Calculus of Variations

Transportation onto log-Lipschitz perturbations

Max Fathi[1,2,3] · Dan Mikulincer[4] · Yair Shenfeld[5]

Abstract

We establish sufficient conditions for the existence of globally Lipschitz transport maps between probability measures and their log-Lipschitz perturbations, with dimension-free bounds. Our results include Gaussian measures on Euclidean spaces and uniform measures on spheres as source measures. More generally, we prove results for source measures on manifolds satisfying strong curvature assumptions. These seem to be the first examples of dimension-free Lipschitz transport maps in non-Euclidean settings, which are moreover sharp on the sphere. We also present some applications to functional inequalities, including a new dimension-free Gaussian isoperimetric inequality for log-Lipschitz perturbations of the standard Gaussian measure. Our proofs are based on the Langevin flow construction of transport maps of Kim and Milman.

1 Introduction

This work concerns the problem of transporting a probability measure $d\mu = e^{-V} d\,\mathrm{Vol}$ on a Riemannian manifold (M, g) onto a measure $d\nu = e^{-(V+W)} d\,\mathrm{Vol}$ which is an L-log-

---

✉ Max Fathi
mfathi@lpsm.paris

Dan Mikulincer
danmiku@mit.edu

Yair Shenfeld
Yair_Shenfeld@Brown.edu

1   Laboratoire Jacques Louis Lions & Laboratoire de Probabilités Statistique et Modélisation, Université Paris Cité and Sorbonne Université and CNRS, 75013 Paris, France

2   Département de Mathématiques et applications, Ecole Normale Supérieure, Université PSL, 75005 Paris, France

3   Institut Universitaire de France (IUF), Paris, France

4   Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

5   Division of Applied Mathematics, Brown University, Providence, RI, USA

Lipschitz perturbation of μ, i.e., $|\nabla W| \leq L$. We will show that the *Langevin transport map* (also known as the heat flow transport map of Kim and Milman) between μ and ν is Lipschitz in various Euclidean and manifold settings. This construction of transport maps was introduced in [1]. Our main results are Theorem 1 in the Euclidean setting and Theorem 3 in the Riemannian setting (with sharp results in the special case of the sphere, Theorem 2). We shall also discuss applications to functional inequalities and optimal transport.

The question of proving Lipschitz bounds for transport maps goes back to the seminal work of Caffarelli [2], who proved that the quadratic optimal transport map (or Brenier map) from a standard Gaussian measure onto a uniformly log-concave measure on $\mathbb{R}^n$ is globally Lipschitz, with a dimension-free bound (see [3, 4] for alternative proofs based on entropic optimal transport and [5] for Sobolev bounds). The existence of such maps allows to transfer various functional, isoperimetric, and concentration inequalities from the source measure to the target measure [2, 6, 7]. For example, Caffarelli's result immediately recovers the classical results of Bakry and Émery on sharp functional inequalities for uniformly log-concave measures. Moreover, Milman [8] showed that such Lipschitz estimates imply bounds on higher eigenvalues of certain differential operators, for which no non-transport proofs are known at this time. What is crucial for many applications, such as the correlation inequalities [7] and quantitative central limit theorems [9], is to ensure that the Lipschitz estimates are dimension-free.

Several extensions of Caffarelli's theorem on the Brenier maps have been proven since then, such as [10], which showed that the optimal transport map from a Gaussian measure onto any log-compactly-supported perturbation is globally Lipschitz. Beyond these extensions not much is known about the Lipschitz properties of the optimal transport map, and there are many remaining questions [11]. However, for most applications, it is not particularly important that the map is optimal and any globally Lipschitz map will suffice. An emerging line of research has focused on the construction and analysis of Lipschitz transport maps beyond the setting of optimal transport. Our starting point is the paper [1], in which Kim and Milman introduced a new construction of transport maps and recovered, as well as extended, Caffarelli's result. The construction is based on a time reversal of an overdamped Langevin (or drift-diffusion) SDE, and we shall call it the Langevin transport map in the sequel. As shown by Tanana [12], this construction does not coincide with the Brenier map in general (although they do coincide in dimension one), see also [13]. It is on this construction that the present work is based.

More recently, there have been several works investigating Lipschitz properties of the Langevin transport map in the Euclidean setting when the target measure satisfies certain convexity conditions [14–17]. Our focus here is on a different class of target measures, with a first-order condition on the target rather than a second-order condition (but still using strong convexity assumptions on the source measure). As a general motivation for relaxing regularity assumptions, we mention that reversing diffusion processes to construct transport maps has recently gained a lot of traction in the machine learning community [18, 19], where such constructions are used to generate samples from unknown distributions.

The study of the Langevin transport map led to new results that were previously unattainable for the optimal transport map. However, all results mentioned above are limited to measures on Euclidean spaces, while the question of Lipschitz transport maps is also interesting for non-Euclidean geometries. Indeed, for Riemannian manifolds there are some motivating questions of E. Milman [8] and of Beck and Jerison [20]. In light of this, an additional goal in this work is to establish sufficient conditions on weighted Riemmnanian manifolds to ensure the existence of Lipschitz transport maps. While the existence of a Lipschitz optimal transport map can be established in some manifold settings [21, Theorem 3.1],

the bounds are dimension-dependent and the necessary assumptions are highly restrictive. In contrast, under the first-order condition described above, our results yield, for the first time, a construction of globally Lipschitz transport maps with explicit dimension-free bounds in a general manifold setting.

Main results

As explained above, our results concern Lipschitz estimates for the Langevin transport map between a source measure (on $\mathbb{R}^n$ or on a Riemannian manifold) and a target measure on the same space, whose density with respect to the original measure is log-Lipschitz.

Euclidean spaces

Our main result in the Euclidean setting, which shall be proved in Sect. 3, is

Theorem 1 *Let $\kappa > 0$ and suppose that $V : \mathbb{R}^n \to \mathbb{R}$ is a thrice-differentiable function such that*

$$\nabla^2 V(x) \succeq \kappa I \quad \forall\, x \in \mathbb{R}^n \quad and \quad |\nabla^3 V(x)(u, u)| \leq K \quad \forall\, u \in \mathbb{S}^{n-1}.$$

Let $d\mu = e^{-V} dx$ and $d\nu = e^{-(V+W)} dx$ with

$$|\nabla W(x)| \leq L \quad \forall\, x \in \mathbb{R}^n.$$

Then, the Langevin transport map between μ and ν is Lipschitz with constant $\exp\left(10\left[\frac{L}{\sqrt{\kappa}} + \frac{L^2}{\kappa} + \frac{LK}{\kappa^2}\right]\right)$.

Remark 1 The log-Lipschitz assumption implies similar results for measures ν whose relative density with respect to μ is Lipschitz and strictly bounded from below. Indeed, if $|\nabla e^{-W}| \leq C$ and $e^{-W} > c > 0$, then $|\nabla W| \leq L := \frac{C}{c}$.

The dependence on L in Theorem 1 is sharp as can be evidenced by the standard Gaussian. Indeed, if μ is the standard Gaussian measure on $\mathbb{R}$, then $\kappa = 1$ and $K = 0$. We now choose $W(x) = L|x| + \log Z$, where Z is a normalizing constant, and suppose that there is an M-Lipschitz transport map from μ to ν. Then, as in [6, 11], the Gaussian isoperimetric inequality can be transported to ν, up to a factor of $\frac{1}{M}$. More precisely, given a set $A \subset \mathbb{R}$ satisfying $\nu(A) = \frac{1}{2}$, we have that $\nu^+(A) \geq \frac{1}{\sqrt{2\pi} M}$ where $\nu^+(A) := \lim_{\varepsilon \downarrow 0} \frac{\nu(A^\varepsilon) - \nu(A)}{\varepsilon}$ with A^ε standing for the ε-neighborhood of A. We now take $A = [0, \infty)$ and note that $\nu(A) = \frac{1}{2}$, by symmetry, and that $\nu^+(A) = \frac{d\nu}{dx}(0)$. We have

$$Z = \int e^{L|x|} d\mu \geq \int e^{-Lx} d\mu = e^{\frac{L^2}{2}}$$

so

$$\frac{d\nu}{dx}(0) = e^{-W(0)} \frac{d\mu}{dx}(0) = \frac{1}{Z} \frac{1}{\sqrt{2\pi}} \leq \frac{e^{-\frac{L^2}{2}}}{\sqrt{2\pi}}.$$

It follows that $M \geq e^{\frac{L^2}{2}}$.

Riemannian manifolds

We now discuss the manifold setting in which our first result applies to the uniform measure on the round sphere $\mathbb{S}^{n-1}$, with radius scaled as $\sqrt{n-2}$. In this case, we obtain a sharp estimate, similar to Theorem 1.

Theorem 2 *Let $n \geq 3$. Then, the Langevin transport map from the uniform measure on $\mathbb{S}^{n-1}(\sqrt{n-2})$ onto a measure with L-log-Lipschitz density is Lipschitz with constant* $\exp\left(35(L + L^2)\right)$.

The scaling in $\sqrt{n}$ of the radius is essentially sharp for dimension-free behavior, as well as the behavior in $\exp(L + L^2)$ of the Lipschitz constant, since if we let n go to infinity, the distribution of a single coordinate converges to a standard Gaussian measure, so we should recover the same type of estimates as in the Gaussian case in the limit, which is indeed the case with our estimate. When compared to Theorem 1, one may see that the convexity term κ is replaced by the Ricci curvature of the sphere, which is of order 1 when the diameter scales like $\sqrt{n}$. The other difference is the existence of the regularity parameter K, which does not appear in Theorem 2. Indeed, the source measure is uniform and its density is constant.

In Sect. 5 we shall consider a more general setting of weighted Riemannian manifolds and our bounds will be given both in terms of the regularity of the density and the curvature of the manifold, as in Theorem 2. In particular, our results shall apply to manifolds that satisfy the Bakry-Émery curvature-dimension condition, see [22, Chapter 1.6], as well as some extra conditions. We state here an informal version of the theorem and defer the exact statement and definitions to Theorem 5 in Sect. 5.

Theorem 3 *Let (M, g, μ) be a weighted Riemannian manifold with $\mu = e^{-V} d\mathrm{Vol}$, and let $\nu = e^{-W} d\mu$, with W an L-Lipschitz function. Then, under appropriate assumptions on the curvature of M and on V, the Langevin transport map from μ onto ν is Lipschitz with constant $e^{e^{cL^2}}$ for large L, where c can be made explicit, and depends on the curvature and on V, but not on the dimension.*

The estimate of Theorem 3 is still dimension-free, although significantly worse than what we derived in Theorem 1 and Theorem 2. We believe that the estimate of Theorem 3 is suboptimal and that the double exponential can be omitted, as in the sphere.

