DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **36**, Number **12**, December **2016**

doi:10.3934/dcds.2016097

pp. **X–XX**

QUANTITATIVE LOGARITHMIC SOBOLEV INEQUALITIES AND STABILITY ESTIMATES

Max Fathi

Université Pierre et Marie Curie
Paris, France

Emanuel Indrei

Carnegie Mellon University
Pittsburgh, USA

Michel Ledoux

University of Toulouse and Institut Universitaire de France
Toulouse, France

(Communicated by Alessio Figalli)

Abstract. We establish an improved form of the classical logarithmic Sobolev inequality for the Gaussian measure restricted to probability densities which satisfy a Poincaré inequality. The result implies a lower bound on the deficit in terms of the quadratic Kantorovich-Wasserstein distance. We similarly investigate the deficit in the Talagrand quadratic transportation cost inequality this time by means of an L^1-Kantorovich-Wasserstein distance, optimal for product measures, and deduce a lower bound on the deficit in the logarithmic Sobolev inequality in terms of this metric. Applications are given in the context of the Bakry-Émery theory and the coherent state transform. The proofs combine tools from semigroup and heat kernel theory and optimal mass transportation.

1. **Introduction.** The classical logarithmic Sobolev inequality of L. Gross [27] for the standard Gaussian measure

$$d\gamma(x) = d\gamma^n(x) = e^{-|x|^2/2} \frac{dx}{(2\pi)^{n/2}}$$

on the Borel sets of $\mathbb{R}^n$ (cf. e.g. [42, 43, 6]) states that if $d\nu = f d\gamma$ is a probability measure with density f with respect to γ,

$$\mathrm{H}(\nu) \leq \frac{1}{2} \mathrm{I}(\nu) \tag{1.1}$$

where

$$\mathrm{H}(\nu) = \mathrm{H}\big(\nu \,|\, \gamma\big) = \int_{\mathbb{R}^n} f \log f \, d\gamma$$

2010 *Mathematics Subject Classification.* Primary: 58J35, 60J60.

Key words and phrases. Logarithmic Sobolev inequalities, deficit estimates, transportation inequalities, optimal transport theory, semigroup theory.

The second author is supported by US NSF Grants OISE-0967140 (PIRE), DMS-0405343, and DMS-0635983 administered by the Center for Nonlinear Analysis in Pittsburgh, PA.

is the relative entropy of ν with respect to γ and

$$\mathrm{I}(\nu) \,=\, \mathrm{I}(\nu\,|\,\gamma) \,=\, \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f}\, d\gamma$$

is the Fisher information of ν with respect to γ.

Logarithmic Sobolev inequalities (LSI) are a useful tool in analysis and probability in the study of convergence to equilibrium, large deviations, and measure concentration. They are also equivalent to hypercontractivity for their associated semigroup (cf. [42, 43, 6]). To ensure that the various terms of the LSI are well-defined, some smoothness and positivity properties of the density f of ν have to be considered. These may be handled by approximation and regularization (see e.g. [6]). When dealing with entropy $\mathrm{H}(\nu)$ and Fisher information $\mathrm{I}(\nu)$ (and below the LSI deficit $\delta_{\mathrm{LSI}}(\nu)$ (1.3)), it will be usually implicitly understood that they are well-defined (and finite) for suitable density functions f.

The constant $1/2$ in the Gaussian LSI (1.1) is known to be optimal, and it was first shown in [13] that the cases of equality are exactly the measures of the form

$$d\gamma_b(x) \,=\, e^{b\cdot x - \frac{|b|^2}{2}}\, d\gamma(x), \quad b \in \mathbb{R}^n. \tag{1.2}$$

In other words, the extremal densities f are exponential functions. (Note that b is the barycenter of γ_b, so that in particular the only centered extremal measure is γ itself.)

However, the study of the *logarithmic Sobolev deficit*

$$\delta_{\mathrm{LSI}}(\nu) \,=\, \frac{1}{2}\,\mathrm{I}(\nu) - \mathrm{H}(\nu) \tag{1.3}$$

to quantify proximity with the extremal measures is still largely open in spite of recent developments for classical Sobolev and related isoperimetric inequalities. In the broader context of stability results for functional inequalities, when looking at a functional inequality with known optimal constants and optimizers, a natural question is indeed whether functions that are close to achieving the optimum are close to some optimizer. The task is to bound from below the deficit by some functional that measures how far we are from some optimizer (typically, a distance). Examples of such results are the recent quantitative stability estimates for Sobolev [15, 23], Brunn-Minkowski [21, 20], and isoperimetric inequalities [25, 22, 19, 30].

The first main result of this note is to propose a (strict) strengthening of the Gaussian LSI (1.1) within a subclass of probability measures ν which in turn produces a lower bound on the deficit $\delta_{\mathrm{LSI}}(\nu)$. Denote by $\mathcal{P}(\lambda)$ the class of probability measures ν on the Borel sets of $\mathbb{R}^n$ satisfying a Poincaré inequality with constant $\lambda > 0$ in the sense that for every smooth $g : \mathbb{R}^n \to \mathbb{R}$ such that $\int_{\mathbb{R}^n} g d\nu = 0$,

$$\lambda \int_{\mathbb{R}^n} g^2 d\nu \,\le\, \int_{\mathbb{R}^n} |\nabla g|^2 d\nu. \tag{1.4}$$

Note that under such a Poincaré inequality, the measure ν necessarily has a second moment.

Theorem 1. *For any centered ($\int_{\mathbb{R}^n} x d\nu = 0$) probability measure $d\nu = f d\gamma$ in the class $\mathcal{P}(\lambda)$,*

$$\mathrm{H}(\nu) \,\le\, \frac{c(\lambda)}{2}\,\mathrm{I}(\nu),$$

where

$$c(\lambda) \,=\, \frac{1-\lambda+\lambda\log\lambda}{(1-\lambda)^2} \,<\, 1 \qquad \big(c(1) \,=\, \tfrac{1}{2}\big).$$

The constant is sharp, as can be seen when taking ν with density $f(x) = \sqrt{\lambda}\, e^{(1-\lambda)x^2/2}$, $\lambda > 0$, on the line. Of course, since the constant $1/2$ in the Gaussian LSI is optimal, such a strengthening can only be expected to hold on a subset of probability measures.

In dimension $n = 1$, the class of probability measures satisfying a Poincaré inequality (1.4) has been completely characterized. A probability measure ν with density p with respect to the Lebesgue measure and median m satisfies a Poincaré inequality if and only if the following holds (see [8, 6]):

$$A^+ \,=\, \sup_{x\geq m} \nu\big([x,+\infty[\big) \int_m^x \frac{1}{p(t)}\, dt \,<\, \infty,$$

$$A_- \,=\, \sup_{x\leq m} \nu\big(]-\infty,x]\big) \int_x^m \frac{1}{p(t)}\, dt \,<\, \infty.$$

Moreover, the optimal Poincaré constant $\lambda_{\rm opt}$ for ν satisfies

$$\frac{1}{2}\, \max(A^+,A^-) \,\leq\, \lambda_{\rm opt} \,\leq\, 4\max(A^+,A^-).$$

In higher dimension, there is no such simple characterization, but fairly general sufficient conditions are available. For example, if ν has a density of the form e^{-V} with respect to the Lebesgue measure, a sufficient condition is the existence of $a \in]0,1[$ such that $a|\nabla V|^2 - \Delta V$ is bounded from below by some positive constant outside of some ball (see [2]). A more classical condition is the Bakry-Émery criterion

$$\mathrm{Hess}(V) \geq \eta\, \mathrm{Id} \qquad \text{for some } \ \eta > 0 \tag{1.5}$$

on the potential V ([3, 42, 6]) ensuring a Poincaré inequality with constant $\lambda = \eta$.

As an equivalent formulation of Theorem 1, for ν centered in $\mathcal{P}(\lambda)$,

$$\delta_{\rm LSI}(\nu) \,\geq\, c_1(\lambda)\, \mathrm{I}(\nu) \tag{1.6}$$

where $c_1(\lambda) = \frac{1}{2}\,(1 - c(\lambda))$. The non-centered version of (1.6), and thus of Theorem 1, reads as follows.

Corollary 2. *For any probability measure $d\nu = f d\gamma$ in the class $\mathcal{P}(\lambda)$ with barycenter $b = b(\nu)$,*

$$\delta_{\rm LSI}(\nu) \,\geq\, c_1(\lambda) \int_{\mathbb{R}^n} \big|\nabla(\log f) - b\big|^2 d\nu.$$

Corollary 2 follows by a rescaling argument involving the barycenter. For $d\nu = f d\gamma$ with mean b, define

$$d\nu_b(x) \,=\, f(x+b) e^{-\left(b\cdot x + \frac{|b|^2}{2}\right)}\, d\gamma(x). \tag{1.7}$$

The probability measure ν_b has mean 0 and, as is easily checked, satisfies $\mathrm{H}(\nu_b) = \mathrm{H}(\nu) - \frac{|b|^2}{2}$ and $\mathrm{I}(\nu_b) = \mathrm{I}(\nu) - |b|^2$, so that $\delta_{\rm LSI}(\nu_b) = \delta_{\rm LSI}(\nu)$. The conclusion then easily follows.

