 6, which concludes the proof.

F.5 Proof of Proposition 13

Let $\gamma \in (0, \bar{\gamma}]$. Using the fact that Π_X is non expansive, the Log-Sobolev inequality [4, Theorem 5.5] and the Jensen inequality we obtain for any $x \in \mathbb{R}^d$

$$
\begin{aligned}
\mathrm{R}_\gamma V(x) &\leq \exp\left[\mathtt{m}_3\mathrm{R}_\gamma\phi(x) + \gamma\mathtt{m}_3^2/2\right] \leq \exp\left[\mathtt{m}_3\sqrt{1 + \mathrm{R}_\gamma \|x\|^2} + \gamma\mathtt{m}_3^2/2\right] \\
&\leq \exp\left[\mathtt{m}_3\sqrt{1 + \|\mathcal{T}_\gamma(x)\|^2 + \gamma d} + \gamma\mathtt{m}_3^2/2\right] .
\end{aligned}
\tag{63}
$$

Let $x \in \mathbb{R}^d$. The rest of the proof is divided in two parts.

(a) In the first case, $\|x\| \geq R_4$. Since $\|x\| \geq R_3$ and $\gamma \leq 2\mathtt{m}_2$, we have using $\mathbf{C}3$

$$\|\mathcal{T}_\gamma(x)\|^2 \leq \|x\|^2 - 2\gamma\mathtt{m}_1\|x\| + \gamma(\gamma - 2\mathtt{m}_2)\|b(x)\|^2 + \gamma c \leq \|x\|^2 - 2\gamma\mathtt{m}_1\|x\| + \gamma c \,. \quad (64)$$

Since $\|x\| \geq 1$ we have $2\|x\| \geq \phi(x)$ and therefore, using that $\|x\| \geq (d+c)/\mathtt{m}_1$, $2\mathtt{m}_1\|x\| \geq 2\mathtt{m}_3\phi(x) + d + c$. This inequality, combined with the fact that for any $t \in (-1, +\infty)$, $\sqrt{1+t} \leq 1 + t/2$, yields

$$\sqrt{1 + \|x\|^2 + \gamma(-2\mathtt{m}_1\|x\| + d + c)} - \phi(x) \leq \gamma(-2\mathtt{m}_1\|x\| + d + c)/(2\phi(x)) \leq -\mathtt{m}_3\gamma \,. \quad (65)$$

Combining (63), (64) and (65) we get

$$\mathrm{R}_\gamma V(x) \leq \lambda^\gamma V(x) \,.$$

(b) In the second case $\|x\| \leq R_4$. We have the following inequality using $\mathbf{C}3$ and that $\gamma \leq 2\mathtt{m}_2$

$$\|\mathcal{T}_\gamma(x)\|^2 \leq \|x\|^2 + \gamma(\gamma - 2\mathtt{m}_2)\|b(x)\|^2 + \gamma c \leq \|x\|^2 + \gamma c \,. \quad (66)$$

Combining (63), (66) and the fact that for any $t \in (-1, +\infty)$, $\sqrt{1+t} \leq 1 + t/2$ we get

$$\mathrm{R}_\gamma V(x) \leq \exp\left[\gamma\mathtt{m}_3(d+c)/(2\phi(x)) + \gamma\mathtt{m}_3^2/2\right] V(x) \leq \exp\left[\gamma(\mathtt{m}_3(d+c) + \mathtt{m}_3^2)/2\right] V(x) \,. \quad (67)$$

Note that for any $c_1 \geq c_2$ and $t \in [0, \bar{t}]$ we have the following inequality

$$\mathrm{e}^{c_1 t} \leq \mathrm{e}^{c_2 t} + \mathrm{e}^{c_1 \bar{t}}(c_1 - c_2)t \,. \quad (68)$$

Combining (67) and (68) we get

$$\mathrm{R}_\gamma V(x) \leq \lambda^\gamma V(x) + \exp\left[\bar{\gamma}(\mathtt{m}_3(d+c) + \mathtt{m}_3^2)/2\right](\mathtt{m}_3(d+c)/2 + \mathtt{m}_3^2)\exp(\mathtt{m}_3(1 + R_4^2)^{1/2})\gamma \,,$$

which concludes the proof.

