Caffarelli's Contraction Theorem of Gaussian measures

Gaussian measures satisfy a number of functional and measure-theoretic inequalities, such as Poincaré, logarithmic Sobolev, transportation cost, isoperimetric or concentration inequalities, as illustrated in e.g. [1, 2, 3, 4]. These Gaussian inequalities remarkably are independent from the dimension of the underlying state space, allowing in particular for infinite-dimensional extensions.

As a prototype example, the Poincaré inequality [1] for the standard Gaussian probability measure $d\gamma_n(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}|x|^2} dx$ on $\mathbb{R}^n$ expresses that for every smooth function $f : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{Var}_{\gamma_n}(f) \le \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n.$$

A natural question is whether other (non-Gaussian) measures also satisfy such a Poincaré inequality (up to a constant), and how to produce classes of such measures. A simple observation in this regard is that if ν is a Borel probability measure on $\mathbb{R}^n$ such that there is a Lipschitz transport map $T : \mathbb{R}^n \to \mathbb{R}^n$ that pushes forward γ_n on ν (in the sense that for every positive measurable function $\psi : \mathbb{R}^n \to \mathbb{R}$,

$$\int_{\mathbb{R}^n} \psi \, d\nu = \int_{\mathbb{R}^n} \psi \circ T \, d\gamma_n), \tag{1}$$

then

$$\mathrm{Var}_\nu(f) \le \|T\|_{\mathrm{Lip}}^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\nu$$

Drafted by Max Fathi **v1** November 2025

for every smooth function $f : \mathbb{R}^n \to \mathbb{R}$. For a proof,

$$\begin{aligned}
\mathrm{Var}_\nu(f) \,=\, \mathrm{Var}_{\gamma_n}(f \circ T) \,&\leq\, \int_{\mathbb{R}^n} \big|\nabla (f \circ T)\big|^2 \, d\gamma_n \\
&= \int_{\mathbb{R}^n} \big|\nabla T \big((\nabla f) \circ T \big)\big|^2 \, d\gamma_n \\
&\leq \, \|T\|_{\mathrm{Lip}}^2 \int_{\mathbb{R}^n} \big| (\nabla f) \circ T\big|^2 \, d\gamma_n \,=\, \|T\|_{\mathrm{Lip}}^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\nu.
\end{aligned}$$

Clearly, the scheme similarly applies to the other functional and isoperimetric inequalities satisfied by γ_n. For example, the (Gaussian) isoperimetric profile I_ν of ν (cf. [3]) satisfies $\mathrm{I}_\nu \geq \frac{1}{\|T\|_{\mathrm{Lip}}} \mathrm{I}_{\gamma_n}$.

A remarkable theorem by L. Caffarelli in optimal transport [8] emphasizes that uniformly log-concave probability measures on $\mathbb{R}^n$ can be realized as the image of the standard Gaussian measure by a globally Lipschitz map. Precisely, a non-degenerate (not supported on some hyperplane) probability measure $d\nu = e^{-V} dx$ on the Borel sets of $\mathbb{R}^n$, for some measurable $V : \mathbb{R}^n \to \mathbb{R}$, is said to be κ-uniformly log-concave for some $\kappa > 0$ if $V - \kappa \frac{|x|^2}{2}$ is convex. Then, for such a ν, there is a Lipschitz map $T : \mathbb{R}^n \to \mathbb{R}^n$ that pushes γ_n on ν and moreover $\|T\|_{\mathrm{Lip}} \leq \frac{1}{\sqrt{\kappa}}$.

As such, functional and measure-theoretic inequalities for ν immediately follow from the corresponding ones for γ_n, with in addition dimension-free constants, inequalities which were mostly established before by specific tools (see [24, 5]). Other applications include correlation inequalities [19] and weighted spectral inequalities for domains [6]. Note that in each of these examples, it is of course necessary to start from a known functional inequality for the Gaussian measure. It is often possible to use optimal transport to prove Gaussian functional inequalities [24], but such proofs do not need any quantitative regularity estimate.

The post is devoted to a sketch of Caffarelli's Contraction Theorem and its proof, relying on the Brenier optimal transport map recalled in the first section. The last paragraph describes a few extensions to other measures than the Gaussian and to other target measures.

Table of contents

1 Brenier's optimal transport map

Brenier's theorem [7] on existence and structure of optimal transport maps for the quadratic cost is summarized in the following statement.