As far as we know, in the Riemannian setting, the above-mentioned results are the first results about Lipschitz transport maps with a dimension-free behavior (under appropriate scaling). The existence of such maps paves the way to several applications of interest. In Sect. 7, we prove new functional inequalities, both in the Euclidean and Riemannian settings, and we shall also discuss open problems on optimal transport maps.

The inverse map

In all of the examples above, we have taken a well-conditioned weighted manifold that satisfies some desirable combination of convexity and curvature assumptions. In these cases, we showed that log-Lipschitz perturbations can be realized as push-forward by Lipschitz mappings. We also investigate the analogous question in the reverse direction. Namely, we show that there exists a Lipschitz transport map from the perturbation, which is often not as well-behaved, back to the source measure.

Theorem 4 *In all of the above settings, the inverses of the Langevin transport maps are Lipschitz with dimension-free constants.*

The exact value of the Lipschitz constants turn out to be comparable to the Lipschitz constants of the Langevin transport maps. In Sect. 6 we give the exact dependence of the Lipschiz constants on all parameters. Theorem 4 has an additional interesting consequence: since maps can be composed, the theorem implies the existence of Lipschitz maps from one log-Lipschitz perturbation onto any other.

2 Preliminaries

Before explaining the construction of the Langevin transport map, we introduce the Langevin dynamics on $\mathbb{R}^n$. Let $d\mu = e^{-V}dx$ be a probability measure on $\mathbb{R}^n$ with $V : \mathbb{R}^n \to \mathbb{R}$ twice-differentiable. The *Langevin dynamics*, associated to μ, starting at $x \in \mathbb{R}^n$ is the solution of the stochastic differential equation

$$dX_t^x = -\nabla V(X_t^x)dt + \sqrt{2}d\omega_t, \quad X_0^x = x, \tag{1}$$

where (ω_t) is a standard Brownian motion in $\mathbb{R}^n$. We denote by (P_t) the Langevin semigroup,

$$P_t f(x) := \mathbb{E}[f(X_t^x)],$$

for any $f : \mathbb{R}^n \to \mathbb{R}$ for which the above is defined. It is well known that under appropriate regularity assumptions (P_t) is ergodic. Consequently, for any $x \in \mathbb{R}^n$, the law of X_t^x converges weakly to μ as $t \to \infty$.

The definition of the Langevin dynamics can readily be extended to Riemannian manifolds (see [22, 23] for in-depth discussions). Let (M, g) be a complete n-dimensional Riemannian manifold, and let $O(M)$ be its orthonormal frame bundle. Let (B_t) be a standard Brownian motion in $\mathbb{R}^n$ and let (Φ_t) be the horizontal Brownian motion on $O(M)$ defined by

$$d\Phi_t = \sum_{i=1}^{n} H^i(\Phi_t) \circ dB_t,$$

where $\{H^i\}_{i\in[n]}$ are the horizontal lifts of the standard basis $\{e_i\}_{i\in[n]}$, and $\circ$ stands for the Stratonovitch integral. We assume that (Φ_t) does not explode in finite time. The process (ω_t) given by $\omega_t = \pi\Phi_t$, where $\pi : O(M) \to M$ is the canonical projection, is what we refer to as the Brownian motion on M.

Now, as in the Euclidean case, given a probability measure $d\mu = e^{-V}d\text{Vol}$ on M, where $d\text{Vol}$ is the volume measure on M and $V : M \to \mathbb{R}$ is twice-differentiable, we define the *Langevin dynamics* starting at $x \in M$ as the solution of the stochastic differential equation

$$dX_t^x = -\nabla_M V(X_t^x)dt + \sqrt{2} \circ d\omega_t, \quad X_0^x = x, \tag{2}$$

where ∇_M is the gradient in M. Note that the use of the Stratonovitch integral instead of the Itô integral is immaterial in our setting because the coefficient of the Brownian motion is a constant.

The Langevin transport map

We now briefly explain the construction of the Langevin transport map. The reader is referred to [14] for further details on the constructions introduced in this section. Let $d\nu = e^{-W}d\mu$ be a probability measure on M and consider the Langevin dynamics, associated to μ, and starting at ν,

$$X_t := X_t^x \text{ with } x \sim \nu. \tag{3}$$

The flow of probability measures $(\rho_t) := (\mathrm{Law}(X_t))$ forms an interpolation between $\rho_0 = \nu$ and $\rho_\infty = \mu$. In particular, it satisfies the transport equation

$$\partial_t \rho_t = -\nabla_M \cdot (\rho_t \nabla_M \log P_t e^{-W}),$$

with ∇_M standing for the gradient in M. We define a flow of diffeomorphisms $\{S_t\}_{t \geq 0}$ as a solution to the following integral curve equation

$$\partial_t S_t = -\nabla_M \log P_t e^{-W}(S_t), \quad S_0(x) = x \ \forall x \in M. \tag{4}$$

It turns out that for any $t \geq 0$, S_t transports ν to ρ_t. Setting $T_t := S_t^{-1}$, and letting

$$T := \lim_{t \to \infty} T_t,$$

we see that T transports μ to ν. This is the *Langevin transport map* between μ and ν.[1] The particular form of the equation in (4) is particularly amenable to establishing Lipschitz properties of T from estimates of $\nabla_M^2 \log P_t e^{-W}$, where ∇_M^2 is the Hessian in M, since differentiating (4) yields

$$\partial_t \nabla_M S_t = -\nabla_M^2 \log P_t e^{-W}(S_t) \nabla_M S_t, \quad \nabla_M S_0 = \mathrm{I}.$$

This is the content of the following Lemma. The proof is an adaption of [14, Lemma 3] and [15, Lemma 3.2] which only consider Euclidean spaces.

Proposition 1 *If for all $t \geq 0$,*

$$\nabla_M^2 \log P_t e^{-W} \preceq \theta_t g,$$

then T is $\exp\left(\int_0^\infty \theta_t \, dt\right)$-Lipschitz.

Proof Consider $F_t : M \longrightarrow \mathbb{R}$ a time-dependent family of smooth functions on (M, g), such that for any $t > 0$ we have

$$\nabla_M^2 F_t \geq -\theta_t g.$$

As is classical (e.g., [25, Proposition 16.2 (iv')]), this is equivalent to saying that for any $t > 0$, $x, y \in M$ and constant speed geodesic $\alpha : [0, 1] \longrightarrow M$ connecting x to y, we have

$$\langle \dot{\alpha}(1), \nabla_M F_t(y) \rangle - \langle \dot{\alpha}(0), \nabla_M F_t(x) \rangle \geq -\theta_t d(x, y)^2.$$

Therefore, if x_t and y_t are time-dependent gradient flows of F_t with different initial data, and α^t is a constant-speed geodesic connecting x_t to y_t, we have (with d standing for the distance induced by the Riemannian metric),

$$
\begin{aligned}
\frac{d}{dt} d(x_t, y_t)^2 &= 2 \left(\langle \dot{y}_t, \dot{\alpha}^t(1) \rangle - \langle \dot{x}_t, \dot{\alpha}^t(0) \rangle \right) \\
&= 2 \left(\langle \nabla_M F(y_t), \dot{\alpha}^t(1) \rangle - \langle \nabla_M F(x_t), \dot{\alpha}^t(0) \rangle \right) \\
&\geq -2\theta_t d(x_t, y_t)^2.
\end{aligned}
$$

The rest of the proof proceeds as in the proof of [15, Lemma 3.2], with the manifold setting making no difference. $\square$

---

[1] The existence of solutions to the above equations, and of the limits, is non-trivial in general. These issues are discussed in length in [1, 14, 24] for the various settings we consider. In particular, the convexity and curvature assumptions we shall enforce in the coming proofs are enough to guarantee that all maps under consideration are well-defined.

As is clear from Proposition 1, our goal will be to estimate the Hessian along the Langevin dynamics. Such estimates have previously been studied in the literature, in particular by Elworthy and Li [26]. We also mention [27], which contains a pedagogical exposition of second derivative computations and estimates in the manifold setting. For the general manifold setting, our results will be based on the recent work of Cheng, Thalmaier, and Wang [28]. However, for the Euclidean and Sphere setting, our sharp results require new Hessian estimates, which we shall provide. All of the above methods are based on Bismut's integration by parts formula [29].

3 Euclidean spaces

This section is devoted to the proof of Theorem 1. For the rest of this section, we fix two probability measures $\mu = e^{-V(x)}dx$ and $\nu = e^{-W(x)}d\mu$ on $\mathbb{R}^n$. We require that μ, the source, satisfies the assumptions of Theorem 1,

$$\nabla^2 V(x) \succeq \kappa I \quad \forall x \in \mathbb{R}^n \quad \text{and} \quad |\nabla^3 V(x)(u,u)| \leq K \quad \forall u \in \mathbb{S}^{n-1},$$

for some $\kappa > 0$ and $K \geq 0$. We also require that ν, the target, is an L-log-Lipschitz perturbation of μ,

$$|\nabla W(x)| \leq L \quad \forall x \in \mathbb{R}^n.$$

Further, as in Sect. 2, we let X_t stand for the Langevin dynamics, associated to μ with $X_0 \sim \nu$, and P_t is its semigroup. As suggested by Proposition 1 we shall require the following global bound of $\nabla^2 \log P_t e^{-W(x)}$.

Proposition 2 *For every $t \geq 0$ and $x \in \mathbb{R}^n$,*

$$\nabla^2 \log P_t e^{-W(x)} \preceq L e^{-\kappa t} \left[5L + \frac{5}{\sqrt{t}} + \frac{Kt}{2} \right] I.$$

The proof of Proposition 2 will be the main focus of this section. Before delving into the proof, let us show how Theorem 1 follows.