Theorem 1 improves upon the recent [29] where stronger conditions on the Hessian of the density f are considered (in particular parts of the class $\mathcal{P}(\lambda)$), with weaker dependence of the constant. The work [29] actually investigates how far an

admissible density is from saturating the logarithmic Sobolev inequality as measured with Wasserstein distance, providing a control of the deficit $\delta_{\mathrm{LSI}}(\nu)$ in the logarithmic Sobolev inequality by the (quadratic) Kantorovich-Wasserstein distance $\mathrm{W}_2(\nu,\gamma)$. Within the class $\mathcal{P}(\lambda)$, this is easily achieved via Theorem 1 together with the Talagrand quadratic transportation cost inequality [41] (cf. [42, 43, 6])

$$\mathrm{W}_2(\nu,\gamma)^2 \,\leq\, 2\,\mathrm{H}(\nu) \tag{1.8}$$

holding for all probability measures ν (absolutely continuous with respect to γ). Recall that the Kantorovich-Wasserstein distance $\mathrm{W}_2(\nu,\mu)$ between two probability measures ν and μ is given by

$$\mathrm{W}_2(\nu,\mu) \,=\, \inf\left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |x-y|^2 d\pi(x,y)\right)^{1/2}$$

where the infimum is over all couplings π of probability measures on $\mathbb{R}^n\times\mathbb{R}^n$ with respective marginals ν and μ. Note that if $\nu\in\mathcal{P}(\lambda)$, it has necessarily a second moment so that the Kantorovich-Wasserstein distance $\mathrm{W}_2(\nu,\gamma)$ is finite.

Corollary 3. *For any centered probability measure $d\nu = f d\gamma$ in the class $\mathcal{P}(\lambda)$,*

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, c_2(\lambda)\,\mathrm{W}_2(\nu,\gamma)^2,$$

where $c_2(\lambda) = \frac{1}{2}\big(\frac{1}{c(\lambda)} - 1\big)$ and $c(\lambda)$ is as in Theorem 1.

This corollary may be compared to the Otto-Villani HWI inequality [37] (cf. [42, 43, 6]), valid for any probability ν,

$$\mathrm{H}(\nu) \,\leq\, \mathrm{W}_2(\nu,\gamma)\,\sqrt{\mathrm{I}(\nu)} - \frac{1}{2}\,\mathrm{W}_2(\nu,\gamma)^2. \tag{1.9}$$

It should be mentioned that one cannot expect

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, c\,\mathrm{W}_2(\nu,\gamma)^2$$

to hold for some $c>0$ and all probability measures ν. Indeed, such an inequality combined with the HWI inequality would then imply the logarithmic Sobolev inequality $\mathrm{H}(\nu) \leq \frac{1+c}{2+4c}\,\mathrm{I}(\nu)$ (by applying Young's inequality to the term $\mathrm{W}\sqrt{\mathrm{I}}$ in the HWI inequality) with therefore a constant strictly better than the optimal $1/2$. A complete stability result for the Gaussian LSI therefore requires a distance weaker than W_2.

In this direction, Theorem 1 may also be used to provide a lower bound on the deficit δ_{LSI} in terms of the total variation distance (which is the L^1 distance between densities). Indeed, as the standard Gaussian measure γ satisfies a $(1,1)$-Poincaré inequality (cf. e.g. [31])

$$\int_{\mathbb{R}^n} |g|\,d\gamma \,\leq\, 2\int_{\mathbb{R}^n} |\nabla g|\,d\gamma \tag{1.10}$$

for every smooth $g:\mathbb{R}^n\to\mathbb{R}$ with mean zero, if $d\nu = f d\gamma$,

$$\int_{\mathbb{R}^n} |f-1|\,d\gamma \,\leq\, 2\int_{\mathbb{R}^n} |\nabla f|\,d\gamma \,\leq\, 2\,\sqrt{\mathrm{I}(\nu)}$$

by the Cauchy-Schwarz inequality. We then only state the consequence of (1.6) in the centered case.

Corollary 4. *For any centered probability measure $d\nu = f d\gamma$ in the class $\mathcal{P}(\lambda)$,*

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, \frac{c_1(\lambda)}{4}\left(\int_{\mathbb{R}^n} |f-1|\,d\gamma\right)^2 \,=\, \frac{c_1(\lambda)}{4}\,\|\nu-\gamma\|_{\mathrm{TV}}^2.$$

While Corollaries 3 and 4 are strictly weaker than Theorem 1, they have the advantage of providing a lower bound on the deficit in the Gaussian LSI in terms of a metric.

A one-dimensional stability result of the same kind as Corollary 3 is proven in Corollary 4.4 of [9], however with a worse constant of proportionality. The main assumption is uniform log-concavity of ν (i.e. (1.5)) which is used to apply a $(1,1)$-Poincaré inequality. As far as we know, the argument of [9] does not extend to higher dimensions. Nevertheless, the one-dimensional result may be combined with a tensorization argument to cover the case of n-dimensional random vectors with uniformly log-concave distributions whose one-dimensional projections form a martingale. Such an assumption is not the same as simply assuming that the mean of ν is zero. More generally,

$$\delta_{\mathrm{LSI}}(\nu) \;\geq\; c\,\mathrm{W}_2(\bar{\nu},\gamma)^2,$$

where $\bar{\nu}$ is the law of a random vector $\bar{X}$ obtained by modifying a random vector X with law ν in such a way that its one-dimensional marginals $X_1,\dots,X_n$ form a martingale [9]. For unconditional random variables, this is the same as assuming the mean to be zero, but in general it does not seem like $\mathrm{W}_2(\bar{\nu},\gamma)$ and $\mathrm{W}_2(\nu,\gamma)$ can be easily compared. The contribution [9] also contains deficit estimates for general ν, but with lower bounds that are either not a power of a distance, are dimension-dependent, or involve $\bar{\nu}$. For example, there is a universal constant $c>0$ such that for all smooth probability measures ν on $\mathbb{R}^n$,

$$\delta_{\mathrm{LSI}}(\nu) \;\geq\; c\,\frac{\mathcal{T}(\bar{\nu},\gamma)^2}{\mathrm{H}(\bar{\nu})} \tag{1.11}$$

where $\bar{\nu}$ is the previously discussed martingale rearrangement of ν and $\mathcal{T}$ is a transportation cost associated to the function $t\mapsto t-\log(1+t)$.

The second main result of this note investigates the deficit in the Talagrand quadratic transportation cost inequality (1.8). A result of Otto and Villani [37] states that a measure satisfying a logarithmic Sobolev inequality automatically satisfies a Talagrand-type inequality (1.8). It is easy to see, using the HWI inequality (1.9), that the cases of equality for Talagrand's inequality are exactly the same as for the Gaussian LSI. Therefore, it is natural to investigate lower bounds on the *Talagrand deficit*

$$\delta_{\mathrm{Tal}}(\nu) \;=\; 2\,\mathrm{H}(\nu) - \mathrm{W}_2(\nu,\gamma)^2.$$

In dimension one, it was shown by Barthe and Kolesnikov [7] that the deficit $\delta_{\mathrm{Tal}}(\nu)$ satisfies

$$\delta_{\mathrm{Tal}}(\nu) \;\geq\; c\,\inf_{\pi}\int_{\mathbb{R}^n}\varphi\big(|x-y|\big)\pi(dx,dy),$$

where the infimum is over couplings π of ν and γ, and $\varphi(t)=t-\log(1+t)$. Note that the right-hand side in this inequality is an optimal transport cost, with a cost that is quadratic-then-linear in the distance. This inequality immediately yields the weaker version

$$\delta_{\mathrm{Tal}}(\nu) \;\geq\; c\,\min\left(\mathrm{W}_1(\nu,\gamma)^2,\mathrm{W}_1(\nu,\gamma)\right),$$

just by applying the elementary inequality $t-\log(1+t)\geq\frac{1}{6}\min(t^2,t)$, where W_1 is the L^1-Kantorovitch-Wasserstein distance (with ℓ^2-cost function on $\mathbb{R}^n$) between the one-dimensional measures ν and γ.

We establish here the following multi-dimensional version of the Barthe-Kolesnikov result. Let

$$\mathrm{W}_{1,1}(\nu,\mu) \,=\, \inf \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \sum_{i=1}^n |x_i - y_i| d\pi(x,y)$$

be the L^1-Kantorovich-Wasserstein distance with ℓ^1-cost function on $\mathbb{R}^n$ where the infimum is over couplings π of ν and μ.

