ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus
Mathématique

Djalil Chafaï and Max Fathi

On cutoff via rigidity for high dimensional curved diffusions

Volume 363 (2025), p. 1103-1121

Online since: 13 October 2025

https://doi.org/10.5802/crmath.776

CENTRE
MERSENNE

The Comptes Rendus. Mathématique *are a member of the*
Mersenne Center for open scientific publishing
www.centre-mersenne.org — e-ISSN : 1778-3569

ACADÉMIE DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus. Mathématique
2025, Vol. 363, p. 1103-1121
https://doi.org/10.5802/crmath.776

Research article / *Article de recherche*
Functional analysis, Probability theory / *Analyse fonctionnelle, Probabilités*

On cutoff via rigidity for high dimensional curved diffusions

Sur le phénomène de convergence abrupte via la rigidité pour des diffusions courbées en grande dimension

Djalil Chafaï [a,b] and Max Fathi [a,c]

[a] DMA, École normale supérieure, 45 rue d'Ulm, 75230 Paris, France

[b] CEREMADE, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris, France
URL: https://djalil.chafai.net/

[c] Université Paris Cité and Sorbonne Université, CNRS, Laboratoire Jacques-Louis Lions and Laboratoire de Probabilités, Statistique et Modélisation, 75013 Paris, France

E-mails: djalil@chafai.net, mfathi@lpsm.paris

Abstract. We consider overdamped Langevin diffusions in Euclidean space, with curvature equal to the spectral gap. This includes the Ornstein–Uhlenbeck process as well as non Gaussian and non product extensions with convex interaction, such as the Dyson process from random matrix theory. We show that a cutoff phenomenon or abrupt convergence to equilibrium occurs in high-dimension, at a critical time equal to the logarithm of the dimension divided by twice the spectral gap. This cutoff holds for Wasserstein distance, total variation, relative entropy, and Fisher information. A key observation is a relation to a spectral rigidity, linked to the presence of a Gaussian factor. A novelty is the extensive usage of functional inequalities, even for short-time regularization, and the reduction to Wasserstein. The proofs are short and conceptual. Since the product condition is satisfied, an Lp cutoff holds for all p. We moreover discuss a natural extension to Riemannian manifolds, a link with logarithmic gradient estimates in short-time for the heat kernel, and ask about stability by perturbation. Finally, beyond rigidity but still for diffusions, a discussion around the recent progresses on the product condition for nonnegatively curved diffusions leads us to introduce a new curvature product condition.

Résumé. Nous considérons des diffusions de Langevin sur-amorties dans l'espace euclidien, avec une courbure égale au trou spectral. Ceci inclut le processus d'Ornstein–Uhlenbeck ainsi que des extensions non gaussiennes et non-produit avec interaction convexe, telles que le processus de Dyson issu de la théorie des matrices aléatoires. Nous montrons qu'un phénomène de convergence abrupte vers l'équilibre se produit en grande dimension, à un temps critique égal au logarithme de la dimension divisé par deux fois le trou spectral. Cela a lieu pour la distance de Wasserstein, la variation totale, l'entropie relative et l'information de Fisher. Une observation clé est une relation à une rigidité spectrale, liée à la présence d'un facteur gaussien. Une nouveauté est l'utilisation extensive d'inégalités fonctionnelles, même pour la régularisation en temps court, et la réduction à Wasserstein. Les preuves sont courtes et conceptuelles. Puisque la condition produit est satisfaite, une coupure Lp est valable pour tout p. Nous discutons également d'une extension naturelle aux variétés riemanniennes, d'un lien avec les estimations de gradient logarithmique en temps court pour le noyau de chaleur, et nous nous interrogeons sur la stabilité par perturbation. Enfin, au-delà de la rigidité, mais toujours pour les diffusions, une discussion autour des progrès récents sur la condition de produit pour les diffusions à courbure positive nous conduit à introduire une nouvelle condition de produit de courbure.

Keywords. Markov diffusion process, curvature-dimension inequality, spectral gap, ergodicity and cutoff, Wasserstein distance, relative entropy, total variation, Fisher information.

Mots-clés. Processus de diffusions markovien, inégalité de courbure dimension, trou spectral, ergodicity et convergence abrupte, distance de Wasserstein, entropie relative, information de Fisher.

2020 Mathematics Subject Classification. 60J60, 58J65, 47D07, 53C21, 58J50.

Funding. This work has been (partially) supported by the Project CONVIVIALITY ANR-23-CE40-0003 of the French National Research Agency.

Manuscript received 20 December 2024, revised 4 June 2025, accepted 4 July 2025.

1. Introduction and main results

1.1. *Diffusions on Euclidean spaces, with convex potential*

Let $(X_t)_{t\geq 0}$ be the Markov diffusion process solving the stochastic differential equation (SDE)

$$\mathrm{d}X_t = -\nabla V(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t, \quad X_0 = x_0 \in \mathbb{R}^d, \tag{1}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion in $\mathbb{R}^d$, $V\colon \mathbb{R}^d \to \mathbb{R}$ is strictly convex and $\mathscr{C}^2$ with $\lim_{|x|\to\infty} V(x) = +\infty$, and $|\cdot|$ is the Euclidean norm of $\mathbb{R}^d$. In Statistical Physics, this drift-diffusion is also known as an overdamped Langevin process with potential V. By adding a constant to V, we can assume without loss of generality that $\mu = \mathrm{e}^{-V}$ namely

$$\mathrm{d}\mu(x) = \mathrm{e}^{-V(x)}\,\mathrm{d}x \tag{2}$$

is a probability measure. It is the unique invariant law of the process, and it is moreover reversible. The associated infinitesimal generator is the linear differential operator

$$\mathscr{L} = \Delta - \nabla V \cdot \nabla \tag{3}$$

acting on smooth functions. It is symmetric in $L^2(\mu)$, and its kernel is the set of constant functions. Moreover, its spectrum is included in $(-\infty, -\lambda_1] \cup \{0\}$, for some $\lambda_1 > 0$ called the spectral gap of $\mathscr{L}$. The Ornstein–Uhlenbeck (OU) process is obtained when $V(x) = \frac{\rho}{2}|x|^2$, $\rho > 0$, for which $\lambda_1 = \rho$ while $\mathrm{Hess}(V)(x) = \rho\,\mathrm{I}_d$ for all $x \in \mathbb{R}^d$.

If V is ρ-convex for some $\rho > 0$, namely if $V - \frac{\rho}{2}|\cdot|^2$ is convex, then the spectral gap is an eigenvalue of $-\mathscr{L}$, since the spectrum is discrete, see [20, Proposition 6.7].

1.2. *Cutoff for high dimensional curved diffusions*

Let us denote by W_2 the L^2 Wasserstein (or Monge–Kantorovich) coupling distance between probability measures on the same metric space with finite second moment, namely

$$\mathrm{W}_2(\nu, \mu) = \inf_{(X,Y)} \sqrt{\mathbb{E}\big(|X - Y|^2\big)} \tag{4}$$

where the infimum runs over all couples (X, Y) of random variables with $X \sim \nu$ and $Y \sim \mu$, see [38]. By an abuse of notation, we also write $\mathrm{W}_2(Z, \mu) = \mathrm{W}_2\big(\mathrm{Law}(Z), \mu\big)$.

Theorem 1 (Main Wasserstein estimate). *Let $(X_t)_{t\geq 0}$ be the process* (1)*, with potential V, spectral gap λ_1, and invariant law $\mu = \mathrm{e}^{-V}$. If, for all $x \in \mathbb{R}^d$, and as quadratic forms,*

$$\mathrm{Hess}(V)(x) \geq \lambda_1\,\mathrm{I}_d, \tag{5}$$

then for all non-empty set of initial conditions $S \subset \mathbb{R}^d$,

$$\frac{\mathrm{e}^{-\lambda_1 t}}{\sqrt{\lambda_1}} \Big(\sup_{x_0 \in S} \Lambda(x_0)\Big)^{1/2} \leq \sup_{x_0 \in S} \mathrm{W}_2(X_t, \mu) \leq \mathrm{e}^{-\lambda_1 t} \sup_{x_0 \in S} \Big(\int |x - x_0|^2\,\mathrm{d}\mu(x)\Big)^{1/2} \tag{6}$$

with, denoting E_1 *the eigenspace of* $-\mathscr{L}$ *associated to* λ_1*, and* $k_1 = \dim(E_1)$*,*

$$\Lambda(x_0) = k_1 + \sup_{(f_1,\dots,f_{k_1})} \sum_{i=1}^{k_1} \left|f_i(x_0)\right|^2, \tag{7}$$

where the supremum runs over the set of orthonormal bases of E_1*. Moreover, if* μ *is centered and* $S = \left\{x \in \mathbb{R}^d : |x| \le R\right\}$*, then* $\sup_{x_0 \in S} \Lambda(x)$ *can be replaced by* $k_1 + \lambda_1 R^2$*.*

Remark 2 (Dimensions). We always have $k_1 \le d$, see [6, Lemma 14].

Corollary 3 (Wasserstein cutoff). *Let* $(X_t^{(d)})$*,* $V^{(d)}$*,* $\lambda_1^{(d)}$*,* $\mu^{(d)}$ *be as in Theorem 1, satisfying* (5) *for any dimension* d*. Let* $m^{(d)}$ *be the mean of* $\mu^{(d)}$*, and assume that*

$$\liminf_{d\to\infty} \lambda_1^{(d)} > 0. \tag{8}$$

Then a cutoff phenomenon occurs at critical time

$$t_* = \frac{\log(d)}{2\lambda_1^{(d)}}, \tag{9}$$

namely for all $\varepsilon > 0$*,*

$$\lim_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{W}_2\big(X_{t_d}^{(d)}, \mu^{(d)}\big) = \begin{cases} +\infty & \text{if } t_d = (1-\varepsilon)t_*, \\ 0 & \text{if } t_d = (1+\varepsilon)t_*, \end{cases} \tag{10}$$

where the set of initial conditions is a ball of the following form

$$S^{(d)} = \left\{x \in \mathbb{R}^d : |x - m^{(d)}| \le c\sqrt{d}\right\}, \quad \text{for an arbitrary constant } c. \tag{11}$$

For probability measures μ and ν on the same space, we denote by

$$\mathrm{d_{TV}}(\mu,\nu) = \sup_A \left|\mu(A) - \nu(A)\right| \in [0,1] \tag{12}$$

their total variation distance, and by

$$\mathrm{H}(\nu \mid \mu) = \int \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \log\Big(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\Big)\,\mathrm{d}\mu = \int \log\Big(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\Big)\,\mathrm{d}\nu \in [0,+\infty] \tag{13}$$

the relative entropy or Kullback–Leibler divergence of ν with respect to μ, with convention $\mathrm{H}(\nu \mid \mu) = +\infty$ if ν is not absolutely continuous with respect to μ. The Fisher information of ν with respect to μ is

$$\mathrm{I}(\nu \mid \mu) = \int \frac{\left|\nabla \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right|^2}{\frac{\mathrm{d}\nu}{\mathrm{d}\mu}}\,\mathrm{d}\mu = \int \left|\nabla \log\Big(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\Big)\right|^2 \mathrm{d}\nu \in [0,+\infty], \tag{14}$$

with the convention $\mathrm{I}(\nu \mid \mu) = +\infty$ if ν is not absolutely continuous with respect to μ. By an abuse of notation, we take the freedom of replacing ν by $X \sim \nu$ in these expressions.