Proof of Theorem 1 Using

$$\int_0^\infty e^{-\kappa t} dt = \frac{1}{\kappa}, \qquad \int_0^\infty \frac{e^{-\kappa t}}{\sqrt{t}} dt = \sqrt{\frac{\pi}{\kappa}}, \qquad \int_0^\infty t e^{-\kappa t} dt = \frac{1}{\kappa^2},$$

the combination of Proposition 1 and Proposition 2 yields that T, the Langevin transport map, is Lipschitz with constant $\exp\left(\frac{5L^2}{\kappa} + \frac{5\sqrt{\pi}L}{\sqrt{\kappa}} + \frac{LK}{2\kappa^2} \right) \leq \exp\left(10\left[\frac{L}{\sqrt{\kappa}} + \frac{L^2}{\kappa} + \frac{LK}{\kappa^2} \right] \right)$. □

3.1 Stochastic representation of the Hessian

In order to establish Lipschitz properties of the Langevin transport map our main task will be to derive global bounds on $\nabla^2 \log P_t e^{-W}$. Toward this goal, our main technical tool is a stochastic representation of $\nabla^2 P_t$, known as *Bismut's formula*. Before stating the formula, we first define the *Jacobi flow* of the Langevin dynamics as $(\nabla X_t^x)_{t \geq 0}$, where $\nabla X_t^x : \mathbb{R}^n \to \mathbb{R}^n$ is the derivative of (the random function) X_t^x with respect to the initial condition x,

$$\nabla_u X_t^x := \lim_{\varepsilon \downarrow 0} \frac{X_t^{x+\varepsilon u} - X_t^x}{\varepsilon} \in \mathbb{R}^n, \quad \forall u \in \mathbb{R}^n.$$

The Jacobi flow $(\nabla X_t^x)_{t \geq 0}$ appears naturally when applying the chain rule in the differentiation of $P_t f(x)$ with respect to x, formally,

$$\nabla P_t f(x) = \nabla \mathbb{E}[f(X_t^x)] = \mathbb{E}[\nabla f(X_t^x) \nabla X_t^x].$$

The process (∇X_t^x) satisfies the differential equation

$$\partial_t \nabla X_t^x = -\nabla^2 V(X_t^x) \nabla X_t^x, \quad \nabla X_0^x = I, \tag{5}$$

which can be seen by differentiating (1) and noting that (ω_t) does not depend on x. Since we require Hessian estimates of $P_t f$, we also consider the *second variation* of (X_t^x),

$$\nabla_{u,v}^2 X_t^x := \lim_{\varepsilon \downarrow 0} \frac{\nabla_v X_t^{x+\varepsilon u} - \nabla_v X_t^x}{\varepsilon} \in \mathbb{R}^n, \quad \forall u, v \in \mathbb{R}^n$$

(which is symmetric with respect to u and v). The equation for the process $(\nabla^2 X_t^x)$ can be obtained by differentiating (5),

$$\partial_t \nabla_{u,v}^2 X_t^x = -\nabla^3 V(X_t^x)(\nabla_u X_t^x, \nabla_v X_t^x) - \nabla^2 V(X_t^x) \nabla_{u,v}^2 X_t^x, \quad \nabla^2 X_0^x = 0, \tag{6}$$

where, for $x \in \mathbb{R}^n$ and $u, v \in \mathbb{R}^n$,

$$[\nabla^3 V(x)(u, v)]_i = \sum_{j,k=1}^n \partial_{ijk}^3 V(x) u_j v_k = \nabla_u \nabla_v (\nabla V)_i(x) \quad \forall i \in [n].$$

With these definitions, and the following notation,

$$\nabla_v f(x) := \langle f(x), v \rangle \quad \text{and} \quad \nabla_{v,u}^2 f(x) = \langle v, \nabla^2 f(x) u \rangle$$

(the definition is extended naturally for vector-valued functions), the stochastic formula now reads:

Lemma 1 ([26, p. 8 in arXiv version]) *For any $x \in \mathbb{R}^n$, $u, v \in \mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}$ twice-differentiable,*

$$\nabla_{u,v}^2 P_t f(x) = \frac{1}{t\sqrt{2}} \mathbb{E}\left[\nabla_v[f(X_t^x)] M_t^{x,u}\right] + \frac{1}{t} \int_0^t \mathbb{E}\left[\langle \nabla P_{t-s} f(X_s^x), \nabla_{u,v}^2 X_s^x \rangle\right] ds$$

where

$$M_t^{x,u} := \int_0^t \langle \nabla_u X_s^x, d\omega_s \rangle \quad x \in \mathbb{R}^n, u \in \mathbb{R}^n.$$

In our setting, $f = e^{-W}$ and we need to bound the various quantities appearing in Lemma 1. The first derivative of $P_t e^{-W}$ can be controlled by the log-Lipschitz assumption. The first variation ∇X_t^x is controlled by the κ-log-concavity assumption, and the second variation $\nabla^2 X_t^x$ is controlled by the κ-log-concavity assumption and the bound on the third derivative of V.

3.2 Estimates

The purpose of this section is to provide the requisite estimates for the proof of Proposition 2. We start with the first variation.

Lemma 2 (First variation estimate) *Suppose that $\nabla^2 V(x) \succeq \kappa I$ for every $x \in \mathbb{R}^n$. Then, a.s., for any fixed $v \in S^{n-1}$,*

$$|\nabla_v X_t^x| \leq e^{-\kappa t} \quad \forall \, t \geq 0.$$

Consequently, for any differentiable $f : \mathbb{R}^n \to \mathbb{R}$,

$$|\nabla[f(X_t^x)]| \leq e^{-\kappa t} |\nabla f(X_t^x)|.$$

Proof By the Cauchy-Schwarz inequality, for any $v \in \mathbb{R}^n$,

$$|\nabla_v[f(X_t^x)]| = |\langle \nabla f(X_t^x), \nabla_v X_t^x \rangle| \leq |\nabla f(X_t^x)||\nabla_v X_t^x|.$$

Hence, taking the supremum over $v \in \mathbb{S}^{n-1}$ it suffices to show that $|\nabla_v X_t^x| \leq e^{-\kappa t}$. To see the latter bound fix $v \in \mathbb{S}^{n-1}$ and let $\beta_v : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be given by $\beta_v(t) := |\nabla_v X_t^x|$. Then, by (5) and the convexity assumption on V,

$$\partial_t \beta_v(t) = \frac{\langle \nabla_v X_t^x, \partial_t \nabla_v X_t^x \rangle}{\beta_v(t)} = -\frac{\langle \nabla_v X_t^x, \nabla^2 V(X_t^x) \nabla_v X_t^x \rangle}{\beta_v(t)} \leq -\kappa \frac{|\nabla_v X_t^x|^2}{\beta_v(t)} = -\kappa \beta_v(t).$$

Since $\beta_v(0) = 1$, Grönwall's inequality yields $|\nabla_v X_t^x| \leq e^{-\kappa t}$. □

We now move to the second variation.

Lemma 3 (Second variation estimate) *Suppose that*

$$\nabla^2 V(x) \succeq \kappa I \quad \forall \, x \in \mathbb{R}^n \quad and \quad |\nabla^3 V(x)(u, u)| \leq K \quad \forall \, u \in \mathbb{S}^{n-1}.$$

Then, a.s., for any $t \geq 0$ and $v \in \mathbb{S}^{n-1}$,

$$|\nabla_{v,v}^2 X_t^x| \leq K t e^{-\kappa t}.$$

Proof Fix $v \in \mathbb{S}^{n-1}$ and let $\beta_v : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be given by $\beta_v(t) := |\nabla_{v,v}^2 X_t^x|$. By (6) and Lemma 2,

$$\begin{aligned}
\partial_t \beta_v(t) &= \frac{\langle \nabla_{v,v}^2 X_t^x, \partial_t \nabla_{v,v}^2 X_t^x \rangle}{\beta_v(t)} \\
&= \frac{\langle \nabla_{v,v}^2 X_t^x, -\nabla^3 V(X_t^x)(\nabla_v X_t^x, \nabla_v X_t^x) - \nabla^2 V(X_t^x) \nabla_{v,v}^2 X_t^x \rangle}{\beta_v(t)} \\
&\leq K \frac{|\nabla_{v,v}^2 X_t^x||\nabla_v X_t^x|}{\beta_v(t)} - \kappa \frac{|\nabla_{v,v}^2 X_t^x|^2}{\beta_v(t)} = K|\nabla_v X_t^x| - \kappa \beta_v(t) \leq K e^{-\kappa t} - \kappa \beta_v(t).
\end{aligned}$$

The solution to the differential equation

$$\partial_t \xi(t) = K e^{-\kappa t} - \kappa \xi(t), \quad \xi(0) = 0$$

is $\xi(t) = K t e^{-\kappa t}$ so by Grönwall's inequality

$$|\nabla_{v,v}^2 X_t^x| = \beta_v(t) \leq K t e^{-\kappa t}.$$

□

We will also use a reverse Hölder inequality for log-Lipschitz functions under a κ-log-concavity assumption:

Lemma 4 *Let $d\mu = e^{-V}dx$ and suppose that $\nabla^2 V(x) \succeq \kappa I$ for every $x \in \mathbb{R}^n$. Then, for every L-log-Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$,*

$$P_t(f^2) \leq \exp\left(L^2 \frac{1 - e^{-2\kappa t}}{\kappa}\right)(P_t f)^2.$$

Proof It is well-known (e.g., [30, Eq. (1.3)]) that if a measure η satisfies a log-Sobolev inequality with constant C then it satisfies a reverse Hölder inequality for L-log-Lipschitz functions,

$$\|f\|_{L^p(\eta)} \leq \exp\left(\frac{C}{2}L^2(p - q)\right)\|f\|_{L^q(\eta)} \quad \forall \, 0 \leq q < p.$$

Since the measure $\eta := P_t^* \delta_x$ satisfies a log-Sobolev inequality with constant $C = \frac{1-e^{-2\kappa t}}{2\kappa}$ [22, Eq. (5.5.5)], and since

$$P_t(f^2)(x) = \|f\|_{L^2(\eta)}^2 \quad \text{and} \quad (P_t f)^2 = \|f\|_{L^1(\eta)}^2,$$

we get

$$P_t(f^2) \leq \exp\left(\frac{1}{4}L^2 \frac{1 - e^{-2\kappa t}}{\kappa}\right)(P_t f)^2 \leq \exp\left(L^2 \frac{1 - e^{-2\kappa t}}{\kappa}\right)(P_t f)^2.$$

$\square$

Remark 2 The statement, and proof, of Lemma 4 can be extended verbatim to the manifold setting by requiring that the measure satisfies the so-called curvature-dimension condition $CD(\kappa, \infty)$. In particular, the statement is for the uniform measure on the unit sphere $\mathbb{S}^{n-1}$, with κ replaced by $n - 2$.

We shall also make use of the following classical lemma about martingales:

Lemma 5 *Let M_t be a continuous martingale, with $M_0 = 0$, and whose quadratic variation is almost surely bounded, that is, $\langle M \rangle_t \leq \varphi(t)$ for all $t \geq 0$. Then*

$$\mathbb{P}[|M_t| \geq \delta] \leq 2\exp\left(-\frac{\delta^2}{2\varphi(t)}\right).$$

Proof From Novikov's condition, for any $\sigma \in \mathbb{R}$ we have

$$\mathbb{E}[\exp(\sigma M_t - \sigma^2 \langle M \rangle_t / 2)] = 1,$$

so

$$\mathbb{E}[\exp(\sigma M_t)] \leq \exp(\sigma^2 \varphi(t)/2).$$

We then apply the Chernoff bound,

$$\mathbb{P}[M_t \geq \delta] \leq \inf_{\sigma \geq 0} \exp(-\sigma\delta)\mathbb{E}[\exp(\sigma M_t)]$$

to conclude. $\square$

3.3 Proof of Proposition 2

By Lemma 1,

$$\nabla_{u,v}^2 P_t f(x) = \frac{1}{t\sqrt{2}} \mathbb{E}\left[\nabla_v[f(X_t^x)]M_t^{x,u}\right] + \frac{1}{t}\int_0^t \mathbb{E}\left[(\nabla P_{t-s}f(X_s^x))^\top \nabla_{u,v}^2 X_s^x\right]ds,$$

and we will apply this identity with $f = e^{-W}$. We start by analyzing the first term.