Theorem 5. *There is a numerical constant $c > 0$ such that for any centered probability measure $d\nu = f d\gamma$ on $\mathbb{R}^n$ with finite second moments and $f > 0$ locally bounded,*

$$\delta_{\mathrm{Tal}}(\nu) \,\geq\, c \,\min\left(\frac{\mathrm{W}_{1,1}(\nu,\gamma)^2}{n}\,,\, \frac{\mathrm{W}_{1,1}(\nu,\gamma)}{\sqrt{n}}\right).$$

One feature of this result is that it is valid for general measures. Moreover, the lower bound is expressed in terms of a metric on the space of probability measures on $\mathbb{R}^n$ and the exponent is independent of the dimension. In general, the deficit in Theorem 5 is only optimal for small perturbations of the Gaussian. For an n-dimensional product measure $\nu^n = \nu^{\otimes n}$, $\delta_{\mathrm{Tal}}(\nu^n) = n\delta_{\mathrm{Tal}}(\nu)$ grows linearly in n. This is also the behavior of

$$\frac{\mathrm{W}_{1,1}(\nu^n,\gamma^n)^2}{n} \,=\, n\,\mathrm{W}_{1,1}(\nu,\gamma^1)^2.$$

When $n >> \mathrm{W}_{1,1}(\nu,\gamma^1)^{-2}$, the expected growth is lost. Nevertheless, for product measures whose one-dimensional marginals are close enough to $\gamma = \gamma^1$ (i.e. such that $\mathrm{W}_{1,1}(\nu,\gamma^1)^2 \leq \frac{c}{n}$), Theorem 5 yields the correct order of magnitude in the dimension.

Theorem 5 furthermore yields a new proof of the equality case for the Gaussian LSI. Indeed, by the HWI inequality,

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, \frac{1}{2}\Big(\sqrt{\mathrm{I}(\nu)} - W_2(\nu,\gamma)\Big)^2.$$

Therefore, if ν is such that $\delta_{\mathrm{LSI}}(\nu) = 0$, then $\mathrm{I}(\nu) = \mathrm{W}_2(\nu,\gamma)^2$. By the conjunction of the Talagrand (1.8) and LSI (1.1) inequalities,

$$\mathrm{W}_2(\nu,\gamma)^2 \,\leq\, 2\,\mathrm{H}(\nu) \,\leq\, \mathrm{I}(\nu),$$

so that there is also equality in Talagrand's inequality and thus $\delta_{\mathrm{Tal}}(\nu) = 0$. Therefore, Theorem 5 implies that the only centered measure satisfying $\delta_{\mathrm{LSI}}(\nu) = 0$ is precisely γ. The non-centered case follows as for Corollary 2.

The preceding argument may be quantified in terms of the $\mathrm{W}_{1,1}$ metric and yields a general stability result for LSI. Recall ν_b from (1.7).

Corollary 6. *There is a numerical constant $c > 0$ such that for any probability measure $d\nu = f d\gamma$ on $\mathbb{R}^n$ with $f > 0$ locally bounded and positive entropy, and with barycenter $b = b(\nu)$,*

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, \frac{c}{\mathrm{H}(\nu)} \,\min\left(\frac{\mathrm{W}_{1,1}(\nu_b,\gamma)^4}{n^2}\,,\, \frac{\mathrm{W}_{1,1}(\nu_b,\gamma)^2}{n}\right).$$

Indeed, as above, by the HWI (1.9), logarithmic Sobolev (1.1) and Talagrand's (1.8) inequalities,

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, \frac{1}{2}\Big(\sqrt{\mathrm{I}(\nu)} - \mathrm{W}_2(\nu,\gamma)\Big)^2 \,\geq\, \frac{1}{2}\Big(\sqrt{2\,\mathrm{H}(\nu)} - \mathrm{W}_2(\nu,\gamma)\Big)^2.$$

Hence

$$\delta_{\mathrm{LSI}}(\nu) \geq \frac{(2\,\mathrm{H}(\nu) - W_2^2(\nu,\gamma))^2}{2(\sqrt{2\,\mathrm{H}(\nu)} + \mathrm{W}_2(\nu,\gamma))^2} \geq \frac{\delta_{\mathrm{Tal}}(\nu)^2}{16\,\mathrm{H}(\nu)}\,.$$

The result then follows from Theorem 5 for a centered ν, and in the general case by recentering as above.

Note that the inequality given by Corollary 6 is of a similar form to (1.11) established in [9] for smooth measures. It does not seem that the measure $\bar{\nu}$ involved in (1.11) is directly comparable to ν in general, whereas ν_b is an explicit transformation of ν. In particular, Corollary 6 immediately implies the equality cases of LSI for general measures without any additional argument.

Finally, there is also a lower bound on the deficit $\delta_{\mathrm{LSI}}(\nu)$ which may be expressed only in terms of Kantorovich-Wasserstein distances. For simplicity, only the centered case is considered.

Corollary 7. *There is a numerical constant $c > 0$ such that for any centered probability measure $d\nu = f d\gamma$ on $\mathbb{R}^n$*

$$\delta_{\mathrm{LSI}}(\nu) \geq \min\left[\frac{c\,\mathrm{W}_{1,1}(\nu,\gamma)^4}{n^2\,\mathrm{W}_2(\nu,\gamma)^2}\,,\frac{1}{2}\left(\sqrt{\mathrm{W}_2(\nu,\gamma)^2 + \frac{c\,\mathrm{W}_{1,1}(\nu,\gamma)}{\sqrt{n}}} - \mathrm{W}_2(\nu,\gamma)\right)^2\right].$$

For the proof, argue as for Corollary 6 combining the HWI, logarithmic Sobolev and Talagrand inequalities to get that

$$\delta_{\mathrm{LSI}}(\nu) \geq \frac{1}{2}\left(\sqrt{\mathrm{W}_2(\nu,\gamma)^2 + \delta_{\mathrm{Tal}}(\mu)} - \mathrm{W}_2(\nu,\gamma)\right)^2.$$

Write $\mathrm{W}_2 = \mathrm{W}_2(\nu,\gamma)$ and $\mathrm{W}_{1,1} = \mathrm{W}_{1,1}(\nu,\gamma)$ to ease the notation. By Theorem 5,

$$\begin{aligned}\delta_{\mathrm{LSI}}(\nu) &\geq \frac{1}{2}\min\left[\left(\sqrt{\mathrm{W}_2^2 + \frac{c\,\mathrm{W}_{1,1}^2}{n}} - \mathrm{W}_2\right)^2, \left(\sqrt{\mathrm{W}_2^2 + \frac{c\,\mathrm{W}_{1,1}}{\sqrt{n}}} - \mathrm{W}_2\right)^2\right] \\ &= \frac{1}{2}\min\left[\mathrm{W}_2^2\left(\sqrt{1 + \frac{c\,\mathrm{W}_{1,1}^2}{n\,\mathrm{W}_2}} - 1\right)^2, \left(\sqrt{\mathrm{W}_2^2 + \frac{c\,\mathrm{W}_{1,1}}{\sqrt{n}}} - \mathrm{W}_2\right)^2\right].\end{aligned}$$

Since $\mathrm{W}_{1,1}^2 \leq n\,\mathrm{W}_2^2$,

$$\sqrt{1 + \frac{c\,\mathrm{W}_{1,1}^2}{n\,\mathrm{W}_2^2}} \geq 1 + \frac{c'\,\mathrm{W}_{1,1}^2}{n\,\mathrm{W}_2^2}$$

for some $c' > 0$ only depending on c, and the claim follows.

This work exploits two techniques to prove deficit estimates (semigroup and transport techniques). Several other methods have been used in the literature to study deficit estimates for functional inequalities (Sobolev inequalities, isoperimetric inequalities):

- Expansion and compactness arguments, used in [5] and extended in [26] to study Sobolev inequalities. This technique yields lower bounds on the deficit in strong metrics, but has the drawback of not giving estimates of the constants involved (since the argument is by contradiction). In particular, it is not clear how the constant obtained in [5] depends on the dimension. For the LSI, mimicking the expansion argument of [5] suggests looking for deficit estimates which do not appear to be true in full generality.