Corollary 4 (I, H, **and** TV **cutoffs).** *Under the setting of Corollary 3, for all* $\varepsilon > 0$*,*

$$\lim_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{I}\big(X_{t_d}^{(d)} \mid \mu^{(d)}\big) = \begin{cases} +\infty & \text{if } t_d = (1-\varepsilon)t_*, \\ 0 & \text{if } t_d = (1+\varepsilon)t_*, \end{cases} \tag{15}$$

$$\lim_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{H}\big(X_{t_d}^{(d)} \mid \mu^{(d)}\big) = \begin{cases} +\infty & \text{if } t_d = (1-\varepsilon)t_*, \\ 0 & \text{if } t_d = (1+\varepsilon)t_*, \end{cases} \tag{16}$$

$$\lim_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{d_{TV}}\big(X_{t_d}^{(d)}, \mu^{(d)}\big) = \begin{cases} 1 & \text{if } t_d = (1-\varepsilon)t_*, \\ 0 & \text{if } t_d = (1+\varepsilon)t_*. \end{cases} \tag{17}$$

We emphasize that in contrast with what is done for instance by [9,35], we obtain the I, H, and TV cutoffs from the Wasserstein cutoff. Moreover, they occur at exactly the same critical time, due to the choice of initial condition we make. In comparison, [35] works under more general curvature assumptions, and uses Fisher information and entropy to derive a differential inequality controlling entropy beyond the mixing time, and then proves cutoff in total variation distance using a reverse Pinsker inequality. All these approaches are crucially based on functional inequalities.

Condition (5) states that the process has curvature at least equal to the spectral gap. It turns out that it is the best possible lower bound on the curvature, as we explain later on in relation with a notion of spectral rigidity. Condition (5) is satisfied by the OU process with $V=\frac{\rho}{2}|\cdot|^2$, and in this case, we have $\lambda_1=\rho$, $k_1=d$, and $f_i(x)=\sqrt{\rho}x_i$. An important class of non-Gaussian and non-product examples beyond pure OU is

$$V(x)=\frac{\rho}{2}|x|^2+W(x),\quad x\in\mathbb{R}^d, \tag{18}$$

where $\rho>0$ and $W\colon\mathbb{R}^d\to\mathbb{R}$ is convex and translation invariant in the direction $(1,\dots,1)\in\mathbb{R}^d$, namely for all $u\in\mathbb{R}$ and all $x\in\mathbb{R}^d$,

$$W\big(x+u(1,\dots,1)\big)=W(x). \tag{19}$$

This is the case for example when for some convex even function $h\colon\mathbb{R}\to\mathbb{R}$,

$$W(x)=\sum_{i<j}h(x_i-x_j),\quad x\in\mathbb{R}^d. \tag{20}$$

If π and $\pi^\perp$ are the orthogonal projections on $\mathbb{R}(1,\dots,1)$ and its orthogonal, respectively, then $|x|^2=\big|\pi(x)\big|^2+\big|\pi^\perp(x)\big|^2$, while the translation invariance of W in the direction $\big(1/\sqrt{n},\dots,1/\sqrt{n}\big)$ gives $W(x)=W\big(\pi(x)+\pi^\perp(x)\big)=W\big(\pi^\perp(x)\big)$, therefore

$$\mathrm{e}^{-V(x)}=\mathrm{e}^{-\frac{\rho}{2}|\pi(x)|^2}\,\mathrm{e}^{-\left(W(\pi^\perp(x))+\frac{\rho}{2}|\pi^\perp(x)|^2\right)} \tag{21}$$

which means that μ is, up to a rotation, a product measure, and splits into a one-dimensional Gaussian factor $\mathcal{N}\big(0,\frac{1}{\rho}\big)$ and a log-concave factor with a ρ-convex potential.

Theorem 5 (Boltzmann–Gibbs probability measure with convex interaction). *Let us consider the Langevin process* (1) *with V as in* (18) *with $\rho>0$ and with W convex and translation invariant in the direction* $(1,\dots,1)$. *Then the following properties hold true:*

(i) *$\lambda_1=\rho$ and the symmetric Hermite polynomial $x_1+\cdots+x_d$ belongs to E_1;*
(ii) *V is ρ-convex and μ has a Gaussian factor $\mathcal{N}\big(0,\frac{1}{\rho}\big)$ in the direction* $(1,\dots,1)$;
(iii) *The curvature condition* (5) *is satisfied.*

This covers as a special degenerate case the Dyson–OU or DOU process studied in [9,11], when

$$h(x)=\begin{cases}-\beta\log(x) & \text{if } x>0,\\ +\infty & \text{if } x\leq 0,\end{cases}\qquad\text{for an arbitrary constant } \beta\geq 0, \tag{22}$$

the degeneracy being equivalent to define the Dyson–OU process on the convex domain $\{x\in\mathbb{R}^d: x_1>\cdots>x_n\}$ instead of on the whole space $\mathbb{R}^d$, to exploit convexity. In this case, the symmetric Hermite polynomial $x_1+\cdots+x_d$ is an eigenfunction associated to the spectral gap, the tip of an iceberg of integrability, as observed in [25].

Corollary 6 (Cutoff for Langevin with convex interactions). *Let $(X_t^{(d)})$, $V^{(d)}$, $\mu^{(d)}$ be as in Theorem 5, for any dimension d, and for a fixed $\rho>0$. Then there is cutoff at critical time*

$$t_*=\frac{\log(d)}{2\rho} \tag{23}$$

in the sense that (10), (15), (16), *and* (17) *occur when* $S^{(d)}$ *can be either* $B\big(m^{(d)}, c\sqrt{d}\big)$ *or* $m^{(d)} + [-c, c]^d$ *for an arbitrary fixed constant* $c > 0$ *with* $m^{(d)}$ *being the mean of* $\mu^{(d)}$.

To summarize, for rigid curved diffusions with bounded below spectral gap, the mixing time in Wasserstein distance, relative entropy, and total variation distance is always $\frac{\log(d)}{2\lambda_1}$, and a cutoff occurs at the mixing time in high dimension, for an explicit class of initial conditions. This covers non-Gaussian non-product examples, beyond the (D)OU case.

1.3. *Product condition and* L^p *cutoff at the mixing time*

It seems that the cutoff phenomenon for diffusions was first explored by [36], notably for Brownian motion on compact Lie groups, using the L^2 decomposition and functional inequalities. This was further explored for manifolds with special symmetries in relation with representation theory by [28]. The cutoff for diffusions on non-compact spaces such as the Dyson–OU process, is considered by [9], for various distances and divergences, in relation with integrability. The OU process is a special Dyson–OU process, and is also a special Gaussian ergodic Markov process with independent components, a tensorized or product situation for which the cutoff was studied earlier notably by [4,5,24]. The role of the eigenfunctions associated to the spectral gap for lower bounds is an old observation that dates back to Persi Diaconis and David Wilson, see for instance [9,36] and references therein for the case of diffusions.

It was shown by [12] that for ergodic Markov processes on arbitrary state spaces, cutoff occurs at the mixing time, in L^p distance, $p > 1$, provided that the product of the spectral gap and the mixing time tends to infinity. This *product condition* was proposed by Yuval Peres. The method relies on a reduction to the Euclidean case $p = 2$ by interpolation. It does not provide the mixing time. Very recently, [35] extended this theorem to $p = 1$ (total variation distance), in the case of nonnegatively curved diffusions. The method relies crucially on a two-sided comparison between total variation and relative entropy, namely Pinsker and reverse Pinsker inequalities, in other words functional inequalities. It is a variant of the method developed by [34] for finite state spaces, with a crucial specificity. Namely, it is shown that in the continuous setting a varentropy functional is controlled by the derivative of the entropy via a local Poincaré inequality, which arises as a consequence of nonnegative curvature. In the discrete setting, chain rule issues are an obstruction, and current known statements have some extra restrictions, see [31].

For curved diffusions, rigidity implies the product condition: if the spectral gap is bounded from below, then the mixing time in total variation goes to infinity by comparing with the projection of the process on an (affine) eigenfunction associated to the spectral gap. This allows to use [12, Theorem 3.3] and [35, Corollary 1] to get cutoff for L^p distance, $p \geq 1$, namely, denoting $\mu_t^{(d)}$ the law of $X_t^{(d)}$, this writes

$$\lim_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \left\| \frac{\mathrm{d}\mu_t^{(d)}}{\mathrm{d}\mu^{(d)}} - 1 \right\|_{\mathrm{L}^p(\mu^{(d)})} = \begin{cases} \max & \text{if } t_d = (1-\varepsilon)t_*, \\ 0 & \text{if } t_d = (1+\varepsilon)t_*, \end{cases} \tag{24}$$

where $\max = 1$ if $p = 1$ and $\max = +\infty$ if $p > 1$, see [10] for a pedagogical presentation. We emphasize that Corollaries 3, 4, and 6 give the mixing time as well as the cutoff.

