WARWICK
THE UNIVERSITY OF WARWICK

Manuscript version: Author's Accepted Manuscript
The version presented in WRAP is the author's accepted manuscript and may differ from the published version or, Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/116726

How to cite:
Please refer to published version for the most recent bibliographic citation information.
If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

Publisher's statement:
Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

warwick.ac.uk/lib-publications

Quantitative contraction rates for Markov chains on general state spaces [*]

Andreas Eberle[†]   Mateusz B. Majka[‡]

Abstract

We investigate the problem of quantifying contraction coefficients of Markov transition kernels in Kantorovich (L^1 Wasserstein) distances. For diffusion processes, relatively precise quantitative bounds on contraction rates have recently been derived by combining appropriate couplings with carefully designed Kantorovich distances. In this paper, we partially carry over this approach from diffusions to Markov chains. We derive quantitative lower bounds on contraction rates for Markov chains on general state spaces that are powerful if the dynamics is dominated by small local moves. For Markov chains on $\mathbb{R}^d$ with isotropic transition kernels, the general bounds can be used efficiently together with a coupling that combines maximal and reflection coupling. The results are applied to Euler discretizations of stochastic differential equations with non-globally contractive drifts, and to the Metropolis adjusted Langevin algorithm for sampling from a class of probability measures on high dimensional state spaces that are not globally log-concave.

1 Introduction

In recent years, convergence bounds for Markov processes in Kantorovich (L^1 Wasserstein) distances have emerged as a powerful alternative to more traditional approaches based on the total variation distance [36], spectral gaps and L^2 bounds [27, 9, 10], or entropy estimates [27, 9, 1]. In particular, Hairer, Mattingly and Scheutzow have

---

[†]Universität Bonn, Institut für Angewandte Mathematik, Endenicher Allee 60, 53115 Bonn, Germany. E-mail: eberle@uni-bonn.de, https://wt.iam.uni-bonn.de/eberle/

[‡]Department of Statistics, University of Warwick, Coventry CV4 7AL, UK. E-mail: mateusz.majka@warwick.ac.uk, https://sites.google.com/site/mateuszbmajka/

developed an analogue to Harris' Theorem assuming only local strict contractivity in a Kantorovich distance on the "small" set and a Lyapunov condition combined with non-strict contractivity outside, cf. [22, 24]. Meanwhile there have been numerous extensions and applications of their result [25, 6, 3, 11].

In [29], Joulin and Ollivier have shown that strict Kantorovich contractivity of the transition kernel implies bounds for the variance and concentration estimates for ergodic averages of a Markov chain. Their results have since been extended to cover more general frameworks by Paulin [38]. More recently, Pillai and Smith [39] as well as Rudolf and Schweizer [40] have developed a perturbation theory for Markov chains that are contractive in a Kantorovich distance, cf. also Huggins and Zou [26] as well as Johndrow and Mattingly [28] for related results. These works show that variants of the results in [29] carry over to perturbations of the original chain, thus paving the way for a much broader range of applications.

All the works mentioned above assume that, at least locally, strict contractivity holds w.r.t. an L^1 Wasserstein distance based on some underlying distance function on the state space of the Markov chain. The contraction rate is the key quantity in the resulting bounds, and it is hence important to develop applicable methods for quantifying contraction rates.

Contractivity with respect to the L^1 Wasserstein distance based on the Euclidean distance in $\mathbb{R}^d$ is sometimes interpreted as non-negative Ricci curvature of the Markov chain w.r.t. this metric [41, 29, 37]. This is a strong condition that is often not satisfied in applications. However, in many cases it is still possible to obtain contractivity with respect to a Kantorovich distance in which the underlying distance function has been modified accordingly. This allows for applying the results from [29] to a significantly broader class of examples. For diffusion processes, a corresponding approach to quantitative contraction rates in appropriately designed metrics has been developed systematically in recent years in a series of papers [15, 17, 44, 19, 20], see also [4, 5, 42] for previous results. The approach has been extended to Lévy driven SDEs in [32, 33], see also [31, 43].

Below we propose a corresponding approach for Markov chains on general metric state spaces. The approach is powerful in situations where the dynamics is dominated by small, local moves. This will be demonstrated below for Euler schemes for non-globally contractive stochastic differential equations, as well as for the Metropolis-adjusted Langevin Algorithm (MALA). In these cases, the Ricci curvature condition required in [29] is not satisfied in the standard L^1 Wasserstein distance and hence the construction of an alternative metric is required. For dynamics dominated by large or global moves, our approach does not apply in the form presented here. Sometimes, related approaches can be used nevertheless, see e.g. [2] for the construction of a contractive distance for Hamiltonian Monte Carlo.

2 Main results

Let $p(x, dy)$ be a Markov transition kernel on a separable metric space (S, d). To study contraction properties of p we construct distance functions $\rho : S \times S \to [0, \infty)$ by transforming the metric d in an appropriate way. Note that if $f_0 : [0, \infty) \to [0, \infty)$ is a concave, increasing function with $f_0(0) = 0$ and $f_0(r) > 0$ for $r \neq 0$, then $\rho(x, y) = f_0(d(x, y))$ is a metric on S. More generally, let $a, \delta \geq 0$ be non-negative constants, and let $V : S \to [0, \infty)$ be a non-negative measurable function on S. We set

$$f := f_0 + a\mathbb{1}_{(0, \infty)},$$

and we consider distance functions of the form

$$
\begin{aligned}
\rho(x,y) &= f(d(x,y)) + \delta\left(V(x)+V(y)\right)\mathbb{1}_{x\neq y} \\
&= f_0(d(x,y)) + (a + \delta V(x) + \delta V(y))\mathbb{1}_{x\neq y}\,.
\end{aligned}
\tag{2.1}
$$

We assume that f_0 is continuous and in some of our results we will choose $a = 0$ and $V \equiv 0$. Similar but slightly different classes of distance functions have been used e.g. in [23, 24, 3, 19] to study properties of Markov chains and diffusion processes.

For probability measures μ and ν on S, the Kantorovich distance (L^1 Wasserstein distance) $\mathcal{W}_\rho(\mu,\nu)$ based on the underlying distance function ρ is defined as

$$
\mathcal{W}_\rho(\mu,\nu) \;=\; \inf_{X\sim\mu,Y\sim\nu} \mathbb{E}[\rho(X,Y)].
\tag{2.2}
$$

Here the infimum is over all couplings of μ and ν, i.e., over all random variables X,Y defined on a common probability space $(\Omega,\mathcal{A},\mathbb{P})$ such that $\mathbb{P}\circ X^{-1} = \mu$ and $\mathbb{P}\circ Y^{-1} = \nu$.

For $f_0 \equiv 0$, $a = 1$ and $V \equiv 0$, $\mathcal{W}_\rho$ coincides with the total variation distance $d_{\mathrm{TV}}(\mu,\nu)$ (or with $d_{\mathrm{TV}}(\mu,\nu)/2$, depending on the convention used in the definition of the total variation distance), whereas for $f_0(r) = r$, $a = 0$ and $V \equiv 0$, $\mathcal{W}_\rho$ is the standard L^1 Wasserstein distance $\mathcal{W}_d$ on (S,d). The distance functions we consider are in between these two extremes. Notice, however, that if $a > 0$ then

$$
d_{\mathrm{TV}}(\mu,\nu) \leq\; a^{-1}\mathcal{W}_\rho(\mu,\nu),
\tag{2.3}
$$

and if $f(r) \geq br$ for some constant $b > 0$ then

$$
\mathcal{W}_d(\mu,\nu) \leq\; b^{-1}\mathcal{W}_\rho(\mu,\nu),
\tag{2.4}
$$

Therefore, in these cases, contraction properties w.r.t. $\mathcal{W}_\rho$ directly imply upper bounds for the total variation and L^1 Wasserstein distances w.r.t. the metric d.

We now assume that we are given a *Markovian coupling* of the transition probabilities $p(x,\cdot)$ ($x \in S$) in the form of measurable maps $X',Y' : \Omega \to S$, defined on a measurable space $(\Omega,\mathcal{A})$, and a probability kernel $(x,y,A) \mapsto \mathbb{P}_{x,y}(A)$ from $S \times S \times \mathcal{A}$ to $[0,1]$ such that for any $x,y \in S$,

$$
X' \sim p(x,\cdot) \qquad \text{and} \qquad Y' \sim p(y,\cdot) \qquad \text{under } \mathbb{P}_{x,y}.
\tag{2.5}
$$

For probability measures μ on S and γ on $S\times S$ let $(\mu p)(B) = \int \mu(dx)p(x,B)$ for $B \in \mathcal{B}(S)$, and $\mathbb{P}_\gamma(A) = \int \gamma(dx\,dy)\mathbb{P}_{x,y}(A)$ for $A \in \mathcal{A}$. Note that if γ is a coupling of two probability measures μ and ν on S, then under $\mathbb{P}_\gamma$ the joint law of (X',Y') is a coupling of the probability measures μp and νp, i.e.,

$$
X' \sim \mu p \qquad \text{and} \qquad Y' \sim \nu p \qquad \text{under } \mathbb{P}_\gamma.
\tag{2.6}
$$

Our goal is to derive explicit bounds of the form

$$
\mathbb{E}_{x,y}[\rho(X',Y')] \leq (1-c)\rho(x,y) \qquad \text{for any } x,y \in S,
\tag{2.7}
$$

where c is a strictly positive constant. Here the choice of the metric ρ is adapted in order to maximize the value of c in our bounds. If (2.7) holds, then the transition kernel p is a strict contraction w.r.t. the distance $\mathcal{W}_\rho$.

Lemma 2.1. Suppose that (2.7) holds for all $x,y \in S$. Then

$$
\mathcal{W}_\rho(\mu p,\nu p) \leq (1-c)\mathcal{W}_\rho(\mu,\nu) \qquad \text{for all } \mu,\nu \in \mathcal{P}(S).
\tag{2.8}
$$

Proof. Let μ and ν be probability measures on S and suppose that γ is a coupling of μ and ν. Then, under $\mathbb{P}_\gamma$, the joint law of (X', Y') is a coupling of μp and νp. Therefore by (2.7),

$$
\begin{aligned}
\mathcal{W}_\rho(\mu p, \nu p) &\leq \mathbb{E}_\gamma[\rho(X', Y')] = \int \mathbb{E}_{x,y}[\rho(X', Y')]\gamma(dx\,dy) \\
&\leq (1-c) \int \rho(x, y)\gamma(dx\,dy).
\end{aligned}
$$

The assertion follows by taking the infimum over all couplings of μ and ν. $\qquad\square$

In the terminology of Joulin and Ollivier [29], (2.8) says that the Markov chain has a Ricci curvature lower bound c on the metric space (S, ρ). By general results, such a bound has many important consequences including quantitative convergence to a unique equilibrium [18], upper bounds on biases and variances as well as concentration inequalities for ergodic averages [29, 38], a central limit theorem for ergodic averages [30], robustness under perturbations [39, 40, 26, 28], etc. However, in applications, it is usually not clear how to choose a distance function ρ such that we have good bounds for c. This is the problem addressed in this paper for the case of a "local dynamics" where the Markov chain is mainly making "small" moves. Depending on whether or not the probability measures $p(x, \cdot)$ and $p(y, \cdot)$ have a significant overlap for x close to y, we suggest two different approaches.

2.1 Contractivity with positive coupling probability

Our first two general results apply in situations where the probability measures $p(x, \cdot)$ and $p(y, \cdot)$ have a significant overlap if x and y are sufficiently close. In this case we can always consider a coupling $((X', Y'), \mathbb{P}_{x,y})$ of the transition probabilities such that $\mathbb{P}_{x,y}[X' = Y'] > 0$ for x close to y. This enables us to obtain strict contractivity in metrics that have a total variation part, i.e., the function f defining the underlying distance has a discontinuity $a > 0$ at 0.

To state the results, we fix a positive constant $\varepsilon > 0$ and couplings $((X', Y'), \mathbb{P}_{x,y})$ as above. For $x, y \in S$ we set

$$
r = d(x, y), \qquad\qquad R' = d(X', Y'), \qquad\qquad \Delta R = R' - r, \qquad (2.9)
$$

and we define

$$
\begin{aligned}
\beta(x, y) &= \mathbb{E}_{x,y}[\Delta R], && (2.10) \\
\alpha(x, y) &= \mathbb{E}_{x,y}\left[|(\Delta R)^- \wedge \varepsilon|^2\right], && (2.11) \\
\pi(x, y) &= \mathbb{P}_{x,y}[R' = 0], && (2.12)
\end{aligned}
$$

where $(\Delta R)^- = \max(-\Delta R, 0)$. In particular,

$$
\alpha(x, y) \geq \mathbb{E}_{x,y}[(\Delta R)^2 \mathbb{1}_{\{R' \in (r-\varepsilon, r)\}}]. \qquad (2.13)
$$

One can think of $\beta(x, y)$ as a drift for the coupling distance, whereas $\alpha(x, y)$ provides a lower bound for fluctuations that decrease the distance and $\pi(x, y)$ is the probability of coupling successfully in the next step. Suppose that there exist functions $\overline{\beta} : (0, \infty) \to \mathbb{R}$ and $\underline{\alpha}, \underline{\pi} : (0, \infty) \to [0, \infty)$ such that for any $r > 0$ and $x, y \in S$ with $d(x, y) = r$,

$$
\beta(x, y) \leq \overline{\beta}(r), \quad \alpha(x, y) \geq \underline{\alpha}(r), \quad \text{and} \quad \pi(x, y) \geq \underline{\pi}(r). \qquad (2.14)
$$

Hence $\overline{\beta}(r)$ is an upper bound for the expectation of the increase ΔR of the distance during a single transition step of coupled Markov chains with initial states x and y such

that $d(x, y) = r$. Similarly, $\underline{\alpha}(r)$ is a lower bound for distance decreasing fluctuations of ΔR, and $\underline{\pi}(r)$ is a lower bound for the coupling probability. We make the following assumptions on $\underline{\alpha}, \overline{\beta}$ and $\underline{\pi}$:

(A1) There exists a positive constant $r_0 \in (0, \infty)$ such that

    (a) $\inf_{r \in (0, r_0]} \underline{\pi}(r) > 0$,   and

    (b) $\inf_{r \in (r_0, s)} \underline{\alpha}(r) > 0$ for any $s \in (r_0, \infty)$.

(A2) $\sup_{r \in (0, s)} \overline{\beta}(r) < \infty$ for any $s \in (0, \infty)$.

(A3) $\limsup_{r \to \infty} r^{-1} \overline{\beta}(r) < 0$.

Theorem 2.2. Suppose that (A1), (A2) and (A3) are satisfied, and let

$$\rho(x, y) = f(d(x, y)), \tag{2.15}$$

where $f : [0, \infty) \to [0, \infty)$ is the concave increasing function defined in (3.8) below. Then for any $x, y \in S$,

$$\mathbb{E}_{x,y}[\rho(X', Y')] \le (1 - c)\rho(x, y), \tag{2.16}$$

where c is an explicit strictly positive constant defined in (3.13) below.

The proof is given in Section 3. Explicit expressions for the function f and the contraction rate c depending only on $\underline{\alpha}, \overline{\beta}, \underline{\pi}$ and ε are given in Subsection 3.1. Although these expressions are somehow involved, they can be applied to derive quantitative bounds in concrete models. In particular, the asymptotic dependence of the contraction rate on parameters of the model can often be made explicit. This will be demonstrated for the Euler scheme in Section 2.4.

By Lemma 2.1, Theorem 2.2 implies that the transition kernel p is contractive with rate c w.r.t. the $\mathcal{W}_\rho$ distance on probability measures on S. Since the function f defined in (3.8) is bounded from below by a multiple of both $\mathbb{1}_{(0,\infty)}$ and of the identity, the theorem yields quantitative bounds for convergence to equilibrium both w.r.t. the total variation and the standard L^1 Wasserstein distance.

The assumption (A3) imposed in Theorem 2.2 is sometimes too restrictive. By a modification of the metric, it can be replaced by the following Lyapunov condition:

(A4) There exist a measurable function $V : S \to [0, \infty)$ and $C, \lambda \in (0, \infty)$ s.t.

    (a) $pV \le (1 - \lambda)V + C$,   and

    (b) $\inf_{d(x,y)=r} \dfrac{V(x)+V(y)}{\overline{\beta}(r)^+} \longrightarrow \infty$ as $r \to \infty$.

In ((A4)b) we use the convention that the value of the fraction is $+\infty$ if $\overline{\beta}(r) \le 0$.

Theorem 2.3. Suppose that (A1), (A2) and (A4) are satisfied, and let

$$\rho(x, y) = f(d(x, y)) + \frac{M}{2C}(V(x) + V(y))\mathbb{1}_{x \ne y}, \tag{2.17}$$

where $f \colon [0, \infty) \to [0, \infty)$ is the concave increasing function defined in (3.17) below, and the constant $M \in \mathbb{R}_+$ is defined in (3.19). Then for any $x, y \in S$,

$$\mathbb{E}_{x,y}[\rho(X', Y')] \le (1 - c)\rho(x, y), \tag{2.18}$$

where c is an explicit strictly positive constant defined in (3.24) below.

The proof of the theorem is given in Section 3 and explicit expressions for the function f and the constants M and c in terms of $\underline{\alpha}, \overline{\beta}, \underline{\pi}, \varepsilon, V, C$ and λ are provided in Subsection 3.2.

The idea of adding a Lyapunov function to the metric appears for example in [21] and has been further worked out in the diffusion case in [19]. Theorem 2.3 can be seen as a more quantitative version of Theorem 4.8 in [24], which is an extension of the classical Harris' Theorem. Note, however, that contractivity in our result is expressed in an additive metric ρ, as opposed to the multiplicative semimetric used in [24]; see also [19] for a more detailed discussion on these two types of metrics. An application of Theorem 2.3 to the Euler scheme is given in Theorem 6.1 below.

2.2 Contractivity without positive coupling probability

The assumption that there is a significant overlap between the measures $p(x, \cdot)$ and $p(y, \cdot)$ for x close to y is sometimes too restrictive. For example, it may cause a bad dimension dependence of the resulting bounds in high dimensional applications. Therefore, we now state an alternative contraction result that applies even when $\pi(x, y) = 0$ for all x and y.