- Stochastic calculus techniques, used by Eldan in [18] to study the Gaussian isoperimetric inequality. For the LSI, one could follow the proof of the LSI of [32] to obtain a representation of the deficit in terms of the quadratic variation of a certain stochastic process, but deducing from this representation a lower bound in terms of a useful distance seems like a challenging problem.
- Recently, Barchiesi, Brancolini and Julin [4] gave a proof of a sharp, dimension-free quantitative Gaussian isoperimetric inequality, using techniques from calculus of variations by reducing the problem to analyzing the first and second variations of a suitable functional. Such arguments were also used in [24] to prove (nonquantitative) transport inequalities.
- A technique involving a Taylor expansion on a compact manifold has been employed in [30, 28] to address the stability of the polygonal isoperimetric inequality. The spectral theory of circulant matrices shows up in the analysis and is utilized to compare the associated quadratic forms. In the context of LSI, one could consider Gaussian manifolds and perform a Taylor expansion of the deficit to identify suitable lower bounds involving classical L^p metrics. Such an approach was recently employed by Christ in the context of the sharp Hausdorff-Young inequality [14] and may be relevant for LSI via the Beckner-Hirschman entropic uncertainty principle.

The rest of the paper is organized as follows. In Section 2, we prove the main results. In Section 3, we establish several one-dimensional results. Lastly, in Section 4, we present an improvement of the Bakry-Émery theorem for symmetric measures satisfying a Poincaré inequality and obtain quantitative versions of the Wehrl conjectures established by Lieb [33] and Carlen [12] in the context of the coherent state transform.

2. **Proofs of Theorems 1 and 5.** We start with the proof of Theorem 1. The results in [29] rely on mass transportation tools. The arguments here are based on the standard semigroup interpolation along the Ornstein-Uhlenbeck semigroup going back to [3] (cf. [1, 6]), together with heat kernel inequalities as developed in [6] (to which we refer for the necessary background).

Proof of Theorem 1. Recall the Ornstein-Uhlenbeck semigroup $(P_t)_{t\geq 0}$ given on suitable functions $g : \mathbb{R}^n \to \mathbb{R}$ by

$$P_t g(x) \,=\, \int_{\mathbb{R}^n} g\big(e^{-t}x + \sqrt{1-e^{-2t}}\, y\big) d\gamma(y), \quad t \geq 0,\ x \in \mathbb{R}^n.$$

The Ornstein-Uhlenbeck semigroup $(P_t)_{t\geq 0}$ is invariant and symmetric with respect to γ and, on smooth functions, $\nabla P_t g = e^{-t} P_t(\nabla g)$ (as vectors). For each $t \geq 0$, set $d\nu_t = P_t f d\gamma$. The classical de Brujin's formula indicates that

$$\mathrm{H}(\nu) \,=\, \int_0^\infty \mathrm{I}(\nu_t) dt. \tag{2.12}$$

This identity follows from the fact that the Fisher information $\mathrm{I}(\nu_t)$ is the time-derivative of the entropy along the Ornstein-Uhlenbeck flow.

In the first step of the argument, we show that for any $t \geq 0$, ν_t satisfies a Poincaré inequality (1.4) with constant

$$\lambda_t \,=\, \frac{1}{\lambda^{-1} e^{-2t} + 1 - e^{-2t}}\,.$$

To prove this, consider a smooth function g with

$$\int_{\mathbb{R}^n} g\, d\nu_t \,=\, \int_{\mathbb{R}^n} g\, P_t f\, d\gamma \,=\, \int_{\mathbb{R}^n} P_t g\, d\nu \,=\, 0$$

(by symmetry of P_t). First, by the local Poincaré inequalities for $(P_t)_{t\geq 0}$ (cf. [6]), for every $t \geq 0$,

$$P_t(g^2) \,\leq\, (P_t g)^2 + (1 - e^{-2t}) P_t\big(|\nabla g|^2\big).$$

Hence,

$$\int_{\mathbb{R}^n} g^2 d\nu_t \,=\, \int_{\mathbb{R}^n} P_t(g^2) d\nu \,\leq\, \int_{\mathbb{R}^n} (P_t g)^2 d\nu + (1 - e^{-2t}) \int_{\mathbb{R}^n} P_t\big(|\nabla g|^2\big) d\nu.$$

Then, by the Poincaré inequality applied to $P_t g$, since $\int_{\mathbb{R}^n} P_t g d\nu = 0$,

$$\begin{aligned}
\int_{\mathbb{R}^n} g^2 d\nu_t \,&\leq\, \frac{1}{\lambda} \int_{\mathbb{R}^n} |\nabla P_t g|^2 d\nu + (1 - e^{-2t}) \int_{\mathbb{R}^n} P_t\big(|\nabla g|^2\big) d\nu \\
&\leq\, \Big(\frac{e^{-2t}}{\lambda} + 1 - e^{-2t}\Big) \int_{\mathbb{R}^n} P_t\big(|\nabla g|^2\big) d\nu \\
&\leq\, \Big(\frac{e^{-2t}}{\lambda} + 1 - e^{-2t}\Big) \int_{\mathbb{R}^n} |\nabla g|^2 d\nu_t
\end{aligned}$$

where we used the heat kernel inequality $|\nabla P_t g|^2 \leq e^{-2t} P_t(|\nabla g|^2)$ and again the symmetry of P_t. The claim follows.

Towards the second step of the argument, recall that by integration by parts, for every $t > 0$,

$$\mathrm{I}(\nu_t) \,=\, \int_{\mathbb{R}^n} \frac{|\nabla P_t f|^2}{P_t f}\, d\gamma \,=\, \int_{\mathbb{R}^n} P_t f |\nabla \log P_t f|^2 d\gamma \,=\, \int_{\mathbb{R}^n} |\nabla \log P_t f|^2 d\nu_t.$$

As is classical (cf. [1, 6]),

$$\frac{d}{dt}\, \mathrm{I}(\nu_t) \,=\, -2 \int_{\mathbb{R}^n} P_t f\, \Gamma_2(\log P_t f) d\gamma \,=\, -2 \int_{\mathbb{R}^n} \Gamma_2(\log P_t f) d\nu_t \tag{2.13}$$

where $\Gamma_2(v) = |\mathrm{Hess}(v)|^2 + |\nabla v|^2$.

Since ν has a first moment, $|\nabla P_t f| \in \mathrm{L}^1(\gamma)$ for every $t > 0$. Then, if $v_t = \log P_t f$, by the Gaussian integration by parts formula,

$$\int_{\mathbb{R}^n} \nabla v_t\, d\nu_t \,=\, \int_{\mathbb{R}^n} \nabla P_t f\, d\gamma \,=\, \int_{\mathbb{R}^n} x P_t f\, d\gamma.$$

By symmetry,

$$\int_{\mathbb{R}^n} x P_t f\, d\gamma \,=\, \int_{\mathbb{R}^n} P_t x\, f\, d\gamma \,=\, e^{-t} \int_{\mathbb{R}^n} x f\, d\gamma = 0.$$

Since ν_t satisfies a Poincaré inequality with constant λ_t, applied to $v_t = \log P_t f$ for which therefore $\int_{\mathbb{R}^n} \nabla v_t d\nu_t = 0$,

$$\lambda_t \int_{\mathbb{R}^n} |\nabla v_t|^2 d\nu_t \,\leq\, \int_{\mathbb{R}^n} \big|\mathrm{Hess}(v_t)\big|^2 d\nu_t.$$

As a consequence,

$$\frac{d}{dt}\, \mathrm{I}(\nu_t) \,\leq\, -2(\lambda_t + 1)\, \mathrm{I}(\nu_t).$$

Integrating this differential inequality, for every $t \geq 0$,

$$\mathrm{I}(\nu_t) \,\leq\, \mathrm{I}(\nu)\, e^{-4t}\, \frac{\lambda_t}{\lambda}\,.$$

Finally, by de Brujin's formula (2.12), the conclusion follows. The proof of Theorem 1 is complete. □

We now turn to the proof of Theorem 5, which is based on mass transportation arguments.

Proof of Theorem 5. The starting point is Cordero-Erausquin's mass transportation proof of Talagrand's inequality [16]. Let $d\nu = f d\gamma$ be centered and

$$T \,=\, (T_1, \ldots, T_n) : \mathbb{R}^n \,\to\, \mathbb{R}^n$$

be the Brenier map pushing γ onto ν. It satisfies the Monge-Ampère equation

$$e^{-|x|^2/2} \,=\, f\big(T(x)\big) e^{-|T(x)|^2/2} \det\big(\nabla T(x)\big),$$

$d\gamma$-a.e. in the sense of Alexandrov [35, 11]. Following [16],

$$\mathrm{H}(\nu) \,\geq\, \frac{1}{2}\, \mathrm{W}_2(\nu,\gamma)^2 + \int_{\mathbb{R}^n} \big[\Delta\theta - \log\det\big(\mathrm{Id} + \mathrm{Hess}(\theta)\big)\big]\, d\gamma$$

where $\nabla\theta(x) \,=\, T(x) - x$. Since the Laplacian is the sum of the eigenvalues of the Hessian, and since by the Brenier theorem T is given by the gradient $\nabla\phi$ of a convex function $\phi : \mathbb{R}^n \to \mathbb{R}$ (cf. [42, 43]), denoting by $\lambda_1, \ldots, \lambda_n$ the non-negative eigenvalues of ∇T, we have

$$\delta_{\mathrm{Tal}}(\nu) \,\geq\, \int_{\mathbb{R}^n} \sum_{i=1}^n [\lambda_i - 1 - \log\lambda_i]\, d\gamma \,\geq\, \frac{1}{6} \int_{\mathbb{R}^n} \sum_{i=1}^n \min\big(|\lambda_i - 1|^2, |\lambda_i - 1|\big) d\gamma.$$