The product condition can be seen as weak rigidity, connecting trend to equilibrium and spectral gap. It is well-known that in the discrete setting, the product condition does not always imply cutoff, see for instance [12] for detailed counter examples. Discrete state space processes can be more subtle than random walks on locally Euclidean graphs, which are the discrete analogue of diffusions on manifolds. In particular they can allow complete graphs moves producing quick convergence to equilibrium. For diffusions, the analogue of this type of degeneracy would be a non-local term.

In another direction, for e^{-V} log-concave on $\mathbb{R}^d$, the Kannan, Lovász and Simonovitz (KLS) conjecture states that the inverse spectral gap λ_1^{-1} of $\Delta - \nabla V \cdot \nabla$ is up to universal multiplicative constants equal to the spectral radius $r(K)$ of the covariance matrix

$$K_{ij} = \int x_i x_j \,\mathrm{e}^{-V(x)}\,\mathrm{d}\mu(x) - |m|^2 \quad \text{where} \quad m_i = \int x_i \,\mathrm{e}^{-V(x)}\,\mathrm{d}x.$$

The best available bound is logarithmic: $\lambda_1^{-1} \asymp \sqrt{\log(d)}\, r(K)$, see [23].

1.4. *Positively curved diffusions, curvature product condition, and cutoff*

Recall that the recent result [35, Corollary 1] states that for nonnegatively curved diffusions, and for total variation distance, the product condition implies cutoff at the mixing time (a direct corollary of a theorem bounding the mixing time window). By combining the approaches used in the proof of [35, Corollary 1], the proof of our Corollary 4, and the proof of [12, Theorem 3.3], we arrive at the following theorem. We state it in the restricted Euclidean setting for simplicity, but it holds in the Riemannian setting, since it is the curvature bound and the diffusion property that matter.

Theorem 7 (Cutoff for positively curved diffusions with curvature product condition). *Let $(X_t^{(d)})$ be a sequence of diffusion processes on $\mathbb{R}^d$ of the form* (1)*, with convex potential $V^{(d)}$, initial condition $x_0^{(d)}$, and invariant measure $\mu^{(d)}$. Let $S^{(d)} \subset \mathbb{R}^d$ be an arbitrary non-empty set of initial conditions. Let $\eta \in (0,1)$ be an arbitrary fixed threshold which does not dependent on d. Suppose that there exists positive constants $\kappa^{(d)}$ such that* $\mathrm{Hess}(V^{(d)}) \geq \kappa^{(d)} I_d$ *and that the following curvature product condition holds:*

$$\lim_{d\to\infty} \kappa^{(d)} t_0^{(d)} = +\infty \quad \textit{where} \quad t_0^{(d)} := \inf\Big\{t \geq 0 : \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{d_{TV}}(X_t^{(d)}, \mu^{(d)}) \leq \eta\Big\}.$$

Then there is cutoff at the critical time $t_0^{(d)}$ in the sense that for all $\varepsilon \in (0,1)$,

$$\lim_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{d_{TV}}\big(X_{t_d}^{(d)}, \mu^{(d)}\big) = \begin{cases} 1 & \textit{if } t_d = (1-\varepsilon)\, t_0^{(d)}, \\ 0 & \textit{if } t_d = (1+\varepsilon)\, t_0^{(d)}, \end{cases} \tag{25}$$

$$\lim_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{I}\big(X_{t_d}^{(d)} \mid \mu^{(d)}\big) = \begin{cases} +\infty & \textit{if } t_d = (1-\varepsilon)\, t_0^{(d)}, \\ 0 & \textit{if } t_d = (1+\varepsilon)\, t_0^{(d)}, \end{cases} \tag{26}$$

$$\lim_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{H}\big(X_{t_d}^{(d)} \mid \mu^{(d)}\big) = \begin{cases} +\infty & \textit{if } t_d = (1-\varepsilon)\, t_0^{(d)}, \\ 0 & \textit{if } t_d = (1+\varepsilon)\, t_0^{(d)}, \end{cases} \tag{27}$$

$$\lim_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{W}_2\big(X_{t_d}^{(d)}, \mu^{(d)}\big) = \begin{cases} +\infty & \textit{if } t_d = (1-\varepsilon)\, t_0^{(d)}, \\ 0 & \textit{if } t_d = (1+\varepsilon)\, t_0^{(d)}. \end{cases} \tag{28}$$

1.5. *Rigidity and Gaussian factorization for curved diffusions*

It turns out that the presence of a Gaussian factor, as well as the λ_1 uniform lower bound on the Hessian of the potential (5), are both equivalent to a notion of rigidity.

Let us start with the notion of Bakry–Émery curvature. For a Langevin process in the Euclidean space as in (1), for all $\rho > 0$, the following properties are equivalent:

- (C1) $V - \frac{\rho}{2}|\cdot|^2$ is convex, in other words V is ρ-convex;
- (C2) $\mathrm{Hess}(V)(x) \geq \rho\,\mathrm{I}_d$ as quadratic forms, for all $x \in \mathbb{R}^d$;
- (C3) $\langle \mathrm{Hess}(V)\nabla f, \nabla f\rangle \geq \rho|\nabla f|^2$ for all smooth $f : \mathbb{R}^d \to \mathbb{R}$;
- (C4) the curvature-dimension inequality $\mathrm{CD}(\rho,\infty)$ is satisfied, see Section 1.6.

Following [2], when these properties hold true, we say that the process has (Bakry–Émery) curvature bounded below by ρ. It follows then that

$$\lambda_1 \geq \rho. \tag{29}$$

To see it, we observe that from the Bochner formula $\nabla\mathscr{L} - \mathscr{L}\nabla = -\operatorname{Hess}(V)\nabla$, we get

$$\int\Big(\big\langle \operatorname{Hess}(V)\nabla, \nabla f\big\rangle - \rho|\nabla f|^2\Big)\,\mathrm{d}\mu = \int\Big(\big\langle \mathscr{L}\nabla f - \nabla\mathscr{L}f, \nabla f\big\rangle - \rho|\nabla f|^2\Big)\,\mathrm{d}\mu, \tag{30}$$

while on the other hand, by integration by parts, we get[1]

$$-\int \langle \mathscr{L}\nabla f, \nabla f\rangle\,\mathrm{d}\mu = \sum_{i=1}^{d}\int \big|\nabla\partial_i f\big|^2\,\mathrm{d}\mu = \sum_{i,j=1}^{d}\int (\partial^2_{ij}f)^2\,\mathrm{d}\mu = \int \|\operatorname{Hess}(f)\|^2_{\mathrm{HS}}\,\mathrm{d}\mu, \tag{31}$$

therefore, by specializing to f such that $-\mathscr{L}f = \lambda_1 f$, we obtain

$$\int\Big(\underbrace{\big\langle \operatorname{Hess}(V)\nabla, \nabla f\big\rangle - \rho|\nabla f|^2}_{\geq 0 \text{ by (C3)}}\Big)\,\mathrm{d}\mu = (\lambda_1 - \rho)\int |\nabla f|^2\,\mathrm{d}\mu - \int \|\operatorname{Hess}(f)\|^2_{\mathrm{HS}}\,\mathrm{d}\mu, \tag{32}$$

which gives (29).[2] Equality in (29) is obviously achieved for the OU case $V = \frac{\rho}{2}|\cdot|^2$. More generally, beyond the OU case, we say that the process is *rigid* when the spectral gap matches the curvature lower bound: $\lambda_1 = \rho$. Rigidity can be reformulated, and it turns out that for all $\rho > 0$, the following items are equivalent:

- (R1) the process is rigid: it has curvature bounded below by ρ and $\lambda_1 = \rho$;
- (R2) up to a rotation and translation (change of coordinates), the probability measure $\mu = \mathrm{e}^{-V}$ is product, and splits into a one-dimensional Gaussian factor $\mathscr{N}\big(0, \frac{1}{\rho}\big)$ and a second factor which is log-concave with a ρ-convex potential; in which case, we get $\lambda_1 = \rho$ and all the eigenfunctions associated to λ_1 are affine.

It is immediate to check that (R2) gives (R1). Conversely, if (R1) is satisfied, then for f such that $-\mathscr{L}f = \lambda_1 f$, we obtain that both sides of (32) are identically zero. This gives $\operatorname{Hess}(f) \equiv 0$, hence f is affine, and this also gives that the constant vector ∇f is in the kernel of the symmetric positive semidefinite matrix $\operatorname{Hess}(V) - \rho\,\mathrm{I}_d$. Moreover $\nabla f \neq 0$ since f is not constant, leading to (R2). We observe that by iterating the procedure, we get a Gaussian factor with same dimension as the eigenspace of λ_1.

The equivalence between the rigidity properties (R1) and (R2) appears in [14] in the traditional broader Riemannian setting of Bakry–Émery, discussed in Section 3.8. An alternative proof in the Euclidean case, based on optimal transport, was found by [16]. These works were motivated by the more general problem of understanding manifolds optimizing a certain geometric quantity (namely, the spectral gap) under a curvature constraint. This type of question goes back to a theorem of [29], which states that positively curved smooth Riemannian manifolds with minimal spectral gap are isometric to spheres. Rigidity is typically used to reinforce statements in comparison geometry. See for example [3] for applications of Gamma calculus to comparison geometry under curvature assumptions.

1.6. *Riemannian manifolds*

We now discuss Theorem 1 in a broader geometric context. A weighted Riemannian manifold is a triplet (M, g, μ), where (M, g) is a Riemannian manifold with metric tensor g, and μ is a measure on M. Here we shall assume that μ is a probability measure, absolutely continuous with respect to

[1] We denote by $\|\cdot\|_{\mathrm{HS}}$ the Hilbert–Schmidt or trace or Frobenius matrix norm.