For any $r \in (0, \infty)$ we consider an interval near r given by

$$I_r = (r - l(r), r + u(r)) \tag{2.19}$$

where $l(r), u(r) \geq 0$ and $l(r) \leq r$. Similarly as in (2.10) and (2.11), we define

$$\beta(x, y) = \mathbb{E}_{x,y}[\Delta R], \tag{2.20}$$

$$\alpha(x, y) = \mathbb{E}_{x,y}\left[|(\Delta R \wedge u(r)) \vee (-l(r))|^2 \right], \tag{2.21}$$

where r, R' and ΔR are defined by (2.9). In particular,

$$\alpha(x, y) \geq \mathbb{E}_{x,y}[(\Delta R)^2 \mathbb{1}_{\{R' \in I_r\}}]. \tag{2.22}$$

In Subsection 2.1, we have chosen $l(r) = \varepsilon$ and $u(r) = 0$, i.e., $I_r = (r - \varepsilon, r)$. Now, we will assume instead that there is a finite constant $r_0 > 0$ such that

$$u(r) = 0 \quad \text{for } r \geq r_0, \qquad \text{and} \qquad u(r) = r_0 \quad \text{for } r < r_0. \tag{2.23}$$

As above, we assume that there exist functions $\overline{\beta} : (0, \infty) \to \mathbb{R}$ and $\underline{\alpha} : (0, \infty) \to (0, \infty)$ such that for any $r > 0$ and $x, y \in S$ with $d(x, y) = r$,

$$\beta(x, y) \leq \overline{\beta}(r) \quad \text{and} \quad \alpha(x, y) \geq \underline{\alpha}(r). \tag{2.24}$$

We now impose the following conditions on $\underline{\alpha}$ and $\overline{\beta}$:

(B1) $\displaystyle \inf_{r \in (0,s)} \frac{\underline{\alpha}(r)}{r} > 0$ for any $s \in (0, \infty)$,

(B2) $\displaystyle \sup_{r \in (0,s)} \frac{\overline{\beta}(r)}{\underline{\alpha}(r)} < \infty$ for any $s \in (0, \infty)$,

(B3) $\limsup_{r \to \infty} r^{-1} \overline{\beta}(r) < 0$.

Thus we no longer assume a positive coupling probability for $r < r_0$. Instead, we require in (B1) and (B2) that $\underline{\alpha}(r) = \Omega(r)$ and $\overline{\beta}(r)/\underline{\alpha}(r) = O(1)$ as $r \downarrow 0$. These assumptions can be verified for example for Euler schemes if the coupling is constructed carefully. We will do this in Section 2.4 for Euler discretizations of SDEs with contractive drifts, whereas for more general drifts we will follow a slightly different approach.

Theorem 2.4. Suppose that (B1), (B2) and (B3) are satisfied, and let

$$\rho(x,y) = f(d(x,y)), \tag{2.25}$$

where $f \colon [0, \infty) \to [0, \infty)$ is a continuous concave increasing function satisfying $f(0) = 0$ which is defined explicitly in (4.8) below. Then for any $x, y \in S$,

$$\mathbb{E}_{x,y}[\rho(X', Y')] \le (1 - c)\rho(x, y), \tag{2.26}$$

where c is an explicit strictly positive constant defined in (4.12) below.

The proof is given in Section 4. Notice that in contrast to Theorem 2.2 and Theorem 2.3, the function f in Theorem 2.4 does not have a jump at 0, i.e., the Kantorovich metric $\mathcal{W}_\rho$ does not contain a total variation part. This corresponds to the fact that under Assumptions (B1), (B2) and (B3), it can not be expected in general that the coupled Markov chains meet in finite time.

2.3 Stability under perturbations

Contractions in Kantorovich distances can sometimes be carried over to small perturbations of a given Markov chain. For instance, in Subsection 2.5 we will deduce contractivity for the Metropolis adjusted Langevin algorithm from corresponding properties of the Euler proposal chain. Suppose as above that $((X', Y'), \mathbb{P}_{x,y})$ is a Markovian coupling of the transition probabilities $p(x, \cdot)$ and $p(y, \cdot)$. Moreover, let $((\widetilde{X}, \widetilde{Y}), \mathbb{P}_{x,y})$ be a corresponding coupling of $\widetilde{p}(x, \cdot)$ and $\widetilde{p}(y, \cdot)$ for another (perturbed) Markov transition kernel $\widetilde{p}$ on S. Here we assume that for given $x, y \in S$, (X', Y') and $(\widetilde{X}, \widetilde{Y})$ are defined on a common probability space. We start with a simple observation. If there exists a metric ρ on S and a constant $c \in (0, \infty)$ such that for $x, y \in S$,

$$\mathbb{E}_{x,y}[\rho(X', Y')] \quad \le \quad (1 - c)\rho(x, y), \qquad \text{and} \tag{2.27}$$

$$\mathbb{E}_{x,y}[\rho(\widetilde{X}, \widetilde{Y})] \quad \le \quad \mathbb{E}_{x,y}[\rho(X', Y')] + \frac{c}{2}\rho(x, y), \qquad \text{then} \tag{2.28}$$

$$\mathbb{E}_{x,y}[\rho(\widetilde{X}, \widetilde{Y})] \quad \le \quad (1 - c/2)\,\rho(x, y). \tag{2.29}$$

In applications it is often difficult or even impossible to verify Condition (2.28) for x very close to y. If $\mathbb{P}_{x,y}[\widetilde{X} = \widetilde{Y}] > 0$ for x close to y, then this condition can be relaxed.

Theorem 2.5. Suppose that $\rho(x, y) = f(d(x, y))$ for a concave increasing contraction $f \colon [0, \infty) \to [0, \infty)$ satisfying $f(0) = 0$. Suppose that there exist constants $c, b, p, r_0 \in [0, \infty)$ such that for all $x, y \in S$,

$$\mathbb{E}_{x,y}[\rho(X', Y')] \le (1 - c)\rho(x, y), \tag{2.30}$$

$$\mathbb{E}_{x,y}\left[(d(\widetilde{X}, \widetilde{Y}) - d(X', Y'))^+ \right] \le b + \frac{c}{2}\rho(x, y), \qquad \text{and} \tag{2.31}$$

$$\mathbb{P}_{x,y}[\widetilde{X} = \widetilde{Y}] \ge p \qquad \text{if } d(x, y) < r_0. \tag{2.32}$$

Assume that $p > 0$, $b \le cf(r_0)/4$, and let $\widetilde{\rho}$ be the metric defined by

$$\widetilde{\rho}(x, y) = \rho(x, y) + \frac{2b}{p}\mathbb{1}_{x \ne y}. \tag{2.33}$$

Then

$$\mathbb{E}_{x,y}\left[\widetilde{\rho}(\widetilde{X}, \widetilde{Y}) \right] \le \left(1 - \frac{1}{8}\min(c, 2p) \right) \widetilde{\rho}(x, y) \qquad \text{for all } x, y \in S. \tag{2.34}$$

The proof is given in Section 5. In Section 7 we will apply Theorem 2.5 to our results for the Euler scheme in order to obtain contractivity for the Metropolis adjusted Langevin algorithm (MALA).

Note that Theorem 2.5 is related to the perturbation results in [39, 40, 28]. In all these papers, a Kantorovich contraction in some metric is assumed for the initially given unperturbed Markov chain. Then, in [39], the authors obtain bounds on the distance to equilibrium of a perturbed Markov chain in the same Kantorovich metric. In [40, 28], the metric also remains unchanged, but the object of interest is a bound on the distance between a perturbed and the unperturbed chain. A related result in continuous time, giving bounds on the distance between invariant measures of a perturbed and an unperturbed diffusion, has been obtained in [26]. In contrast to these results, we consider a perturbed metric in Theorem 2.5, but we obtain a stronger result showing that the perturbed Markov chain is again contractive w.r.t. the modified metric.

2.4 Application to Euler schemes

We now show how to apply the general methods developed above to Euler discretizations of stochastic differential equations of the form

$$dX_t = b(X_t)\,dt + dB_t\,, \tag{2.35}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion in $\mathbb{R}^d$, and $b : \mathbb{R}^d \to \mathbb{R}^d$ is a Lipschitz continuous vector field. Quantifying contraction rates for Euler discretizations is important in connection with the derivation of error bounds for the unadjusted Langevin algorithm (ULA), cf. [7, 14, 13, 12, 8] for corresponding results. Such applications of the techniques presented below will be discussed in detail in the upcoming paper [34] by the second author. The transitions of the Markov chain for the Euler scheme with step size $h > 0$ are given by

$$x \mapsto \hat{x} + \sqrt{h}Z, \quad \text{where} \quad \hat{x} := x + hb(x) \quad \text{and} \quad Z \sim N(0, I_d). \tag{2.36}$$

The corresponding transition probabilities are given by

$$p(x, \cdot) \;=\; N(\hat{x}, hI_d) \qquad \text{for any } x \in \mathbb{R}^d, \tag{2.37}$$

i.e., the transition density from x is

$$p(x, x') \;=\; \phi_{\hat{x}, hI_d}(x') \;=\; (2\pi h)^{-d/2} \exp\left(-\frac{1}{2h}|x' - \hat{x}|^2 \right). \tag{2.38}$$

In the case of $b \equiv 0$, the Markov chain is a Gaussian random walk with transitions $x \mapsto x + \sqrt{h}Z$.

The coupling

For $x, y \in \mathbb{R}^d$ let

$$\hat{e} \;=\; \frac{\hat{x} - \hat{y}}{|\hat{x} - \hat{y}|} \qquad \text{if } \hat{x} \neq \hat{y}, \qquad \hat{e} = 0 \quad \text{otherwise.} \tag{2.39}$$

We consider the coupling of two transitions of the Euler chain from x and y respectively given by

$$X' \;=\; \hat{x} + \sqrt{h}Z, \qquad Y' \;=\; \begin{cases} X' & \text{if } U \leq \phi_{\hat{y}, hI}(X')/\phi_{\hat{x}, hI}(X'), \\ Y'_{\text{refl}} & \text{otherwise,} \end{cases} \tag{2.40}$$

where $Z \sim N(0, I_d)$ and $U \sim \text{Unif}(0, 1)$ are independent random variables, and

$$Y'_{\text{refl}} \;=\; \hat{y} + \sqrt{h}(I_d - 2\hat{e}\hat{e}^T)Z \tag{2.41}$$

is obtained by adding to $\hat{y}$ the increment $\sqrt{h}Z$ added to $\hat{x}$, reflected at the hyperplane between $\hat{x}$ and $\hat{y}$.
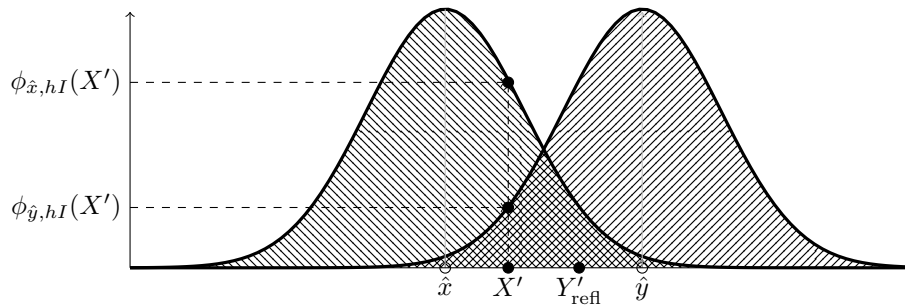
Figure 1: Construction of the coupling of $p(x, \cdot)$ and $p(y, \cdot)$: Given the value of X', we set $Y' = X'$ with the maximal probability $\min(1, p(y, X')/p(x, X'))$, and $Y' = Y'_{\text{refl}}$ otherwise.

Both (X', Y'_{refl}) and (X', Y') are couplings of the probability measures $p(x, \cdot)$ and $p(y, \cdot)$. For the coupling (2.40), $Y' = X'$ with the maximal probability $\min(1, p(y, X')/p(x, X'))$. Furthermore, in the case where $Y' \neq X'$, the coupling coincides with the reflection coupling, i.e., $Y' = Y'_{\text{refl}}$. The resulting combination of reflection coupling and maximal coupling is an optimal coupling of the Gaussian measures $p(x, \cdot)$ and $p(y, \cdot)$ w.r.t. any Kantorovich distance based on a metric $\rho(x, y) = f(|x - y|)$ with f concave, cf. [35] for the one-dimensional case. We will not use the optimality here, but it shows that (2.40) is an appropriate coupling to consider if we are interested in contraction properties for single transition steps of the Markov chain.

Note that an analogous coupling construction still works if the Gaussian measure above is replaced with an isotropic (i.e., rotationally symmetric) measure $p(x, \cdot)$ such that for any distinct points $x, y \in \mathbb{R}^d$ the measures $p(x, \cdot)$ and $p(y, \cdot)$ have sufficient overlap. See [32] for a discussion on similar coupling constructions in the context of stochastic differential equations driven by isotropic Lévy processes.

Remark 2.6 (Relation to reflection coupling of diffusion processes). A reflection coupling of two copies of a diffusion process satisfying a stochastic differential equation of the form (2.35) is given by

$$
\begin{aligned}
dX_t &= b(X_t)\, dt + dB_t, \\
dY_t &= b(Y_t)\, dt + (I - 2e_t e_t^T)\, dB_t \quad \text{for } t < T, \quad X_t = Y_t \quad \text{for } t \geq T,
\end{aligned}
\tag{2.42}
$$

where $e_t = (X_t - Y_t)/|X_t - Y_t|$ and $T = \inf\{t \geq 0 : X_t = Y_t\}$ is the coupling time. Hence the noise increment is reflected up to the coupling time, whereas after time T, X_t and Y_t move synchronously. Our coupling in discrete time has a similar effect. If $\hat{x}$ and $\hat{y}$ are far apart then the transition densities $\phi_{\hat{x},hI}$ and $\phi_{\hat{y},hI}$ have little overlap, and hence reflection coupling is applied with very high probability. If, on the other hand, $\hat{x}$ and $\hat{y}$ are sufficiently close, then with a non-negligible probability, $X' = Y'$. Once both Markov chains have reached the same position, they stick together since their transition densities coincide subsequently. In this sense, the coupling (2.40) is a natural discretization of reflection coupling. Indeed, we would expect that as $h \downarrow 0$, the coupled Markov chains with time rescaled by a factor h converge in law to the reflection coupling (2.42) of the diffusion processes. On the other hand, a coupling of Markov chains in which jumps are always reflected (i.e., a coupling without the positive probability of jumping to the same point) would converge as $h \downarrow 0$ to a reflection coupling of diffusions in which the coupled processes do not follow the same path after the coupling time.

We assume that under the probability measure $\mathbb{P}_{x,y}$, (X', Y') is the coupling of $p(x, \cdot)$

and $p(y, \cdot)$ introduced above. We set

$$r_0 := \sqrt{h}, \tag{2.43}$$

and we consider the intervals

$$I_r \;=\; \begin{cases} (0, r + \sqrt{h}) & \text{for } r < r_0, \\ (r - \sqrt{h}, r) & \text{for } r \geq r_0. \end{cases} \tag{2.44}$$

Thus in the notation from Section 2.2, we set

$$u(r) := \sqrt{h}\,\mathbb{1}_{r < r_0}, \qquad l(r) := \sqrt{h}\,\mathbb{1}_{r \geq r_0}. \tag{2.45}$$

For given $x, y \in \mathbb{R}^d$ let $r = |x - y|$, $\hat{r} = |\hat{x} - \hat{y}|$, $R' = |X' - Y'|$,

$$\hat{\beta}(x, y) \;=\; \mathbb{E}_{x,y}[R' - \hat{r}], \tag{2.46}$$
$$\hat{\alpha}(x, y) \;=\; \mathbb{E}_{x,y}\left[|((R' - \hat{r}) \wedge u(\hat{r})) \vee (-l(\hat{r}))|^2\right], \qquad \text{and} \tag{2.47}$$
$$\pi(x, y) \;=\; \mathbb{P}_{x,y}[R' = 0]. \tag{2.48}$$

In particular,

$$\hat{\alpha}(x, y) \;\geq\; \mathbb{E}_{x,y}[(R' - \hat{r})^2 \mathbb{1}_{\{R' \in I_{\hat{r}}\}}]. \tag{2.49}$$

Notice that the definitions of $\hat{\beta}$ and $\hat{\alpha}$ differ from those of β and α given in (2.20) and (2.21), since $\hat{\beta}$ and $\hat{\alpha}$ take into account only the coupled random walk transition step from $(\hat{x}, \hat{y})$ to (X', Y'), but not the deterministic transition from (x, y) to $(\hat{x}, \hat{y})$. We also consider

$$\beta(x, y) \;=\; \mathbb{E}_{x,y}[R' - r] \;=\; \hat{\beta}(x, y) + \hat{r} - r, \qquad \text{and} \tag{2.50}$$
$$\alpha(x, y) \;=\; \mathbb{E}_{x,y}\left[|((R' - r) \wedge u(r)) \vee (-l(r))|^2\right]. \tag{2.51}$$

Assumptions

In our main result for the Euler scheme we assume that there exist constants $J \in [0, \infty)$ and $K, L, \mathcal{R} \in (0, \infty)$ such that the following conditions hold:

(C1) *One-sided Lipschitz condition:*

$$(x - y) \cdot (b(x) - b(y)) \leq J|x - y|^2 \qquad \text{for any } x, y \in \mathbb{R}^d.$$

(C2) *Strict contractivity outside a ball:*

$$(x - y) \cdot (b(x) - b(y)) \leq -K|x - y|^2 \qquad \text{if } |x - y| \geq \mathcal{R}.$$

(C3) *Global Lipschitz condition:*

$$|b(x) - b(y)| \leq L|x - y| \qquad \text{for any } x, y \in \mathbb{R}^d.$$

Notice that by (C2) and (C3), $L \geq K$. Of course, (C3) implies (C1) with $J = L$. Note, however, that we can often choose J much smaller than L, e.g., we can even choose $J = 0$ if $b = -\nabla U$ for a convex function $U \in C^2(\mathbb{R}^d)$. The global Lipschitz condition is required for the stability of the Euler scheme, but the constant L will affect our lower bound for the contraction rate only in a marginal way. On the other hand, our bound on the contraction rate will depend in an essential way on the one-sided Lipschitz constant J.