Let $I = \{1 \leq i \leq n\,;\, |\lambda_i - 1| \leq 1\}$. Then

$$\begin{aligned}\sum_{i=1}^n \min\big(|\lambda_i - 1|^2, |\lambda_i - 1|\big) \,&=\, \sum_{i\in I} |\lambda_i - 1|^2 + \sum_{i\in I^c} |\lambda_i - 1| \\ &\geq\, \sum_{i\in I} |\lambda_i - 1|^2 + \sqrt{\sum_{i\in I^c} |\lambda_i - 1|^2}.\end{aligned}$$

Hence

$$\begin{aligned}\delta_{\mathrm{Tal}}(\nu) \,&\geq\, \frac{1}{6} \int \sum_{i\in I} |\lambda_i - 1|^2 d\gamma + \frac{1}{6} \int \sqrt{\sum_{i\in I^c} |\lambda_i - 1|^2}\, d\gamma \\ &\geq\, \frac{1}{6} \bigg(\int \sqrt{\sum_{i\in I} |\lambda_i - 1|^2}\, d\gamma \bigg)^2 + \frac{1}{6} \int \sqrt{\sum_{i\in I^c} |\lambda_i - 1|^2}\, d\gamma\end{aligned}$$

by Jensen's inequality. Assuming that $\delta_{\mathrm{Tal}}(\nu) \leq \alpha$ for some $\alpha > 0$,

$$\delta_{\mathrm{Tal}}(\nu) \,\geq\, \frac{1}{6} \bigg(\int \sqrt{\sum_{i\in I} |\lambda_i - 1|^2}\, d\gamma \bigg)^2 + \frac{1}{36\alpha} \bigg(\int \sqrt{\sum_{i\in I^c} |\lambda_i - 1|^2}\, d\gamma \bigg)^2.$$

Then

$$\begin{aligned}\delta_{\mathrm{Tal}}(\nu) \,&\geq\, \frac{1}{72\max(\alpha,1)} \bigg(\int \sqrt{\sum_{i\in I} |\lambda_i(x) - 1|^2}\, d\gamma + \int \sqrt{\sum_{i\in I^c} |\lambda_i - 1|^2}\, d\gamma \bigg)^2 \\ &\geq\, \frac{1}{72\max(\alpha,1)} \bigg(\int \sqrt{\sum_{i=1}^n |\lambda_i - 1|^2}\, d\gamma \bigg)^2.\end{aligned}$$

Now, by the Cauchy-Schwarz inequality,

$$\sqrt{\sum_{i=1}^n |\lambda_i - 1|^2} \;=\; \sqrt{\sum_{i,j=1}^n \big|(\nabla T)_{ij} - \delta_{ij}\big|^2} \;\geq\; \frac{1}{\sqrt{n}} \sum_{i=1}^n \big|\nabla (T_i - x_i)\big|.$$

The characterization $T = \nabla \phi$, where $\phi : \mathbb{R}^n \to \mathbb{R}$ is convex, implies that ϕ is an Alexandrov solution to

$$\det\big(\mathrm{Hess}(\phi)\big) \;=\; \frac{e^{-|x|^2/2}}{f(T(x)) e^{-|T(x)|^2/2}}\,.$$

Since $f > 0$ and T are locally bounded, the right-hand side is bounded away from zero and infinity on every compact set. In particular, ϕ is $W^{2,1}$ [36] (see also Remark 8 below). The $(1,1)$-Poincaré inequality (1.10) holds for mean zero $W^{1,1}$ functions. Observing that $\int_{\mathbb{R}^n} [T_i(x) - x_i] d\gamma = 0$, $i = 1, \ldots, n$, we thus obtain that

$$\delta_{\mathrm{Tal}}(\nu) \;\geq\; \frac{1}{288 n \max(\alpha, 1)} \bigg(\int_{\mathbb{R}^n} \sum_{i=1}^n |T_i - x_i| d\gamma \bigg)^2 \;\geq\; \frac{1}{288 n \max(\alpha, 1)}\, \mathrm{W}_{1,1}(\nu, \gamma)^2.$$

As a result, for every $\alpha > 0$,

$$\delta_{\mathrm{Tal}}(\nu) \;\geq\; \min\bigg(\frac{\mathrm{W}_{1,1}(\nu,\gamma)^2}{288 n \max(\alpha, 1)}\,, \alpha \bigg).$$

Optimizing in $\alpha > 0$ concludes the proof of Theorem 5. □

Remark 8. In the proof of Theorem 5, [36] was employed to deduce $W^{2,1}$-regularity of the potential function ϕ. In our framework, one may also infer the regularity in a different way. Indeed, from [10] it follows that if ϕ is not strictly convex at a point, then it is affine on a line. Since ϕ is globally convex, this implies that it only depends on $(n-1)$ variables. In particular, $\nabla\phi(\mathbb{R}^n)$ is contained in an $(n-1)$-dimensional subspace, and this contradicts that $\nabla\phi$ pushes $d\gamma$ onto $f d\gamma$. Hence, ϕ is strictly convex on $\mathbb{R}^n$, and the desired regularity follows from [17].

3. **One dimensional estimates via mass transfer.** The proof of Theorem 1 relies on heat kernel theory. In this section, we establish an L^1 estimate via mass transfer theory for measures satisfying a $(1,1)$-Poincaré inequality on the real line

$$\lambda \int_{\mathbb{R}} |g|\, d\nu \;\leq\; \int_{\mathbb{R}} |\nabla g|\, d\nu \tag{3.14}$$

for some $\lambda > 0$ and every smooth mean zero $g : \mathbb{R} \to \mathbb{R}$. Sufficient conditions to guarantee the (1,1)-Poincaré are given in [2] (see e.g. Theorem 1.5 there). In general, the L^1 Poincaré is stronger than the standard L^2 inequality (1.4), which makes Theorem 9 below weaker than Theorem 1. However, the emphasis here is on the method of proof.

Theorem 9. *Let $d\nu = f d\gamma$ be a probability measure on $\mathbb{R}$ with barycenter $b = b(\nu)$ satisfying a $(1,1)$-Poincaré inequality with constant $\lambda > 0$. Then there exists $\tilde{c} = \tilde{c}(\lambda) > 0$ such that if $\delta_{\mathrm{LSI}}(\nu) \leq 1$,*

$$\delta_{\mathrm{LSI}}(\nu) \;\geq\; \tilde{c} \bigg(\int_{\mathbb{R}} \big|(\log f)' - b\big| d\nu \bigg)^2.$$

Proof. Let T be the optimal transport map between $d\nu = f d\gamma$ and $d\gamma$. Note that $T = G^{-1} \circ F$, where F and G are the cumulative distribution functions of $d\nu$ and $d\gamma$, respectively. In particular,

$$T'(x) \,=\, \frac{f(x) e^{-|x|^2/2}}{e^{-|T(x)|^2/2}} \,.$$

From Cordero-Erausquin's mass transportation proof of the logarithmic Sobolev inequality [16], we extract the estimates

$$2\delta_{\mathrm{LSI}}(\nu) \,\geq\, \int_{\mathbb{R}} \big|T - x + (\log f)'\big|^2 d\nu(x) \tag{3.15}$$

and

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, \int_{\mathbb{R}} \big[T' - 1 - \log\big(1 + (T' - 1)\big)\big] d\nu(x) \tag{3.16}$$

where T is the optimal transport map between $d\nu = f d\gamma$ and $d\gamma$. Recall $\varphi : (-1, \infty) \to \mathbb{R}$ defined by $\varphi(t) = t - \log(1+t)$ and set

$$\widetilde{\varphi}(t) \,=\, \begin{cases} \frac{t^2}{6}\,, & -1 \leq t \leq 1, \\ \varphi(t) - \frac{5}{6} + \log 2, & t \geq 1. \end{cases} \tag{3.17}$$

Note that $\widetilde{\varphi}(t) = \widetilde{\varphi}(|t|)$ is convex and $\varphi(t) \geq \frac{1}{10}\widetilde{\varphi}(t)$. By (3.16), Jensen's inequality and the fact that $T' \geq 0$, we obtain

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, \frac{1}{10} \int_{\mathbb{R}} \widetilde{\varphi}\big(|T' - 1|\big) d\nu \,\geq\, \frac{1}{10} \widetilde{\varphi}\bigg(\int_{\mathbb{R}} |T' - 1| \, d\nu \bigg). \tag{3.18}$$