Lemma 6 *For any $v \in \mathbb{S}^{n-1}$,*

$$\frac{1}{t\sqrt{2}}\mathbb{E}\left[\nabla_v[e^{-W(X_t^x)}]M_t^{x,u}\right] \leq 5e^{-\kappa t}\left[L^2 + \frac{L}{\sqrt{t}}\right]P_t e^{-W(x)}.$$

Proof For any $\delta \geq 0$ and f differentiable,

$$
\begin{aligned}
&|\mathbb{E}\left[\nabla_v[f(X_t^x)]M_t^{x,u}\right]| \\
&\quad \leq \mathbb{E}\left[|\nabla_v[f(X_t^x)]||M_t^{x,u}|\right] \\
&\quad = \mathbb{E}\left[1_{|M_t^{x,u}|<\delta}|\nabla_v[f(X_t^x)]||M_t^{x,u}|\right] + \mathbb{E}\left[1_{|M_t^{x,u}|\geq\delta}|\nabla_v[f(X_t^x)]||M_t^{x,u}|\right] \\
&\quad \leq \delta\mathbb{E}\left[|\nabla_v[f(X_t^x)]|\right] + \mathbb{E}\left[|\nabla_v[f(X_t^x)]|^2\right]^{1/2}\mathbb{E}\left[1_{|M_t^{x,u}|\geq\delta}|M_t^{x,u}|^2\right]^{1/2} \\
&\quad \leq \delta\mathbb{E}\left[|\nabla_v[f(X_t^x)]|\right] + \mathbb{E}\left[|\nabla_v[f(X_t^x)]|^2\right]^{1/2}\mathbb{E}\left[|M_t^{x,u}|^4\right]^{1/4}\mathbb{P}[|M_t^{x,u}| \geq \delta]^{1/4}.
\end{aligned}
$$

We will now analyze the terms above one-by-one. For $v \in \mathbb{S}^{n-1}$, Lemma 2 yields

$$|\nabla_v[f(X_t^x)]| \leq |\nabla[f(X_t^x)]| \leq e^{-\kappa t}|\nabla f(X_t^x)| = e^{-\kappa t}|f(X_t^x)||\nabla \log f(X_t^x)|,$$

so taking $f = e^{-W}$ yields

$$|\nabla_v[e^{-W(X_t^x)}]| \leq Le^{-\kappa t}e^{-W(X_t^x)}. \tag{7}$$

It follows that

$$\mathbb{E}[|\nabla_v[e^{-W(X_t^x)}]|] \leq Le^{-\kappa t}P_t e^{-W(x)}. \tag{8}$$

For the second term, (7) gives

$$\mathbb{E}[|\nabla_v[e^{-W(X_t^x)}]|^2]^{1/2} \leq Le^{-\kappa t}\left(P_t(e^{-W(x)})^2\right)^{1/2},$$

so applying Lemma 4 yields

$$\mathbb{E}[|\nabla_v[e^{-W(X_t^x)}]|^2]^{1/2} \leq L\exp\left(L^2\frac{1-e^{-2\kappa t}}{2\kappa} - \kappa t\right)P_t e^{-W(x)}. \tag{9}$$

Next we turn to analyze the martingale (M_t). By Lemma 2, $|\nabla_u X_t^x| \leq e^{-\kappa t}$ a.s. for any $u \in \mathbb{S}^{n-1}$ so the quadratic variation of (M_t) satisfies

$$\langle M^{x,u}\rangle_t = \int_0^t |\nabla_u X_s^x|^2 ds \leq \frac{1-e^{-2\kappa t}}{2\kappa}. \tag{10}$$

On the other hand, the Burkholder-Davis-Gundy inequalities yield that, for any $0 < p < \infty$,

$$\mathbb{E}[|M_t^{x,u}|^p] \leq C_p\mathbb{E}[\langle M^{x,u}\rangle_t^{p/2}], \tag{11}$$

for a constant C_p depending only on p. Combining (10) and (11), with $p = 4$, and using $C_4 \leq 360$ (see [31, proof of Proposition 3.26]), yields

$$\mathbb{E}[|M_t^{x,u}|^4]^{1/4} \leq 5\sqrt{\frac{1 - e^{-2\kappa t}}{2\kappa}}. \tag{12}$$

Moreover, applying Lemma 5, we have

$$\mathbb{P}[|M_t^{x,u}| \geq \delta] \leq 2 \exp\left(-\delta^2 \frac{\kappa}{1 - e^{-2\kappa t}}\right).$$

It follows that

$$\mathbb{P}[|M_t^{x,u}| \geq \delta]^{1/4} \leq 2^{1/4} \exp\left(-\frac{\delta^2}{4} \frac{\kappa}{1 - e^{-2\kappa t}}\right). \tag{13}$$

Combining (8), (9), (12), (13) yields

$$\frac{1}{t\sqrt{2}}|\mathbb{E}\left[\nabla_v[f(X_t^x)]M_t^{x,u}\right]|$$

$$\leq \frac{1}{t\sqrt{2}}\left[\delta L e^{-\kappa t} + L \exp\left(L^2\frac{1 - e^{-2\kappa t}}{2\kappa} - \kappa t\right)5\sqrt{\frac{1 - e^{-2\kappa t}}{2\kappa}}2^{1/4}\exp\left(-\frac{\delta^2}{4}\frac{\kappa}{1 - e^{-2\kappa t}}\right)\right]P_t e^{-W(x)}$$

$$= \frac{Le^{-\kappa t}}{t\sqrt{2}}\left[\delta + 5 \cdot 2^{1/4}\sqrt{\frac{1 - e^{-2\kappa t}}{2\kappa}}\exp\left(L^2\frac{1 - e^{-2\kappa t}}{2\kappa}\right)\exp\left(-\frac{\delta^2}{8}\frac{2\kappa}{1 - e^{-2\kappa t}}\right)\right]P_t e^{-W(x)}.$$

We choose δ so that the term e^{L^2} vanishes. In particular, we take

$$\delta := 2\sqrt{2}L\frac{1 - e^{-2\kappa t}}{2\kappa}$$

to get

$$\frac{1}{t\sqrt{2}}|\mathbb{E}\left[\nabla_v[f(X_t^x)]M_t^{x,u}\right]| \leq \frac{Le^{-\kappa t}}{t\sqrt{2}}\left[2\sqrt{2}L\frac{1 - e^{-2\kappa t}}{2\kappa} + 5 \cdot 2^{1/4}\sqrt{\frac{1 - e^{-2\kappa t}}{2\kappa}}\right]P_t e^{-W(x)}$$

$$\leq \frac{Le^{-\kappa t}}{t\sqrt{2}}\left[2\sqrt{2}L\frac{2\kappa t}{2\kappa} + 5 \cdot 2^{1/4}\sqrt{\frac{2\kappa t}{2\kappa}}\right]P_t e^{-W(x)}$$

$$= \left[2L^2 e^{-\kappa t} + 5 \cdot 2^{-1/2} \cdot 2^{1/4}\frac{Le^{-\kappa t}}{\sqrt{t}}\right]P_t e^{-W(x)}$$

$$\leq 5e^{-\kappa t}\left[L^2 + \frac{L}{\sqrt{t}}\right]P_t e^{-W(x)}.$$

$\square$

We now analyze the second term.

Lemma 7

$$\frac{1}{t}\int_0^t\left|\mathbb{E}\left[\langle\nabla P_{t-s}e^{-W(X_s^x)}, \nabla_{u,u}^2 X_s^x\rangle\right]\right|ds \leq LKt\frac{e^{-\kappa t}}{2}P_t e^{-W(x)}.$$

Proof By [22, Eq. (5.5.4)], for any $y \in \mathbb{R}^n$,

$$|\nabla P_{t-s} e^{-W(y)}| \leq e^{-\kappa(t-s)} P_{t-s} \left(|\nabla e^{-W(y)}| \right) = e^{-\kappa(t-s)} P_{t-s} \left(|\nabla W| e^{-W(y)} \right)$$
$$\leq L e^{-\kappa(t-s)} P_{t-s} e^{-W(y)} \tag{14}$$

so by (14) and Lemma 3,

$$\frac{1}{t} \int_0^t \left| \mathbb{E}\left[\langle \nabla P_{t-s} e^{-W(X_s^x)}, \nabla_{u,u}^2 X_s^x \rangle \right] \right| ds$$

$$\leq \frac{1}{t} \int_0^t \mathbb{E}\left[|\nabla P_{t-s} e^{-W(X_s^x)}| |\nabla_{u,u}^2 X_s^x| \right] ds$$

$$\leq \frac{1}{t} \int_0^t \mathbb{E}\left[e^{-\kappa(t-s)} L (P_{t-s} e^{-W(X_s^x)}) K s e^{-\kappa s} \right] ds = L K \frac{e^{-\kappa t}}{t} P_t e^{-W(x)} \int_0^t s\, ds$$

$$= L K t \frac{e^{-\kappa t}}{2} P_t e^{-W(x)}.$$

$\square$

We can now complete the proof of Proposition 2. By Lemma 6 and Lemma 7,

$$\nabla_{u,u}^2 \log P_t e^{-W(x)} = \frac{\nabla_{u,u}^2 P_t e^{-W(x)}}{P_t e^{-W(x)}} - |\nabla_u \log P_t e^{-W(x)}|^2 \leq \frac{\nabla_{u,u}^2 P_t e^{-W(x)}}{P_t e^{-W(x)}}$$

$$\leq 5 e^{-\kappa t} \left[L^2 + \frac{L}{\sqrt{t}} \right] + \frac{L K t}{2} e^{-\kappa t}$$

$$= L e^{-\kappa t} \left[5L + \frac{5}{\sqrt{t}} + \frac{K t}{2} \right].$$

4 The round sphere

This section is devoted to the proof of Theorem 2. Throughout this section, we let μ stand for the uniform probability measure on $\mathbb{S}^{n-1}$, the unit sphere in $\mathbb{R}^n$, and let $\nu = e^{-W(x)} d\mu$ be an L-log-Lipschitz perturbation of μ. In other words, if ∇_S is the spherical gradient, then for almost every $x \in \mathbb{S}^{n-1}$,

$$|\nabla_S W(x)| \leq L.$$

As before, X_t^x is the Langevin process associated with μ, with semigroup (P_t). Since μ is uniform, X_t^x takes a particularly simple form $X_t^x = \omega_t$, where ω_t is a standard Brownian motion on $\mathbb{S}^{n-1}$, with $\omega_0 = x$. We also write X_t when $X_0 \sim \nu$.