F.6 Proof of Theorem 14

Let $\bar{\gamma} \in (0, 2\mathtt{m}_2]$. Using Proposition 11 we obtain that for any $\gamma \in (0, \bar{\gamma}]$, K_γ satisfies $\mathbf{D}_\mathrm{d}(W, \lambda^\gamma, A\gamma, \mathsf{X})$, with λ and A explicitly given in Proposition 11. Let

$$K_\mathrm{d} = 2A(1 + \bar{\gamma})(1 + \log^{-1}(1/\lambda)) \quad \text{and} \quad M_\mathrm{d} = 2\log(2K_\mathrm{d})/\mathtt{m}_3 \,.$$

Note that for any $x, y \in \mathsf{X}$ such that $\|x - y\| \geq M_\mathrm{d}$, either $V(x) \geq 2K_\mathrm{d}$ or $V(y) \geq 2K_\mathrm{d}$, and therefore $W(x, y) > K_\mathrm{d}$. The rest of the proof is similar to the one of Theorem 12.

F.7 Proof of Proposition 19

Let $p \in \mathbb{N}^\star$ and $V \in \mathrm{C}^2(\mathbb{R}^d, [1, +\infty))$ be defined for any $x \in \mathbb{R}^d$ by $V(x) = 1 + \|x\|^{2p}$. We have that $\nabla V(x) = 2p \|x\|^{2(p-1)} x$ and $\Delta V(x) = (4p(p-1) + 2pd) \|x\|^{2(p-1)}$. Therefore, using **B**3(m) and the definition of $\mathcal{A}$ we obtain that for any $x \in \mathbb{R}^d$

$$\mathcal{A}V(x) \le [2p(p-1) + p(d - 2\mathtt{m})] V(x) . \tag{69}$$

Hence, using (69) and [42, Theorem 3.5], we obtain that $(\mathbf{X}_t)_{t \ge 0}$ is defined for any $t \ge 0$, for any starting point $\mathbf{X}_0 = x \in \mathbb{R}^d$. In addition, we have:

(a) if there exists $\varepsilon_b > 0$ such that $\sup_{x \in \mathbb{R}^d} \|b(x)\|^{2(1+\varepsilon_b)} (1 + \|x\|^{2p})^{-1} < +\infty$, using (69) and Lemma 15-(a) we obtain that **L**3 holds.

(b) if there exists $\varepsilon_b > 0$ such that $\sup_{x \in \mathbb{R}^d} \|b(x)\| \, \mathrm{e}^{-\mathtt{m}_2^+ \|x\|^2} < +\infty$, and **C**2 holds, then we set $V \in \mathrm{C}^2(\mathbb{R}^d, [1, +\infty))$ where for any $x \in \mathbb{R}^d$, $V(x) = \mathrm{e}^{\mathtt{m}_2^+ \|x\|^2}$. We have the following inequalities for any $x \in \bar{\mathrm{B}}(0, R_2)^c$,

$$\mathcal{A}V(x) \le \mathtt{m}_2^+ \left[d + (2\mathtt{m}_2^+ - 4\mathtt{m}_2^+/2) \|x\|^2 \right] V(x) \le \mathtt{m}_2^+ dV(x) . \tag{70}$$

Setting $\zeta = (\mathtt{m}_2^+ d) \vee \sup_{x \in \bar{\mathrm{B}}(0, R_2)} \mathcal{A}V(x)$, we obtain that V satisfies $\mathbf{D}_c(V, \zeta, 0)$. Therefore using (70) and Lemma 15-(a) , we obtain that **L**3 holds.

F.8 Proof of Proposition 20

(a) It is easy to check that **L**2 holds with $\beta = 2\alpha$. **L**4 is trivial. It only remains to show that **L**5 holds. Let $W(x) = \|x\|$, $n \in \mathbb{N}$, $\gamma \in (0, \bar{\gamma}]$, $x \in \mathbb{R}^d$ and $X = x + \gamma b_{\gamma,n}(x) + \sqrt{\gamma} Z$, where Z is a d-dimensional Gaussian random variable with zero mean and covariance identity. We have the following inequality using **B**3(m)

$$\mathbb{E}\left[\|X\|^2 \right] \le \|x\|^2 - 2\gamma \mathtt{m}_0 \Phi_n(x) \|x\|^2 + \gamma^2 \Phi_n(x)^2 \|b(x)\|^2 + \gamma d , \tag{71}$$

with $\Phi_n(x) = \varphi_n(x) + (1 - \varphi_n(x))(1 + \gamma^\alpha \|b(x)\|)^{-1} \le 1$. We also have that for any $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$

$$\Phi_n(x) \|b(x)\| \le \mathtt{L}_{n+1} \|x\| + \gamma^{-\alpha} . \tag{72}$$

Combining (71) and (72) we obtain

$$\mathbb{E}\left[1 + \|X\|^2 \right] \le (1 + \|x\|^2) \left[1 - 2\gamma \mathtt{m}_0 + 2\gamma^2 \mathtt{L}_{n+1}^2 \right] + 2\gamma^{2-2\alpha} + \gamma d , \tag{73}$$

which concludes the proof of Proposition 20-(a) upon taking $\tilde{A} = \exp(2 |\mathtt{m}| + d + 1)$, $\varepsilon_n = 1 - 2\alpha$ and $E_n = 2(\mathtt{L}_{n+1} \bar{\gamma}^{2\alpha} + 1)$.