[2] We could also invoque the Brascamp–Lieb inequality $\int f^2\,\mathrm{d}\mu - \big(\int f\,\mathrm{d}\mu\big)^2 \leq \int\big\langle \operatorname{Hess}(V)^{-1}\nabla f, \nabla f\big\rangle\,\mathrm{d}\mu$, or equivalently the covariance representation of Hörmander–Helffer–Sjöstrand.

the volume measure. We shall write $\mu = \mathrm{e}^{-V}$, so that V plays the same role of a potential as in the Euclidean setting. The Markov process $(X_t)_{t\geq 0}$ we consider is the drift diffusion that combines a Brownian motion $(B_t)_{t\geq 0}$ on (M,g) and drift $-\nabla V$, given by the SDE

$$\mathrm{d}X_t = -\nabla V(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t, \quad X_0 = x_0 \in M, \tag{33}$$

where ∇ is the gradient on (M,g). We refer to [2,22] for more background. A key role is played by the Ricci curvature tensor of the manifold, which we shall denote by Ric. We refer to [30] for an introduction to curvature on Riemannian manifolds. Under the assumption that the Ricci curvature is bounded from below, and that V is geodesically semi-convex, then solutions to (33) exist for all times [21, Theorem 11.8]. We shall not make use of (33), and only rely on the generator of the process

$$\mathscr{L} = \Delta - \nabla V \cdot \nabla \tag{34}$$

where Δ is the Laplace–Beltrami operator on (M,g). The weighted manifold (M,g,μ) has Bakry–Émery curvature bounded below by ρ when

$$\mathrm{Ric} + \mathrm{Hess}(V) \geq \rho g. \tag{35}$$

A celebrated result due to [1], see also [2], states that this implies a lower bound on the spectral gap of $\mathscr{L}$ in the sense that

$$\lambda_1 \geq \rho. \tag{36}$$

When the manifold is unweighted, this is the dimension-free version of the famous spectral gap bound due to [26]. Actually the Bakry–Émery approach implies stronger functional inequalities, such as a logarithmic Sobolev inequality, and a Gaussian isoperimetric inequality. Note that following for instance [20, Proposition 6.7], when $\rho > 0$ the spectrum of $\mathscr{L}$ is discrete and the spectral gap λ_1 is always an eigenvalue of $-\mathscr{L}$.

In the Bakry–Émery approach, the bound (35) is formulated in a more abstract way involving the Γ and Γ_2 functional quadratic forms, defined by

$$\Gamma(f,g) = \tfrac{1}{2}\big(\mathscr{L}(fg) - f\mathscr{L}g - g\mathscr{L}f\big), \tag{37}$$

$$\Gamma_2(f,g) = \tfrac{1}{2}\big(\mathscr{L}\Gamma(f,g) - \Gamma(f,\mathscr{L}g) - \Gamma(g,\mathscr{L}f)\big). \tag{38}$$

They are fully characterized by their diagonal $\Gamma(f) = \Gamma(f,f)$ and $\Gamma_2(f) = \Gamma_2(f,f)$, by polarization. In the case of the Langevin operator (34), they simply boil down to

$$\Gamma(f) = |\nabla f|^2 \qquad \text{and} \qquad \Gamma_2(f) = \mathrm{Ric}(\nabla f,\nabla f) + \big\|\mathrm{Hess}(f)\big\|_{\mathrm{HS}}^2 + \big\langle \mathrm{Hess}(V)\nabla f,\nabla f\big\rangle. \tag{39}$$

The formula for Γ_2 follows from the Bochner formula $\nabla\mathscr{L} = \mathscr{L}\nabla - \mathrm{Hess}(V)\nabla - \mathrm{Ric}(\nabla,\nabla)$. The Bakry–Émery curvature-dimension inequality $\mathrm{CD}(\rho,n)$, where $\rho \in \mathbb{R}$ is the curvature and $n \in \mathbb{R}$ is the dimension, writes, for all f in a rich enough class of functions,

$$\Gamma_2(f) \geq \rho\Gamma(f) + \frac{1}{n}(\mathscr{L}f)^2. \tag{40}$$

Now $\mathrm{CD}(\rho,\infty)$ is equivalent to the curvature lower bound (35). Moreover, in the Euclidean case, the Ricci tensor is zero and $\mathrm{CD}(\rho,\infty)$ is equivalent to say that V is ρ-convex.

In order to formulate a neat abstract analogue of Theorem 1, we shall use vector-valued functions whose coordinates are eigenfunctions, as defined in the following.

Definition 8 (Multi-eigenfunction). *Let E_1 be the eigenspace of $\mathscr{L}$ associated with the eigenvalue λ_1, in $L^2(\mu)$. Let $k_1 = \dim E_1$. We define a multi-eigenfunction as being a map $M \to \mathbb{R}^{k_1}$ whose coordinates are orthogonal elements of E_1 in $L^2(\mu)$. We denote by F_1 the set of all multi-eigenfunctions, and by $\mathbb{S}_{F_1}$ the set of multi-eigenfunctions whose coordinates are elements of the unit sphere of E_1 in $L^2(\mu)$, thus orthonormal.*

The link between this definition and Theorem 1 is that the component $\sum_{i=1}^{d}\left|f_i(x_0)\right|^2$ in the definition of Λ in Theorem 1 is exactly the ℓ^2 norm of $T(x_0)$ where T is some element of $\mathbb{S}_{F_1}$. The Riemannian analog of Theorem 1 is now the following.

Theorem 9 (Wasserstein estimate on weighted Riemannian manifold). *Let $(X_t)_{t\geq 0}$ be the diffusion* (33)*, and let $\mathscr{L}$ be its generator* (34)*. Assume that for some $\rho>0$,*

$$\mathrm{CD}(\rho,\infty) \text{ is satisfied and } \lambda_1=\rho. \tag{41}$$

Then for all non-empty set of possible initial conditions $S\subset M$, and all $t\geq 0$,

$$\frac{\mathrm{e}^{-\rho t}}{\sqrt{\rho}}\Big(\sup_{x_0\in S}\Lambda(x_0)\Big)^{1/2}\leq \mathrm{W}_2(X_t,\mu)\leq \mathrm{e}^{-\rho t}\sup_{x_0\in S}\Big(\int d_M(x_0,x)^2\,\mathrm{d}\mu(x)\Big)^{1/2},$$

where d_M is the Riemannian distance on M, and where

$$\Lambda(x)=k_1+\sup_{T\in\mathbb{S}_{F_1}}\left|T(x)\right|_M^2. \tag{42}$$

Moreover, if $S=\{x\in M: d_M(x,m)\leq R\}$, then we can replace $\sup_{x_0\in S}\Lambda(x_0)$ by

$$\sup_{T\in\mathbb{S}_{F_1}}\Big(\left|T(m)\right|_M^2+\rho R^2\Big). \tag{43}$$

1.7. *About stability*

Our study is under the strong assumption $\lambda_1=\rho$. For the d-dimensional unit sphere, we have $\rho=d-1$ and $\lambda_1=d$, so the abstract theorem does not apply. Yet cutoff does occur, and moreover $\lambda_1/\rho\to 1$. This naturally leads to asking the following.

Question 10. *Is there a (useful) analog of Corollary 3 under the weaker assumption*

$$\frac{\lambda_1^{(d)}}{\rho^{(d)}}\xrightarrow[d\to\infty]{}1? \tag{44}$$

Some results on positively curved manifolds when λ_1 is close to ρ were obtained by [7,15,18,27]. However, the quantitative estimates on eigenfunctions seem to be too weak to easily generalize Theorem 9. The estimate on the Wasserstein distance strongly relies on the fact that the eigenfunction is Lipschitz, and the estimate is sharp because it is actually affine, namely that $\mathrm{Hess}(f)=0$. When λ_1 is close to ρ, then $\mathrm{Hess}(f)$ is small, but for example the smallness estimates in [7] are only in L^2 norm, so we do not actually control the Lipschitz norm of f. It is unclear if a stronger L^∞ estimate can be expected in general, so maybe some extra assumptions are needed.

2. Further comments

This section gathers several remarks, mostly about the Euclidean case for simplicity.

2.1. *Fisher information and the heat kernel*

It is natural to try to directly upper bound the Fisher information by using the heat kernel, instead of using a comparison with relative entropy. Let us examine first the exactly solvable case of the OU process, for which the Mehler formula $\mu_t=\mathcal{N}\big(\mathrm{e}^{-\rho t}x_0,(1-\mathrm{e}^{-2\rho t})\mathrm{I}_d\big)$ gives, for $t>0$,

$$\nabla\log\frac{\mathrm{d}\mu_t}{\mathrm{d}\mu}(x)=-\frac{x-\mathrm{e}^{-\rho t}x_0}{1-\mathrm{e}^{-2\rho t}}+x=a_tx+b_tx_0 \tag{45}$$

where a_t and b_t do not depend on the dimension d, leading to

$$\mathrm{I}(\mu_{t_0} \mid \mu) \leq 2a_{t_0}^2 \int |x|^2 \,\mathrm{d}\mu(x) + 2b_{t_0}^2 |x_0|^2 = 2a_{t_0}^2 d + 2b_{t_0}^2 |x_0|^2. \tag{46}$$

This gives the upper bound (82). Beyond the OU process, denoting, for all $t > 0$, by p_{t,x_0} the density of μ_t with respect to the Lebesgue measure, which is known as the heat kernel of $\mathscr{L}$, we have, for all x and all $t > 0$,

$$\nabla \log \frac{\mathrm{d}\mu_t}{\mathrm{d}\mu}(x) = \nabla \log p_{t,x_0}(x) + \nabla V(x). \tag{47}$$

Inspired by the OU example, we could expect that under $\mathrm{CD}(\rho,\infty)$, for t small enough,

$$\left|\nabla \log p_{t,x_0}(x)\right|^2 \leq a_t^2 |x|^2 + b_t^2 |x_0|^2, \tag{48}$$

which recalls the parabolic logarithmic Harnack and Li–Yau inequalities, leading to

$$\mathrm{I}(\mu_{t_0} \mid \mu) \leq 2 \int \left(a_{t_0}^2 |x|^2 + \left|\nabla V(x)\right|^2\right) \mathrm{d}\mu_t(x) + 2b_{t_0}^2 |x_0|^2. \tag{49}$$

Here a_t and b_t are constants with respect to x, that may depend on t but not on d and x_0. The explicit dependence over the dimension and initial condition is crucial here for our purposes. Unfortunately, we were not able to locate (48) in the literature on short time logarithmic gradient bounds for the heat kernel on non-compact manifolds with a lower curvature bound. Close works are [17] where a constant is not controlled with respect to the dimension, and [13] which is curvature free but is restricted to a compact subset. It would not be a surprise to assume, in addition to positive curvature, an additional condition regarding eigenfunctions associated to the spectral gap.