The bounds provided in the next lemma are crucial to apply the techniques developed above to the Euler scheme.

Lemma 2.7. Let $x, y \in \mathbb{R}^d$, and let $r = |x - y|$ and $\hat{r} = |\hat{x} - \hat{y}|$. Then

(i) $\hat{\beta}(x, y) = 0$,

(ii) $\hat{\alpha}(x, y) \geq c_0 \min(\hat{r}, \sqrt{h})\sqrt{h}$,   and

(iii) $\pi(x, y) \geq p_0 \mathbb{1}_{\hat{r} \leq 2\sqrt{h}}$,

where $c_0, p_0 \in (0, 1)$ are explicit universal constants ($c_0 \geq 0.007$, $p_0 \geq 0.15$). Furthermore, if the assumptions (C1), (C2) and (C3) hold true, then

(iv) $\beta(x, y) \leq \min(L, J + L^2 h/2)hr$,

(v) $\beta(x, y) \leq -(K - L^2 h/2)hr$ if $r \geq \mathcal{R}$,

(vi) $\alpha(x, y) \geq \widetilde{c}_0 h \mathbb{1}_{r \geq \sqrt{h}}$ if $r \leq 1/(4L\sqrt{h})$,   and

(vii) $\pi(x, y) \geq p_0 \mathbb{1}_{r \leq \sqrt{h}}$ if $h \leq 1/L$.

Here $\widetilde{c}_0$ is an explicit universal constant ($\widetilde{c}_0 \geq 0.0005$).

The proof of the lemma is contained in Section 6.

Contractive case

At first, we consider the case where the deterministic part of the Euler transition is a contraction, i.e.,

$$\hat{r} = |\hat{x} - \hat{y}| \leq |x - y| = r \qquad \text{for any } x, y \in \mathbb{R}^d. \tag{2.52}$$

In this simple case, we can prove a rather sharp result. We choose a metric $\rho_a \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ of type

$$\rho_a(x, y) = a\mathbb{1}_{x \neq y} + f_a(|x - y|), \qquad f_a(r) = \int_0^r g_a(s \wedge \mathcal{R})\, ds. \tag{2.53}$$

Here a is a non-negative constant, $\mathcal{R}$ is chosen as in Assumption (C2), and $g_a \colon [0, \mathcal{R}] \to \mathbb{R}$ is an appropriately chosen decreasing function (see (6.14) for $a = 0$ and (6.20) for $a \neq 0$) satisfying

$$g_a(0) = 1 \qquad \text{and} \qquad g_a(s) \in [1/2, 1] \quad \text{for any } s \in [0, \mathcal{R}]. \tag{2.54}$$

Hence f_a is a concave increasing function satisfying $r/2 \leq f_a(r) \leq r$, and thus

$$a\mathbb{1}_{x \neq y} + |x - y|/2 \leq \rho_a(x, y) \leq a\mathbb{1}_{x \neq y} + |x - y| \qquad \text{for any } x, y \in \mathbb{R}^d. \tag{2.55}$$

In particular, the distance ρ_0 is equivalent to the Euclidean distance.

Theorem 2.8 (Euler scheme, contractive case)**.** Suppose that Conditions (C1), (C2), (C3) and (2.52) are satisfied, and let $h_0 = \frac{1}{L} \min\left(\frac{K}{L}, \frac{1}{2}\right)$. Suppose that $a = 0$ or $a \geq \sqrt{h}$, and let ρ_a be defined by (2.53) with g_a specified in (6.14), (6.20), respectively. Let

$$c_1(0) = \frac{1}{4} \min\left(K, \frac{2c_0}{\mathcal{R}^2 + 2\sqrt{h}\mathcal{R} + 12\,h}\right) \qquad \text{and} \tag{2.56}$$

$$c_1(a) = \frac{1}{4} \min\left(\frac{K}{1 + a/\mathcal{R}}, \frac{2c_0}{\mathcal{R}^2 + 2(a + \sqrt{h})\mathcal{R}}, \frac{2p_0}{h}\right) \qquad \text{for } a > 0, \tag{2.57}$$

where c_0 is the explicit constant in Lemma 2.7. Then (2.54) and (2.55) hold, and if $h \in (0, h_0)$, then

$$\mathbb{E}_{x,y}[\rho_a(X', Y')] \leq (1 - c_1(a)h)\, \rho_a(x, y) \quad \text{for all } x, y \in \mathbb{R}^d. \tag{2.58}$$

The proof, based on Theorem 2.2 for $a > 0$ and Theorem 2.4 for $a = 0$, is given in Section 6.

Remark 2.9 (Dependence on parameters and dimension)**.** The lower bound for the contraction rate in (2.58) is of the correct order $\Omega(h \min(\mathcal{R}^{-2}, K))$. This corresponds to the optimal contraction rate $\Theta(\min(\mathcal{R}^{-2}, K))$ for the corresponding diffusion process, see [17, Lemma 1 and Remark 5]. Note also that the lower bound for the contraction rate does not depend on the dimension d provided the parameters $\mathcal{R}, K$ and L can be chosen independent of d.

General case

We now turn to the general, not globally contractive case. Here it is no longer possible to obtain contractivity w.r.t. a metric satisfying (2.55), but we can still choose a metric that is comparable to the Euclidean distance, and apply the theorems above. We illustrate this at first by applying Theorem 2.2. Let

$$\Lambda = \min(L, J + L^2 h/2). \tag{2.59}$$

We now choose a metric $\rho_a \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ of type

$$\rho_a(x, y) = a \mathbb{1}_{x \neq y} + f_a(|x - y|), \qquad f_a(r) = \int_0^r g_a(s \wedge r_2) \varphi(s \wedge r_2) \, ds. \tag{2.60}$$

Here a is a non-negative constant,

$$\varphi(r) = \exp\left(-\widetilde{c}_0^{-1} \Lambda \left((r \wedge \mathcal{R})^2 + 2\sqrt{h}\, r \wedge \mathcal{R}\right)\right) \tag{2.61}$$

with $\mathcal{R}$ and $\widetilde{c}_0$ chosen as in Assumption (C2) and Lemma 2.7, respectively,

$$r_2 = \mathcal{R} + \sqrt{2\widetilde{c}_0/K}, \tag{2.62}$$

and $g_a \colon [0, r_2] \to \mathbb{R}$ is an appropriately chosen decreasing function (see (6.20)) satisfying

$$g_a(0) = 1 \qquad \text{and} \qquad g_a(s) \in [1/2, 1] \quad \text{for any } s \in [0, r_2]. \tag{2.63}$$

Theorem 2.10 (Euler scheme, general case I)**.** Suppose that Conditions (C1), (C2) and (C3) are satisfied, and let $h_0 = \frac{1}{L} \min(\frac{p_0}{2}, \frac{K}{L}, \frac{1}{64\, L r_2^2})$ with r_2 specified in (2.62). Let $a \in [2\sqrt{h}, \Phi(\mathcal{R})]$ where $\Phi(\mathcal{R}) := \int_0^{\mathcal{R}} \varphi(r) \, dr$, let ρ_a be defined by (2.60) with φ and g_a specified in (2.61) and (6.20), respectively, and let

$$c_2(a) = \frac{1}{8} \min\left(\frac{K\, \varphi(\mathcal{R})}{1 + (a + /\sqrt{h})\sqrt{2K/\widetilde{c}_0}}, \frac{2\, \widetilde{c}_0\, \varphi(\mathcal{R})}{\mathcal{R}^2 + 2(a + \sqrt{h})\mathcal{R}}, \frac{4p_0}{h}\right). \tag{2.64}$$

Then (2.63) holds, and if $h \in (0, h_0)$, then

$$\mathbb{E}_{x,y}[\rho_a(X', Y')] \leq (1 - c_2(a)h)\, \rho_a(x, y) \quad \text{for all } x, y \in \mathbb{R}^d.$$

Note that except for the additional factor $\varphi(\mathcal{R}) = \exp(-\widetilde{c}_0^{-1} \Lambda(\mathcal{R}^2 + 2\sqrt{h}\mathcal{R}))$, the expression for the contraction rate $c_2(a)$ is similar to the one for the rate $c_1(a)$ in the contractive case. The proof based on Theorem 2.2 is given in Section 6. If the interval $[2\sqrt{h}, \Phi(\mathcal{R})]$ is empty, the theorem can still be applied with $\mathcal{R}$ replaced by a slightly larger value. It is also possible to replace Condition (C2) by a Lyapunov condition and apply Theorem 2.3 instead of Theorem 2.2. A corresponding result for the Euler scheme is given in Section 6, cf. Theorem 6.1.

Remark 2.11 (Dependence on parameters and dimension). The lower bound for the contraction rate in (2.64) does not depend on the dimension d provided the parameters $\mathcal{R}, K$ and Λ can be chosen independent of d. Moreover, by choosing h sufficiently small, we can ensure that Λ is close to the one-sided Lipschitz constant J. Hence the global Lipschitz constant L is only required for controlling the step size h, whereas the contraction properties for sufficiently small h can be controlled essentially by one-sided Lipschitz bounds. This is important since in many applications, only a one-sided Lipschitz condition is satisfied globally. In this case, our approach can still be applied on a large ball if the step size is chosen sufficiently small depending on the radius of the ball and the growth of the local Lipschitz constant.

The explicit expression for the metric in Theorem 2.10 is a bit complicated. As an alternative, we can use a simplified metric without a discontinuity that is sufficient to derive bounds of similar order as for the metric used above, whenever condition (C2) is satisfied. We assume $hL \leq 1/6$, and we set

$$r_1 := (1 + hL)\mathcal{R} \leq \frac{7}{6}\mathcal{R} \tag{2.65}$$

with $\mathcal{R}$ and L as in Assumptions (C2) and (C3). The choice of r_1 ensures that

$$\hat{r} = |\hat{x} - \hat{y}| \leq (1 + hL)r \leq r_1 \quad \text{whenever } r = |x - y| \leq \mathcal{R}. \tag{2.66}$$

Let c_0 denote the explicit constant in Lemma 2.7, and let

$$q = 7c_0^{-1}\Lambda\mathcal{R}. \tag{2.67}$$

We now consider a simplified metric of the form

$$\rho(x, y) = f(|x - y|), \qquad f(r) = \int_0^r \exp(-q(s \wedge r_1))\, ds. \tag{2.68}$$

Theorem 2.12 (Euler scheme, general case II). Suppose that Conditions (C1), (C2) and (C3) are satisfied, and let ρ be defined by (2.68) with q specified in (2.67). Let

$$c_2 = \min\left(\frac{K}{2}, \frac{245}{24c_0}\Lambda^2\mathcal{R}^2\right) \exp\left(-\frac{49}{6c_0}\Lambda\mathcal{R}^2\right) \quad \text{and} \tag{2.69}$$

$$h_0 = \frac{1}{L}\min\left(\frac{1}{6}, \frac{K}{L}, \frac{1}{3}L\mathcal{R}^2, \frac{c_0^2}{970}\frac{1}{L\mathcal{R}^2}\right), \tag{2.70}$$

where c_0 is chosen as in Lemma 2.7. Then

$$\mathbb{E}_{x,y}[\rho(X', Y')] \leq (1 - c_2 h)\rho(x, y) \quad \text{for any } x, y \in \mathbb{R}^d \text{ and } h \in (0, h_0]. \tag{2.71}$$

The proof of the theorem is contained in Section 6.

Remark 2.13. Again, the lower bound c_2 for the contraction rate only depends on $\mathcal{R}$, K and Λ. Furthermore, note that $r\exp(-qr_1) \leq f(r) \leq r$ for all $r \geq 0$, and hence the metric ρ is comparable to the Euclidean distance. As a consequence, Theorem 2.12 implies weak contractivity in the standard L^1 Wasserstein distance. Note also that the function f depends on the discretization parameter h via q and r_1. It is, however, possible to modify the definition of f so that it no longer depends on h, at the cost of getting a worse constant c_2. We refer the interested reader to [34], where similar bounds are used with a metric independent of h.

Theorem 2.12 can be extended to cover pseudo metrics based on functions that are strictly convex at infinity. This allows us to obtain upper L^2 bounds for Euler schemes under similar assumptions as above. Such bounds are applied to the analysis of Multi-level Monte Carlo algorithms in the upcoming paper [34].

2.5 Application to MALA

The Metropolis-adjusted Langevin Algorithm is a Metropolis-Hastings method for approximate sampling from a given probability measure μ where the proposals are obtained by an Euler discretization of an overdamped Langevin SDE. In [16], the dimension dependence of contraction rates of MALA chains w.r.t. standard Kantorovich distances has been studied for a class of strictly log-concave probability measures that have a density w.r.t. a Gaussian reference measure. Our goal is to provide a partial extension of these results to non log-concave measures. By considering the MALA transition step as a perturbation of the Euler proposals, we obtain contraction rates w.r.t. a modified Kantorovich distance provided the discretization time step is of order $h = O(d^{-1})$.

We consider a similar setup as in [16]: μ is a probability measure on $\mathbb{R}^d$ given by

$$\mu(dx) = \mathcal{Z}^{-1} \exp(-U(x))\, dx = (2\pi)^{d/2} \mathcal{Z}^{-1} \exp(-V(x))\gamma^d(dx)\,, \qquad (2.72)$$

where V is a function in $C^4(\mathbb{R}^d)$,

$$U(x) = \frac{1}{2}|x|^2 + V(x), \qquad (2.73)$$

γ^d denotes the d-dimensional standard normal distribution, and

$$\mathcal{Z} = \int \exp(-U(x))\, dx.$$

We assume that we are given a norm $\|\cdot\|_-$ on $\mathbb{R}^d$ such that

$$\|x\|_- \ \leq \ |x| \ \leq \ d\|x\|_- \qquad \text{for any } x \in \mathbb{R}^d, \qquad (2.74)$$

as well as finite constants $C_n \in [0, \infty)$, $p_n \in \{0, 1, 2, \ldots\}$, and $K_c, \mathcal{R}_c \in (0, \infty)$ such that the following conditions hold for any $n \in \{1, \ldots, 4\}$:

$$|\partial^n_{\xi_1, \ldots, \xi_n} U(x)| \ \leq \ C_n \max(1, \|x\|_-^{p_n})\|\xi_1\|_- \cdots \|\xi_n\|_- \qquad \forall\, x, \xi_1, \ldots, \xi_n \in \mathbb{R}^d. \quad (2.75)$$

$$(\partial_{\xi\xi} U)(x) \ \geq \ K_c|\xi|^2 \qquad \forall\, x, \xi \in \mathbb{R}^d : |x| \geq \mathcal{R}_c. \qquad (2.76)$$

Here (2.76) can be interpreted as strict convexity of U outside a Euclidean ball.

Remark 2.14. (i) For discretizations of infinite-dimensional models, $\|\cdot\|_-$ is typically a finite-dimensional approximation of a norm that is almost surely finite w.r.t. the limit measure in infinite dimensions, see for instance [16, Example 1.6]. Correspondingly, we may assume that the measure concentrates on a ball of a fixed radius w.r.t. $\|\cdot\|_-$. This will be relevant for the application of Theorem 2.16 below, which states uniform contractivity on such balls.

(ii) Condition (2.75) is the same condition that has been assumed in the strictly convex case in [16].

(iii) In (2.76), we assume strict convexity outside a ball of fixed radius w.r.t. the Euclidean norm and not w.r.t. $\|\cdot\|_-$. Such a bound can be expected to hold with $\mathcal{R}_c$ independent of the dimension if, for example, the non-convexity occurs only in finitely many directions. The application of a coupling approach in situations where (2.76) does not hold requires more advanced techniques, see e.g. [44].

The transition step of a Metropolis-Hastings chain with proposal density $p(x, y)$ and target distribution $\mu(dx) = \mu(x)\, dx$ is given by

$$\widetilde{x} = \begin{cases} X' & \text{if } \widetilde{U} \leq \alpha(x, X') \\ x & \text{otherwise.} \end{cases}, \qquad (2.77)$$

where x is the previous position, X' is the proposed move,

$$\alpha(x,y) = \min\left(1, \frac{\mu(y)p(y,x)}{\mu(x)p(x,y)}\right)$$

is the Metropolis-Hastings acceptance probability, and $\widetilde{U} \sim \mathrm{Unif}(0,1)$ is a uniform random variable that is independent of X'. We consider the proposal

$$X' = x - \frac{h}{2}x - \frac{h}{2}\nabla V(x) + \sqrt{h - \frac{h^2}{4}}Z, \quad Z \sim N(0, I_d), \tag{2.78}$$

where $h \in (0,2)$ is the step size of the time discretization. The corresponding proposal kernel is $p_h(x,\cdot) = N(x - \frac{h}{2}x - \frac{h}{2}\nabla V(x), (h - h^2/4)I_d)$. Substituting $h = \frac{\varepsilon}{1+\varepsilon/4}$, we see that the proposal is a transition step of the semi-implicit Euler discretization

$$X' = x - \frac{\varepsilon}{2}\frac{X'+x}{2} - \frac{\varepsilon}{2}\nabla V(x) + \sqrt{\varepsilon}Z \tag{2.79}$$

for the Langevin SDE $dX_t = -\frac{1}{2}X_t\,dt - \frac{1}{2}\nabla V(X_t)\,dt + dB_t$ with invariant measure μ. The reason for considering the semi-implicit instead of the explicit Euler approximation is that under appropriate conditions, the acceptance probability

$$\alpha_h(x,y) = \min\left(1, \frac{\mu(y)p_h(y,x)}{\mu(x)p_h(x,y)}\right) \tag{2.80}$$

for the corresponding Metropolis-Hastings scheme has a better dimension dependence. Indeed, if V vanishes, then $\alpha_h(x,y) = 1$. More generally, if (2.75) holds, then the average rejection probability is of order $O(h^{\frac{3}{2}})$.