Since it is asumed that $\delta_{\mathrm{LSI}}(\nu) \leq 1$, it follows from the properties of $\widetilde{\varphi}$ that

$$\widetilde{\varphi}\bigg(\int_{\mathbb{R}} |T' - 1| d\nu \bigg) \,\geq\, c \bigg(\int_{\mathbb{R}} |T' - 1| \, d\nu \bigg)^2$$

for a universal $c > 0$. Hence

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, c \bigg(\int_{\mathbb{R}} |T' - 1| \, d\nu \bigg)^2 . \tag{3.19}$$

By the push-forward condition, $\int_{\mathbb{R}} (x - T) d\nu = b - \int_{\mathbb{R}} T d\nu = b$. Thus, combining this information with (3.15), the Cauchy-Schwarz inequality and the $(1,1)$-Poincaré inequality (3.14),

$$\begin{aligned} \int_{\mathbb{R}} \big|(\log f)' - b\big| d\nu \,&\leq\, \int_{\mathbb{R}} \big|(\log f)' - (x - T)\big| d\nu + \int_{\mathbb{R}} \big|(x - T) - b\big| d\nu \\ &\leq\, \sqrt{2\delta_{\mathrm{LSI}}(\nu)} + \frac{1}{\lambda} \int_{\mathbb{R}} |T' - 1| \, d\nu. \end{aligned}$$

Together with (3.19), the claim is easily completed. □

The next corollary is achieved as Corollary 2.

Corollary 10. *Let $d\nu = f d\gamma$ be a centered probability measure on $\mathbb{R}$ satisfying a $(1,1)$-Poincaré inequality with constant $\lambda > 0$. Then there exists $\tilde{c} = \tilde{c}(\lambda) > 0$ such that if $\delta_{\mathrm{LSI}}(\nu) \leq 1$,*

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, \tilde{c} \, ||\nu - \gamma||_{\mathrm{TV}}^2 .$$

As already mentioned, since Theorem 1 cannot hold for all probability measures, one may not hope to generalize Corollary 2 by enlarging the function space. However, this does not prevent the weaker estimates in Theorem 9 and Corollary 10 from being true in general. If these estimates held in full generality, without the assumption that ν satisfies some Poincaré inequality, then they would automatically recover the equality cases of the Gaussian logarithmic Sobolev inequality.

We conclude this section by proving a version of Corollaries 2 and 3 on the real line for probability measures satisfying a second moment bound (without assuming a Poincaré inequality). The proof is again based on mass transfer. Recall the function $\widetilde{\varphi}$ (3.17) from the proof of Theorem 9.

Theorem 11. *Let $d\nu = f d\gamma$ be a probability measure on $\mathbb{R}$ with barycenter $b = b(\nu)$ such that $\mathrm{Var}_\nu(x) \leq 1$. Then, for some $C > 0$,*

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, \widetilde{\varphi}\Big(C \int_{\mathbb{R}} \big|(\log f)' - b\big|^2 d\nu \Big).$$

In particular, for some numerical $c > 0$,

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, c\, \mathrm{W}_2(\nu, \gamma_b)^4$$

where γ_b is given in (1.2).

A multidimensional version of this result was proved in [9], with a smoothness assumption on f. The proof there is based on a rescaling property of the LSI. The contribution here is an alternative technique of proof. It would be of interest to see if the multidimensional version can be similarly obtained using transport arguments.

Proof. By approximation, it may be assumed that f has compact support and is smooth enough with derivative at least in $\mathrm{L}^1(\gamma)$. Letting as above $T : \mathbb{R} \to \mathbb{R}$ be the increasing map pushing ν onto γ, we have

$$\int_{\mathbb{R}} \big|(\log f)' - b\big|^2 d\nu \,=\, \int_{\mathbb{R}} \big|(\log f)' + (T-x) - (T-x) - b\big|^2 d\nu$$

$$= \int_{\mathbb{R}} \big|(\log f)' + (T-x)\big|^2 d\nu + b^2 - \int_{\mathbb{R}} |T-x|^2 d\nu - 2\int_{\mathbb{R}} (T - x + b)(\log f)' d\nu$$

$$= \int_{\mathbb{R}} \big|(\log f)' + (T-x)|^2 d\nu + b^2 - \int_{\mathbb{R}} |T-x|^2 d\nu - 2\int_{\mathbb{R}} (T-x) f' d\gamma - 2b \int_{\mathbb{R}} f' d\gamma.$$

By Gaussian integration by parts, $\int_{\mathbb{R}} f' d\gamma = \int_{\mathbb{R}} x f d\gamma = b$ and similarly

$$\int_{\mathbb{R}} (T-x) f' d\gamma \,=\, \int_{\mathbb{R}} x(T-x) d\nu - \int_{\mathbb{R}} (T' - 1) d\nu.$$

After some algebra, it follows that

$$\int_{\mathbb{R}} \big|(\log f)' - b\big|^2 d\nu \,=\, \int_{\mathbb{R}} \big|(\log f)' + (T-x)\big|^2 d\nu + 2\int_{\mathbb{R}} (T' - 1) d\nu + \mathrm{Var}_\nu(x) - 1.$$

Using (3.15) and (3.18), we get that

$$\int_{\mathbb{R}} \big|(\log f)' - b\big|^2 d\nu \,\leq\, 2\delta_{\mathrm{LSI}}(\nu) + 2\widetilde{\varphi}^{-1}\big(\delta_{\mathrm{LSI}}(\nu)\big) + \mathrm{Var}_\nu(x) - 1,$$

where $\widetilde{\varphi}^{-1}$ is the inverse of $\widetilde{\varphi}$ on $\mathbb{R}^+$. Since $\widetilde{\varphi}^{-1}(x) \geq Cx$ for some $C > 0$,

$$\delta_{\mathrm{LSI}}(\nu) \,\geq\, \widetilde{\varphi}\Big(C \int_{\mathbb{R}} \big|(\log f)' - b\big|^2 d\nu \Big).$$

But $\int_{\mathbb{R}} |(\log f)' - b|^2 d\nu$ is the relative Fisher information of ν with respect to the non-centered Gaussian $d\gamma_b = e^{b.x - b^2/2} d\gamma$ which satisfies a logarithmic Sobolev inequality with constant $\frac{1}{2}$. Therefore, together with Talagrand's inequality (1.8),

$$\int_{\mathbb{R}} \big|(\log f)' - b\big|^2 d\nu \;\geq\; \mathrm{H}\big(\nu \,|\, \gamma_b\big) \;\geq\; \mathrm{W}_2(\nu, \gamma_b)^2$$

and hence

$$\delta_{\mathrm{LSI}}(\nu) \;\geq\; \widetilde{\varphi}\big(C\, \mathrm{W}_2(\nu, \gamma_b)^2\big).$$

By definition of the Wassertein distance W_2,

$$\mathrm{W}_2(\nu, \gamma_b)^2 \;\leq\; 2\,\mathrm{Var}_\nu(x) + 2\,\mathrm{Var}_{\gamma_b}(x) \;\leq\; 4$$

under the assumption $\mathrm{Var}_\nu(x) \leq 1$. Since $\widetilde{\varphi}$ behaves quadratically near the origin, it finally follows that for some numerical $c > 0$,

$$\delta_{\mathrm{LSI}}(\nu) \;\geq\; c\, \mathrm{W}_2(\nu, \gamma_b)^4.$$

□

4. **Extensions and applications.**

4.1. **The Bakry-Émery theorem for symmetric measures in $\mathcal{P}(\lambda)$.** In what follows we describe an extension of Theorem 1 to families of log-concave measures. Let $d\mu = e^{-V} dx$ where $V : \mathbb{R}^n \to \mathbb{R}$ a smooth potential be a probability measure on $\mathbb{R}^n$ satisfying the convexity condition (1.5), that is $\mathrm{Hess}(V) \geq \eta\, \mathrm{Id}$ for some $\eta > 0$. The Gaussian case corresponds to the quadratic potential $V(x) = \frac{|x|^2}{2}$ with $\eta = 1$.