To summarize, the upper bound on $\mathrm{I}(\mu_t \mid \mu)$ follows from a logarithmic gradient bound for the heat kernel in short time, which expresses a regularization property of the evolution equation driven by $\mathscr{L}$ when its starts from a Dirac mass. Such an upper bound on the Fisher information I implies all the others. Indeed, it would give an upper bound on H by using the logarithmic Sobolev inequality, giving in turn an upper bound on W_2 by using the Talagrand inequality, as well as an upper bound on d_{TV} by using the Pinsker inequality.

An estimate on I is at the heart of the varentropy approach in [34]. In [35], the control of Fisher information is circumvented by a direct comparison with varentropy.

2.2. *Optimal constants*

An observation that dates back to [1], see also [2], is that $\mathrm{CD}(\rho,\infty)$ implies a logarithmic Sobolev inequality (LSI) with constant $2/\rho$ as well as a Poincaré inequality (PI) of constant $1/\rho$ (directly as well as by linearization). Recall that the optimal Poincaré constant is precisely the inverse of the spectral gap $1/\lambda_1$. It follows that rigidity implies that the optimal LSI constant is twice the optimal PI constant, just like for the OU process. By the way, a famous alternative due to Oscar Rothaus states that in the compact setting, if the optimal LSI constant is not twice the optimal PI constant, then there exists an extremal function for LSI, see for instance [33, Theorem at bottom of p. 107] as well as [37, Theorem 2.2.3, p. 333] in the lecture notes by Laurent Saloff-Coste for a discrete version. Note also that in the case (18), the function $x \mapsto \mathrm{e}^{\alpha(x_1+\cdots+x_d)}$ is extremal for LSI, as observed in [11], while $x \mapsto x_1 + \cdots + x_d$ is extremal for PI.

2.3. *Normalization*

For any real $\alpha > 0$ which may depend on d, the time-changed process $X^{(\alpha)} := (X_{\alpha t})_{t\geq 0}$ solves $\mathrm{d}X_t^{(\alpha)} = -\alpha\nabla V(X_t^{(\alpha)})\,\mathrm{d}t + \sqrt{2\alpha}\,\mathrm{d}B_t$ and has generator $\alpha\mathscr{L}$. The process $X^{(\alpha)}$ has cutoff at critical time t_* iif $X^{(\alpha)}$ has cutoff at critical time t_*/α. Such scaled processes play a role with respect to mean-field limits of interacting particle systems related to McKean–Vlasov semilinear PDE, see [9]. Note that the product condition is invariant under such a scaling.

2.4. *Temperature*

For any real $\sigma > 0$ which may depend on d, playing the role of the temperature, let us consider the Markov diffusion process $(Y_t^{(\sigma)})_{t\geq 0}$ solving the stochastic differential equation $\mathrm{d}Y_t^{(\sigma)} = -\nabla V(Y_t^{(\sigma)})\,\mathrm{d}t + \sqrt{2\sigma^2}\,\mathrm{d}B_t$. This is (1) when $\sigma = 1$. Its invariant law is $\mu_\sigma = \mathrm{e}^{-\frac{1}{\sigma^2}V}$ while its generator is $\sigma^2\Delta - \nabla V\cdot\nabla$. The Bakry–Émery operators are

$$\Gamma(f) = \sigma^2|\nabla f|^2 \qquad \text{and} \qquad \Gamma_2(f) = \sigma^4\big\|\mathrm{Hess}(f)\big\|_{\mathrm{HS}}^2 + \sigma^2\big\langle \mathrm{Hess}(V)\nabla f, \nabla f\big\rangle. \tag{50}$$

Thus, for all $\rho > 0$, $\mathrm{CD}(\rho,\infty)$ of $\mathscr{L}$ is equivalent to $\mathrm{Hess}(V) \geq \rho\,\mathrm{I}_d$ as quadratic form, which is free of σ. In particular, in the rigid case $\rho = \lambda_1$, the cutoff critical time is free of σ. This can be also obtained from Section 2.3 by using $\alpha = \sigma^2$ and $\frac{1}{\sigma^2}V$ instead of V.

3. Proofs

3.1. *Preliminaries on rigidity*

As already mentioned, our starting point is a Gaussian splitting or factorization theorem of [14], which in the Euclidean setting takes the form of Theorem 11 below. See also [11, Remark 3.2].

Theorem 11 (Gaussian factorization in the Euclidean space). *Consider the operator* (3) *and assume that for some constant $\rho > 0$, V is ρ-convex and $\lambda_1 = \rho$. Then there is an orthonormal basis $(e_1,\dots,e_d)$ of $\mathbb{R}^d$ and a vector $m\in\mathbb{R}^d$ such that V is of the form*

$$V(x) = \frac{\rho}{2}\big((x_1-m_1)^2+\cdots+(x_k-m_{k_1})^2\big) + \widetilde{V}(x_{k_1+1},\dots,x_d) \tag{51}$$

where k_1 is the dimension of the eigenspace associated with λ_1, and $\widetilde{V}$ is ρ-convex on $\mathbb{R}^{d-k_1}$. Moreover, all eigenfunctions with eigenvalue λ_1 are affine, and only depend on the first k_1 coordinates in the above basis.

The vector m actually is the center of mass of the probability measure e^{-V}.

Remark 12 (Eigenfunctions structure). Since eigenfunctions satisfy

$$\int|\nabla f|^2\,\mathrm{d}\mu = -\int(\mathscr{L}f)f\,\mathrm{d}\mu = \lambda_1\int f^2\,\mathrm{d}\mu, \tag{52}$$

any eigenfunction with eigenvalue λ_1 is of the form $\langle a, p(x)\rangle_{\mathbb{R}^k} + b$ with p the projection on the k-dimensional Gaussian factor and $|a|^2 = \lambda_1\|f\|^2_{L^2(\mu)}$.

Another way of stating this result is that up to a rotation (the change of basis) and a translation by the vector m, the law e^{-V} is a product measure, with a centered Gaussian factor with variance ρ on the first k coordinates, and a log-concave factor with a ρ-convex potential on the last $d-k$ coordinates. Note that the result can only be true up to a rotation and translation, since the assumptions are stable by isometries of $\mathbb{R}^d$.

An alternative proof in the Euclidean setting, based on a rigidity property for regularity of solutions to the Monge–Ampère PDE for optimal transport maps, was given by [16]. We shall discuss some elements of proof in the full Riemannian setting in Section 3.8.

Lemma 13 (OU process associated to an eigenfunction). *Under the setting of Theorem 11 and Definition 8, for all $T\in\mathbb{S}_{F_1}$, the process $\big(T(X_t)\big)_{t\geq 0}$ is a k_1-dimensional OU process scaled by a factor ρ, that is a process on $\mathbb{R}^{k_1}$ with generator $\rho\Delta - \rho x\cdot\nabla$.*

Proof. We shall show that $T(X)$ is a Markov process and recognize it as an OU process by computing the generator. If $\vec{v}$ is a vector-valued function, then $\vec{\mathscr{L}}\vec{v}$ is the vector obtained by applying $\mathscr{L}$ to each coordinate, and $\Gamma(\vec{v})$ the matrix whose coefficients are $\Gamma(\vec{v}_i, \vec{v}_j)$. Since the coordinates of T are orthogonal normalized eigenfunctions, $\vec{\mathscr{L}}T = -\lambda_1 T$ and $\Gamma(T) = \rho\,\mathrm{Id}$, as per Remark 12. By the diffusion property, we have

$$\mathscr{L}(g\circ T) = \nabla g\circ T\cdot\vec{\mathscr{L}}T + \langle\nabla^2 g\circ T, \Gamma(T)\rangle \tag{53}$$

$$= -\lambda_1\nabla g\circ T\cdot T + \rho\Delta g\circ T \tag{54}$$

$$= -\rho\nabla g\circ T\cdot T + \rho\Delta g\circ T. \tag{55}$$

This is a function of T, so $T(X)$ is a Markov process, and when viewing it as such it is indeed the generator of an OU process with variance ρ^{-1}, applied to a function g. □

Hence, in the rigid case, the full process contains an OU subprocess, and hence cannot converge to equilibrium faster than it. This will yield the lower bound in Theorem 1.

3.2. *Proof of Theorem 1*

The upper bound is an immediate well-known consequence of the exponential convergence to equilibrium in Wasserstein distance under the $\mathrm{CD}(\rho,\infty)$ condition, see for instance [32]: for all X_0, Y_0 and $t \geq 0$,

$$\mathrm{W}_2(X_t, Y_t) \leq \mathrm{e}^{-\rho t}\mathrm{W}_2(X_0, Y_0). \tag{56}$$

The desired upper bound follows by taking $Y_0 \sim \mu$ and X_0 deterministic.

Let us now prove the lower bound. Consider a multi-eigenmap T, see Definition 8. Up to an isometry, it is of the form $T(x) = Ax + b$. Moreover, since the coordinates of T are orthogonal eigenfunctions, the columns of A are orthogonal vectors. If moreover we assume that $T \in \mathbb{S}_{F_1}$, then each column of A has norm $\sqrt{\rho} = \sqrt{\lambda_1}$. Hence we get

$$\mathrm{W}_2\big((T(X_t), T(Y_t)\big) \leq \sqrt{\lambda_1}\,\mathrm{W}_2(X_t, Y_t). \tag{57}$$

Applying Lemma 13, $Z_t := T(X_t)$ and $T(Y_t)$ are OU processes with speed accelerated by a factor ρ. Suppose now that Y_0 is distributed according to the equilibrium measure: $Y_0 \sim \mu$. Then $T(Y_t)$ follows a standard Gaussian law γ_{k_1} on $\mathbb{R}^{k_1}$, for all t. Thus using the formula $\mathrm{W}_2\big(\mathcal{N}(m_1,\Sigma_1),\mathcal{N}(m_2,\Sigma_2)\big)^2 = (m_1 - m_2)^2 + \mathrm{Tr}\big((\sqrt{\Sigma_1} - \sqrt{\Sigma_2})^2\big)$, the Mehler formula $Z_t \sim \mathcal{N}\big(z_0\,\mathrm{e}^{-\rho t}, (1-\mathrm{e}^{-2\rho t})I_{k_1}\big)$, and $\gamma_{k_1} = \mathcal{N}(0, I_{k_1})$, we find

$$\mathrm{W}_2(Z_t, \gamma_{k_1})^2 = \mathrm{e}^{-2\rho t}\big(|z_0|^2 + k_1\big). \tag{58}$$

Hence

$$\mathrm{W}_2(X_t, \mu) \geq \frac{\mathrm{e}^{-\rho t}}{\sqrt{\rho}}\sqrt{\big|T(x_0)\big|^2 + k_1}. \tag{59}$$

Optimizing over x_0 and T concludes the proof.