Lemma 2.15 (Upper bounds for rejection probability). Suppose that (2.75) holds and let $k \in \mathbb{N}$. Then there exists an explicit polynomial $P_k\colon \mathbb{R}^2 \to \mathbb{R}_+$ of degree $\max(p_3+3, 3p_2+2)$ such that for any $x \in \mathbb{R}^2$ and $h \in (0,2)$,

$$\mathbb{E}[(1 - \alpha_h(x,X'))^k]^{\frac{1}{k}} \le P_k(\|x\|_-, \|x + \nabla V(x)\|_-)h^{\frac{3}{2}}.$$

The proof of the lemma is given in [16, Proposition 1.7]. The polynomials P_k are explicit. Their coefficients depend only on the constants C_2, C_3, p_2 and p_3 in (2.75) and on the moments

$$m_n = \mathbb{E}[\|Z\|_-^n],\ n \le k\max(p_3 + 3, 2p_2 + 2).$$

Apart from replacing $\sqrt{h}$ by $\sqrt{h - \frac{h^2}{4}}$, (2.78) coincides with the explicit Euler discretization of the SDE $dX_t = b(X_t)\,dt + dB_t$, where $b(x) = -\frac{1}{2}x - \frac{1}{2}\nabla V(x)$.

Therefore, the results in the last section apply to the proposal chain, thus yielding a contraction rate of order $\Omega(h)$. Since the rejection probability is of higher order, we can then apply the perturbation result in (2.5) to prove a corresponding contractivity for the MALA chain. To this end, we consider the coupling $(\widetilde{X}, \widetilde{Y})$ of transition steps of the MALA chain from positions $x, y \in \mathbb{R}^d$ given by (2.77) and

$$\widetilde{Y} = \begin{cases} Y' & \text{if } \widetilde{U} \le \alpha(y, Y') \\ y & \text{otherwise.} \end{cases} \tag{2.81}$$

where (X', Y') is the (optimal) coupling for the proposal steps considered in (2.40), and $\widetilde{U} \sim \mathrm{Unif}(0,1)$ is independent of both X' and Y'. Hence, the proposals are coupled optimally and the same uniform random variable $\widetilde{U}$ is used to decide about acceptance or rejection for each of the steps. Nevertheless, in general $(\widetilde{X}, \widetilde{Y})$ is not an optimal coupling of the corresponding MALA transition probabilities.

Theorem 2.16 (Contraction rates for MALA)**.** Suppose that conditions (2.75) and (2.76) hold and fix $R \in (0, \infty)$. Then there exists a concave strictly increasing function $\widetilde{f} \colon [0, \infty) \to [0, \infty)$ with $\widetilde{f}(0) = 0$ and constants $c_3, h_0 \in (0, \infty)$ such that for any $h \in (0, h_0 d^{-1})$ and for any $x, y \in \mathbb{R}^d$ with $\|x\|_- \le R$, $\|y\|_- \le R$,

$$\mathbb{E}_{x,y}[\widetilde{f}(|\widetilde{X} - \widetilde{Y}|)] \le (1 - c_3 h)\widetilde{f}(|x - y|) \,. \tag{2.82}$$

The function $\widetilde{f}$ and the constants c_3 and h_0 depend only on R and on the values of the constants C_n, p_n, K_c, $\mathcal{R}_c$ in assumptions (2.75), (2.76).

The proof of Theorem 2.16 is given in Section 7.

Remark 2.17. The theorem shows that by choosing the step size of order $\Theta(d)$, a contraction rate of the same order holds on balls w.r.t. $\|\cdot\|_-$ provided conditions (2.75) and (2.76) are satisfied. In the strictly convex case, it has been shown in [16] by a synchronous coupling that a corresponding result holds even for step sizes of order $\Theta(1)$ if the Euclidean norm in (2.82) is replaced by $\|\cdot\|_-$. One could hope for a similar result in the not globally convex case, but the combination of reflection coupling with a different norm leads to further difficulties. A possibility to overcome these difficulties might be the two-scale approach developed in [44].

3 Proofs of Theorems 2.2 and 2.3

In this section, we prove the first two theorems. We first specify the explicit choice of the metric and the explicit values of the contraction rate c. The reason for choosing the metric this way will become clear by the subsequent proofs of the theorems.

For $r, s > 0$, we consider the intervals

$$I_r = ((r - \varepsilon)^+, r) \tag{3.1}$$

and the dual intervals

$$\hat{I}_s = \{r > r_0 : s \in I_r\} = (s, s + \varepsilon) \cap (r_0, \infty) \,. \tag{3.2}$$

For $r \in (r_0, \infty)$ we set

$$\overline{\gamma}(r) = 2\overline{\beta}(r)/\underline{\alpha}(r) \,. \tag{3.3}$$

Let $\widetilde{\gamma} \colon [0, \infty) \to [0, \infty)$ be a function satisfying

$$\sup_{r \in \hat{I}_s} \overline{\gamma}(r) \quad \le \quad \widetilde{\gamma}(s) \quad \text{for any } s \in [0, \infty), \qquad \text{i.e.,} \tag{3.4}$$

$$\overline{\gamma}(r) \quad \le \quad \widetilde{\gamma}(s) \quad \text{for any } r > r_0, \ s \in I_r \,. \tag{3.5}$$

By assumptions (A1) and (A2), such a function exists. If (A3) holds, then we may assume w.l.o.g. that $\widetilde{\gamma}(s) = 0$ for large s.

3.1 Choice of the metric in Theorem 2.2

Suppose conditions (A1), (A2) and (A3) hold. We set

$$r_1 := \sup\{r > 0 : \widetilde{\gamma}(r) > 0\}, \tag{3.6}$$

where $\sup \emptyset = 0$. By Assumption (A3) we can choose $\widetilde{\gamma}$ such that r_1 is finite. We have

$$\widetilde{\gamma}(r) = 0 \,, \qquad \overline{\gamma}(r) \le 0 \,, \qquad \overline{\beta}(r) \le 0 \quad \text{for any } r \ge r_1 \,. \tag{3.7}$$

We also fix a constant $r_2 \in (r_1, \infty)$. The value of r_2 will be determined in condition (3.12) below. The underlying metric we consider is given by (2.15), where $f \colon [0, \infty) \to [0, \infty)$ is a concave increasing function defined by

$$f(r) = a\mathbb{1}_{r>0} + \int_0^r \varphi(s \wedge r_2) g(s \wedge r_2)\, ds \tag{3.8}$$

with decreasing differentiable functions φ and g such that $\varphi(0) = g(0) = 1$ and a constant $a \in (0, \infty)$ that are all specified below. Hence, f is twice differentiable except at 0, $f(0+) - f(0) = a$, $f' = \varphi g$ on $(0, r_2)$, and f' is constant on $[r_2, \infty)$.

The function φ and the constant a are chosen such that

$$\varphi(r) = \exp\left(-\int_0^r \widetilde{\gamma}(s)\, ds\right), \qquad \text{and} \tag{3.9}$$

$$a \geq r_0 + 2 \sup_{|x-y| \leq r_0} \frac{\beta(x,y)}{\pi(x,y)}. \tag{3.10}$$

Notice that by (3.7), the function $\varphi(r)$ is constant for $r \geq r_1$. Setting

$$\Phi(r) = \int_0^r \varphi(s)\, ds, \tag{3.11}$$

the constant r_2 is chosen such that

$$-\frac{\overline{\beta(r)}}{a + \Phi(r)} \geq \frac{1}{2}\left(\int_{r_1}^{r_2} \frac{\Phi(s)}{\underline{\alpha}(s)}\, ds\right)^{-1} \quad \text{for all } r \geq r_2. \tag{3.12}$$

Assumption (A3) ensures that such a constant exists. Indeed, for $r \geq r_1$, $\widetilde{\gamma}$ vanishes, whence φ is constant and Φ is linear. By definition of α, we see that $\underline{\alpha}$ is uniformly bounded by ε^2. Therefore, the value on the right hand side of (3.12) goes to zero as $r_2 \to \infty$, and (3.12) holds for large r_2 by (A3).

The contraction rate is now given by

$$c = \min\left(\frac{1}{2} \inf_{r \leq r_0} \pi(r), \frac{1}{4}\left(\int_0^{r_2} \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{a + \Phi(u)}{\underline{\alpha}(u)}\, ds\right)^{-1}\right) \tag{3.13}$$

and the function g is defined as

$$g(r) = 1 - 2c \int_0^r \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{a + \Phi(u)}{\underline{\alpha}(u)}\, ds. \tag{3.14}$$

Note that (3.13) guarantees that $g(r) \geq \frac{1}{2}$ for $r \leq r_2$.

3.2 Choice of the metric in Theorem 2.3

Now suppose that (A1), (A2) and (A4) hold. In this case we set

$$r_1 := \sup\{d(x,y) : x, y \in S,\, V(x) + V(y) < 4C/\lambda\}. \tag{3.15}$$

By (A4)b and (A2), r_1 is finite. Moreover, by (A4)a,

$$\mathbb{E}_{x,y}[V(X') + V(Y')] \leq \left(1 - \frac{\lambda}{2}\right)(V(x) + V(y)) \qquad \text{if } d(x,y) \geq r_1. \tag{3.16}$$

We also fix a constant $r_2 \in (r_1, \infty)$. The value of r_2 will be determined by condition (3.21) below.

The function $f \colon [0, \infty) \to [0, \infty)$ determining the metric in (2.17) is now defined by

$$f(r) = a \mathbb{1}_{r>0} + \int_0^r \varphi(s \wedge r_2) g(s \wedge r_2) \, ds \qquad (3.17)$$

with decreasing differentiable functions φ and g such that $\varphi(0) = g(0) = 1$, and a constant $a \in (0, \infty)$ that are all specified below. Hence, f is twice differentiable except at 0, $f(0+) - f(0) = a$, $f' = \varphi g$ on $(0, r_2)$ and f' is constant on $[r_2, \infty)$.

The function φ and the constants a and M in (2.17) are chosen such that

$$\varphi(r) = \exp\left(- \int_0^r \widetilde{\gamma}(s) \, ds \right), \qquad (3.18)$$

$$M \leq \frac{1}{4} \left(\int_0^{r_1} \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{1}{\underline{\alpha}(u)} \, ds \right)^{-1}, \qquad (3.19)$$

$$a \geq r_0 + 2 \sup_{|x-y| \leq r_0} \frac{\beta(x,y) + M}{\pi(x,y)}, \qquad (3.20)$$

$$\overline{\beta}(r)\varphi(r) \leq \frac{\lambda M}{16\,C}(V(x) + V(y)) \ \text{ if } |x - y| \geq r_2. \qquad (3.21)$$

By (A4)b and since $\varphi \leq 1$, there always exists a finite r_2 such that (3.21) holds. To optimize the estimates, we choose r_2 as small as possible, i.e., we set

$$r_2 = r_1 \vee \sup\left\{ d(x,y) = r : x, y \in S, V(x) + V(y) < \frac{16\,C}{\lambda M}\overline{\beta}(r)\varphi(r) \right\}. \qquad (3.22)$$

Setting

$$\Phi(r) = \int_0^r \varphi(s) \, ds, \qquad (3.23)$$

the contraction rate c is given by

$$c = \min\left(\frac{1}{2} \inf_{r \leq r_0} \underline{\pi}(r), \ \frac{\lambda}{4}, \ \frac{\lambda M}{16\,C} \inf_{r \geq r_2} \frac{V(x) + V(y)}{\Phi(r)}, \ \frac{1}{8}\left(\int_0^{r_2} \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{a + \Phi(u)}{\underline{\alpha}(u)} \, ds \right)^{-1} \right). \qquad (3.24)$$

and the function g is defined as

$$g(r) = 1 - 2c \int_0^r \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{a + \Phi(u)}{\underline{\alpha}(u)} \, ds - M \int_0^{r \wedge r_1} \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{1}{\underline{\alpha}(u)} \, ds. \qquad (3.25)$$

Note that (3.24) and (3.19) guarantee that $g(r) \geq \frac{1}{2}$ for $r \leq r_2$. In the minimum defining c, the first term guarantees contractivity for $r \leq r_0$, the second term is used for all r, the third term guarantees contractivity for $r \geq r_2$ and the last term ensures contractivity with rate c for $r_0 < r \leq r_1$.

3.3 Proof of Theorem 2.2 and Theorem 2.3

Since the arguments are similar, we prove both theorems simultaneously, distinguishing cases where needed. In the situation of Theorem 2.2, we set $M = 0$. Let $x, y \in \mathbb{R}^d$ such that $r = |x - y| > 0$. Since $f''(t) \leq 0$ for all $t > 0$,

$$
\begin{aligned}
f(R') - f(r) &= -a\mathbb{1}_{R'=0} + \int_r^{R'} f'(s) \, ds \\
&= -a\mathbb{1}_{R'=0} + (R' - r)f'(r) + \int_r^{R'} \int_r^s f''(t) \, dt \, ds \\
&\leq -a\mathbb{1}_{R'=0} + (R' - r)f'(r) + \frac{1}{2}((R' - r)^- \wedge \varepsilon)^2 \sup_{u \in I_r} f''(u),
\end{aligned}
$$

where $I_r = ((r - \varepsilon)^+, r)$. By taking expectations, we conclude that

$$\mathbb{E}_{x,y}[f(R') - f(r)] \leq -a\pi(x,y) + \beta(x,y)f'(r) + \frac{1}{2}\alpha(x,y) \sup_{u \in I_r} f''(u). \tag{3.26}$$

Our goal is to compensate the second term by the first term for $r \leq r_0$ and by the last term for $r_0 < r \leq r_2$ (and possibly by a Lyapunov part for $r \geq r_2$). In order to verify (2.16) and (2.18), we now distinguish three cases.

Case $r \in (r_0, r_2)$. Since $f' = g\varphi$ on $(0, r_2)$, we have

$$\sup_{I_r} f'' \leq \sup_{I_r}(g'\varphi) + \sup_{I_r}(g\varphi'). \tag{3.27}$$

Note that both summands are negative since g and φ are decreasing. Now we note first that our choice of φ guarantees that

$$\frac{1}{2}\alpha(x,y) \sup_{I_r}(g\varphi') + \beta(x,y)f'(r) \leq 0. \tag{3.28}$$

Indeed, (3.28) is satisfied provided

$$\sup_{I_r}(g\varphi') \leq -\overline{\gamma}(r)g(r)\varphi(r). \tag{3.29}$$

Since $\varphi' \leq 0$ and g is decreasing, we have

$$\sup_{I_r}(g\varphi') \leq \inf_{I_r} g \sup_{I_r} \varphi' \leq g(r) \sup_{I_r} \varphi'.$$

Hence, (3.29) is satisfied if

$$\sup_{s \in I_r} \varphi'(s) \leq -\overline{\gamma}(r)\varphi(r). \tag{3.30}$$

But indeed, by definition of φ and $\widetilde{\gamma}$, we have for $s \in I_r$

$$\varphi'(s) = -\widetilde{\gamma}(s)\varphi(s) \leq -\overline{\gamma}(r)\varphi(r).$$

Next, we observe that our choice of g (in particular, $g \leq 1$) guarantees that

$$\frac{1}{2}\alpha(x,y) \sup_{I_r}(g'\varphi) + M\mathbb{1}_{r<r_1} \leq -cf(r). \tag{3.31}$$

Indeed, since $f(r) \leq a + \Phi(r)$ and $f'(r) \leq 1$, it is sufficient to show

$$\sup_{s \in I_r}(g'(s)\varphi(s)) \leq -2c\frac{a + \Phi(r)}{\underline{\alpha}(r)} - \frac{M\mathbb{1}_{r<r_1}}{\underline{\alpha}(r)} \quad \text{if } r_0 < r < r_2, \tag{3.32}$$

or

$$g'(s)\varphi(s) \leq -2c\sup_{r \in \hat{I}_s}\frac{a + \Phi(r)}{\underline{\alpha}(r)} - \sup_{r \in \hat{I}_s}\frac{M\mathbb{1}_{r<r_1}}{\underline{\alpha}(r)} \quad \text{if } 0 < s \leq r_2. \tag{3.33}$$

In (3.14), (3.25) respectively, the function g has been defined in such a way that this condition is satisfied. Now, by combining (3.26), (3.28) and (3.31), and bounding the term $-a\pi(x,y)$ in (3.26) by zero, we obtain for $r \in (r_0, r_2)$:

$$\mathbb{E}_{x,y}[f(R') - f(r)] \leq -M\mathbb{1}_{r<r_1} - cf(r). \tag{3.34}$$

In the setup of Theorem 2.2, we have chosen $M = 0$ and $\rho(x,y) = f(d(x,y))$. Hence, (3.34) implies the assertion

$$\mathbb{E}_{x,y}[\rho(X', Y')] \leq (1 - c)\rho(x,y) \quad \text{for } r = d(x,y) \in (r_0, r_2). \tag{3.35}$$

In the setup of Theorem 2.3, by Assumption (A4),

$$\mathbb{E}_{x,y}[V(X') + V(Y')] \leq (1-\lambda)(V(x) + V(y)) + 2C. \qquad (3.36)$$

Since $\rho(x,y) = f(d(x,y)) + \frac{M}{2C}(V(x) + V(y))\mathbb{1}_{x \neq y}$, we obtain for $r \in (r_0, r_1)$:

$$
\begin{aligned}
\mathbb{E}_{x,y}[\rho(X', Y')] & \qquad\qquad (3.37) \\
&\leq \quad -M + (1-c)f(d(x,y)) + \frac{M}{2C}(1-\lambda)(V(x) + V(y)) + M \\
&\leq \quad (1-c)\rho(x,y).
\end{aligned}
$$

Here the last inequality holds since $\lambda \geq c$. On the other hand, for $r \in [r_1, r_2)$, we have $V(x) + V(y) \geq 4C/\lambda$ by (3.15). Hence in this case, by (3.36),

$$\mathbb{E}_{x,y}[V(X') + V(Y')] \leq (1 - \lambda/2)(V(x) + V(y)). \qquad (3.38)$$

Since $c \leq \lambda/2$, (3.34) and (3.38) then again imply

$$
\begin{aligned}
\mathbb{E}_{x,y}[\rho(X', Y')] &\leq \quad (1-c)f(d(x,y)) + \left(1 - \frac{\lambda}{2}\right)\frac{M}{2C}(V(x) + V(y)) \qquad (3.39) \\
&\leq \quad (1-c)\rho(x,y).
\end{aligned}
$$