Given a probability measure $d\nu = f d\mu$ with density f with respect to μ, the relative entropy and Fisher information with respect to μ are defined as in the Gaussian case by

$$\mathrm{H}\big(\nu \,|\, \mu\big) \;=\; \int_{\mathbb{R}^n} f \log f \, d\mu \qquad \text{and} \qquad \mathrm{I}\big(\nu \,|\, \mu\big) \;=\; \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f}\, d\mu,$$

and the Bakry-Émery LSI (see [3, 42, 43, 6]) ensures that

$$\mathrm{H}\big(\nu \,|\, \mu\big) \;\leq\; \frac{1}{2\eta}\, \mathrm{I}\big(\nu \,|\, \mu\big).$$

As for the Gaussian LSI, the proof relies on the semigroup $(P_t^V)_{t \geq 0}$ with infinitesimal generator $\mathcal{L}^V = \Delta - \nabla V \cdot \nabla$ for which the analogues of (2.12) and (2.13) read, with $d\nu_t = P_t^V f d\mu$,

$$\mathrm{H}\big(\nu \,|\, \mu\big) \;=\; \int_0^\infty \mathrm{I}\big(\nu_t \,|\, \mu\big) dt$$

and

$$\frac{d}{dt}\, \mathrm{I}\big(\nu_t \,|\, \mu\big) \;=\; -2 \int_{\mathbb{R}^n} \Gamma_2(P_t^V \log f) d\nu_t$$

where, this time,

$$\Gamma_2(v) \;=\; \big|\mathrm{Hess}(v)\big|^2 + \langle \mathrm{Hess}(V) \nabla v, \nabla v \rangle \;\geq\; \big|\mathrm{Hess}(v)\big|^2 + \eta\, |\nabla v|^2.$$

If we try to mimic the proof of Theorem 1 in this context, it should be proved that as soon as ν belongs to $\mathcal{P}(\lambda)$, ν_t belongs to $\mathcal{P}(\lambda_t)$ with

$$\lambda_t \;=\; \frac{1}{\lambda^{-1} e^{-2\eta t} + \eta^{-1}(1 - e^{-2\eta t})}$$

(which is proved as in the Gaussian case), and that, whenever ν is centered, $\int_{\mathbb{R}^n} \nabla v_t d\nu_t = 0$ for all $t \geq 0$ where $v_t = \log P_t^V f$. The latter requirement is however not true in this general context. It can nevertheless hold in some more restricted setting, for example as soon as V is even and ν is symmetric (i.e. if f is also even) in which case $\int_{\mathbb{R}^n} \nabla v_t d\nu_t = \int_{\mathbb{R}^n} \nabla V d\nu_t = 0$.

These observations lead to the following improvement of the Bakry-Émery theorem for symmetric measures in $\mathcal{P}(\lambda)$.

Theorem 12. *Assume that $d\mu = e^{-V} dx$ is a symmetric probability measure such that $\mathrm{Hess}(V) \geq \eta \,\mathrm{Id}$ for some $\eta > 0$, and let $d\nu = f d\mu$ be a symmetric probability measure in the class $\mathcal{P}(\lambda)$ for some $\lambda > 0$. Then, for every $t \geq 0$,*

$$\mathrm{I}\big(\nu_t \,|\, \mu\big) \;\leq\; e^{-4\eta t}\, \frac{\lambda_t}{\lambda}\, \mathrm{I}\big(\nu \,|\mu\big).$$

Consequently, if $\lambda \neq \eta$,

$$\mathrm{H}\big(\nu \,|\mu\big) \;\leq\; \frac{\eta - \lambda - \lambda(\ln \eta - \ln \lambda)}{2(\eta - \lambda)^2}\, \mathrm{I}\big(\nu \,|\mu\big)$$

and, if $\lambda = \eta$,

$$\mathrm{H}\big(\nu \,|\mu\big) \;\leq\; \frac{1}{4\eta}\, \mathrm{I}\big(\nu \,|\mu\big).$$

Note that this result is not a stability result, since the constant given by the Bakry-Émery theorem is not optimal in general. Theorem 12 nevertheless yields improved estimates on the speed of convergence to equilibrium for the semigroup, of interest for example in the context of Monte Carlo Markov Chain sampling of the measure μ.

Similar estimates can obtained for measures which are given by bounded perturbations of uniformly convex potentials, using the Holley-Stroock approach. This includes the important example of the quartic double-well potential $V(x) = (x^2 - 1)^2$ (which is used in statistical physics for continuous versions of the Ising model).

4.2. **Coherent state transform.** For $h > 0$, let $d\mu_h$ denote h^{-n} times the Lebesgue measure on $\mathbb{C}^n$ viewed as $\mathbb{R}^{2n}$. The coherent state transform is an integral transform mapping $(\mathrm{L}^2(\mathbb{R}^n), dx)$ isometrically onto a subspace of $(\mathbb{R}^{2n}, d\mu_h)$ and given explicitly by

$$\psi \;\mapsto\; \mathcal{L}\psi(p,q) \;=\; e^{ip\cdot q/2h^*} \int_{\mathbb{R}^n} e^{ip\cdot x/h^*} e^{-|x-y|^2/2h^*} \psi(x) dx$$

with $h^* = \frac{h}{2\pi}$. The map $\mathcal{L}$ is built out of Weyl's representation of the Heisenberg group and has applications in quantum mechanics, where $|\mathcal{L}\psi|^2$ is interpreted as the phase space density in the state ψ. Bounds on $|\mathcal{L}\psi|^2$ are useful in estimating, e.g., the ground state energy of a Schrödinger operator (see [34, 12]).

The concentration of a density ρ can be measured via the entropy functional S defined by

$$\mathrm{S}(\rho) \;=\; -\int_{\mathbb{R}^{2n}} \rho \log \rho \, d\mu_h.$$

Note that this is the *physical* entropy, which is the negative of the mathematical entropy. Wehrl [44] conjectured n to be a lower bound on the entropy of phase space densities induced by $\mathcal{L}$ acting on $(\mathrm{L}^2(\mathbb{R}^n), dx)$, that is

$$\mathrm{S}(\rho) \;\geq\; n$$

whenever $\rho = |\mathcal{L}\psi|^2$ and $\psi \in (\mathrm{L}^2(\mathbb{R}^n), dx)$. Lieb [33] established this inequality with a method based on the sharp Young and Haussdorf-Young inequalities. Carlen [12] recovered Lieb's result via an approach based on the logarithmic Sobolev inequality and also settled the problem of characterizing the cases of equality.

In what follows we apply our results from the previous sections to show that in some configurations, one can obtain positive lower bounds on the *Wehrl deficit*

$$\delta_{\mathrm{Wehrl}}(\rho) \,=\, \mathrm{S}(\rho) - n$$

in terms of well-known metrics. The method of proof is based on Carlen's approach.

Theorem 13. *Suppose $\rho = |\mathcal{L}\psi|^2$ is a probability density on $(\mathbb{R}^{2n}, d\mu_h)$ with barycenter $b = b_\rho \in \mathbb{R}^{2n}$. Let*

$$d\nu_\rho(z) \,=\, e^{\frac{|z|^2}{2}} \rho\Big(\sqrt{\frac{h}{2\pi}}\, z\Big) d\gamma(z),$$

$$d\nu_{\rho,b}(z) \,=\, e^{\frac{|z|^2}{2}} \rho\Big(\sqrt{\frac{h}{2\pi}}\, z + b\Big) d\gamma(z)$$

where γ is the standard Gaussian measure on $\mathbb{R}^{2n}$. There exists $c > 0$ such that if ρ is not identically $e^{-\frac{\pi}{h}|z|^2}$, then

$$\delta_{\mathrm{Wehrl}}(\rho) \,\geq\, \frac{c}{\mathrm{H}(\nu_\rho)} \,\min\left(\frac{\mathrm{W}_{1,1}(\nu_{\rho,b},\gamma)^4}{n^2}, \frac{\mathrm{W}_{1,1}(\nu_{\rho,b},\gamma)^2}{n}\right).$$

Moreover, in the class of probability densities ρ with finite second moments, $\delta_{\mathrm{Wehrl}}(\rho) = 0$ exactly when $\rho = e^{-\frac{\pi}{h}|z-z_0|^2}$ for some $z_0 \in \mathbb{R}^{2n}$ or alternatively, when $\psi_{p_0,q_0}(x) = e^{ip_0\cdot x}\phi_0(x-q_0)$ for some $(p_0,q_0) \in \mathbb{R}^{2n}$ and $\phi_0(x) = \left(\frac{2}{h}\right)^{\frac{n}{2}} e^{-\frac{|x|^2}{2}}$.

Proof. Let f_h be the density of ν_ρ with respect to γ so that $\int_{\mathbb{R}^{2n}} f_h d\gamma = \int_{\mathbb{R}^{2n}} \rho\, d\mu_h = 1$ and

$$\mathrm{H}(\nu_\rho) \,=\, \int_{\mathbb{R}^{2n}} f_h \log f_h d\gamma \,=\, \int_{\mathbb{R}^{2n}} \Big(\frac{\pi}{h}\rho\,|x|^2 + \rho\log\rho\Big) d\mu_h.$$

Since f_h is not identically 1, the strict convexity of the function $t \to t\log t$ implies (via Jensen) that $\mathrm{H}(\nu_\rho) > 0$. Since ρ has finite first moment, $\mathrm{W}_{1,1}(\nu_{\rho,b},\gamma) < \infty$. Thus, if $\mathrm{H}(\nu_\rho) = \infty$, there is nothing to prove, so we may assume without loss that ρ has finite second moments.