Theorem 1 (Brenier's optimal transport map)**.** *Let μ and ν be two probability measures on $\mathbb{R}^n$ with a finite second moment, and assume that μ is absolutely continuous with respect to the Lebesgue measure. There exists a convex function φ such that $\nabla\varphi$ is a map that pushes forward μ onto ν.*

This map arises as a solution to the optimal transport problem with quadratic cost. Background about optimal transport may be found in [24, 17] among others.

The following is the main Contraction Theorem by L. Caffarelli [8].

Theorem 2 (Caffarelli's Contraction Theorem)**.** *Let γ_n be the standard Gaussian measure on $\mathbb{R}^n$, and let $d\nu = e^{-V}dx$ be a probability measure with V κ-uniformly convex with $\kappa > 0$. Then the Brenier optimal transport map sending γ_n onto ν is $\frac{1}{\sqrt{\kappa}}$-Lipschitz.*

The dependence on κ is optimal, as can be shown by testing the bound for non-standard Gaussian targets. The dimension-free nature of the bound is quite remarkable, and crucial for many applications.

2 Proof of Caffarelli's Contraction Theorem via the maximum principle

Caffarelli's proof of his contraction theorem relies on the maximum principle for fully non-linear PDE. A formal sketch of proof, disregarding some regularity issues, is presented in this section, referring to [8, 24, 17]... for complete details.

From the pushforward definition (1), the Brenier map $\nabla\varphi$ from Theorem 1 solves the Monge-Ampère PDE

$$\frac{1}{(2\pi)^{\frac{n}{2}}}\, e^{-|x|^2/2} \,=\, e^{-V(\nabla\varphi)} \det(\nabla^2\varphi).$$

It will be assumed in the following that φ is smooth (say C^4). Taking the logarithm of this equation and differentiating twice in a direction $e \in \mathbb{R}^n$, it holds true that

$$1 \,=\, \nabla\partial_e^2\varphi(x)\cdot\nabla V\big(\nabla\varphi(x)\big) + \nabla^2 V\big(\nabla\varphi(x)\big)\partial_e\nabla\varphi(x)\cdot\partial_e\nabla\varphi(x) - \partial_e^2\big[\log\det\nabla^2\varphi\big](x).$$

At a point x_0 and a direction e_0 that maximizes $(x,e) \to \nabla_e^2 \varphi(x)$, so that $\|\nabla\varphi\|_{\mathrm{Lip}} = \partial_{e_0}^2 \varphi(x_0)$, the preceding relation yields

$$\nabla \partial_{e_0}^2 \varphi(x_0) \,=\, 0, \quad \nabla^2 \partial_{e_0}^2 \varphi \,\leq\, 0. \tag{2}$$

Using concavity of log det, the inequality on the Hessian in (2) implies that

$$\partial_{e_0}^2 [\log\det \nabla^2 \varphi](x_0) \,\leq\, 0,$$

so that

$$1 \,\geq\, \nabla^2 V\big(\nabla\varphi(x)\big)\partial_e \nabla\varphi(x) \cdot \partial_e \nabla\varphi(x) \,\geq\, \kappa \left(\partial_{e_0}^2 \varphi(x_0)\right)^2$$

which is the desired bound.

There are of course at least two un-rigorous elements in the above argument: φ is not necessarily so smooth, and it is not justified that $(x,e) \to \nabla_e^2 \varphi(x)$ attains its global maximum at some point. To get around these two issues, the proof in [8] assumes that ν has compact support (which can be ensured by approximation) and considers the second-order finite difference

$$\frac{1}{t^2}\left(\varphi(x+te) + \varphi(x-te) - 2\varphi(x)\right)$$

to estimate its maximum via the maximum principle strategy. To do so, only the second derivatives of φ have to be considered, which is fine by Caffarelli's regularity theory for the Monge-Ampère PDE (cf. [8]). Moreover, whenever ν has compact support, $\nabla\varphi$ is bounded and the second-order finite difference does reach a maximum at a point.

An alternative proof, via entropic regularization of optimal transport maps, was obtained in [15], later simplified in [10].

3 Extensions to other measures

Caffarelli's original proof allows for more general starting distributions than standard Gaussians, requiring log-concavity instead. Uniform log-concavity of the target distribution is crucial in the maximum principle proof. Some extensions to slightly non-log-concave targets have been considered in [12], but the bounds do not behave so well in high dimension in practice. Lipschitz estimates for heavy-tailed distributions have been studied in [9].

Other quantitative regularity estimates have been considered in the literature, including Sobolev norm bounds [21] and non-linear bounds on the modulus of continuity [21, 18]. In another direction, growth estimates for optimal transport maps have been studied in [9] using the maximum principle, and in [11, 13] using concentration inequalities.