As suggested by Proposition 1, the proof of Theorem 2 will require bounding $\nabla_S^2 \log P_t e^{-W(x)}$. The bound of $\nabla_S^2 \log P_t e^{-W(x)}$ will be obtained, as in the Euclidean setting, by using a stochastic representation via Bismut's formula. The main result of this section is the following result.

Proposition 3 *For every $t \geq 0$ and $x \in \mathbb{S}^{n-1}$,*

$$\nabla_S^2 \log P_t e^{-W(x)} \preceq 12 \left(L + \frac{L^2}{\sqrt{n-2}} \right) e^{-(n-2)t} \left(\frac{1}{\sqrt{t}} + 1 \right) g.$$

Before delving into the proof of Proposition 3, let us show how Theorem 2 follows.

Proof of Theorem 2 Using

$$\int_0^\infty e^{-(n-2)t} dt = \frac{1}{n-2}, \qquad \int_0^\infty \frac{e^{-(n-2)t}}{\sqrt{t}} dt = \sqrt{\frac{\pi}{n-2}},$$

the combination of Proposition 1 and Proposition 3 yields that T, the Langevin transport map, is Lipschitz with constant $\exp\left(12\left(\frac{L}{\sqrt{n-2}} + \frac{L^2}{n-2}\right)\left(\frac{1}{\sqrt{n-2}} + \sqrt{\pi}\right)\right) \leq \exp\left(35\left(\frac{L}{\sqrt{n-2}} + \frac{L^2}{n-2}\right)\right)$. The result now follows by re-scaling the sphere. $\square$

4.1 Stochastic representation of the Hessian

The proof of Proposition 3 will rely on a stochastic representation of the Hessian of the heat semigroup on the sphere, analogous to Lemma 1. Before introducing the particular form of this representation, we recall a few facts about the curvature of the sphere. Let g be the metric on $\mathbb{S}^{n-1}$ induced from $\mathbb{R}^n$. Given vector fields X, Y, Z, the Riemannian curvature tensor on the sphere is given by

$$\mathrm{Riem}_S(X, Y)Z = g(Y, Z)X - g(X, Z)Y, \tag{15}$$

and its trace, the Ricci curvature, is given by

$$\mathrm{Ric}_S(Y, Z) = (n-2)g(Y, Z).$$

For $t \geq 0$ and $x \in \mathbb{S}^{n-1}$, we use $//_t : T_x \mathbb{S}^{n-1} \to T_{X_t^x} \mathbb{S}^{n-1}$ for the (random) *parallel transport* operator between the tangent spaces at x and at X_t^x. The role of the Jacobi flow will be played by the operators $Q_t : T_x \mathbb{S}^{n-1} \to T_{X_t^x} \mathbb{S}^{n-1}$ defined as,

$$Q_t := e^{-(n-2)t} //_t. \tag{16}$$

An elementary calculation using the above formula for Ric_S shows that

$$\partial_t Q_t = -\mathrm{Ric}_S^\sharp Q_t; \quad Q_0 = \mathrm{id},$$

which is analogous to (5) for the Jacobi flow with a curvature term. Here, $\mathrm{Ric}_S^\sharp$ is defined by the canonical isomorphism $g(\mathrm{Ric}_S^\sharp(u), v)_x = \mathrm{Ric}_S(u, v)$ for all $x \in M$, and $u, v \in T_x M$.

Lemma 8 ([28, Lemma 2.6][2]) *Let $f : \mathbb{S}^{n-1} \to \mathbb{R}$ and $t > 0$. For any continuously differentiable $k : [0, t] \to \mathbb{R}$ such that $k(0) = 1$ and $k(t) = 0$, we have, for any $x \in \mathbb{S}^{n-1}$ and $v, u \in T_x \mathbb{S}^{n-1}$,*

$$\nabla_S^2 P_t f(v, u) = -\sqrt{2}\mathbb{E}\left[g(\nabla_S f(X_t^x), Q_t(v)) \int_0^t g(Q_s(\dot{k}(s)u), //_s dB_s)\right]$$

$$+ \sqrt{2}\mathbb{E}\left[g\left(\nabla_S f(X_t^x), Q_t \int_0^t Q_s^{-1} \mathrm{Riem}_S(//_s dB_s, Q_s(v))Q_s(k(s)u)\right)\right],$$

where B_s is a standard Brownian motion on $T_x \mathbb{S}^{n-1}$.

Lemma 8 allows for some flexibility with the choice of the function k. For our use, we fix the following specific choice

$$k(s) := 1 - \frac{s}{t} \quad \forall s \in [0, t]. \tag{17}$$

---

[2] We rescale the generator of the heat equation so our t corresponds to $2t$ in [28].

4.2 Proof of Proposition 3

As in the Euclidean setting, we need to bound the first and second terms of Lemma 8 separately. The first term is analogous to the first term of Lemma 1 where we need to bound the martingale $\int_0^t \langle Q_s(\dot{k}(s)w), //_s dB_s \rangle$. The second term of Lemma 1 is absent in the sphere setting since μ is the uniform measure. However, due to the curvature of the sphere, we get the second term in Lemma 8. As we will show, this term *also* involves a martingale. More precisely, the term

$$Q_t \int_0^t Q_s^{-1} \operatorname{Riem}_S(//_s dB_s, Q_s(v))Q_s(k(s)v)$$

is a martingale, which is a feature of the constant curvature of the sphere, absent in the general manifold setting. In light of this, to bound $\nabla_S^2 P_t e^{-W}$, we shall require the following martingale argument, which extends the bounds obtained in the Euclidean setting.

Lemma 9 *Let $f : \mathbb{S}^{n-1} \to \mathbb{R}$ be an L-log-Lipschitz function and let (M_t) be a martingale on $T_x \mathbb{S}^{n-1}$, for some $x \in \mathbb{S}^{n-1}$. Assume that almost surely, M_t has bounded quadratic variation, that is*

$$\langle M \rangle_t \le \varphi(t) \quad a.s.,$$

for some $\varphi(t) > 0$. Then

$$\mathbb{E}[|g(\nabla_S f(X_t^x), //_t M_t)|] \le \left(6L + 2\frac{L^2}{\sqrt{n-2}} \right) \sqrt{\varphi(t)} \mathbb{E}[f(X_t^x)].$$

Proof Since μ is the uniform measure on $\mathbb{S}^{n-1}$, Lemma 4 and Remark 2 yield

$$\mathbb{E}[|\nabla_S f(X_t^x)|^2] \le L^2 \mathbb{E}[f(X_t^x)^2] \le L^2 \exp\left(L^2 \frac{1 - e^{-2(n-2)t}}{n-2} \right) \mathbb{E}[f(X_t^x)]^2,$$

where the first inequality is the L-log-Lipschitz condition coupled with the chain rule. As in the Euclidean setting, we have, for any $\delta > 0$,

$$\mathbb{E}[|\nabla_S f(X_t^x)||M_t|] \le \delta \mathbb{E}[|\nabla_S f(X_t^x)|] + \mathbb{E}[|\nabla_S f(X_t^x)|^2]^{1/2} \mathbb{E}[|M_t|^4]^{1/4} \mathbb{P}[|M_t| \ge \delta]^{1/4},$$

as well as the bound $\mathbb{E}[|M_t|^4]^{1/4} \le 5\varphi(t)^{1/2}$ which follows from the Burkholder-Davis-Gundy inequality [31, Proposition 3.26] and the bound on the quadratic variation. Moreover, using Lemma 5, $\mathbb{P}[|M_t| \ge \delta] \le 2 \exp\left(-\frac{\delta^2}{2\varphi(t)} \right)$. It follows that

$$\mathbb{E}[|\nabla_S f(X_t^x)||M_t|] \le \delta \mathbb{E}[|\nabla_S f(X_t^x)|]$$
$$+ \left(5 \cdot 2^{1/4} L \exp\left(L^2 \frac{1 - e^{-2(n-2)t}}{2(n-2)} \right) \varphi(t)^{1/2} \left(-\frac{\delta^2}{8\varphi(t)} \right) \right) \mathbb{E}[f(X_t)]$$
$$\le \left(\delta L + 5 \cdot 2^{1/4} L \exp\left(L^2 \frac{1 - e^{-2(n-2)t}}{2(n-2)} \right) \varphi(t)^{1/2} \right.$$
$$\left. \exp\left(-\frac{\delta^2}{8\varphi(t)} \right) \right) \mathbb{E}[f(X_t)],$$

where the second inequality used $\mathbb{E}[|\nabla_S f(X_t^x)|] \leq L\mathbb{E}[f(X_t^x)]$ since f is L-log-Lipschitz. We choose δ so that the term e^{L^2} vanishes. In particular, we take $\delta = 2\sqrt{2}L\sqrt{\varphi(t)}\sqrt{\frac{1-e^{-2(n-2)t}}{2(n-2)}}$, we get

$$\mathbb{E}[|\nabla_S f(X_t^x)||M_t|] \leq \sqrt{\varphi(t)}\left(2\sqrt{2}L^2\sqrt{\frac{1-e^{-2(n-2)t}}{2(n-2)}} + 5 \cdot 2^{1/4}L\right)\mathbb{E}[f(X_t^x)]$$

$$\leq \left(6L + 2\frac{L^2}{\sqrt{n-2}}\right)\sqrt{\varphi(t)}\mathbb{E}[f(X_t^x)].$$

$\square$

We now bound the two terms appearing in Lemma 8 using Lemma 9. For brevity, let us use $f := e^{-W}$. We start by analyzing the first term.

Lemma 10 *Fix $x \in \mathbb{S}^{n-1}$. For every unit vectors $v, u \in T_x\mathbb{S}^{n-1}$ and $t > 0$,*

$$\left|\mathbb{E}\left[g(\nabla_S f(X_t^x), Q_t(v))\int_0^t g(Q_s(\dot{k}(s)u), //_s dB_s)\right]\right|$$

$$\leq 6\frac{e^{-(n-2)t}}{\sqrt{t}}\left(L + \frac{L^2}{\sqrt{n-2}}\right)P_t f(x).$$

Proof Using (16) we see that the term $\mathbb{E}\left[g(\nabla_S f(X_t^x), Q_t(v))\int_0^t g(Q_s(\dot{k}(s)u), //_s dB_s)\right]$ can be written as $e^{-(n-2)t}\mathbb{E}[g(\nabla_S f(X_t^x), //_t M_t)]$ where

$$M_t = \left(\int_0^t g(Q_s(\dot{k}(s)u), //_s dB_s)\right)v$$

is a martingale. The quadratic variation of M_t is equal to, using (17),

$$\int_0^t e^{-2(n-2)s}\dot{k}(s)^2 ds = \frac{1}{t^2}\int_0^t e^{-2(n-2)s} = \frac{1}{t^2}\frac{1-e^{-2(n-2)t}}{2(n-2)} \leq \frac{1}{t} =: \varphi(t)$$

where the inequality follows from $1 - e^{-2(n-2)t} \leq 2(n-2)t$. The proof is complete by Lemma 9. $\square$

We now analyze the second term where we take $v = u$.