Case $r \leq r_0$. Noting that $f'' \leq 0$ and $f' \leq 1$ and applying (A2), we see that for $r \in (0, r_0)$, (3.26) implies

$$\mathbb{E}_{x,y}[f(R') - f(r)] \leq -a\pi(x,y) + \beta(x,y) \leq -M - c(a + r_0) \leq -M - cf(r) \qquad (3.40)$$

provided $c \leq \frac{1}{2}\pi(x,y)$ and

$$\frac{a}{2}\pi(x,y) \geq \frac{r_0}{2}\pi(x,y) + M + \beta(x,y). \qquad (3.41)$$

This condition is satisfied by our choice of a, cf. (3.20). Again, using (3.36), and since $c \leq \lambda/2$, we obtain

$$
\begin{aligned}
&\mathbb{E}_{x,y}[\rho(X', Y')] \\
&\qquad \leq -M + (1-c)f(r) + (1-\lambda)\frac{M}{2C}(V(x) + V(y)) + 2C\frac{M}{2C} \leq (1-c)\rho(x,y).
\end{aligned}
$$

Case $r \geq r_2$. Here, we use the bound $f(R') - f(r) \leq (R' - r)f'(r)$ yielding

$$\mathbb{E}_{x,y}[f(R') - f(r)] \leq \beta(x,y)f'(r). \qquad (3.42)$$

Now, we consider first the setup of Theorem 2.2. Here, for $r \geq r_2$, we have

$$\mathbb{E}_{x,y}[f(R') - f(r)] \leq \beta(r)f'(r) \leq \beta(r)\frac{\varphi(r_1)}{2},$$

where we have used that by (3.7), $\beta(r) \leq 0$, $f' \geq \varphi/2$ and φ is constant on $[r_1, \infty)$. To prove that the right hand side is bounded from above by $-cf(r)$, it is sufficient to show

$$c(a + \Phi(r)) \leq -\beta(r)\frac{\varphi(r_1)}{2} \quad \text{for any } r \geq r_2.$$

We claim that this holds by the definition of r_2. Indeed, by the definition of c,

$$c^{-1} \geq 4\int_0^{r_2} \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{a + \Phi(u)}{\underline{\alpha}(u)} \, ds \geq 4\int_{r_1}^{r_2} \frac{\Phi(s)}{\varphi(s)\underline{\alpha}(s)} \, ds = \frac{4}{\varphi(r_1)}\int_{r_1}^{r_2} \frac{\Phi(s)}{\underline{\alpha}(s)} \, ds.$$

Hence, by (3.12),

$$c(a + \Phi(r)) \leq \frac{1}{4} \left(\int_{r_1}^{r_2} \frac{\Phi(s)}{\underline{\alpha}(s)} \, ds \right)^{-1} (a + \Phi(r))\varphi(r_1) \leq -\beta(r)\frac{\varphi(r_1)}{2},$$

and thus

$$\mathbb{E}_{x,y}[\rho(X', Y') - \rho(x,y)] \leq -c(a + \Phi(r)) \leq -c\rho(x,y).$$

Finally, we now show contractivity for $r \geq r_2$ under the conditions in Theorem 2.3. Here, by (3.38) and (3.26),

$$\begin{aligned}
\mathbb{E}_{x,y}[\rho(X', Y')] &= \mathbb{E}_{x,y}[f(R')] + \frac{M}{2C}\mathbb{E}_{x,y}[V(X') + V(Y')] &\text{(3.43)} \\
&\leq f(r) + \overline{\beta}(r)f'(r) + \frac{M}{2C}\left(1 - \frac{\lambda}{2}\right)(V(x) + V(y)).
\end{aligned}$$

Since $c \leq \lambda/4$ by its definition, we obtain

$$\mathbb{E}_{x,y}[\rho(X', Y')] \leq (1 - c)\left(f(r) + \frac{M}{2C}(V(x) + V(y)) \right) = (1 - c)\rho(x,y) \qquad \text{(3.44)}$$

provided

$$cf(r) + \overline{\beta}(r)f'(r) \leq \frac{M}{2C}\frac{\lambda}{4}(V(x) + V(y)). \qquad \text{(3.45)}$$

However, due to our choice of r_2 in (3.22) and since $f' \leq \varphi$, we have

$$\overline{\beta}(r)f'(r) \leq \frac{\lambda M}{16C}(V(x) + V(y)) \quad \text{if } |x - y| \geq r_2.$$

Moreover, due to our choice of c in (3.24) and since $f \leq \Phi$, we get

$$cf(r) \leq \frac{\lambda M}{16C}(V(x) + V(y)) \quad \text{if } |x - y| \geq r_2.$$

Hence (3.45) is indeed satisfied for $r \geq r_2$ and the proof is complete.

4 Proof of Theorem 2.4

For proving Theorem 2.4, we proceed in a similar way as in the proofs of Theorem 2.2 and Theorem 2.3 above. Suppose that conditions (B1), (B2) and (B3) hold. Now, the intervals I_r, $r \in (0, \infty)$ are given by (2.19) and we consider the dual intervals $\hat{I}_s$, $s \in (0, \infty)$ defined by

$$\hat{I}_s = \{r \in (0, \infty) : s \in I_r\}. \qquad \text{(4.1)}$$

By (2.23), $I_r = (r - \ell(r), r)$ for $r \geq r_0$ and $I_r \subseteq (0, 2r_0)$ for $r < r_0$. Therefore

$$\hat{I}_s = \{r > s : r - \ell(r) < s\} \qquad \text{for } s \geq 2r_0, \text{ and} \qquad \text{(4.2)}$$

$$\hat{I}_s \subseteq \{r > s : r - \ell(r) < 2r_0\} \qquad \text{for } s < 2r_0. \qquad \text{(4.3)}$$

Let $\overline{\gamma}(r) = 2\overline{\beta}(r)/\underline{\alpha}(r)$ as in (3.3). Similarly as in the proof of Theorem 2.2 and Theorem 2.3, we assume that $\widetilde{\gamma} \colon [0, \infty) \to [0, \infty)$ is a function satisfying

$$\sup_{\hat{I}_s} \overline{\gamma} \leq \widetilde{\gamma}(s) \quad \text{for any } s \in (0, \infty), \qquad \text{(4.4)}$$

$$4\sup_{\hat{I}_s} \overline{\gamma} \leq \widetilde{\gamma}(s) \quad \text{for any } s \in (0, 2r_0), \quad \text{and} \qquad \text{(4.5)}$$

$$\int_0^{2r_0} \widetilde{\gamma}(s) \, ds \leq \log 2. \qquad \text{(4.6)}$$

Note the additional factor 4 that has been introduced for technical reasons for $s < 2r_0$. In applications, this will usually not affect the bounds too much, as typically r_0 is a small constant. Condition (4.6) can always be satisfied by choosing r_0 small enough. As in (3.6), we set

$$r_1 := \sup\{r > 0 : \widetilde{\gamma}(r) > 0\}, \tag{4.7}$$

where $\sup \emptyset = 0$. Similarly as below (3.6), by Assumption (B3), we can choose $\widetilde{\gamma}$ such that r_1 is finite. The metric is chosen similarly as in the proof of Theorem 2.2 above, where now $a = 0$. We define

$$f(r) = \int_0^r \varphi(r \wedge r_2) g(s \wedge r_2) \, ds. \tag{4.8}$$

Here

$$\varphi(r) = \exp\left(-\int_0^r \widetilde{\gamma}(s) \, ds\right), \qquad \Phi(r) = \int_0^r \varphi(s) \, ds, \tag{4.9}$$

the constant r_2 is chosen such that

$$\frac{-\overline{\beta}(r)}{\Phi(r)} \geq \frac{1}{8}\left(\int_{r_1}^{r_2} \frac{\Phi(s)}{\underline{\alpha}(s)} \, ds\right)^{-1} \quad \text{for } r \geq r_2, \tag{4.10}$$

and

$$g(r) = 1 - 2c \int_0^r \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{\Phi(u)}{\underline{\alpha}(u)} \, ds, \tag{4.11}$$

where the contraction rate c is given by

$$c = \frac{1}{4}\left(\int_0^{r_2} \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{\Phi(u)}{\underline{\alpha}(u)} \, ds\right)^{-1}. \tag{4.12}$$

Proof of Theorem 2.4. Let $x, y \in \mathbb{R}^d$ and $r = d(x, y)$.

For $r \geq r_2$, (2.26) follows in the same way as in the proof of Theorem 2.2 (with $a = 0$). The crucial assumption for this is (4.10), which holds due to (B3), by analogy to (3.12) in the proof of Theorem 2.2, which holds due to (A3).

Now assume that $r < r_2$. To prove (2.26), we show that

$$\mathbb{E}_{x,y}[f(R') - f(r)] \leq \overline{\beta}(r) f'(r) + \frac{1}{2} \underline{\alpha}(r) \sup_{I_r} f'' \leq -cf(r). \tag{4.13}$$

The first inequality follows similarly as in the proof of Theorem 2.2, cf. (3.26). To prove the second inequality, note that on $(0, r_2)$,

$$f' = g\varphi, \qquad f'' = g\varphi' + g'\varphi \qquad \text{and} \qquad f \leq \Phi. \tag{4.14}$$

By (4.14) it is sufficient to show that φ, g and c have been chosen in such a way that

$$\sup_{I_r}(g\varphi') \leq -2\frac{\overline{\beta}(r)}{\underline{\alpha}(r)} g(r)\varphi(r) \tag{4.15}$$

$$\sup_{I_r}(g'\varphi) \leq -2c\frac{\Phi(r)}{\underline{\alpha}(r)}. \tag{4.16}$$

Then, by (4.13) we can conclude that

$$\mathbb{E}_{x,y}[\rho(X', Y') - \rho(x, y)] = \mathbb{E}_{x,y}[f(R') - f(r)] \leq -c\Phi(r) \leq -cf(r) = -c\rho(x, y).$$

We first verify (4.15). This condition is satisfied provided

$$g(s)\varphi'(s) \leq -\sup_{\hat{I}_s}(\overline{\gamma}g\varphi) \quad \text{for any } s \leq r_2. \tag{4.17}$$

For $s \geq 2r_0$ we have

$$\sup_{\hat{I}_s}(\overline{\gamma}g\varphi) \leq \left(\sup_{\hat{I}_s}\overline{\gamma}\right)\left(\sup_{\hat{I}_s} g\varphi\right) \leq \widetilde{\gamma}(s)g(s)\varphi(s)\,,$$

because $\hat{I}_s \subseteq (s,\infty)$ by (4.2) and since $g\varphi$ is decreasing. Hence, (4.17) holds by definition of φ.

For $s < 2r_0$ we have to argue differently, since, in general, $\hat{I}_s$ is not contained in (s,∞) in this case. Observe first that if $\sup_{\hat{I}_s}(\overline{\gamma}g\varphi) \leq 0$, then (4.17) holds trivially since φ is decreasing. Hence it is sufficient to consider the case of $\sup_{\hat{I}_s}(\overline{\gamma}g\varphi) > 0$. Noting that $g\varphi \leq 1$, we have by (4.5)

$$\sup_{\hat{I}_s}(\overline{\gamma}g\varphi) \leq \sup_{\hat{I}_s}\overline{\gamma}^+ \leq \frac{1}{4}\widetilde{\gamma}(s)$$

and hence, since $g \geq \frac{1}{2}$,

$$g(s)\varphi'(s) \leq \frac{1}{2}\varphi'(s) = -\frac{1}{2}\widetilde{\gamma}(s)\varphi(s) \leq -2\varphi(s)\sup_{\hat{I}_s}(\overline{\gamma}g\varphi)\,.$$

Thus, (4.17) holds for $s < 2r_0$ since by (4.6),

$$\varphi(s) = \exp\left(-\int_0^s \widetilde{\gamma}(u)\,du\right) \geq \frac{1}{2}\,.$$

We thus have shown that (4.17) and hence (4.15) are satisfied. It remains to verify (4.16). This condition holds provided

$$g'(s)\varphi(s) \leq \inf_{\hat{I}_s}\frac{-2c\Phi}{\underline{\alpha}}, \quad \text{for any } s \leq r_2 \tag{4.18}$$

or, equivalently,

$$g'(s) \leq -2c\frac{1}{\varphi(s)}\sup_{\hat{I}_s}\frac{\Phi}{\underline{\alpha}}, \quad \text{for any } s \leq r_2\,. \tag{4.19}$$

The function g has been chosen in (4.11) in such a way that this condition is satisfied. $\qquad\square$

5 Proof of perturbation result

We now prove the perturbation result in Theorem 2.5. Let $x, y \in S$, $x \neq y$. By (2.33), (2.30), (2.31) and (2.32),

$$\begin{aligned}
\mathbb{E}_{x,y}[\widetilde{\rho}(\widetilde{X},\widetilde{Y}) - \widetilde{\rho}(x,y)] &\leq \mathbb{E}_{x,y}[\rho(\widetilde{X},\widetilde{Y}) - \rho(x,y)] - \frac{2b}{p}\mathbb{P}_{x,y}[\widetilde{X} = \widetilde{Y}]\\
&\leq \mathbb{E}_{x,y}[(d(\widetilde{X},\widetilde{Y}) - d(X',Y'))^+] + \mathbb{E}_{x,y}[\rho(X',Y') - \rho(x,y)] - \frac{2b}{p}\mathbb{P}_{x,y}[\widetilde{X} = \widetilde{Y}]\\
&\leq b - \frac{c}{2}\rho(x,y) - 2b\mathbb{1}_{d(x,y)<r_0}\,.
\end{aligned}$$

Note that in the second inequality we have used that f is a contraction. For $d(x,y) < r_0$ we obtain

$$\mathbb{E}_{x,y}[\widetilde{\rho}(\widetilde{X},\widetilde{Y}) - \widetilde{\rho}(x,y)] \leq -\frac{p}{2}\frac{2b}{p} - \frac{c}{2}\rho(x,y) \leq -\frac{1}{2}\min(c,p)\widetilde{\rho}(x,y)\,. \tag{5.1}$$

For $d(x,y) \geq r_0$, we use the fact that $b = cf(r_0)/4$. Hence,

$$\widetilde{\rho}(x,y) = \rho(x,y) + \frac{2b}{p} \leq \left(1 + \frac{c}{2p}\right)\rho(x,y) \leq \max\left(2,\frac{c}{p}\right)\rho(x,y), \quad \text{and}$$

$$\mathbb{E}_{x,y}[\widetilde{\rho}(\widetilde{X},\widetilde{Y}) - \widetilde{\rho}(x,y)] \leq b - \frac{c}{2}\rho(x,y) \leq -\frac{c}{4}\rho(x,y) \leq -\frac{1}{8}\min(c,2p)\widetilde{\rho}(x,y). \tag{5.2}$$

The assertion of Theorem 2.5 follows from (5.1) and (5.2).

6 Proof of results for the Euler scheme

In this section, we prove the contraction results for the Euler scheme.

Proof of Lemma 2.7 (i), (ii) and (iii). We start with reduction steps. At first, we observe that the definitions of $\hat{\beta}(x,y), \hat{\alpha}(x,y)$ and $\hat{\pi}(x,y)$ only depend on $\hat{r} = |\hat{x} - \hat{y}|$ and $R' = |X' - Y'|$. Thus, the assertions (i) (ii), (iii) are statements about the coupled random walk transition step $(\hat{x}, \hat{y}) \to (X', Y')$ defined by (2.40), and we may assume w.l.o.g. that $(\hat{x}, \hat{y}) = (x, y)$. Furthermore, $\hat{r}$ and the law of R' under $\mathbb{P}_{x,y}$ are invariant under translations and rotations of the underlying state space $\mathbb{R}^d$. Therefore, we may even assume w.l.o.g. that $\hat{x} = x = 0$ and $\hat{y} = y = re_1$, where $r = \hat{r}$ and $e_1, \ldots, e_d$ denotes the canonical basis of $\mathbb{R}^d$. Then

$$X' = \sqrt{h}Z, \qquad Y'_{\text{refl}} = re_1 + \sqrt{h}(I_d - 2e_1e_1^T)Z, \qquad \text{and} \qquad (6.1)$$

$$\phi_{\hat{y},hI}(X')/\phi_{\hat{x},hI}(X') = \phi_{r,h}(X'_1)/\phi_{0,h}(X'_1), \qquad (6.2)$$

where $X'_i = e_i^T X'$. Thus, by (2.40), $Y'_i = X'_i$ for $i \geq 2$, and

$$Y'_1 = \begin{cases} X'_1 & \text{if } U \leq \phi_{r,h}(X'_1)/\phi_{0,h}(X'_1), \\ r - \sqrt{h}Z & \text{otherwise.} \end{cases} \qquad (6.3)$$

In particular, $R' = |X' - Y'| = |(X' - Y')e_1| = |X'_1 - Y'_1|$. Since this is distributed as in the one-dimensional case, we may assume w.l.o.g. $d = 1$.

We are now left with a simple one-dimensional problem where $x = 0$, $y = r$, and $\hat{r} = r = |x - y|$. The coupling is given by

$$X' = \sqrt{h}Z, \qquad Y' = \begin{cases} X' & \text{if } U \leq \phi_{r,h}(X')/\phi_{0,h}(X'), \\ r - X' & \text{otherwise,} \end{cases} \qquad (6.4)$$

where $Z \sim N(0,1)$ and $U \sim \text{Unif}(0,1)$ are independent. Hence $X' \sim N(0,h)$, the conditional probability given Z that $Y' = X'$ is $\min(1, \phi_{r,h}(X')/\phi_{0,h}(X'))$, and if $Y' \neq X'$, then $R' = |X' - Y'| = |r - 2X'|$. Since $\phi_{r,h}(t) \leq \phi_{0,h}(t)$ if and only if $t \leq r/2$, we obtain

$$\begin{aligned} \mathbb{E}_{x,y}[R'] &= \int_{-\infty}^{\infty} |r - 2t| \, (1 - \phi_{r,h}(t)/\phi_{0,h}(t))^+ \, \phi_{0,h}(t) \, dt \\ &= \int_{-\infty}^{r/2} (r - 2t)(\phi_{0,h}(t) - \phi_{r,h}(t)) \, dt \\ &= \frac{1}{2} \int_{-\infty}^{\infty} (r - 2t)(\phi_{0,h}(t) - \phi_{r,h}(t)) \, dt = r. \end{aligned}$$

Here we have used in the third step that the integrand is symmetric w.r.t. $t = r/2$, i.e., invariant under the transformation $t \mapsto r - t$. Thus $\hat{\beta}(x,y) = \mathbb{E}_{x,y}[R' - r] = 0$, which proves Assertion (i).