We now consider the case where μ is centered and $S = B(0,R)$. For any $T \in \mathbb{S}_{F_1}$, writing $T = Ax + b$, since $\int T\,\mathrm{d}\mu = 0$ we see that $b = 0$. Therefore $T(0) = 0$. Moreover, since the columns of A are orthogonal and have norm $\sqrt{\lambda_1}$, we get $\sup_{x\in S}|Ax|^2 = \rho R^2$.

3.3. *Proof of Corollary 3*

Without loss of generality, we can assume that $m^{(d)} = 0$ by translating $V^{(d)}$. We start by using the lower bound of Theorem 1 to prove the convergence to infinity when $t_d = (1-\epsilon)t_*$. Since we are in the centered setting and the set of initial conditions is a centered ball of radius $c\sqrt{d}$, the lower bound is

$$\frac{\mathrm{e}^{-\rho t}}{\sqrt{\rho}}(\rho c^2 d + k_1)^{1/2} \le \mathrm{W}_2(X_t,\mu). \tag{60}$$

Evaluating at $t = t_d$ and neglecting k_1, we get

$$cd^{\epsilon/2} \le \mathrm{W}_2(X_{t_d},\mu) \tag{61}$$

and letting d go to infinity concludes the proof of the lower bound. Let us now prove the case $t_d = (1+\epsilon)t_*$ via the upper bound in Theorem 1. Since $\mu^{(d)}$ is centered,

$$\int |x-x_0|^2\,\mathrm{d}\mu(x) = |x_0|^2 + \int |x|^2\,\mathrm{d}\mu(x). \tag{62}$$

From the spectral gap, for any centered f we have the Poincaré inequality

$$\int f^2\,\mathrm{d}\mu \le \frac{1}{\lambda_1^{(d)}}\int |\nabla f|^2\,\mathrm{d}\mu. \tag{63}$$

Applying this inequality to each coordinate yields

$$\int |x|^2\,\mathrm{d}\mu \le \frac{d}{\lambda_1^{(d)}}. \tag{64}$$

Therefore we have the upper bound

$$\mathrm{W}_2(X_t,\mu) \le \mathrm{e}^{-\lambda_1^{(d)}t}\Big(\sup_{x_0\in S}|x_0|^2 + \frac{d}{\lambda_1^{(d)}}\Big)^{1/2}. \tag{65}$$

Since $\sup_{x_0\in S}|x_0|^2 = c^2 d$, evaluating at $t_d = (1+\epsilon)t_*$ yields

$$\mathrm{W}_2(X_{t_d},\mu) \le d^{-\epsilon/2}\big(c^2 + (\lambda_1^{(d)})^{-1}\big)^{1/2}. \tag{66}$$

Since $\liminf_{d\to\infty}\lambda_1^{(d)} > 0$, letting d go to infinity concludes the proof.

Remark 14. In the proof, the assumption $\liminf_{d\to\infty}\lambda_1^{(d)} > 0$ could have been replaced by a slow enough growth of $(\lambda_1^{(d)})^{-1}$, for instance $(\lambda_1^{(d)})^{-1} \le (\log d)^\alpha$ for some $\alpha > 0$.

3.4. *Proof of Corollary 4*

Let us prove first the cutoffs in relative entropy and in total variation distance by using Corollary 4 on the Wasserstein distance. Let us drop the superscript (d) to simplify the notation.

Upper bound. It is well-known that $\mathrm{CD}(\rho,\infty)$ with $\rho > 0$ implies that for all $t \ge t_0 > 0$,

$$\mathrm{H}(X_t \mid \mu) \le \mathrm{e}^{-\rho(t-t_0)}\,\mathrm{H}(X_{t_0} \mid \mu). \tag{67}$$

The role of $t_0 > 0$ is to benefit from the fact that the law of X_{t_0} is absolutely continuous with respect to μ, while the law of X_0 is a Dirac mass. On the other hand, following for instance [8, Lemma 4.2],[3] $\mathrm{CD}(\rho,\infty)$ with $\rho \in \mathbb{R}_+$ implies that for all $t > 0$,

$$\mathrm{H}(X_t \mid \mu) \le \frac{\rho\,\mathrm{e}^{-2\rho t}}{1-\mathrm{e}^{-2\rho t}}\mathrm{W}_2(X_0,\mu)^2 \le \frac{1}{2t}\mathrm{W}_2(X_0,\mu)^2. \tag{68}$$

Using (68) with $t = t_0 = 1$, and combining with (67), we obtain, for all $t \ge 1$,

$$2\,\mathrm{H}(X_t \mid \mu) \le \mathrm{e}^{-\rho(t-1)}\,\mathrm{W}_2(X_0,\mu)^2. \tag{69}$$

By combining with the general Csiszár–Kullback–Pinsker inequality

$$\mathrm{d_{TV}}(\mu,\nu)^2 \le 2\,\mathrm{H}(\nu \mid \mu), \tag{70}$$

[3] In [8], it is at the heart of a derivation of the Otto–Villani HWI inequality, which contains the logarithmic Sobolev inequality as well as the Talagrand transportation inequality.

we get finally, for all $t \geq 1$,

$$\mathrm{d_{TV}}(X_t,\mu)^2 \leq 2\,\mathrm{H}(X_t \mid \mu) \leq \mathrm{e}^{-\rho(t-1)}\,\mathrm{W}_2(X_0,\mu)^2. \tag{71}$$

Now, since $t_* \to +\infty$ as $d \to \infty$ in Corollary 3, we get, for all $c > 0$ and $\varepsilon > 0$,

$$\lim_{d\to\infty} \sup_{x_0\in B(m,c\sqrt{d})} \mathrm{H}(X_{(1+\epsilon)t_*} \mid \mu) = 0 \quad \text{and} \quad \lim_{d\to\infty} \sup_{x_0\in B(m,c\sqrt{d})} \mathrm{d_{TV}}(X_{(1+\epsilon)t_*},\mu) = 0. \tag{72}$$

The approach differs from the one in [9], in the way we regularize in (68) as well as in the way we control relative entropy, here via the Wasserstein distance.

Lower bound. Both total variation distance and relative entropy decrease by mappings,

$$\mathrm{d_{TV}}(\mu\circ T^{-1},\nu\circ T^{-1}) \leq \mathrm{d_{TV}}(\mu,\nu) \quad \text{and} \quad \mathrm{H}(\nu\circ T^{-1} \mid \mu\circ T^{-1}) \leq \mathrm{H}(\nu \mid \mu). \tag{73}$$

This contractibility argument is also at the heart of the lower bounds in [9]. It follows that we can bound from below the relative entropy and total variation mixing times by those of a suitable OU process, by taking a multi-eigenfunction map. The lower bounds for OU processes have been established for example in [5,24], and [9, Theorem 1.2]. As a consequence, we have, in the setting of Corollary 4, for all $\varepsilon > 0$,

$$\lim_{d\to\infty} \sup_{x_0\in B(m,c\sqrt{d})} \mathrm{d_{TV}}(X_{(1-\epsilon)t_*},\mu) = 1 \quad \text{and} \quad \lim_{d\to\infty} \sup_{x_0\in B(m,c\sqrt{d})} \mathrm{H}(X_{(1-\epsilon)t_*} \mid \mu) = \infty, \tag{74}$$

when $X_0 = x_0$, and for an arbitrary constant $c > 0$. This is compatible with (70).

Another viewpoint for relative entropy is to get the Talagrand inequality

$$\mathrm{W}_2(X_t,\mu)^2 \leq \frac{2}{\rho}\,\mathrm{H}_\mu(X_t) \tag{75}$$

from $\mathrm{CD}(\rho,\infty)$, and deduce the relative entropy lower bound from the W_2 lower bound.

Let us prove the cutoff for the Fisher information. Following [1], see also [2], $\mathrm{CD}(\rho,\infty)$ gives the logarithmic Sobolev inequality

$$\mathrm{H}(\nu \mid \mu^{(d)}) \leq \frac{1}{2\rho}\,\mathrm{I}(\nu \mid \mu^{(d)}). \tag{76}$$

Taking $\nu = \mathrm{Law}(X_t^{(d)})$ and using the cutoff lower bound for H give, for all $\varepsilon > 0$,

$$\lim_{d\to\infty} \sup_{x_0^{(d)}\in S^{(d)}} \mathrm{I}\big(X_{t_d}^{(d)} \,\big|\, \mu^{(d)}\big) = +\infty \quad \text{where} \quad t_d = (1-\varepsilon)t_*. \tag{77}$$

For an upper bound, we can use another side of the Bakry–Émery theorem which states that $\mathrm{CD}(\rho,\infty)$ gives the monotonicities

$$\partial_t \mathrm{H}_t(\mu_t \mid \mu) = -\mathrm{I}(\mu_t \mid \mu) \leq 0 \quad \text{and} \quad \partial_t \mathrm{I}(\mu_t \mid \mu) \leq -2\rho\,\mathrm{I}(\mu_t \mid \mu) \leq 0 \tag{78}$$

where μ_t and μ stand for $\mathrm{Law}(X_t^{(d)})$ and $\mu^{(d)}$. The Grönwall lemma implies then the exponential decay of Fisher information: for all $t \geq t_1 > 0$,

$$\mathrm{I}(\mu_t \mid \mu) \leq \mathrm{e}^{-2\rho(t-t_0)}\,\mathrm{I}(\mu_{t_1} \mid \mu) \tag{79}$$

(recall that μ_0 is a Dirac mass). To upper bound $\mathrm{I}(\mu_{t_1} \mid \mu)$, we write, for all $0 < t_0 < t_1$,

$$\mathrm{H}(\mu_{t_0} \mid \mu) - \mathrm{H}(\mu_{t_1} \mid \mu) = \int_{t_0}^{t_1} \mathrm{I}(\mu_s \mid \mu)\,\mathrm{d}s \geq (t_1 - t_0)\,\mathrm{I}(\mu_{t_1} \mid \mu) \tag{80}$$

where we have used the monotonicity of I. This gives, combined with (69), when $t_0 > 1$,

$$\mathrm{I}(\mu_{t_1} \mid \mu) \leq \frac{\mathrm{H}(\mu_{t_0} \mid \mu)}{t_1 - t_0} \leq \frac{\mathrm{e}^{-\rho(t_0-1)}}{2(t_1 - t_0)}\,\mathrm{W}_2(X_0,\mu)^2. \tag{81}$$