Lemma 11 *Fix $x \in \mathbb{S}^{n-1}$. For every unit vector $v \in T_x\mathbb{S}^{n-1}$ and $t > 0$,*

$$\left|\mathbb{E}\left[g\left(\nabla_S f(X_t^x), Q_t\int_0^t Q_s^{-1}\operatorname{Riem}_S(//_s dB_s, Q_s(v))Q_s(k(s)v)\right)\right]\right|$$

$$\leq 6e^{-(n-2)t}\left(L + \frac{L^2}{\sqrt{n-2}}\right)P_t f(x).$$

Proof We use the expressions (15) and (16) to write

$$
\begin{aligned}
& Q_s^{-1} \operatorname{Riem}_S(//_s dB_s, Q_s(v)) Q_s(k(s)v) \\
&= \left(e^{-(n-2)s}\right)^{-1} //_s^{-1} \operatorname{Riem}_S(//_s dB_s, e^{-(n-2)s} //_s v) e^{-(n-2)s} k(s) //_s v \\
&= e^{-(n-2)s} k(s) //_s^{-1} \operatorname{Riem}_S(//_s dB_s, //_s v) //_s v \\
&= e^{-(n-2)s} k(s) //_s^{-1} \left[g(//_s v, //_s v) //_s dB_s - g(//_s dB_s, //_s v,) //_s v \right] \\
&= e^{-(n-2)s} k(s) //_s^{-1} \left[//_s dB_s - g(//_s dB_s, //_s v) //_s v \right] \\
&= e^{-(n-2)s} k(s) \left[dB_s - g(dB_s, v) v \right],
\end{aligned}
$$

where in the last equation we used that $//_s$ is a linear isometry. The term $dB_s - g(dB_s, v) v$ is a Brownian motion in the hyperplane orthogonal to v inside $T_x \mathbb{S}^{n-1}$, so it follows that $\int_0^t Q_s^{-1} \operatorname{Riem}_S(//_s dB_s, Q_s(v)) Q_s(k(s)v)$ is a martingale in $T_x \mathbb{S}^{n-1}$ whose quadratic variation is bounded from above by (using $0 \le k(s) \le 1$ for all $s \in [0, t]$),

$$
(n-2) \int_0^t e^{-2(n-2)s} k(s)^2 dr \le (n-2) \int_0^t e^{-2(n-2)s} dr \le \frac{n-2}{2(n-2)} = \frac{1}{2},
$$

Therefore, we can view the term

$$
\mathbb{E}\left[g\left(\nabla_S f(X_t^x), Q_t \int_0^t Q_s^{-1} \operatorname{Riem}_S(//_s dB_s, Q_s(v)) Q_s(k(s)v) \right) \right]
$$

as of the form

$$
e^{-(n-2)t} \mathbb{E}[g(\nabla_S f(X_t^x), //_t M_t)],
$$

where M_t is a martingale on $T_x \mathbb{S}^{n-1}$ with quadratic variation bounded by $\frac{1}{2}$. The proof is complete by Lemma 9. $\square$

We can now complete the proof of Proposition 3. By Lemma 10 and Lemma 11,

$$
\begin{aligned}
\nabla_S^2 \log P_t e^{-W(x)}(u, u) &= \frac{\nabla_S^2 P_t e^{-W(x)}(u, u)}{P_t e^{-W(x)}} - g(\nabla_S \log P_t e^{-W(x)}, u)^2 \\
&\le \frac{|\nabla_S^2 P_t e^{-W(x)}(u, u)|}{P_t e^{-W(x)}} \\
&\le 12 \left(L + \frac{L^2}{\sqrt{n-2}} \right) e^{-(n-2)t} \left(\frac{1}{\sqrt{t}} + 1 \right).
\end{aligned}
$$

5 General manifolds

Let (M, g, μ) with $\mu = e^{-V} d\operatorname{Vol}$ be a weighted Riemannian manifold, with tangent bundle TM. We start by introducing the relevant concepts from Riemannian geometry.

Notions from Riemannian geometry

- If S is a symmetric tensor on TM, we denote $S^\#: TM \longrightarrow TM$ the operator such that $g(S^\#(u), v)_x = S(u, v)$, $\forall x \in M, u, v \in T_x M$ where we suppress the dependence of S on x.

- Riem denotes the Riemannian curvature 4-tensor on (M, g). Its norm is defined as

$$\| \operatorname{Riem} \|_\infty := \sup \left\{ \sum_{i=i}^{d} g(\operatorname{Riem}(x)(e_i, u)v, S^\#(e_i)), x \in M, |u|, |v| \le 1, |S|_{\mathrm{op}} \le 1 \right\}$$

where $u, v \in T_x M$, $\{e_i\}$ is any basis of $T_x M$, and S is a symmetric 2-tensor.
- Ric denotes the Ricci curvature tensor on (M, g).
- $d^* \operatorname{Riem}$ is defined as

$$g(d^* \operatorname{Riem}(u, v), w) = g((\nabla_w \operatorname{Ric}^\#)(u), v) - g((\nabla_v \operatorname{Ric}^\#)(w), u).$$

- Ric_V denotes the weighted Ricci curvature tensor of Bakry-Émery on (M, g, μ), that is,

$$\operatorname{Ric}_V := \operatorname{Ric} + \nabla_M^2 V$$

where ∇_M and ∇_M^2 are, respectively, the gradient and Hessian on the manifold M.

The manifold (M, g, μ) satisfies the $\mathrm{CD}(\kappa, \infty)$ *curvature-dimension*, for $\kappa \in \mathbb{R}$, if

$$\operatorname{Ric}_V \succeq \kappa g.$$

With these definitions, we state our main result, which is the precise form of Theorem 3.

Theorem 5 *Let (M, d, μ) be a weighted Riemannian manifold with $\mu = e^{-V} d\mathrm{Vol}$, and let $\nu = e^{-W} d\mu$ be another measure on M, with W an L-Lipschitz function. Assume that $\| \operatorname{Riem} \|_\infty < \infty$, $\operatorname{Ric}_V \succeq \kappa g$ with $\kappa > 0$, and that*

$$\beta := \| \nabla \operatorname{Ric}_V^\# + d^* \operatorname{Riem} + \operatorname{Riem}(\nabla V) \|_\infty < \infty.$$

Then, the Langevin transport map from μ onto ν is Lipschitz with constant $\exp$
$$\left(L \left(\frac{e^{\frac{L^2}{2\kappa}}}{\sqrt{\kappa}} + e^{\frac{L^2}{2\kappa}} \frac{\| \operatorname{Riem} \|_\infty}{\kappa^{\frac{3}{2}}} + \frac{\beta}{\kappa^2} \right) \right).$$

On the assumptions, the lower bound on the Ricci curvature tensor is the natural analogue of the uniform convexity assumption in the Euclidean setting, while a bound on $\nabla \operatorname{Ric}_V^\#$ is the natural counterpart to the third derivative bound on V in the Euclidean setting. The other terms we have to control are purely geometric and vanish if we consider a Euclidean space or the uniform measure on the sphere.

As in the previous section, we shall require a bound on the Hessian along the Langevin semigroup. The following result is the analog of Proposition 2 and Proposition 3 but for general manifolds. The estimate of Proposition 4 is not strong enough to yield the sharp results of Theorem 1 and Theorem 2 which necessitate Proposition 2 and Proposition 3. It is this estimate which leads to the double-exponential estimate on the Lipschitz constant of the transport maps.

Proposition 4 *For every $t \ge 0$ and $x \in M$,*

$$\nabla_M^2 \log P_t e^{-W(x)} \preceq e^{-\kappa t} L \left(\left(\frac{\sqrt{\kappa}}{\sqrt{e^{\kappa t} - 1}} + \frac{\| \operatorname{Riem} \|_\infty}{\sqrt{\kappa}} \right) e^{L^2 \left(\frac{1 - e^{-2\kappa t}}{2\kappa} \right)} + \frac{\beta}{\kappa} \right) g.$$

Proof By Lemma 4 and Remark 2, if f is positive and L-log-Lipschitz then

$$P_t(|\nabla_M f|^2) \le \exp \left(L^2 \frac{1 - e^{-2\kappa t}}{\kappa} \right) (P_t |\nabla_M f|)^2 \le L^2 \exp \left(L^2 \frac{1 - e^{-2\kappa t}}{\kappa} \right) (P_t f)^2. \quad (18)$$

Hence, by [28, Theorem 2.5][3] and (18),

$$\nabla_M^2 P_t e^{-W(x)} \preceq e^{-\kappa t} \left(\left(\frac{\sqrt{\kappa}}{\sqrt{e^{2\kappa t} - 1}} + \frac{\|\operatorname{Riem}\|_\infty}{\sqrt{\kappa}} \right) (P_t |\nabla_M e^{-W}|^2)^{1/2} + \frac{\beta}{\kappa} P_t |\nabla_M e^{-W}| \right) g$$

$$\preceq e^{-\kappa t} \left(\left(\frac{\sqrt{\kappa}}{\sqrt{e^{2\kappa t} - 1}} + \frac{\|\operatorname{Riem}\|_\infty}{\sqrt{\kappa}} \right) L e^{L^2 \left(\frac{1 - e^{-2\kappa t}}{2\kappa} \right)} P_t e^{-W} + \frac{\beta L}{\kappa} P_t e^{-W} \right) g$$

$$(19)$$

so (19) yields

$$\nabla_M^2 \log P_t e^{-W(x)} = \frac{\nabla_M^2 P_t e^{-W(x)}}{P_t e^{-W(x)}} - \left(\nabla_M \log P_t e^{-W(x)} \right)^{\otimes 2} \preceq \frac{\nabla_M^2 P_t e^{-W(x)}}{P_t e^{-W(x)}}$$

$$\preceq e^{-\kappa t} L \left(\left(\frac{\sqrt{\kappa}}{\sqrt{e^{2\kappa t} - 1}} + \frac{\|\operatorname{Riem}\|_\infty}{\sqrt{\kappa}} \right) e^{L^2 \left(\frac{1 - e^{-2\kappa t}}{2\kappa} \right)} + \frac{\beta}{\kappa} \right) g.$$

$\square$

Proof of Theorem 5 Proposition 4 implies, for every $x \in M$ and unit vector $u \in T_x M$,

$$\int_0^\infty \left| \nabla_M^2 \log P_t e^{-W(x)}(u, u) \right| dt$$

$$\leq \int_0^\infty e^{-\kappa t} L \left(\left(\frac{\sqrt{\kappa}}{\sqrt{e^{2\kappa t} - 1}} + \frac{\|\operatorname{Riem}\|_\infty}{\sqrt{\kappa}} \right) e^{\frac{L^2}{2\kappa}} + \frac{\beta}{\kappa} \right) dt$$

$$= L \left(\sqrt{\kappa} e^{\frac{L^2}{2\kappa}} \int_0^\infty \frac{e^{-\kappa t}}{\sqrt{e^{2\kappa t} - 1}} dt + \left(e^{\frac{L^2}{2\kappa}} \frac{\|\operatorname{Riem}\|_\infty}{\sqrt{\kappa}} + \frac{\beta}{\kappa} \right) \int_0^\infty e^{-\kappa t} dt \right)$$

$$= L \left(\frac{e^{\frac{L^2}{2\kappa}}}{\sqrt{\kappa}} + e^{\frac{L^2}{2\kappa}} \frac{\|\operatorname{Riem}\|_\infty}{\kappa^{\frac{3}{2}}} + \frac{\beta}{\kappa^2} \right).$$

The proof is complete by Proposition 1.