Next, we are going to prove the lower bound for $\hat{\alpha}(x,y)$. Recall from (2.43) and (2.44) that $I_r = (0, r + \sqrt{h})$ for $r < \sqrt{h}$ and $I_r = (r - \sqrt{h}, r)$ for $r \geq \sqrt{h}$. We first consider the

case $r \geq \sqrt{h}$. Similarly as above, we obtain

$$
\begin{aligned}
\hat{\alpha}(x, y) &\geq \mathbb{E}_{0,r}[(R' - r)^2 ; R' \in I_r] \\
&\geq \int_{-\infty}^{r/2} (r - 2t - r)^2 \mathbb{1}_{I_r}(r - 2t) \left(\phi_{0,h}(t) - \phi_{r,h}(t)\right) dt \\
&= 4 \int_0^{\sqrt{h}/2} t^2 \left(1 - e^{(rt - r^2/2)/h}\right) \phi_{0,h}(t) \, dt \\
&= 4h \int_0^{1/2} u^2 \left(1 - e^{\frac{r}{\sqrt{h}}\left(u - \frac{r}{2\sqrt{h}}\right)}\right) \phi_{0,1}(u) \, du \\
&\geq 4h \int_0^{1/2} u^2 \left(1 - e^{u - 1/2}\right) \phi_{0,1}(u) \, du.
\end{aligned}
\tag{6.5}
$$

Here we have used in the last step that $s \mapsto s(u - s/2)$ is decreasing for $s \geq u$, and $r/\sqrt{h} \geq 1 \geq u$ for $u \in [0, 1/2]$. Note that in the second step we only use the reflection behaviour of the coupling. This is due to the fact that the contribution from jumping to the same point would be of negligible order in h. Now assume $r < \sqrt{h}$. Then $r - 2t \in I_r$ if and only if $t \in (-\frac{\sqrt{h}}{2}, \frac{r}{2})$. Thus,

$$
\begin{aligned}
\hat{\alpha}(x, y) &\geq \int_{-\sqrt{h}/2}^0 (r - 2t - r)^2 \left(\phi_{0,h}(t) - \phi_{r,h}(t)\right) dt \\
&= 4h \int_{-1/2}^0 u^2 \left(1 - e^{\frac{r}{\sqrt{h}}\left(u - \frac{r}{2\sqrt{h}}\right)}\right) \phi_{0,1}(u) \, du \\
&\geq 4(1 - e^{-1}) h \frac{r}{\sqrt{h}} \int_0^{1/2} u^3 \phi_{0,1}(u) \, du.
\end{aligned}
\tag{6.6}
$$

Here, we have used in the last step that for $r < \sqrt{h}$ and $u \in [-1/2, 0]$, we have $s := \frac{r}{\sqrt{h}}\left(u - \frac{r}{2\sqrt{h}}\right) \in [-1, 0]$ and hence $e^s - 1 \leq (1 - e^{-1})s$. By combining (6.5) and (6.6), we obtain $\hat{\alpha}(x, y) \geq c_0 \min(r, \sqrt{h})\sqrt{h}$, where

$$
c_0 = 4 \min \left(\int_0^{1/2} u^2 (1 - e^{u - 1/2}) \phi_{0,1}(u) \, du, \; (1 - e^{-1}) \int_0^{1/2} u^3 \phi_{0,1}(u) \, du \right) \geq 0.007.
$$

This proves Assertion (ii).

Finally, for $X' \geq r/2$, we have $\phi_{r,h}(X') \geq \phi_{0,h}(X')$, and hence $Y' = X'$. Thus,

$$
\begin{aligned}
\pi(x, y) &= \mathbb{P}_{x,y}[R' = 0] = \mathbb{P}_{x,y}[X' = Y'] \geq \mathbb{P}_{x,y}[X' \geq r/2] \\
&= \int_{r/2}^\infty \phi_{0,h}(t) \, dt = \int_{\frac{r}{2\sqrt{h}}}^\infty \phi_{0,1}(t) \, dt \geq \int_1^\infty \phi_{0,1}(t) \, dt \geq 0.15
\end{aligned}
$$

provided $\hat{r} = r \leq 2\sqrt{h}$. Therefore, Assertion (iii) holds as well. □

Proof of Lemma 2.7 (iv), (v), (vi) and (vii). Note that unlike in the proof of assertions (i)-(iii), here it is important to consider $\hat{r} \neq r$. Assertions (iv) and (v) are straightforward consequences of Assertion (i). Indeed, by (2.50) and Lemma 2.7(i),

$$
\beta(x, y) = \hat{\beta}(x, y) + \hat{r} - r = \hat{r} - r.
\tag{6.7}
$$

Assuming (C1) and (C3), this implies (iv), because

$$
\begin{aligned}
|\hat{r} - r| &\leq |(\hat{x} - \hat{y}) - (x - y)| = |h(b(x) - b(y))| \leq hLr, \quad \text{and} \tag{6.8} \\
\hat{r} &= \sqrt{|x - y|^2 + 2h(x - y) \cdot (b(x) - b(y)) + h^2 |b(x) - b(y)|^2} \tag{6.9} \\
&\leq r\sqrt{1 + 2hJ + h^2 L^2} \leq r(1 + hJ + h^2 L^2/2),
\end{aligned}
$$

where we use $\sqrt{1+x} \leq 1 + x/2$ for $x \geq -1$. Similarly, assuming (C2) and (C3), (6.7) implies (v), since $K \leq L$ and thus $-2hK + h^2L^2 \geq -1$ and

$$\hat{r} \;\leq\; r\sqrt{1 - 2hK + h^2L^2} \;\leq\; r\left(1 - hK + h^2L^2/2\right) \quad \text{for } r \geq \mathcal{R}. \tag{6.10}$$

In order to prove (vi) we assume $\sqrt{h} \leq r \leq 1/(4L\sqrt{h})$. Then by (6.8), $|\hat{r} - r| \leq \sqrt{h}/4$. Therefore, by a similar computation as in (6.5),

$$
\begin{aligned}
\alpha(x,y) &= \mathbb{E}_{0,r}[(R'-r)^2; R' \in (r - \sqrt{h}, r)] \;\geq\; \frac{h}{16}\,\mathbb{P}_{0,r}[R' \in (r - \sqrt{h}, r - \frac{\sqrt{h}}{4})] \\
&\geq\; \frac{h}{16}\,\mathbb{P}_{0,r}[R' \in (\hat{r} - \frac{3\sqrt{h}}{4}, \hat{r} - \frac{\sqrt{h}}{2})] \;\geq\; \frac{h}{16}\int_{\sqrt{h}/4}^{3\sqrt{h}/8}(\phi_{0,h}(t) - \phi_{\hat{r},h}(t))\,dt \\
&=\; \frac{h}{16}\int_{1/4}^{3/8}\left(1 - e^{\frac{\hat{r}}{\sqrt{h}}(u - \frac{\hat{r}}{2\sqrt{h}})}\right)\phi_{0,1}(u)\,du \;\geq\; \frac{h}{16}\int_{1/4}^{3/8}\left(1 - e^{u - \frac{1}{2}}\right)\phi_{0,1}(u)\,du.
\end{aligned}
$$

This shows that (vi) holds with $\widetilde{c}_0 := \frac{1}{16}\int_{1/4}^{3/8}(1 - e^{u-1/2})\phi_{0,1}(u)\,du \geq 0.0005$.

Finally, Assertion (vii) is a direct consequence of Assertion (iii), since by (6.8), $|\hat{r} - r| \leq \sqrt{h}$ if $r \leq \sqrt{h}$ and $h \leq 1/L$. $\qquad\square$

The following proof of Theorem 2.8 follows the argumentation in the proofs of Theorems 2.2 and 2.4 in the case $r \leq \mathcal{R}$. For $r > \mathcal{R}$, the contractivity is shown by a direct argument based on Lemma 2.7 (v).

Proof of Theorem 2.8. Let $x, y \in \mathbb{R}^d$, $h \in (0, h_0]$ and $a \in \{0\} \cup [\sqrt{h}, \infty)$.

(i). We first consider the case where $r = |x - y| > \mathcal{R}$. By the choice of h_0 in the statement of the theorem, $h \leq K/L^2$. Therefore, by Lemma 2.7,

$$\mathbb{E}_{x,y}[R' - r] \;=\; \beta(x,y) \;\leq\; -\left(K - L^2h/2\right)hr \;\leq\; -Khr/2.$$

Since f_a is concave with $f_a' \geq 1/2$, we immediately obtain

$$\mathbb{E}_{x,y}[f_a(R') - f_a(r)] \;\leq\; f_a'(r)\mathbb{E}_{x,y}[R' - r] \;\leq\; -Khr/4, \tag{6.11}$$

and hence, as $f_a(r) \leq r$ and $r > \mathcal{R}$,

$$\mathbb{E}_{x,y}[\rho_a(X',Y') - \rho_a(x,y)] \;\leq\; \frac{-Khr/4}{a + f_a(r)}\rho_a(x,y) \;\leq\; \frac{-Kh/4}{1 + a/\mathcal{R}}\rho_a(x,y). \tag{6.12}$$

(ii). Now suppose $r \leq \mathcal{R}$. Since $\hat{r} \leq r$ by (2.52), we have

$$\mathbb{E}_{x,y}[\rho_a(X',Y') - \rho_a(x,y)] \;\leq\; \mathbb{E}_{x,y}[\rho_a(X',Y') - \rho_a(\hat{x},\hat{y})]. \tag{6.13}$$

We can now apply the arguments in the proofs of Theorems 2.2 and 2.4 with α and β replaced by the corresponding quantities $\hat{\alpha}$ and $\hat{\beta}$ for the coupled random walk transition $(\hat{x}, \hat{y}) \mapsto (X', Y')$, with r_2 and r_1 replaced by $\mathcal{R}$. Indeed, note that since the case of $r > \mathcal{R}$ has already been considered above, we only need to use the parts of the proofs of Theorems 2.2 and 2.4 concerned with the case of $r \leq \mathcal{R}$ and thus Assumptions (A3) and (B3) are not required.

We consider first $a = 0$. In this case, we can proceed as in the proof of Theorem 2.4 with $r_0 = \sqrt{h}$. By Lemma 2.7, we can choose $\underline{\alpha}(\hat{r}) = c_0 \min(\hat{r}, \sqrt{h})\sqrt{h}$, $\overline{\beta} \equiv 0$, $\overline{\gamma} \equiv 0$, $\widetilde{\gamma} \equiv 0$, $\varphi \equiv 1$,

$$\Phi(u) = u, \quad g_0(u) = 1 - 2c\int_0^u \sup_{\hat{I}_s}\frac{\Phi}{\underline{\alpha}}\,ds, \quad \text{and} \quad c = \frac{1}{4}\left(\int_0^{\mathcal{R}}\sup_{\hat{I}_s}\frac{\Phi}{\underline{\alpha}}\,ds\right)^{-1} \tag{6.14}$$

in order to satisfy (4.4), (4.5), (4.6), (4.9), (4.11) and (4.12). Here $\hat{I}_s$ is defined by (4.1). With these choices we obtain as in the proof of Theorem 2.4

$$\mathbb{E}_{x,y}[f_0(R') - f_0(\hat{r})] \leq -cf_0(\hat{r}) \qquad \text{for } \hat{r} \leq \mathcal{R}, \tag{6.15}$$

where f_0 is defined by (2.53). Noting that $\hat{r} \leq r$ by (2.52), the bounds in (6.13) and (6.15) now imply that for $r \leq \mathcal{R}$,

$$\mathbb{E}_{x,y}[\rho_0(X', Y')] = \mathbb{E}_{x,y}[f_0(R')] \leq (1 - c)f_0(\hat{r}) \leq (1 - c)f_0(r) = (1 - c)\rho_0(x, y). \tag{6.16}$$

It only remains to show $c \geq c_1(0)h$. Suppose first that $s < 2\sqrt{h} = 2r_0$. Then $\hat{I}_s \subseteq (0, 3\sqrt{h})$. Since $\Phi(u) = u$ and $\underline{\alpha}(u) = c_0 \min(u, \sqrt{h})\sqrt{h} \geq c_0 u\sqrt{h}/3$ for $u < 3\sqrt{h}$, we obtain

$$\sup_{\hat{I}_s} \frac{\Phi}{\underline{\alpha}} \leq \sup_{u < 3\sqrt{h}} \frac{u}{\underline{\alpha}(u)} \leq 3c_0^{-1}h^{-1/2} \quad \text{for any } s < 2\sqrt{h}. \tag{6.17}$$

For $s \geq 2\sqrt{h}$, $\hat{I}_s = (s, s + \sqrt{h})$. Hence $\underline{\alpha} \equiv c_0 h$ on $\hat{I}_s$, and

$$\sup_{\hat{I}_s} \frac{\Phi}{\underline{\alpha}} = c_0^{-1}h^{-1}(s + \sqrt{h}) \quad \text{for any } s \geq 2\sqrt{h}. \tag{6.18}$$

By (6.14), (6.17), (6.18) we see that

$$c^{-1} \leq 24c_0^{-1} + 2c_0^{-1}h^{-1}\mathcal{R}^2 + 4c_0^{-1}h^{-1/2}\mathcal{R} = 2c_0^{-1}h^{-1}(\mathcal{R}^2 + 2h^{1/2}\mathcal{R} + 12h). \tag{6.19}$$

The assertion for $a = 0$ now follows by (6.11), (6.16) and (6.19).

Now consider the case $a \geq \sqrt{h}$. Here we can proceed as in the proof of Theorem 2.2 with $r_0 = \varepsilon = \sqrt{h}$. We now choose the intervals I_r and the dual intervals $\hat{I}_s$ according to (3.1) and (3.2), i.e., $I_r = ((r - \sqrt{h})^+, r)$ and $\hat{I}_s = (\max(s, \sqrt{h}), s + \sqrt{h})$. By Lemma 2.7, we can choose $\underline{\alpha}$, $\overline{\beta}$, $\overline{\gamma}$, $\widetilde{\gamma}$, φ and Φ as above so that conditions (3.3), (3.4), (3.5), (3.7), (3.9), (3.10), (3.11), (3.13) and (3.14) are satisfied. In particular, choosing $a \geq \sqrt{h} = r_0$ guarantees that (3.10) is satisfied since $\beta(x, y) \leq 0$ for all $x, y \in \mathbb{R}^d$. Note that for $u \in \hat{I}_s$ we have $\underline{\alpha}(u) \geq c_0 h$, because $u \geq \sqrt{h}$. Setting

$$g_a(u) = 1 - 2c \int_0^u \sup_{\hat{I}_s} \frac{a + \Phi}{\underline{\alpha}} \, ds, \quad c = \min\left(\frac{p_0}{2}, \frac{1}{4}\left(\int_0^{\mathcal{R}} \sup_{\hat{I}_s} \frac{a + \Phi}{\underline{\alpha}} \, ds\right)^{-1}\right), \tag{6.20}$$

we obtain

$$\mathbb{E}_{x,y}[\rho_a(X', Y')] \leq (1 - c)\rho_a(\hat{x}, \hat{y}) \leq (1 - c)\rho_a(x, y), \tag{6.21}$$

where ρ_a is defined by (2.53). The bound $c \geq c_1(a)$ follows as in (6.18) and (6.19). $\qquad\square$

Proof of Theorem 2.10. In order to apply Theorem 2.2, we set $\varepsilon = r_0 = \sqrt{h}$, and hence $I_r = ((r - \sqrt{h})^+, r)$ and $\hat{I}_s = (s \vee \sqrt{h}, s + \sqrt{h})$ for all r, $s > 0$. By Lemma 2.7, condition (2.14) is satisfied for $h \leq h_0$ with

$$\underline{\alpha}(r) = \widetilde{c}_0 h \, \mathbb{1}_{\sqrt{h} \leq r \leq 1/(4L\sqrt{h})}, \quad \underline{\pi}(r) = p_0 \, 1_{r \leq \sqrt{h}}, \quad \text{and} \tag{6.22}$$

$$\overline{\beta}(r) = \begin{cases} \Lambda h r & \text{for } r < \mathcal{R}, \\ -Khr/2, & \text{for } r \geq \mathcal{R}. \end{cases} \tag{6.23}$$

Here we have used that by the assumptions, $h_0 L \leq 1$ and $h_0 L^2 \leq K$. Moreover, the assumption on h_0 implies $1/(4L\sqrt{h_0}) \geq \mathcal{R}$ since $r_2 > \mathcal{R}$. Hence by (3.3),

$$\overline{\gamma}(r) = 2\overline{\beta}(r)/\underline{\alpha}(r) = 2\widetilde{c}_0^{-1}\Lambda r \qquad \text{for } \sqrt{h} \leq r < \mathcal{R},$$

$\overline{\gamma}(r) \leq 0$ for $r \geq \mathcal{R}$, and thus (3.4), (3.6) and (3.9) are satisfied with

$$\begin{aligned}
\widetilde{\gamma}(r) &= 2\widetilde{c}_0^{-1}\Lambda(r + \sqrt{h})\, \mathbb{1}_{r<\mathcal{R}}, \quad r_1 = \mathcal{R}, \quad \text{and} \\
\varphi(r) &= \exp\left(-\widetilde{c}_0^{-1}\Lambda\left((r \wedge \mathcal{R})^2 + 2\sqrt{h}(r \wedge \mathcal{R})\right)\right).
\end{aligned} \tag{6.24}$$