This gives finally a cutoff upper bound: for all $\varepsilon > 0$,

$$\lim_{d\to\infty} \sup_{x_0^{(d)}\in S^{(d)}} \mathrm{I}\big(X_{t_d}^{(d)} \,\big|\, \mu^{(d)}\big) = 0 \quad \text{where} \quad t_d = (1+\varepsilon)t_*. \tag{82}$$

3.5. *Proof of Theorem 5*

Let $h(x) = x_1 + \cdots + x_d$ be the symmetric Hermite polynomial of first degree. Its gradient is the constant vector $\nabla h = (1,\ldots,1)$. Since (5) holds, we have $\lambda_1 \geq \rho$. Hence, to get $\lambda_1 = \rho$, it suffices to show that h is an eigenfunction of $-\mathscr{L}$ with eigenvalue ρ, and then apply the splitting theorem. The generator is

$$\mathscr{L} f = \Delta f - \rho x \cdot \nabla f - \nabla W \cdot \nabla f. \tag{83}$$

Since W is invariant along $\mathbb{R}(1,\ldots,1)$, we have $\nabla W \cdot (1,\ldots,1) = 0$, therefore,

$$\mathscr{L} h = 0 - \rho h - \nabla W \cdot (1,\ldots,1) = -\rho h. \tag{84}$$

This gives $h \in E_1$ and $\lambda_1 = \rho$. Note that by rigidity, all the elements of E_1 are affine.

3.6. *Proof of Corollary 6*

First, note that $\sum_{i=1}^d m_i^{(d)} = \int \sum_{i=1}^d x_i \,\mathrm{d}\mu^{(d)}(x) = 0$ since the image law of $\mu^{(d)}$ by $(x_1,\ldots,x_d) \mapsto x_1 + \cdots + x_d$ is a centered Gaussian.

Lower bound when $t_d = (1-\epsilon)t_*$. As per Theorem 5, $x_1 + \cdots + x_d$ is an eigenfunction. Since its squared $L^2(\mu)$ norm is $d\rho^{-1}$, $f_d(x) = \sqrt{\rho/d}(x_1 + \cdots + x_d)$ is a normalized eigenfunction. Now, for both $S^{(d)} = B\big(m^{(d)}, c\sqrt{d}\big)$ and $S^{(d)} = m^{(d)} + [-c,c]^d$ we have

$$\sup_{x \in S^{(d)}} \big|f_d(x)\big| = c\sqrt{d}. \tag{85}$$

Hence the lower bound in Theorem 1 yields

$$\mathrm{W}_2\big(X_{t_d}^{(d)}, \mu^{(d)}\big) \geq \mathrm{e}^{-\rho t_d}\, c\sqrt{d} = c d^{\epsilon/2} \xrightarrow[d\to\infty]{} +\infty. \tag{86}$$

Upper bound when $t_d = (1+\epsilon)t_*$. As in the proof of Corollary 3, the uniform convexity yields, via the Poincaré inequality, for all $1 \leq i \leq d$,

$$\int (x_i - m_i^{(d)})^2 \,\mathrm{d}\mu^{(d)}(x) \leq \rho^{-1}. \tag{87}$$

Hence when applying Theorem 1 we get

$$\mathrm{W}_2\big(X_{t_d}^{(d)}, \mu^{(d)}\big) \leq \mathrm{e}^{-\lambda_1^{(d)} t_d}\Big(\sup_{x_0 \in S^{(d)}} |x_0|^2 + d/\rho\Big)^{1/2} \leq (c^2 + \rho^{-1})^{1/2} d^{-\epsilon/2} \xrightarrow[d\to\infty]{} 0. \tag{88}$$

It remains to combine Corollaries 3 and 4 to conclude the proof.

3.7. *Proof of Theorem 7*

The spectral gap $\lambda_1^{(d)}$ of $X^{(d)}$ satisfies $\lambda_1^{(d)} \geq \kappa^{(d)}$, thus

$$\lim_{d\to\infty} \lambda_1^{(d)} t_0^{(d)} = +\infty, \tag{89}$$

which is the product condition in [35, Corollary 1], hence the cutoff in total variation distance. It remains to prove cutoff for the other cases. Let us start with the relative entropy lower bound. The Csiszár–Kullback–Pinsker inequality (70) gives

$$\liminf_{d\to\infty} \sup_{x_0^{(d)} \in S^{(d)}} \mathrm{H}\Big(X^{(d)}_{(1-\varepsilon/2)t_0^{(d)}} \,\Big|\, \mu^{(d)}\Big) \geq \frac{\eta^2}{2}. \tag{90}$$

Moreover, from the exponential decay of entropy, for all $t > 0$ and $t' \geq t$,

$$\mathrm{H}\big(X_{t'}^{(d)} \mid \mu^{(d)}\big) \leq \mathrm{e}^{-2\kappa^{(d)}(t'-t)}\,\mathrm{H}\big(X_t^{(d)} \mid \mu^{(d)}\big). \tag{91}$$

Using $t' = \left(1-\frac{\varepsilon}{2}\right)t_0^{(d)}$, $t = (1-\varepsilon)\,t_0^{(d)}$, and $\lim_{d\to\infty}\kappa^{(d)}\,t_0^{(d)} = +\infty$, we get

$$\liminf_{d\to\infty} \mathrm{H}\Big(X^{(d)}_{(1-\varepsilon)\,t_0^{(d)}} \,\Big|\, \mu^{(d)}\Big) \geq \mathrm{e}^{\varepsilon \limsup_{d\to\infty}\kappa^{(d)}\,t_0^{(d)}}\,\frac{\eta^2}{2} = +\infty. \tag{92}$$

For the upper bound, a careful reading of the proof of [35, Theorem 1] shows that

$$\limsup_{d\to\infty} \mathrm{H}\Big(X^{(d)}_{\left(1+\frac{\varepsilon}{2}\right)t_0^{(d)}} \,\Big|\, \mu^{(d)}\Big) \leq C_\varepsilon < \infty. \tag{93}$$

Using the exponential decay of the relative entropy and $\lim_{d\to\infty}\kappa^{(d)}\,t_0^{(d)} = +\infty$, we get

$$\limsup_{d\to\infty} \mathrm{H}\Big(X^{(d)}_{(1+\varepsilon)\,t_0^{(d)}} \,\Big|\, \mu^{(d)}\Big) \leq \mathrm{e}^{-\varepsilon \liminf_{d\to\infty}\kappa^{(d)}\,t_0^{(d)}}\,C_\varepsilon = 0. \tag{94}$$

For Wasserstein distance, the upper bound comes from the one for relative entropy via the Talagrand inequality $\kappa^{(d)}\,\mathrm{W}_2(X,\mu^{(d)})^2 \leq 2\,\mathrm{H}(X \mid \mu^{(d)})$, while the lower bound comes from the regularization (68) with $t = \varepsilon\, t_0^{(d)}$ and the semigroup property.

Finally, for Fisher information, the lower bound comes from one for the relative entropy via the logarithmic Sobolev inequality $2\kappa^{(d)}\,\mathrm{H}(X \mid \mu^{(d)}) \leq \mathrm{I}(X \mid \mu^{(d)})$, while the upper bound comes from the regularization (81) and the semigroup property.

3.8. *Proof of Theorem 9*

The proof is exactly the same as in the Euclidean setting (Theorem 1), up to the use of the Riemannian splitting (Theorem 15) below, and the fact that first eigenfunctions are affine through the splitting (Lemma 16). We shall hence only discuss these two elements, and omit the repetition of the proof. In this setting, the rigidity theorem of [14] when $\lambda_1 = \rho$ is the following.

Theorem 15 (Riemannian splitting with Gaussian factor). *Let (M,g,μ) be a weighted Riemannian manifold with probability measure $\mu = \mathrm{e}^{-V}$. If for some $\rho > 0$,*

$$\mathrm{CD}(\rho,\infty) \text{ is satisfied and } \lambda_1 = \rho, \tag{95}$$

then (M,g,μ) is isometric to a product weighted Riemannian manifold

$$\left(\mathbb{R}^k, |\cdot|_2, \gamma_{k,\rho^{-1}}\right) \times (M', g', \mu') = \left(\mathbb{R}^k \times M', |\cdot|_2 \oplus g', \gamma_{k,\rho^{-1}} \otimes \mu'\right) \tag{96}$$

where:

- *k is the dimension of the eigenspace of $-\mathscr{L}$ associated with the eigenvalue λ_1;*
- *$\gamma_{k,\rho^{-1}}$ is the centered Gaussian law on $\mathbb{R}^k$ with covariance matrix $\rho^{-1}\,\mathrm{Id}_k$;*
- *(M', g', μ') is a weighted Riemannian manifold satisfying* $\mathrm{CD}(\rho,\infty)$.

This was proved by [14] in the setting of smooth weighted manifolds, and by [19] in the more general setting of RCD spaces. Splitting theorems for manifolds satisfying a curvature constraint and optimizing certain geometric quantities is a well-studied problem in Riemannian geometry, going back to the Cheeger–Gromoll splitting theorem for nonnegatively curved manifolds containing infinite geodesics.

We now state a structure lemma which enters the proof of Theorem 15.

Lemma 16 (Rigidity for eigenfunctions when $\lambda_1 = \rho$). *Under the setting of Theorem 15, and through the isometry that it provides, any element of E_1 is of the form $a \cdot p(x)$ with p the projection on the Euclidean factor of dimension k, $a \in \mathbb{R}^k$, $|a| = \sqrt{\rho}\|f\|_{L^2(\mu)}$. Moreover, if $f_i = a_i \cdot p(x)$, for $i = 1,2$, are two orthogonal eigenfunctions, then $a_1 \perp a_2$.*

Proof. What follows is a broad sketch, were we focus on justifying the form of the eigenfunctions, but do not discuss in too much detail the splitting of the space, that was established in [14]. See also [19] for a full proof in the non-smooth setting of RCD spaces.