$\square$

6 The inverse of the Langevin transport map

In this section, we focus on the inverse of the Langevin transport map. Recall, from (4), that $(S_t)_{t \geq 0}$ stands for forward maps along the Langevin flow. Thus we have that $S := \lim_{t \to \infty} S_t = T^{-1}$, when T is the Langevin transport map. We shall prove the following precise form of Theorem 4.

Theorem 6 1. *Let μ and ν be as in Theorem 1. Then S is Lipschitz with constant* $\exp\left(\frac{21}{2} \frac{L^2}{\kappa} + \frac{5\sqrt{\pi} L}{\sqrt{\kappa}} + \frac{LK}{2\kappa^2} \right)$

2. *Let μ and ν be as in Theorem 2. Then S is Lipschitz with constant* $\exp\left(35 \frac{L}{\sqrt{n-2}} + \frac{71}{2} \frac{L^2}{n-2} \right)$.

3. *Let μ and ν be as in Theorem 5. Then S is Lipschitz with constant* $\exp$
$\left(L \left(\frac{e^{\frac{L^2}{2\kappa}}}{\sqrt{\kappa}} + e^{\frac{L^2}{2\kappa}} \frac{\|\operatorname{Riem}\|_\infty}{\kappa^{\frac{3}{2}}} + \frac{\beta}{\kappa^2} \right) + \frac{L^2}{2\kappa} \right)$.

---

[3] We rescale the generator of the heat equation so our t corresponds to $2t$ in [28].

The proof of Theorem 6 is analogous to the proofs of the Lipschitz properties of the Langevin transport map T from μ to ν. With the same argument as in the proof of Proposition 1, without reversing the map, we get:

Proposition 5 *If for all $t \geq 0$*

$$\nabla^2_M \log P_t e^{-W} \succeq -\ell_t g,$$

then S is $\exp\left(\int_0^\infty \ell_t dt\right)$*-Lipschitz.*

The next result shows how to combine log-Lipschitz properties together with Hessian estimates to get a lower bound on the Hessian as in Proposition 5.

Proposition 6 *Let (M, g, μ) be weighted Riemannian manifold with the associated Langevin semigroup (P_t) and let $|\cdot| := g(\cdot, \cdot)$. Suppose that we have:*

1. *There exists $\kappa > 0$ such that for every test function f,*

$$|\nabla_M P_t f| \leq e^{-\kappa t} P_t |\nabla_M f|.$$

2. *There exists $\alpha : [0, \infty) \times [0, \infty) \to \mathbb{R}$ such that for every L-log-Lipschitz nonnegative function f,*

$$\frac{\nabla^2_M P_t f}{P_t f} \succeq -\alpha(t, L)g.$$

Then, for every L-log-Lipschitz function f,

$$\nabla^2_M \log P_t f \succeq -(\alpha(t, L) + L^2 e^{-2\kappa t})g.$$

Proof Item 1 implies that

$$|\nabla_M \log P_t f| = \frac{|\nabla_M P_t f|}{|P_t f|} \leq e^{-\kappa t} \frac{P_t |\nabla_M f|}{|P_t f|} = e^{-\kappa t} \frac{P_t |f \nabla_M \log f|}{|P_t f|} \leq L e^{-\kappa t}$$

so, by item 2, for every tangent vector u,

$$\nabla^2_M \log P_t f(u, u) = \frac{\nabla^2_M P_t f(u, u)}{P_t f} - g(\nabla_M \log P_t f, u)^2 \geq (\alpha(t, L) + L^2 e^{-2\kappa t}).$$

$\square$

We can now complete the proof of Theorem 6.

1. The assumption $\nabla^2 V(x) \succeq \kappa I$ together with the classical Bakry-Émery gradient estimate [22, Eq. (5.5.4)] shows that item 1 of Proposition 6 holds with κ. As for item 2, it follows from the proof of Proposition 2 that it holds with

$$\alpha(t, L) = Le^{-\kappa t}\left[5L + \frac{5}{\sqrt{t}} + \frac{Kt}{2}\right].$$

Hence, the condition in Proposition 5 is satisfied with

$$\ell_t = Le^{-\kappa t}\left[5L + \frac{5}{\sqrt{t}} + \frac{Kt}{2}\right] + L^2 e^{-2\kappa t},$$

which completes the proof by integration from $t = 0$ to $t = \infty$ (see the proof of Theorem 1).

2. The sphere $\mathbb{S}^{n-1}$ satisfies the curvature-dimension condition $CD(n-2, \infty)$ so the Bakry-Émery gradient estimate [22, Eq. (5.5.4)] can be applied to show that item 1 of Proposition 6 holds with $n-2$. As for item 2, it follows from the proof of Proposition 3 that it holds with

$$\alpha(t, L) = 45 \left(L + \frac{L^2}{\sqrt{n-2}} \right) e^{-(n-2)t} \left(\frac{1}{\sqrt{t}} + 1 \right).$$

Hence, the condition in Proposition 5 is satisfied with

$$\ell_t = 35 \left(L + \frac{L^2}{\sqrt{n-2}} \right) e^{-(n-2)t} \left(\frac{1}{\sqrt{t}} + 1 \right) + L^2 e^{-2(n-2)t},$$

which completes the proof by integration from $t = 0$ to $t = \infty$ (see the proof of Theorem 2).

3. By assumption, M satisfies the curvature-dimension condition $CD(\kappa, \infty)$ so the Bakry-Émery gradient estimate [22, Eq. (5.5.4)] can be applied to show that item 1 of Proposition 6 holds with κ. As for item 2, it follows from the proof of Proposition 4 that it holds with

$$\alpha(t, L) = e^{-\kappa t} L \left(\left(\frac{\sqrt{\kappa}}{\sqrt{e^{2\kappa t} - 1}} + \frac{\| \text{Riem} \|_\infty}{\sqrt{\kappa}} \right) e^{L^2 \left(\frac{1 - e^{-2\kappa t}}{2\kappa} \right)} + \frac{\beta}{\kappa} \right).$$

Hence, the condition in Proposition 5 is satisfied with

$$\ell_t = e^{-\kappa t} L \left(\left(\frac{\sqrt{\kappa}}{\sqrt{e^{2\kappa t} - 1}} + \frac{\| \text{Riem} \|_\infty}{\sqrt{\kappa}} \right) e^{L^2 \left(\frac{1 - e^{-2\kappa t}}{2\kappa} \right)} + \frac{\beta}{\kappa} \right) + L^2 e^{-2\kappa t},$$

which completes the proof by integration from $t = 0$ to $t = \infty$ (see the proof of Theorem 5).

7 Applications

As is classical, Lipschitz transport maps can be used to prove functional inequalities, by transferring them from the source measure to the target measure. In this section, we discuss some results on functional inequalities for Lispchitz perturbations that appear to be new.

7.1 Dimension-free Gaussian isoperimetric inequalities

One of the strongest functional inequalities for the Gaussian measure is the Gaussian isoperimetric inequality, which states that among all sets with given Gaussian mass, the minimal Gaussian perimeter is achieved for half-spaces. The general functional inequality is defined as follows:

Definition 7 Let μ be a probability measure on a metric space. Given a Borel set A, we define its t-enlargement as $\{x; d(x, A) \leq t\}$ and its boundary measure (or Minkowski content) as

$$\mu^+(A) := \liminf_{t \to 0} t^{-1} \mu(A^t \backslash A).$$

The probability measure μ is said to satisfy a Gaussian isoperimetric inequality with constant $\alpha > 0$ if for any Borel set we have

$$\mu^+(A) \geq \alpha I(\mu(A)), \quad I = \varphi \circ \phi^{-1}$$

where $\varphi(x) = (2\pi)^{-1/2} \exp(-x^2/2)$ and $\phi(x) = \int_{-\infty}^{x} \varphi(t)dt$.

The function I corresponds to the perimeter of a half-space with the right Gaussian mass, and the Gaussian measure satisfies this inequality with constant 1 in all dimensions. This inequality was generalized to uniformly log-concave measures in [32]. The functional form of the Gaussian isoperimetric inequality is Bobkov's inequality [33]. It is immediate that Gaussian isoperimetric inequalities can be transferred by Lipschitz transport maps (see [6]), so we get the following as a corollary of Theorem 1:

Theorem 8 *If μ is a log-Lipschitz perturbation of a standard Gaussian measure on $\mathbb{R}^d$, then it satisfies a Gaussian isoperimetric inequality with a dimension-free constant.*

The same result holds in the general context of Theorems 1, 2 and 5. For the sphere, we can transport the sharp isoperimetric inequality on the sphere, which contains a dimensional improvement compared with the Gaussian isoperimetric inequality.

To our knowledge, this is the first result on dimension-free isoperimetric inequalities for log-Lipschitz perturbations. A result of Miclo (see [34, Lemma 2.1]) states that Lipschitz perturbations of uniformly log-concave measures can be recast as bounded perturbations, allowing to use the Holley-Stroock lemma to deduce functional inequalities. However, the constants obtained in that way are dimension-dependent. Some of the results of [35, Section 5] can be used to prove functional inequalities for Lipschitz perturbations of uniformly-convex measures, but the constants depend on exponential moments of the unperturbed measure, and are hence dimensional.

Gaussian isoperimetric inequalities imply other functional inequalities, including in particular logarithmic Sobolev inequalities:

Definition 9 A probability measure μ on a manifold is said to satisfy a logarithmic Sobolev inequality if for any locally Lipschitz function f such that $|\nabla f| \in L^2(\mu)$, we have

$$\int f^2 \log f^2 d\mu - \left(\int f^2 d\mu\right) \log\left(\int f^2 d\mu\right) \leq 2C_{LSI} \int |\nabla f|^2 d\mu.$$

The standard Gaussian satisfies a logarithmic Sobolev inequality, with sharp constant $C_{LSI} = 1$. This inequality, which is the sharpest possible analogue of Sobolev inequalities for the Gauss space, implies for example dimension-free concentration bounds, and has found many applications in statistics and statistical physics. We refer to [22] for an overview.