For $a \geq 2\sqrt{h}$, condition (3.10) is satisfied by (6.22), (6.23), and since by assumption, $\sqrt{h} + 2\Lambda h^{3/2}/p_0 \leq 2\sqrt{h} \leq a$ for $h \leq h_0$. In order to verify (3.12) we need to choose $r_2 \geq r_1 = \mathcal{R}$ such that

$$2\int_{\mathcal{R}}^{r_2} \frac{\Phi(s)}{\underline{\alpha}(s)} ds \geq \frac{a + \Phi(r)}{-\overline{\beta}(r)} \qquad \text{for all } r \geq r_2. \tag{6.25}$$

To this end, note that for $r \geq \mathcal{R}$, we have $\Phi(r) = \Phi(\mathcal{R}) + (r - \mathcal{R})\varphi(\mathcal{R})$. Furthermore, since $1/(4L\sqrt{h}) \leq r_2$ by assumption, on $[\mathcal{R}, r_2]$ we can use the formula for $\underline{\alpha}$ given in (6.22). Hence (6.25) is satisfied if

$$\frac{2\Phi(\mathcal{R})}{\widetilde{c}_0}(r_2 - \mathcal{R}) + \frac{\varphi(\mathcal{R})(r_2 - \mathcal{R})^2}{\widetilde{c}_0} \geq \frac{a + \Phi(\mathcal{R}) + r\varphi(\mathcal{R})}{Kr/2} \qquad \text{for } r \geq r_2. \tag{6.26}$$

Since we assume that $a \leq \Phi(\mathcal{R})$, this condition holds if we choose

$$r_2 = \mathcal{R} + \sqrt{2\widetilde{c}_0/K}. \tag{6.27}$$

Hence from Theorem 2.2 we obtain $\mathbb{E}_{x,y}[\rho_a(X',Y')] \leq (1-c)\rho_a(x,y)$ with c given by (3.13), for $\rho_a(x,y) = \mathbb{1}_{x\neq y} + f_a(|x-y|)$, where

$$f_a(r) = \int_0^r \varphi(s \wedge r_2)g_a(s \wedge r_2)ds \tag{6.28}$$

with φ given by (6.24) and $g_a = g$ given by (3.14). Moreover, we can easily bound the second quantity appearing in the definition (3.13) of c. Indeed, for $s < r_2$ and $u \in \hat{I}_s$ we have $\sqrt{h} < u < r_2 + \sqrt{h} \leq 1/(4L\sqrt{h})$ for $h < h_0$. Therefore, $\underline{\alpha}(u) \geq \widetilde{c}_0 h$ by (6.22). Since $\varphi(s) \geq \varphi(\mathcal{R})$ and $\Phi(u) \leq u$, we obtain

$$\int_0^{r_2} \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{a + \Phi(u)}{\underline{\alpha}(u)} ds \leq \frac{1}{\widetilde{c}_0 h\varphi(\mathcal{R})}\left(ar_2 + \int_0^{r_2}(s + \sqrt{h})ds\right),$$

and hence

$$c \geq \min\left(\frac{1}{2}p_0, \frac{\widetilde{c}_0 h\varphi(\mathcal{R})}{4r_2(a + \sqrt{h}) + 2r_2^2}\right).$$

This implies the assertion, since by (6.27),

$$r_2^2 + 2r_2(a + \sqrt{h}) \leq 2\max(\mathcal{R}^2 + 2(a + \sqrt{h})\mathcal{R}, \, 2\widetilde{c}_0 K^{-1} + 2(a + \sqrt{h})\sqrt{2\widetilde{c}_0/K}).$$

$\square$

In the following variation of Theorem 2.10, Condition (C2) is replaced by a Lyapunov condition:

Theorem 6.1 (Euler scheme, general case with Lyapunov condition)**.**
Suppose that Conditions (C1) and (C3) are satisfied and that the transition kernel p of the Euler scheme satisfies Assumption (A4)a with a Lyapunov function V, i.e., there exist constants $C, \lambda > 0$ such that $pV \leq (1-\lambda)V + C$. Moreover, assume that $\lim_{r\to\infty} \frac{V(x)+V(y)}{r} = \infty$. Let $h_0 = \min\left(\left(\frac{2L}{p_0} + \frac{\widetilde{c}_0\varphi(r_1)}{4(r_1+1)}\right)^{-2}, (16L^2r_2^2)^{-1}\right)$, where $r_1, r_2 > 0$

are constants specified in (6.31) and (6.35). Suppose further that $a \in (2\sqrt{h}, r_2)$ and let $\rho_a(x, y) = (a + \frac{M}{2C}(V(x) + V(y)))\mathbb{1}_{x \neq y} + f_a(|x - y|)$ with M given by (6.32) and f_a defined in (6.36). Let

$$c_2(a) = \frac{1}{4} \min \left(\frac{2p_0}{h}, \frac{\lambda}{h}, 4\varphi(r_1)\Lambda, \frac{\widetilde{c}_0\varphi(r_2)}{2r_2(a + \sqrt{h}) + r_2^2} \right)$$

with φ given by (6.30). Then for all $h \in (0, h_0)$ we have

$$\mathbb{E}_{x,y}[\rho_a(X', Y')] \leq (1 - c_2(a)h)\, \rho_a(x, y) \quad \text{for all } x, y \in \mathbb{R}^d.$$

Example 6.2. It is easy to see that if the drift b satisfies a linear growth condition $|b(x)|^2 \leq L_0(1 + |x|^2)$ for all $x \in \mathbb{R}^d$ with a constant $L_0 > 0$ (which is implied by (C3) with $L_0 = 2 \max(L^2, |b(0)|^2)$) and a dissipativity condition

$$\langle b(x), x \rangle \leq M_1 - M_2|x|^2 \text{ for all } x \in \mathbb{R}^d \tag{6.29}$$

with constants M_1, $M_2 > 0$, then the transition kernel p of the Euler scheme satisfies the Lyapunov condition $pV \leq (1 - \lambda)V + C$ with the Lyapunov function $V(x) = |x|^2$ and constants $\lambda = 2hM_2 - h^2L_0$ and $C = h^2L_0 + 2hM_1 + hd$, whenever $h < 2M_2/L_0$. Since the quadratic function satisfies the growth condition required in Theorem 6.1 and the dissipativity condition (6.29) is significantly weaker than Assumption (C2), we can apply this result to more general cases than the ones covered by Theorems 2.10 and 2.12.

Proof of Theorem 6.1. Here we want to apply Theorem 2.3 and hence we need to verify the conditions listed in Subsection 3.2. Exactly as in the proof of Theorem 2.10, we choose $\varepsilon = r_0 = \sqrt{h}$ and we have the intervals $I_r = ((r - \sqrt{h})^+, r)$ and $\hat{I}_s = (s \vee \sqrt{h}, s + \sqrt{h})$ for all r, $s > 0$. By Lemma 2.15 we get

$$\underline{\alpha}(r) = \widetilde{c}_0 h, \qquad \overline{\beta}(r) = \Lambda h r, \qquad \text{and} \qquad \overline{\gamma}(r) = \frac{2}{\widetilde{c}_0}\Lambda r.$$

Similarly as in the previous proof, the formula for $\underline{\alpha}(r)$ is valid for all $r \in (\sqrt{h}, r_2)$ since $h_0 \leq (16L^2r_2^2)^{-1}$, although here r_2 is given by (6.35). Moreover, we have

$$\widetilde{\gamma}(s) = \frac{2}{\widetilde{c}_0}\Lambda(s + \sqrt{h}), \qquad \text{and} \qquad \varphi(r) = \exp\left(-\frac{\Lambda}{\widetilde{c}_0}\left(r^2 + 2\sqrt{h}r \right) \right). \tag{6.30}$$

Now we choose r_1 as in (3.15), based on Assumption (A4)a. Namely,

$$r_1 := \sup \left\{ |x - y| = r : x, y \in \mathbb{R}^d, V(x) + V(y) < 4C/\lambda \right\}. \tag{6.31}$$

In order for (3.19) to be satisfied, it is sufficient to choose M such that

$$M \leq \frac{1}{4}\left(\int_0^{r_1} \frac{1}{\varphi(s)} \sup_{u \in \hat{I}_s} \frac{1}{\widetilde{c}_0 h} ds \right)^{-1} = \frac{h\widetilde{c}_0}{4}\left(\int_0^{r_1} \frac{1}{\varphi(s)} ds \right)^{-1}.$$

Note, however, that $\varphi(s) \geq \varphi(r_1)$ for all $s > 0$ and hence

$$\frac{h\widetilde{c}_0}{4}\left(\int_0^{r_1} \frac{1}{\varphi(s)} ds \right)^{-1} \geq \frac{h\widetilde{c}_0}{4}\left(\int_0^{r_1} \frac{1}{\varphi(r_1)} ds \right)^{-1} = \frac{h\widetilde{c}_0}{4}\frac{\varphi(r_1)}{r_1} \geq \frac{h\widetilde{c}_0\varphi(r_1)}{4(r_1 + 1)}.$$

Thus if we choose

$$M = \frac{h\widetilde{c}_0\varphi(r_1)}{4(r_1 + 1)}, \tag{6.32}$$

then condition (3.19) is indeed satisfied. Note that $r_1 + 1$ is chosen here instead of r_1 in order to prevent the value of M from being too large when r_1 is very small (or even zero). Now condition (3.20) reads as

$$a \geq \sqrt{h} + \frac{2}{p_0}\left(\Lambda h^{3/2} + \frac{h\widetilde{c}_0\varphi(r_1)}{4(r_1 + 1)}\right).$$ (6.33)

However, since we choose $h \leq h_0 \leq \left(\frac{2L}{p_0} + \frac{\widetilde{c}_0\varphi(r_1)}{4(r_1+1)}\right)^{-2}$, we see that (6.33) holds for all $a \geq 2\sqrt{h}$. It remains to verify condition (3.21), for which we need

$$V(x) + V(y) \geq \frac{16C}{\lambda M}\varphi(r)\Lambda h r$$

to hold for all $r \geq r_2$. Since φ is decreasing, using the choice of M in (6.32), we see that it is sufficient to have

$$V(x) + V(y) \geq \frac{16C}{\lambda}\frac{4(r_1 + 1)}{\widetilde{c}_0}\Lambda r \text{ for all } r \geq r_2.$$ (6.34)

Since we assume that $\limsup_{r\to\infty}\frac{V(x)+V(y)}{r} = \infty$, we can indeed choose r_2 large enough so that (6.34) and hence (3.21) holds. More precisely, we can choose

$$r_2 := \sup\left\{|x - y| = r : x, y \in \mathbb{R}^d, \frac{V(x) + V(y)}{r} < \frac{64C(r_1 + 1)\Lambda}{\lambda\widetilde{c}_0}\right\}.$$ (6.35)

As a consequence, from Theorem 2.3 we get $\mathbb{E}_{x,y}[\rho_a(X', Y')] \leq (1 - c)\rho_a(x, y)$ with $\rho_a(x, y) = (a + \frac{M}{2C}(V(x) + V(y)))\mathbb{1}_{x\neq y} + f_a(|x - y|)$, where c is given by (3.24) and

$$f_a(r) = \int_0^r \varphi(s \wedge r_2)g(s \wedge r_2)ds$$ (6.36)

with φ given by (6.30) and g given by (3.25). Now it only remains to prove the lower bound on the constant c. Similarly as in the proof of Theorem 2.10, we have

$$\frac{1}{8}\left(\int_0^{r_2}\frac{1}{\varphi(s)}\sup_{u\in\hat{I}_s}\frac{a + \Phi(u)}{\underline{\alpha}(u)}ds\right)^{-1} \geq \frac{\widetilde{c}_0 h\varphi(r_2)}{8r_2(a + \sqrt{h}) + 4r_2^2},$$

since $\varphi(s) \geq \varphi(r_2)$ for $s \leq r_2$. Moreover, due to our choice of M in (6.32), using $\Phi(r) \leq r$ and (6.34), we have

$$\frac{\lambda M}{16C}\frac{V(x) + V(y)}{\Phi(r)} \geq \varphi(r_1)h\Lambda \text{ for all } r \geq r_2.$$

This finishes the proof. $\square$

Proof of Theorem 2.12. Let $x, y \in \mathbb{R}^d$ and set $r = |x - y|$ and $\hat{r} = |\hat{x} - \hat{y}|$. We assume $h \in (0, h_0]$ where h_0 is given by (2.70).

We consider at first the case where $r \geq \mathcal{R}$. By the choice of h_0, we have $L^2 h \leq K$ for $h \leq h_0$. Therefore, for $r \geq \mathcal{R}$, the concavity of f and Lemma 2.7 (v) imply

$$\mathbb{E}_{x,y}[f(R') - f(r)] \leq \mathbb{E}_{x,y}[R' - r]f'(r) \leq -\frac{K}{2}hrf'(r) \leq -\frac{K}{2}e^{-qr_1}hf(r).$$ (6.37)

Here, we have used in the last step that $f(r) \leq r$ and $f'(r) \geq \exp(-qr_1)$. The assertion (2.71) now follows by the choice of r_1 and q in (2.65) and (2.67).

From now on, we assume $r < \mathcal{R}$. Recall that

$$
I_{\hat{r}} = \begin{cases} (0, \hat{r} + \sqrt{h}) & \text{if } \hat{r} \leq \sqrt{h}, \\ (\hat{r} - \sqrt{h}, \hat{r}) & \text{if } \hat{r} > \sqrt{h}, \end{cases} \tag{6.38}
$$

cf. (2.44), i.e., $u(\hat{r}) = \sqrt{h}\mathbb{1}_{\hat{r} < \sqrt{h}}$, $l(\hat{r}) = \sqrt{h}\mathbb{1}_{\hat{r} \geq \sqrt{h}}$. Since by Taylor's formula and by concavity of f,

$$
\begin{aligned}
f(R') - f(\hat{r}) &= \int_{\hat{r}}^{R'} f'(s)\, ds = (R' - \hat{r})f'(\hat{r}) + \int_{\hat{r}}^{R'} \int_{\hat{r}}^{s} f''(t)\, dt\, ds \\
&\leq (R' - \hat{r})f'(\hat{r}) + \frac{1}{2}\left[((R' - \hat{r}) \wedge u(\hat{r})) \vee (-l(\hat{r}))\right]^2 \sup_{I_{\hat{r}}} f'',
\end{aligned}
$$

we can conclude by Lemma 2.7 (i) and (ii) that

$$
\begin{aligned}
\mathbb{E}_{x,y}[f(R') - f(r)] &= f(\hat{r}) - f(r) + \mathbb{E}_{x,y}[f(R') - f(\hat{r})] \\
&\leq (\hat{r} - r)f'(r) + \frac{1}{2}c_0 \min(\hat{r}\sqrt{h}, h) \sup_{I_{\hat{r}}} f''. \tag{6.39}
\end{aligned}
$$

We are going to show that the expression on the right hand side of (6.39) is bounded from above by $-c_2 h f(r)$. Note first that by (6.8) and (6.9),

$$
\hat{r} - r \leq \min(L, J + L^2 h/2)\, hr = \Lambda hr. \tag{6.40}
$$

By the choice of q and h_0 in (2.67) and (2.70), $e^{qh_0 L \mathcal{R}} \leq e^{c_0/28} \leq e^{1/28} \leq 3/2$. Therefore,

$$
f'(r) = e^{-qr} = e^{q(\hat{r} - r)}f'(\hat{r}) \leq e^{qLhr}f'(\hat{r}) \leq \frac{3}{2}f'(\hat{r}), \quad \text{and thus}
$$

$$
(\hat{r} - r)f'(r) \leq \frac{3}{2}\Lambda h\hat{r}f'(\hat{r}). \tag{6.41}
$$

Here we have used that $\hat{r} - r \leq \Lambda h\hat{r}$ by (6.40) if $r \leq \hat{r}$, whereas for $r > \hat{r}$ (6.41) is automatically satisfied and hence it holds for all $r < \mathcal{R}$. Furthermore, by (2.70), $hL \leq h_0 L \leq 1/6$. Therefore, $\hat{r} - r \leq Lhr \leq r/6$, and thus

$$
f(\hat{r}) \geq f\left(\frac{5}{6}r\right) \geq \frac{5}{6}f(r), \tag{6.42}
$$

because f is increasing and concave with $f(0) = 0$. Note that our choice of the bound $h_0 L \leq 1/6$ is to some extent arbitrary and a different choice would lead to $5/6$ above being replaced by a different factor. By (6.39), (6.41) and (6.42), we see that the contractivity condition (2.71) holds provided

$$
\frac{3}{2}\Lambda\, h\hat{r}f'(\hat{r}) + \frac{1}{2}c_0 \min(\hat{r}\sqrt{h}, h) \sup_{I_{\hat{r}}} f'' \leq -\frac{6}{5}c_2 h f(\hat{r}). \tag{6.43}
$$

Furthermore, by (6.40), $\hat{r} \leq (1 + Lh)\mathcal{R} = r_1$. Since $f''(r) = -qe^{-qr}\mathbb{1}_{r \leq r_1}$ is increasing, (6.38) implies

$$
\sup_{I_{\hat{r}}} f'' = \begin{cases} -qe^{-q(\hat{r} + \sqrt{h})} & \text{if } \hat{r} \leq \sqrt{h}, \\ -qe^{-q\hat{r}} & \text{if } \hat{r} > \sqrt{h}. \end{cases} \tag{6.44}
$$

We now consider these two cases separately:

(i) $\hat{r} > \sqrt{h}$. Noting that $f'(\hat{r}) = e^{-q\hat{r}}$ and $f(\hat{r}) = (1 - e^{-q\hat{r}})/q \leq 1/q$, we see that (6.43) is satisfied in this case provided

$$
3\Lambda\,\hat{r} - c_0 q \leq -\frac{12c_2}{5q}e^{q\hat{r}} \quad \text{for } \hat{r} < r_1. \tag{6.45}
$$

We have chosen q in (2.67) such that

$$c_0 q \;=\; 7\Lambda\mathcal{R} \;\geq\; 6\Lambda\, r_1.$$

Therefore, the left hand side in (6.45) is bounded from above by $-c_0 q/2$, and thus (6.45) and (6.43) are satisfied if

$$c_2 \;\leq\; \frac{5}{24}c_0 q^2 e^{-q r_1}. \tag{6.46}$$

By (2.65) and (2.67), we see that the constant c_2 has been defined in (2.69) in such a way that (6.46) holds true, and thus the assertion (2.71) is indeed satisfied.