So far, there are no good analogues of Caffarelli's Contraction Theorem for optimal transport maps on Riemannian manifolds. The references [23, 14] point out some open problems and counterexamples.

It should be observed that the afore-mentioned applications do not use the fact that the transport map is optimal, only its regularity properties. It is therefore plausible that some non-optimal transport maps with good regularity properties may be constructed. A particularly successful approach is the Kim-Milman construction of such maps via heat flow [20], which, beyond recovering Caffarelli's theorem, may also be successfully implemented in situations where optimal transport maps are not currently so well understood quantitatively, such as for manifolds satisfying certain curvature assumptions [16]. In another direction, stochastic constructions of transport maps from the Wiener measure onto finite-dimensional spaces may also enjoy good regularity properties [22].

References

[1] The Gaussian Poincaré inequality. *The Gaussian Blog.*

[2] Logarithmic Sobolev and transportation inequalities. *The Gaussian Blog.*

[3] The Gaussian isoperimetric inequality. *The Gaussian Blog.*

[4] Gaussian concentration inequalities. *The Gaussian Blog.*

[5] D. Bakry, I. Gentil, M. Ledoux. *Analysis and geometry of Markov diffusion operators.* Grundlehren der mathematischen Wissenschaften 348. Springer (2014).

[6] T. Beck, D. Jerison. The Friedland-Hayman inequality and Caffarelli's contraction theorem. *J. Math. Phys.* 62, Paper No. 101504, 11pp. (2021).

[7] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Commun. Pure Appl. Math.* 44, 375–417 (1991).

[8] L. Caffarelli. Monotonicity properties of optimal transportation and the FKG and related inequalities. *Commun. Math. Phys.* 214, 547–563 (2000); erratum ibid. 225, 449–450 (2002).

[9] G. Carlier, A. Figalli, F. Santambrogio. On optimal transport maps between $\frac{1}{d}$-concave densities (2024).

[10] S. Chewi, A.-A. Pooladian. An entropic generalization of Caffarelli's contraction theorem via covariance inequalities. *C. R. Math. Acad. Sci. Paris* 361, 1471–1482 (2023).

[11] M. Colombo, M. Fathi. Bounds on optimal transport maps onto log-concave measures. *J. Differential Equations* 271, 1007–1022 (2021).

[12] M. Colombo, A. Figalli, Y. Jhaveri. Lipschitz changes of variables between perturbations of log-concave measures. *Ann. Sc. Norm. Super. Pisa* 17, 1491–1519 (2017).

[13] M. Fathi. Growth estimates on optimal transport maps via concentration inequalities (2024).

[14] M. Fathi, M. Fradelizi, N. Gozlan, S. Zugmeyer. Some obstructions to contraction theorems on the half-sphere (2024).

[15] M. Fathi, N. Gozlan, M. Prod'homme. A proof of the Caffarelli contraction theorem via entropic regularization. *Calc. Var. Partial Differ. Equ.* 59, Paper No. 96, 18 pp. (2020).

[16] M. Fathi, D. Mikulincer, Y. Shenfeld. Transportation onto log-Lipschitz perturbations. *Calc. Var. Partial Differential Equations* 63, Paper No. 61, 25 pp. (2024).

[17] A. Figalli, F. Glaudo. *An invitation to optimal transport, Wasserstein distances, and gradient flows.* 2nd edition. EMS Textbooks in Mathematics. European Mathematical Society (2023).

[18] N. Gozlan, M. Sylvestre. Global regularity estimates for optimal transport via entropic regularisation (2025).

[19] G. Hargé. A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces. *Probab. Theory Related Fields* 130, 415–440 (2004).

[20] Y.-H. Kim, E. Milman. A generalization of Caffarelli's contraction theorem via (reverse) heat flow. *Math. Ann.* 354, 827–862 (2012).

[21] A. V. Kolesnikov. On Sobolev regularity of mass transport and transportation inequalities. *Theory Probab. Appl.* 57, 243–264 (2013).

[22] D. Mikulincer, Y. Shenfeld. The Brownian transport map. *Probab. Theory Related Fields* 190, 379–444 (2024).

[23] E. Milman. Spectral estimates, contractions and hypercontractivity. *J. Spectr. Theory* 8, 669–714 (2018).

[24] C. Villani. *Topics in optimal transportation.* Graduate Studies in Mathematics 58. American Mathematical Society (2003).