For logarithmic Sobolev inequalities, Aida and Shigekawa [36] showed that such Lipschitz perturbations satisfy them, with a dimension-free constant that has not been made explicit. Gaussian isoperimetric inequalities are strictly stronger than logarithmic Sobolev inequalities.

As a corollary, we recover functional inequalities for Gaussian mixtures with compactly supported mixing measures, which can be rewritten as Lipschitz perturbations of a Gaussian measure (see the discussion at the end of [34, Section 2]).

There are some other estimates for which it is difficult to provide direct proofs, but that can easily be extended from one measure to another via Lipschitz transport maps, including general eigenvalue estimates for diffusion generators [8] and sharpened integrability bounds for non-Lipschitz functions [37, 38].

7.2 A growth estimate for optimal transport maps

Our argument does not apply to quadratic optimal transport maps. Nonetheless, it would be natural to expect similar Lipschitz estimates for them. As a hint towards such a result, we

remark that the Gaussian concentration inequality implied by our Lipschitz estimate for non-optimal maps, combined with the sub-Gaussian case of [39, Theorem 1.1], implies controlled linear growth for optimal transport maps, in the following way:

Proposition 7 *Let μ be a centered and L-log-Lipschitz perturbation of a standard Gaussian measure on $\mathbb{R}^d$. Then the quadratic optimal transport map $\nabla\varphi$ sending γ onto μ satisfies*

$$|\nabla\varphi(x)| \leq C_1 \exp(C_2 \max(L, L^2))\sqrt{d + |x|^2}$$

where C_1 and C_2 are universal constants.

In view of our results, it is natural to expect that Lipschitz estimates hold for optimal transport maps. Even the Gaussian case is open at this time:

Conjecture 1 *The Brenier map from the standard Gaussian measure onto a log-Lipschitz perturbation of it is globally Lipschitz, and its norm can be bounded independently of the dimension.*

Note that such a statement would already significantly extend the Gaussian sub-case of [10, Theorem 1.1] when the perturbation is smooth. Similar questions can be raised for more general measures, as well as in the Riemannian setting, but the Gaussian case seems like a good starting point. Since the heat flow map and the Brenier transport map coincide in dimension one, this conjecture does hold on $\mathbb{R}$.

Let us conclude with some comments on what is missing to prove this conjecture. The optimal transport map is the unique weak solution to the Monge-Ampère equation

$$e^{-V} = e^{-V(\nabla\varphi)-W(\nabla\varphi)} \det \nabla^2\varphi$$

among convex functions. This is a fully nonlinear second-order PDE. One can start by investigating the linearized equation. If we replace W by a small perturbation $\varepsilon W - c_\varepsilon$, with c_ε a constant to enforce unit mass for the target distribution, and look for a solution of the form $\nabla\varphi = x + \varepsilon\nabla h$, the linearized equation as ε goes to zero is

$$\Delta h - \nabla V \cdot \nabla h = W - \int W e^{-V} dx. \tag{20}$$

The operator on the left-hand side is precisely the generator of the diffusion process (1). This linearization highlights the connection between quadratic optimal transport and the heat flow map investigated here: the linearization of both constructions gives rise to the same PDE.

To prove a regularity estimate for a nonlinear PDE, it is natural to start from a proof for the linearized equation. Regularity for solutions to Poisson equations such as (20) was investigated in [40–42] using the stochastic representation of solutions, Malliavin calculus and Bismut's formula. These are precisely the kind of tools we used in this work.

Since optimal transport maps do not admit a similar stochastic representation, it would be natural to start by looking for a PDE proof of the regularity estimates of [40] on (20). To our knowledge, this has not been successfully investigated and is at this point a barrier to proving Conjecture 1. The problem is of course a variant of many well-studied elliptic regularity problems. The main difficulties are that we work in a non-compact setting, allowing for unbounded solutions, and seek explicit *dimension-free* estimates. This last point in particular has rarely been investigated with PDE methods.

References

1. Kim, Y.-H., Milman, E.: A generalization of Caffarelli's contraction theorem via (reverse) heat flow. Math. Ann. **354**(3), 827–862 (2012)
2. Caffarelli, L.A.: Monotonicity properties of optimal transportation and the FKG and related inequalities. Comm. Math. Phys. **214**(3), 547–563 (2000)
3. Fathi, M., Gozlan, N., Prod'homme, M.: A proof of the Caffarelli contraction theorem via entropic regularization. Calc. Var. Part. Differ. Equ. **59**(3), 1–18 (2020)
4. Chewi, S., Pooladian, A.-A.: An entropic generalization of Caffarelli's contraction theorem via covariance inequalities (2022). arXiv preprint arXiv:2203.04954
5. Kolesnikov, A.V.: On Sobolev regularity of mass transport and transportation inequalities. Theory Probab. Appl. **57**(2), 243–264 (2013)
6. Cordero-Erausquin, D.: Some applications of mass transport to Gaussian-type inequalities. Arch. Ration. Mech. Anal. **161**(3), 257–269 (2002)
7. Hargé, G.: A particular case of correlation inequality for the Gaussian measure. Ann. Probab. **27**(4), 1939–1951 (1999)
8. Milman, E.: Spectral estimates, contractions and hypercontractivity. J. Spectr. Theory **8**(2), 669–714 (2018)
9. Mikulincer, D.: A CLT in Stein's distance for generalized Wishart matrices and higher-order tensors. Int. Math. Res. Not. IMRN **10**, 7839–7872 (2022)
10. Colombo, M., Figalli, A., Jhaveri, Y.: Lipschitz changes of variables between perturbations of log-concave measures. Ann. Sc. Norm. Super. Pisa Cl. Sci. **17**(4), 1491–1519 (2017)
11. Kolesnikov, A.V.: Mass transportation and contractions (2011). arXiv preprint arXiv:1103.1479
12. Tanana, Anastasiya: Comparison of transport map generated by heat flow interpolation and the optimal transport Brenier map. Commun. Contemp. Math. **23**(6), 2050025 (2021)
13. Lavenant, H., Santambrogio, F.: The flow map of the Fokker-Planck equation does not provide optimal transport. Appl. Math. Lett. **133**, 108225 (2022)
14. Mikulincer, D., Shenfeld, Y.: On the Lipschitz properties of transportation along heat flows. GAFA Seminar Notes, To appear
15. Neeman, J.: Lipschitz changes of variables via heat flow (2022). arXiv preprint arXiv:2201.03403
16. Klartag, B., Putterman, E.: Spectral monotonicity under gaussian convolution. Ann. Fac. Sci. Toulouse Math, To appear
17. Shenfeld, Y.: Exact renormalization groups and transportation of measures (2022). arXiv preprint arXiv:2205.01642
18. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M.F., Lin, H. (eds.) Advances in neural information processing systems, vol. 33, pp. 6840–6851. Curran Associates Inc (2020)
19. De Bortoli, V., Thornton, J., Heng, J., Doucet, A.: Diffusion Schrödinger bridge with applications to score-based generative modeling. In: Beygelzimer, A., Dauphin, Y., Liang, P., Wortman Vaughan, J. (eds.) Advances in Neural Information Processing Systems (2021)
20. Beck, T., Jerison, D.: Jerison, David: The Friedland-Hayman inequality and Caffarelli's contraction theorem. J. Math. Phys. **62**(10), 101504 (2021)
21. Figalli, Alessio: Regularity of optimal transport maps [after Ma-Trudinger-Wang and Loeper]. Number 332, pages Exp. No. 1009, ix, 341–368. Séminaire Bourbaki. Volume 2008/2009. Exposés 997–1011 (2010)
22. Bakry, D., Gentil, I., Ledoux, M.: Analysis and geometry of Markov diffusion operators. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348. Springer, Cham (2014)

23. Hsu, E.P.: Stochastic analysis on manifolds. Graduate Studies in Mathematics, vol. 38. American Mathematical Society, Providence (2002)
24. Otto, F., Villani, C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. **173**(2), 361–400 (2000)
25. Villani, C.: Optimal transport: old and new, vol. 338. Springer, Berlin (2008)
26. David Elworthy, K., Li, X.-M.: Formulae for the derivatives of heat semigroups. J. Funct. Anal. **125**(1), 252–286 (1994)
27. Thompson, J.: Approximation of Riemannian measures by Stein's method (2020). arXiv preprint arXiv:2001.09910
28. Cheng, L.-J., Wang, F.-Y., Thalmaier, A.: Some inequalities on Riemannian manifolds linking entropy, Fisher information, Stein discrepancy and Wasserstein distance (2021). arXiv preprint arXiv:2108.12755
29. Bismut, J.-M.: Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions. Z. Wahrsch. Verw. Gebiete **56**(4), 469–505 (1981)
30. Milman, E.: Reverse Hölder inequalities for log-Lipschitz functions. Pure Appl. Funct. Anal. **8**(1), 297–310 (2023)
31. Karatzas, I., Shreve, S.E.: Brownian motion and stochastic calculus, volume 113 of Graduate Texts in Mathematics, 2nd edn. Springer, New York (1991)
32. Bakry, D., Ledoux, M.: Lévy-Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator. Invent. Math. **123**(2), 259–281 (1996)
33. Bobkov, S.G.: An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space. Ann. Probab. **25**(1), 206–214 (1997)
34. Bardet, J.-B., Gozlan, N., Malrieu, F., Zitt, P.-A.: Functional inequalities for Gaussian convolutions of compactly supported measures: explicit bounds and dimension dependence. Bernoulli **24**(1), 333–353 (2018)
35. Barthe, F., Kolesnikov, A.V.: Mass transport and variants of the logarithmic Sobolev inequality. J. Geom. Anal. **18**(4), 921–979 (2008)
36. Aida, S., Shigekawa, I.: Logarithmic Sobolev inequalities and spectral gaps: perturbation theory. J. Funct. Anal. **126**(2), 448–475 (1994)
37. Ivanisvili, P., Russell, R.: Exponential integrability for log-concave measures (2020). arXiv preprint arXiv:2004.09704
38. Cianchi, A., Musil, V., Pick, L.: Moser inequalities in Gauss space. Math. Ann. **377**(3–4), 1265–1312 (2020)
39. Colombo, M., Fathi, M.: Bounds on optimal transport maps onto log-concave measures. J. Differ. Equ. **271**, 1007–1022 (2021)
40. Gorham, J., Duncan, A.B., Vollmer, S.J., Mackey, L.: Measuring sample quality with diffusions. Ann. Appl. Probab. **29**(5), 2884–2928 (2019)
41. Fang, X., Shao, Q.-M., Lihu, X.: Multivariate approximations in Wasserstein distance by Stein's method and Bismut's formula. Probab. Theory Relat. Fields **174**(3–4), 945–979 (2019)
42. Fang, X., Shao, Q.-M., Lihu, X.: Correction to: multivariate approximations in Wasserstein distance by Stein's method and Bismut's formula. Probab. Theory Relat. Fields **175**(3–4), 1177–1181 (2019)