(ii) $\hat{r} \leq \sqrt{h}$. Noting that $f'(\hat{r}) \leq 1$ and $f(\hat{r}) \leq \hat{r}$, we see by (6.44) that (6.43) is satisfied for $\hat{r} \leq \sqrt{h}$ provided

$$3\Lambda h \;+\; \frac{12}{5}c_2 h \;\leq\; c_0 q\sqrt{h}e^{-2q\sqrt{h}}. \tag{6.47}$$

This condition holds if both

$$2q\sqrt{h} \leq 1/2 \quad\text{and}\quad 3e^{1/2}(\Lambda + c_2)\sqrt{h} \leq c_0 q. \tag{6.48}$$

It can now be easily verified that our choice of h_0 in (2.70) ensures that (6.48) holds for $h \leq h_0$. Indeed, since q is given by (2.67), we have $2q\sqrt{h} = 14c_0^{-1}\Lambda\mathcal{R}\sqrt{h} \leq 1/2$. Moreover, since $c_2 \leq 11c_0^{-1}\Lambda^2\mathcal{R}^2$, we obtain

$$\frac{c_0 q}{3e^{1/2}(\Lambda + c_2)} = \frac{7\Lambda\mathcal{R}}{3e^{1/2}(\Lambda + c_2)} \geq \frac{7}{3e^{1/2}}\frac{\mathcal{R}}{1 + 11c_0^{-1}\Lambda\mathcal{R}^2} \geq \frac{7}{6e^{1/2}}\min\left(\mathcal{R}, \frac{c_0}{22\Lambda\mathcal{R}}\right).$$

Hence (6.43) holds true, and thus the assertion (2.71) is satisfied in this case as well. $\square$

7 Proof of results for MALA

Proof of Theorem 2.16. By (2.78)

$$X' \;=\; x + h\,b(x) + \sqrt{h - h^2/4}\,Z \tag{7.1}$$

where $b(x) = -\frac{1}{2}\nabla U(x)$. By (2.75) and (2.74), b is Lipschitz continuous on $B_R^- := \{x \in \mathbb{R}^d : \|x\|_- \leq R\}$. Therefore conditions (C1) and (C3) in Section 2.4 hold when we restrict to B_R^-. Moreover, by (2.76), condition (C2) is satisfied as well for appropriate values of K and R depending on K_c and $\mathcal{R}_c$. Noting that $h < 2$ implies $h - \frac{h^2}{4} > \frac{h}{2}$, it is not difficult to see that the proof of Theorem 2.12 carries over to our slightly modified setup. Therefore, similarly to Theorem 2.12, we can find for any fixed $R \in (0, \infty)$ a concave strictly increasing function f with $f(0) = 0$ and constants $c_2 > 0$, $h_0 > 0$ such that for $h \in (0, h_0)$,

$$\mathbb{E}_{x,y}[f(|X' - Y'|)] \leq (1 - c_2 h)f(|x - y|), \quad\text{for any } x, y \in B_R^-. \tag{7.2}$$

We now want to apply the perturbation result in Theorem 2.5. Setting $d(x, y) = |x - y|$ and $\rho(x, y) = f(|x - y|)$, we see that condition (2.30) holds with $c = c_2 h$. Moreover, by Lemma 7.1 below, there exists a constant $p > 0$ depending only on R, such that for h_0 sufficiently small and $h \in (0, h_0)$, condition (2.32) is satisfied for any $x, y \in B_R^-$ with $r_0 = \sqrt{h}$. Thus, to apply Theorem 2.5, it remains to show that (2.31) holds with a constant $b \geq 0$ satisfying

$$b \leq cf(r_0)/4 = c_2 hf(\sqrt{h})/4. \tag{7.3}$$

To this end notice that for $x, y \in B_R^-$,

$$\mathbb{E}_{x,y}[(|\widetilde{X} - \widetilde{Y}| - |X' - Y'|)^+] \leq \mathbb{E}_{x,y}[|\widetilde{X} - X'|] + \mathbb{E}_{x,y}[|\widetilde{Y} - Y'|]. \tag{7.4}$$

Furthermore, since $\widetilde{X} = X'$ if the proposal is accepted and $\widetilde{X} = x$ otherwise, we obtain by (7.1) and Lemma 2.15, that for any $x \in B_R^-$ and $h \in (0,2)$,

$$
\begin{aligned}
\mathbb{E}_{x,y}[|\widetilde{X} - X'|] &= \mathbb{E}_{x,y}[|X' - x|; \widetilde{U} > \alpha_h(x, X')] \\
&= \mathbb{E}_{x,y}[|X' - x|(1 - \alpha_h(x, X'))] \\
&\leq h|b(x)|\mathbb{E}_{x,y}[1 - \alpha_h(x, X')] + \sqrt{h}\mathbb{E}_{x,y}[|Z|(1 - \alpha_h(x, X'))] \\
&\leq c'h^{\frac{5}{2}}|b(x)| + c''h^2\mathbb{E}_{x,y}[|Z|^2]^{\frac{1}{2}},
\end{aligned}
$$

where c' and c'' are finite constants. Noting that $\mathbb{E}_{x,y}[|Z|^2] = d$ and

$$
|b(x)| \leq d\|b(x)\|_- = d\|\nabla U(x)\|_-/2
$$

we see that there exists a finite constant c''' such that for $x \in B_R^-$ and $h \in (0,2)$

$$
\mathbb{E}_{x,y}[|\widetilde{X} - X'|] \leq c'''h^{\frac{3}{2}}(dh + d^{\frac{1}{2}}h^{\frac{1}{2}}).
$$

A corresponding bound holds for $\mathbb{E}_{x,y}[|\widetilde{Y} - Y'|]$ with $y \in B_R^-$. Hence, by (7.4), condition (2.31) is satisfied with

$$
b = 2c'''h^{\frac{3}{2}}(dh + d^{\frac{1}{2}}h^{\frac{1}{2}}).
$$

Since the right hand side in (7.3) is of order $\Omega(h^{3/2})$, we conclude that (7.3) holds for $dh < h_1$ provided $h_1 \in (0,\infty)$ is chosen sufficiently small. Hence, Theorem 2.5 applies and by (2.34) we obtain

$$
\mathbb{E}_{x,y}[\widetilde{f}(|\widetilde{X} - \widetilde{Y}|)] \leq (1 - c_3 h)\widetilde{f}(|x - y|)
$$

for any $x, y \in B_R^-$ and $h < h_1 d^{-1}$, where $c_3 = \min(c_2/8, p/4h)$ and $\widetilde{f}(r) = f(r) + 2bp^{-1}\mathbb{1}_{r>0}$. $\qquad\square$

Lemma 7.1. For any fixed $R \in (0,\infty)$ there exist constants $p, h_0 \in (0,\infty)$ such that

$$
\mathbb{P}_{x,y}[\widetilde{X} = \widetilde{Y}] \geq p
$$

for any $h \in (0, h_0)$ and $x, y \in B_R^-$ with $|x - y| \leq \sqrt{h}$.

Proof. Let $\hat{x} = x + hb(x)$ where $b(x) = -\frac{1}{2}\nabla U(x)$. By (2.75) and (2.74), b is Lipschitz continuous on B_R^-. For $x, y \in B_R^-$ and $h \in (0, L_R^{-1})$

$$
|\hat{x} - \hat{y}| \leq (1 + hL_R)|x - y| \leq 2|x - y|,
$$

where $L_R \in (0,\infty)$ is the Lipschitz constant of b on B_R^-. Hence, by Lemma 2.7, there exists a constant $p_0 \in (0,\infty)$ such that for any $x, y \in B_R^-$ and $h \in (0, L_R^{-1})$

$$
\mathbb{P}_{x,y}[X' = Y'] \geq p_0 \mathbb{1}_{|\hat{x}-\hat{y}|\leq 2\sqrt{h}} \geq p_0 \mathbb{1}_{|x-y|\leq\sqrt{h}}. \tag{7.5}
$$

Furthermore,

$$
\mathbb{P}_{x,y}[\widetilde{X} \neq \widetilde{Y}] \leq \mathbb{P}_{x,y}[X' \neq Y'] + \mathbb{P}_{x,y}[\widetilde{X} \neq X'] + \mathbb{P}_{x,y}[\widetilde{Y} \neq Y']. \tag{7.6}
$$

By (7.5), the first probability on the right hand side is bounded by $1 - p_0$ for $|x - y| \leq \sqrt{h}$. Moreover, by Lemma 2.15, there exists a finite constant $c' \in (0,\infty)$ such that for any $x, y \in B_R^-$ and $h \in (0,2)$

$$
\mathbb{P}_{x,y}[\widetilde{X} \neq X'] = \mathbb{E}_{x,y}[1 - \alpha_h(x, X')] \leq c'h^{\frac{3}{2}}. \tag{7.7}
$$

A corresponding upper bound holds for $\mathbb{P}_{x,y}[\widetilde{Y} \neq Y']$. Hence, by combining (7.5), (7.6) and (7.7), we conclude that there exist constants $h_0 > 0$ and $p = \frac{p_0}{2} > 0$ such that

$$
\mathbb{P}_{x,y}[\widetilde{X} \neq \widetilde{Y}] \leq 1 - p
$$

for any $h \in (0, h_0)$ and $x, y \in B_R^-$ with $|x - y| \leq \sqrt{h}$. $\qquad\square$

References

[1] Dominique Bakry, Ivan Gentil, and Michel Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014. MR-3155209

[2] Nawaf Bou-Rabee, Andreas Eberle, and Raphael Zimmer, *Coupling and Convergence for Hamiltonian Monte Carlo*, ArXiv e-prints (2018), arXiv:1805.00452.

[3] Oleg Butkovsky, *Subgeometric rates of convergence of Markov processes in the Wasserstein metric*, Ann. Appl. Probab. **24** (2014), no. 2, 526–552. MR-3178490

[4] Mu Fa Chen and Feng Yu Wang, *Estimation of the first eigenvalue of second order elliptic operators*, J. Funct. Anal. **131** (1995), no. 2, 345–363. MR-1345035

[5] Mu-Fa Chen and Feng-Yu Wang, *Estimation of spectral gap for elliptic operators*, Trans. Amer. Math. Soc. **349** (1997), no. 3, 1239–1267. MR-1401516

[6] Bertrand Cloez and Martin Hairer, *Exponential ergodicity for Markov processes with random switching*, Bernoulli **21** (2015), no. 1, 505–536. MR-3322329

[7] Arnak S. Dalalyan, *Theoretical guarantees for approximate sampling from smooth and log-concave densities*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **79** (2017), no. 3, 651–676. MR-3641401

[8] Arnak S. Dalalyan and Avetik G. Karagulyan, *User-friendly guarantees for the Langevin Monte Carlo with inaccurate gradient*, ArXiv e-prints (2017), arXiv:1710.00095.

[9] Persi Diaconis and Laurent Saloff-Coste, *Separation cut-offs for birth and death chains*, Ann. Appl. Probab. **16** (2006), no. 4, 2098–2122. MR-2288715

[10] Jian Ding, Eyal Lubetzky, and Yuval Peres, *Total variation cutoff in birth-and-death chains*, Probab. Theory Related Fields **146** (2010), no. 1-2, 61–85. MR-2550359

[11] Alain Durmus, Gersende Fort, and Éric Moulines, *Subgeometric rates of convergence in Wasserstein distance for Markov chains*, Ann. Inst. Henri Poincaré Probab. Stat. **52** (2016), no. 4, 1799–1822. MR-3573296

[12] Alain Durmus and Éric Moulines, *Quantitative bounds of convergence for geometrically ergodic Markov chain in the Wasserstein distance with application to the Metropolis adjusted Langevin algorithm*, Stat. Comput. **25** (2015), no. 1, 5–19. MR-3304898

[13] Alain Durmus and Eric Moulines, *High-dimensional Bayesian inference via the Unadjusted Langevin Algorithm*, ArXiv e-prints (2016), arXiv:1605.01559.

[14] Alain Durmus and Éric Moulines, *Nonasymptotic convergence analysis for the unadjusted Langevin algorithm*, Ann. Appl. Probab. **27** (2017), no. 3, 1551–1587. MR-3678479

[15] Andreas Eberle, *Reflection coupling and Wasserstein contractivity without convexity*, C. R. Math. Acad. Sci. Paris **349** (2011), no. 19-20, 1101–1104. MR-2843007

[16] Andreas Eberle, *Error bounds for Metropolis-Hastings algorithms applied to perturbations of Gaussian measures in high dimensions*, Ann. Appl. Probab. **24** (2014), no. 1, 337–377. MR-3161650

[17] Andreas Eberle, *Reflection couplings and contraction rates for diffusions*, Probab. Theory Related Fields **166** (2016), no. 3-4, 851–886. MR-3568041

[18] Andreas Eberle, *Markov processes*, Lecture Notes University of Bonn, http://wt.iam.uni-bonn.de/eberle/skripten/, 2017.

[19] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer, *Quantitative Harris type theorems for diffusions and McKean-Vlasov processes*, To appear in Trans. Am. Math. Soc., 2019, arXiv:1606.06012.

[20] Andreas Eberle and Raphael Zimmer, *Sticky couplings of multidimensional diffusions with different drifts*, To appear in Ann. Inst. Henri Poincaré Probab. Stat., 2019. arXiv:1612.06125.

[21] Martin Hairer, *Convergence of Markov processes*, Lecture Notes University of Warwick, http://hairer.org/notes/Convergence.pdf, 2016.

[22] Martin Hairer and Jonathan C. Mattingly, *Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations*, Ann. Probab. **36** (2008), no. 6, 2050–2091. MR-2478676

[23] Martin Hairer and Jonathan C. Mattingly, *Yet another look at Harris' ergodic theorem for Markov chains*, Seminar on Stochastic Analysis, Random Fields and Applications VI, Progr. Probab., vol. 63, Birkhäuser/Springer Basel AG, Basel, 2011, pp. 109–117. MR-2857021

[24] Martin Hairer, Jonathan C. Mattingly, and Michael Scheutzow, *Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations*, Probab. Theory Related Fields **149** (2011), no. 1-2, 223–259. MR-2773030

[25] Martin Hairer, Andrew M. Stuart, and Sebastian J. Vollmer, *Spectral gaps for a Metropolis-Hastings algorithm in infinite dimensions*, Ann. Appl. Probab. **24** (2014), no. 6, 2455–2490. MR-3262508

[26] Jonathan H. Huggins and James Zou, *Quantifying the accuracy of approximate diffusions and Markov chains*, Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (Fort Lauderdale, FL, USA) (Aarti Singh and Jerry Zhu, eds.), Proceedings of Machine Learning Research, vol. 54, PMLR, 20–22 Apr 2017, pp. 382–391.

[27] Mark Jerrum, Jung-Bae Son, Prasad Tetali, and Eric Vigoda, *Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains*, Ann. Appl. Probab. **14** (2004), no. 4, 1741–1765. MR-2099650

[28] James E. Johndrow and Jonathan C. Mattingly, *Error bounds for Approximations of Markov chains used in Bayesian Sampling*, ArXiv e-prints (2017), arXiv:1711.05382.

[29] Aldéric Joulin and Yann Ollivier, *Curvature, concentration and error estimates for Markov chain Monte Carlo*, Ann. Probab. **38** (2010), no. 6, 2418–2442. MR-2683634

[30] Tomasz Komorowski and Anna Walczuk, *Central limit theorem for Markov processes with spectral gap in the Wasserstein metric*, Stochastic Process. Appl. **122** (2012), no. 5, 2155–2184. MR-2921976

[31] Dejun Luo and Jian Wang, *Refined basic couplings and Wasserstein-type distances for SDEs with Lévy noises*, To appear in Stochastic Process. Appl., 2019, arXiv:1604.07206.

[32] Mateusz B. Majka, *Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes*, Stochastic Process. Appl. **127** (2017), no. 12, 4083–4125. MR-3718107

[33] Mateusz B. Majka, *Transportation inequalities for non-globally dissipative SDEs with jumps via Malliavin calculus and coupling*, To appear in Ann. Inst. Henri Poincaré Probab. Stat., 2019, arXiv:1610.06916.

[34] Mateusz B. Majka, Aleksandar Mijatović, and Lukasz Szpruch, *Non-asymptotic bounds for sampling algorithms without log-concavity*, Preprint (2018), arXiv:1808.07105.

[35] Robert J. McCann, *Exact solutions to the transportation problem on the line*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **455** (1999), no. 1984, 1341–1380. MR-1701760

[36] Sean P. Meyn and Richard L. Tweedie, *Markov chains and stochastic stability*, Communications and Control Engineering Series, Springer-Verlag London, Ltd., London, 1993. MR-1287609

[37] Yann Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. **256** (2009), no. 3, 810–864. MR-2484937

[38] Daniel Paulin, *Mixing and concentration by Ricci curvature*, J. Funct. Anal. **270** (2016), no. 5, 1623–1662. MR-3452712

[39] Natesh S. Pillai and Aaron Smith, *Ergodicity of Approximate MCMC Chains with Applications to Large Data Sets*, ArXiv e-prints (2014), arXiv:1405.0182.

[40] Daniel Rudolf and Nikolaus Schweizer, *Perturbation theory for Markov chains via Wasserstein distance*, Bernoulli **24** (2018), no. 4A, 2610–2639. MR-3779696

[41] Max-K. von Renesse and Karl-Theodor Sturm, *Transport inequalities, gradient estimates, entropy, and Ricci curvature*, Comm. Pure Appl. Math. **58** (2005), no. 7, 923–940. MR-2142879

[42] Feng Yu Wang, *Application of coupling methods to the Neumann eigenvalue problem*, Probab. Theory Related Fields **98** (1994), no. 3, 299–306. MR-1262968

[43] Jian Wang, L^P-*Wasserstein distance for stochastic differential equations driven by Lévy processes*, Bernoulli **22** (2016), no. 3, 1598–1616. MR-3474827

[44] Raphael Zimmer, *Explicit contraction rates for a class of degenerate and infinite-dimensional diffusions*, Stoch. Partial Differ. Equ. Anal. Comput. **5** (2017), no. 3, 368–399. MR-3689216