(b) Let $n \in \mathbb{N}$. Using the Log-Sobolev inequality [4, Theorem 5.5], the fact that Π_{K_n} is non expansive and the Jensen inequality we obtain for any $\gamma \in (0, \bar{\gamma}]$ and $x \in \mathbb{R}^d$,

$$\mathrm{R}_\gamma V(x) \le \exp\left[\mathrm{R}_\gamma \phi(x) + \gamma^2/2\right] \le \exp\left[\sqrt{\mathrm{R}_\gamma(1 + \|x\|^2)} + \gamma^2/2\right] . \qquad (74)$$

We have that $\sup_{x \in \mathbb{R}^d} \Phi_n(x) \|b(x)\| \le (\sup_{x \in \bar{\mathrm{B}}(0,n+1)} \|b(x)\| \bar{\gamma}^\alpha + 1)\gamma^{-\alpha}$. Let $B_n = \sup_{x \in \bar{\mathrm{B}}(0,n+1)} \|b(x)\|$. Using (71) and $\mathbf{C}\,2$, we obtain that for any $\gamma \in (0, \bar{\gamma}]$ and $x \in \bar{\mathrm{B}}(0, R_2)^c$,

$$\mathrm{R}_\gamma(1 + \|x\|^2) \le 1 + \|x\|^2\left[1 - 2\gamma \mathbf{m}_2^+ \Phi_n(x)\right] + \gamma^{2(1-\alpha)}\left(1 + \bar{\gamma}^\alpha B_n\right)^2 + \gamma d . \qquad (75)$$

On the other hand, using (73), we get for any $\gamma \in (0, \bar{\gamma}]$ and $x \in \bar{\mathrm{B}}(0, R_2)$,

$$\mathrm{R}_\gamma(1 + \|x\|^2) \le (1 + \|x\|^2) + 2(\gamma|\mathbf{m}| + \gamma^2 \mathsf{L}_{n+1}^2)(1 + R_2^2) + 2\gamma^{2-2\alpha} + \gamma d . \qquad (76)$$

Therefore, setting $\tilde{A} = \exp(d + |\mathbf{m}|(1 + R^2) + 1)$, $\varepsilon_n = 1 - 2\alpha$ and $E_n = (1 + \bar{\gamma}^\alpha B_n)^2/2 + \mathsf{L}_{n+1}^2(1 + R_2^2)\bar{\gamma}^{2\alpha} + 3/2$ and combining (74), (75) and (76) with the fact that for any $t \in (-1, +\infty)$, $\sqrt{1+t} \le 1 + t/2$, we get for any $\gamma \in (0, \bar{\gamma}]$ and $x \in \mathbb{R}^d$

$$\mathrm{R}_\gamma V(x) \le \exp\left[\log(C_3)\gamma(1 + E_n \gamma^{\varepsilon_n})\right] V(x) . \qquad (77)$$

Setting $\gamma = T/m$ with $m \in \mathbb{N}^\star$, iterating (77) and taking the limit when $m \to +\infty$ concludes for Proposition 20-(b).

F.9 Proof of Proposition 21

We have for any $x \in \mathbb{R}^d$,

$$\nabla \phi(x) = x/\phi(x) , \qquad \nabla^2 \phi(x) = \mathrm{Id}/\phi(x) - xx^\top/\phi(x)^2 .$$

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

$$\nabla V(x) = M\nabla\phi(x)V(x) , \qquad \nabla^2 V(x) = \left(M^2 \nabla\phi(x)(\nabla\phi(x))^\top + M\nabla^2\phi(x)\right)V(x) .$$

Therefore, for any $x \in \mathbb{R}^d$,

$$(\mathcal{A}V(x))/V(x) \le \left[M^2 \|x\|^2/\phi(x)^2 + M(d/\phi(x) - \|x\|^2/\phi(x)^2)\right]/2 + M\sup_{x \in \mathbb{R}^d}\langle b(x), x\rangle .$$

Hence, for any $x \in \mathbb{R}^d$, $\mathcal{A}V(x) \le \zeta V(x)$ with $\zeta = M(\sup_{x \in \mathbb{R}^d}\langle b(x), x\rangle + d/2) + M^2$. We conclude using Lemma 15-(a) and the Doob maximal inequality.