From the integrated Bochner formula, for any g in the domain of $\mathscr{L}$ we have

$$\int -\Gamma(g,\mathscr{L}g)\,\mathrm{d}\mu \geq \rho\int\Gamma(g)\,\mathrm{d}\mu + \int\|\mathrm{Hess}(g)\|_{\mathrm{HS}}^2\,\mathrm{d}\mu. \tag{97}$$

Taking g to be an eigenfunction with eigenvalue $\lambda_1 = \rho$, we get

$$\rho\int\Gamma(f)\,\mathrm{d}\mu \geq \rho\int\Gamma(f)\,\mathrm{d}\mu + \int\|\mathrm{Hess}(f)\|_{\mathrm{HS}}^2\,\mathrm{d}\mu \tag{98}$$

which forces $\mathrm{Hess}(f) = 0$ almost everywhere, and thus everywhere since eigenfunctions are smooth. Therefore f is affine, and non-constant. In particular, ∇f is a non-trivial parallel vector field, which forces the splitting of the manifold in a product form $\mathbb{R}\times M'$, along which f is only an affine function of the first coordinate. Repeating this strategy for successive orthogonal eigenfunctions gives a splitting with a k-dimensional Euclidean factor. See [14] for full details. And once the splitting is established, we can view the eigenfunctions as functions on $\mathbb{R}^k$.

Since for an affine function $f = a\cdot x + b$, we have $\Gamma(f) = |a|^2$, and

$$|a|^2 = \int\Gamma(f)\,\mathrm{d}\mu = -\int f(\mathscr{L}f)\,\mathrm{d}\mu = \rho\|f\|_{L^2(\mu)}^2. \tag{99}$$

Let us now compute $\Gamma(f_1,f_2)$ where f_1, f_2 are orthogonal eigenfunctions. We have

$$\Gamma(f_1,f_2) = \frac{1}{4}\big(\Gamma(f_1+f_2)-\Gamma(f_1-f_2)\big) = \frac{\rho}{4}\big(\|f_1+f_2\|_{L^2(\mu)}^2 - \|f_1-f_2\|_{L^2(\mu)}^2\big) = 0 \tag{100}$$

where we used the fact that $f_1 \pm f_2$ are also eigenfunctions, with the same eigenvalue λ_1. But since $\Gamma(f_1,f_2) = \langle a_1,a_2\rangle$, the vectors driving f_1 and f_2 must be orthogonal. □

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

References

[1] D. Bakry and M. Émery, "Diffusions hypercontractives", in *Séminaire de probabilités, XIX, 1983/84* (J. Azéma and M. Yor, eds.), Lecture Notes in Mathematics, vol. 1123, Springer, 1985, pp. 177–206.

[2] D. Bakry, I. Gentil and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften, vol. 348, Springer, 2014, xx+552 pages.

[3] D. Bakry and Z. Qian, "Some new results on eigenvectors via dimension, diameter, and Ricci curvature", *Adv. Math.* **155** (2000), no. 1, pp. 98–153.

[4] G. Barrera, "Abrupt convergence for a family of Ornstein–Uhlenbeck processes", *Braz. J. Probab. Stat.* **32** (2018), no. 1, pp. 188–199.

[5] J. Barrera, B. Lachaud and B. Ycart, "Cut-off for n-tuples of exponentially converging processes", *Stochastic Processes Appl.* **116** (2006), no. 10, pp. 1433–1446.

[6] F. Barthe and B. Klartag, "Spectral gaps, symmetries and log-concave perturbations", *Bull. Hell. Math. Soc.* **64** (2020), pp. 1–31.

[7] J. Bertrand and M. Fathi, "Stability of eigenvalues and observable diameter in RCD$(1,\infty)$ spaces", *J. Geom. Anal.* **32** (2022), no. 11, article no. 270 (38 pages).

[8] S. G. Bobkov, I. Gentil and M. Ledoux, “Hypercontractivity of Hamilton–Jacobi equations”, *J. Math. Pures Appl. (9)* **80** (2001), no. 7, pp. 669–696.
[9] J. Boursier, D. Chafaï and C. Labbé, “Universal cutoff for Dyson Ornstein Uhlenbeck process”, *Probab. Theory Relat. Fields* **185** (2023), no. 1-2, pp. 449–512.
[10] D. Chafaï, “Cutoff phenomenon for Markov processes”, January 27, 2024. Online at https://djalil.chafai.net/blog/2024/01/27/cutoff-for-markov-processes/ (accessed on October 2, 2025).
[11] D. Chafaï and J. Lehec, “On Poincaré and logarithmic Sobolev inequalities for a class of singular Gibbs measures”, in *Geometric aspects of functional analysis. Vol. I* (B. Klartag and E. Milman, eds.), Lecture Notes in Mathematics, vol. 2256, Springer, 2020, pp. 219–246.
[12] G.-Y. Chen and L. Saloff-Coste, “The cutoff phenomenon for ergodic Markov processes”, *Electron. J. Probab.* **13** (2008), no. 3, pp. 26–78.
[13] X. Chen, X.-M. Li and B. Wu, “Logarithmic heat kernel estimates without curvature restrictions”, *Ann. Probab.* **51** (2023), no. 2, pp. 442–477.
[14] X. Cheng and D. Zhou, “Eigenvalues of the drifted Laplacian on complete metric measure spaces”, *Commun. Contemp. Math.* **19** (2017), no. 1, article no. 1650001 (17 pages).
[15] T. A. Courtade and M. Fathi, “Stability of the Bakry–Émery theorem on $\mathbb{R}^n$”, *J. Funct. Anal.* **279** (2020), no. 2, article no. 108523 (28 pages).
[16] G. De Philippis and A. Figalli, “Rigidity and stability of Caffarelli’s log-concave perturbation theorem”, *Nonlinear Anal., Theory Methods Appl.* **154** (2017), pp. 59–70.
[17] A. Engoulatov, “A universal bound on the gradient of logarithm of the heat kernel for manifolds with bounded Ricci curvature”, *J. Funct. Anal.* **238** (2006), no. 2, pp. 518–529.
[18] M. Fathi, I. Gentil and J. Serres, “Stability estimates for the sharp spectral gap bound under a curvature-dimension condition”, *Ann. Inst. Fourier* **74** (2024), no. 6, pp. 2425–2459.
[19] N. Gigli, C. Ketterer, K. Kuwada and S.-i. Ohta, “Rigidity for the spectral gap on $\mathrm{RCD}(K,\infty)$-spaces”, *Am. J. Math.* **142** (2020), no. 5, pp. 1559–1594.
[20] N. Gigli, A. Mondino and G. Savaré, “Convergence of pointed non-compact metric measure spaces and stability of Ricci curvature bounds and heat flows”, *Proc. Lond. Math. Soc. (3)* **111** (2015), no. 5, pp. 1071–1129.
[21] A. Grigor’yan, *Heat kernel and analysis on manifolds*, AMS/IP Studies in Advanced Mathematics, vol. 47, American Mathematical Society; International Press, 2009, xviii+482 pages.
[22] E. P. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics, vol. 38, American Mathematical Society, 2002, xiv+281 pages.
[23] B. Klartag, “Logarithmic bounds for isoperimetry and slices of convex sets”, 2023. Online at https://arxiv.org/abs/2303.14938v1.
[24] B. Lachaud, “Cut-off and hitting times of a sample of Ornstein–Uhlenbeck processes and its average”, *J. Appl. Probab.* **42** (2005), no. 4, pp. 1069–1080.
[25] M. Lassalle, “Polynômes de Hermite généralisés”, *C. R. Math.* **313** (1991), no. 9, pp. 579–582.
[26] A. Lichnerowicz, *Géométrie des groupes de transformations*, Travaux et recherches mathématiques, vol. III, Dunod, 1958, ix+193 pages.
[27] C. H. Mai and S.-i. Ohta, “Quantitative estimates for the Bakry–Ledoux isoperimetric inequality”, *Comment. Math. Helv.* **96** (2021), no. 4, pp. 693–739.
[28] P.-L. Méliot, “The cut-off phenomenon for Brownian motions on compact symmetric spaces”, *Potential Anal.* **40** (2014), no. 4, pp. 427–509.
[29] M. Obata, “Certain conditions for a Riemannian manifold to be isometric with a sphere”, *J. Math. Soc. Japan* **14** (1962), pp. 333–340.
[30] Y. Ollivier, “A visual introduction to Riemannian curvatures and some discrete generalizations”, in *Analysis and geometry of metric measure spaces* (G. Dafni, R. J. McCann and A.

Stancu, eds.), CRM Proceedings & Lecture Notes, vol. 56, American Mathematical Society, 2013, pp. 197–220.

[31] F. Pedrotti and J. Salez, "A new cutoff criterion for non-negatively curved chains", 2025. Online at https://arxiv.org/abs/2501.13079v1.

[32] M.-K. von Renesse and K.-T. Sturm, "Transport inequalities, gradient estimates, entropy, and Ricci curvature", *Commun. Pure Appl. Math.* **58** (2005), no. 7, pp. 923–940.

[33] O. S. Rothaus, "Diffusion on compact Riemannian manifolds and logarithmic Sobolev inequalities", *J. Funct. Anal.* **42** (1981), no. 1, pp. 102–109.

[34] J. Salez, "Cutoff for non-negatively curved Markov chains", *J. Eur. Math. Soc.* **26** (2024), no. 11, pp. 4375–4392.

[35] J. Salez, "Cutoff for non-negatively curved diffusions", 2025. Online at https://arxiv.org/abs/2501.01304v1.

[36] L. Saloff-Coste, "Precise estimates on the rate at which certain diffusions tend to equilibrium", *Math. Z.* **217** (1994), no. 4, pp. 641–677.

[37] L. Saloff-Coste, "Lectures on finite Markov chains", in *Lectures on probability theory and statistics (Saint-Flour, 1996)* (P. Bernard, ed.), Lecture Notes in Mathematics, vol. 1665, Springer, 1997, pp. 301–413.

[38] F. Santambrogio, *Optimal transport for applied mathematicians*, Progress in Nonlinear Differential Equations and their Applications, vol. 87, Birkhäuser/Springer, 2015, xxvii+353 pages.