A direct calculation shows that

$$\frac{|\nabla f_h|^2}{f_h} \,=\, e^{\frac{|z|^2}{2}} \left(\frac{h}{2\pi}\, \frac{\big|\nabla\rho\big(\sqrt{\frac{h}{2\pi}}\, z\big)\big|^2}{\rho\big(\sqrt{\frac{h}{2\pi}}\big)} + 2\sqrt{\frac{h}{2\pi}}\, \nabla\rho\Big(\sqrt{\frac{h}{2\pi}}\, z\Big)\cdot z + \rho\Big(\sqrt{\frac{h}{2\pi}}\, z\Big)|z|^2\right)$$

and, by changing variables and using the divergence theorem,

$$\begin{aligned}\int_{\mathbb{R}^{2n}} \frac{|\nabla f_h|^2}{f_h} d\gamma \,&=\, \int_{\mathbb{R}^{2n}} \left(\frac{h}{2\pi}\,\frac{|\nabla\rho|^2}{\rho} + 2\nabla\rho\cdot x + \frac{2\pi}{h}\,\rho\,|x|^2\right) d\mu_h \\ &=\, \int_{\mathbb{R}^{2n}} \left(\frac{h}{2\pi}\,\frac{|\nabla\rho|^2}{\rho} + \frac{2\pi}{h}\,\rho\,|x|^2\right) d\mu_h - 4n.\end{aligned}$$

Therefore,

$$\delta_{\mathrm{LSI}}(\nu_\rho) \,=\, \frac{1}{2}\,\mathrm{I}(\nu_\rho) - \mathrm{H}(\nu_\rho) \,=\, \frac{h}{4\pi}\int_{\mathbb{R}^{2n}} \frac{|\nabla\rho|^2}{\rho}\, d\mu_h + \mathrm{S}(\rho) - 2n.$$

Since $\rho = |\mathcal{L}\psi|^2$, an application of [12, Theorem 6] yields

$$\int_{\mathbb{R}^{2n}} \frac{|\nabla\rho|^2}{\rho} d\mu_h \;=\; 4\int_{\mathbb{R}^{2n}} |\nabla\rho^{\frac{1}{2}}|^2 d\mu_h \;=\; \frac{4n\pi}{h}\,.$$

Thus $\delta_{\mathrm{LSI}}(\nu_\rho) = \mathrm{S}(\rho) - n$ and Corollary 6 implies

$$\mathrm{S}(\rho) - n \;\geq\; \frac{c}{\mathrm{H}(\nu_\rho)}\,\min\left(\frac{\mathrm{W}_{1,1}(\nu_{\rho,b},\gamma)^4}{4n^2}\,,\,\frac{\mathrm{W}_{1,1}(\nu_{\rho,b},\gamma)^2}{2n}\right)$$

where

$$\nu_{\rho,b}(dz) \;=\; f_h(z+b_h)e^{(-(b_h\cdot z+\frac{|b_h|^2}{2}))}d\gamma(z) \;=\; e^{|z|^2/2}\rho\left(\sqrt{\frac{h}{2\pi}}(z+b_h)\right)d\gamma(z)$$

and b_h is the barycenter of f_h with respect to the Gaussian. To conclude the proof of the inequality, note that

$$b_h \;=\; \int_{\mathbb{R}^{2n}} z f_h d\gamma \;=\; \sqrt{\frac{2\pi}{h}}\int_{\mathbb{R}^{2n}} z\rho(z)d\mu_h \;=\; \sqrt{\frac{2\pi}{h}}b_\rho.$$

Next, assume that $\mathrm{S}(\rho) = n$. Since ρ has finite second moments, $\mathrm{H}(\nu_\rho) < \infty$. If $\mathrm{H}(\nu_\rho) = 0$, Jensen's inequality ensures that ρ has the desired form. If $0 < \mathrm{H}(\nu_\rho) < \infty$, it follows that $\nu_{\rho,b} = \gamma$. Thus,

$$e^{\frac{|z|^2}{2}}\rho\left(\sqrt{\frac{h}{2\pi}}\,z + b_\rho\right) \;=\; 1$$

for some $b_\rho = (p_0, q_0) \in \mathbb{R}^{2n}$. Consequently,

$$\rho(z) \;=\; e^{-\frac{\pi}{h}|z-b_\rho|^2} \;=\; |\mathcal{L}\Psi_{p_0,q_0}|^2$$

and Lieb [33] has shown that the map $\psi \to |\mathcal{L}\psi|^2$ is injective. □

In a similar way, one may use Corollaries 3 and 4 to obtain dimension-independent lower bounds on the Wehrl deficit for a subclass of probability measures. For instance, Corollary 3 implies the following result.

Theorem 14. *Suppose $\rho = |\mathcal{L}\psi|^2$ is a probability density on $(\mathbb{R}^{2n}, d\mu_h)$ with barycenter $b = b_\rho \in \mathbb{R}^{2n}$, finite second moments, and such*

$$z \;\mapsto\; e^{|z|^2/2}\rho\left(\sqrt{\frac{h}{2\pi}}\,z\right)$$

satisfies a Poincaré inequality with constant $\lambda > 0$. Then

$$\delta_{\mathrm{Wehrl}}(\rho) \;\geq\; c_2(\lambda)\;\mathrm{W}_2(\nu_{\rho,b},\gamma)^2,$$

where $c_2(\lambda)$ is as in Corollary 3.

As an example of illustration, for $M > 0$, let

$$\rho(z) \in \mathcal{F}_M \;=\; \{e^{-\psi(z)} : \mathrm{Hess}(\psi) \geq M\}.$$

Set

$$f_h(z) \;=\; e^{|z|^2/2}\rho\left(\sqrt{\frac{h}{2\pi}}\,z\right)$$

and note that

$$-\mathrm{Hess}\big(\log(f_h)\big) \;=\; \frac{h}{2\pi}\,\mathrm{Hess}(\psi)\left(\sqrt{\frac{h}{2\pi}}\,z\right) - \mathrm{Id} \;\geq\; \frac{Mh}{2\pi} - \mathrm{Id}.$$

Thus, if $M > \frac{3\pi}{h}$, the previous theorem applies in $\mathcal{F}_M$.

It is well known that the range of $\mathcal{L}$ is closely related to the space $\mathcal{A}^2$ of entire function Φ on $\mathcal{C}^n$ such that

$$\int |\Phi(z)|^2 e^{-2\pi|z|^2/h} dp\, dq \, < \, \infty$$

where $z = (q + ip)/\sqrt{2}$. The precise statement is that for every $\psi \in (\mathrm{L}^2(\mathbb{R}^n), dx)$,

$$\mathcal{L}\psi(p,q) \, = \, e^{ip\cdot q/2h^*} \Phi((q-ip)/\sqrt{2}) e^{(p^2+q^2)/4h^*}$$

where $\Phi \in \mathcal{A}^2$. In fact, Segal [38, 39] (see also [40]) proved that the map $\widetilde{\mathcal{L}} : \psi \to \Phi$ is unitary from $(L^2(\mathbb{R}^n), dx)$ onto $\mathcal{A}^2$, and therefore Carlen [12] calls $\widetilde{\mathcal{L}}$ the Segal transform. With this in mind, the Segal transform may be useful in characterizing the subspace of functions ψ in the domain of $\mathcal{L}$ mapping to functions $|\mathcal{L}\Psi|^2$ admitting a Poincaré inequality and hence a dimensionless W_2-estimate via Theorem 14.

REFERENCES

[1] D. Bakry, L'hypercontractivité et son utilisation en théorie des semigroupes, *École d'Été de Probabilités de Saint-Flour, Lecture Notes in Math.* 1581 (1994), 1–114. Springer.

[2] D. Bakry, F. Barthe, P. Cattiaux and A. Guillin, A simple proof of the Poincaré inequality in a large class of probability measures including log-concave cases, *Elec. Comm. Prob.* 13 (2008), 60–66.

[3] D. Bakry and M. Émery, Diffusions hypercontractives, *Séminaire de Probabilités XIX, Lecture Notes in Math.* 1123 (1985), 177–206. Springer.

[4] M. Barchiesi, A. Brancolini and V. Julin, Sharp dimension free quantitative estimates for the Gaussian isoperimetric inequality, *to appear in Ann. Probab.* (2015).

[5] G. Bianchi and H. Egnell, A note on the Sobolev inequality, *J. Funct. Anal.* 100(1), 18–24 (1991).

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

[7] F. Barthe and A. V., Kolesnikov, Mass transport and variants of the logarithmic Sobolev inequality, *J. Geom. Anal.* 18 (2008), 921–979.

[8] S. Bobkov and F. Götze, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, *J. Funct. Anal.* 163 (1999), 1–28.

[9] S. Bobkov, N. Gozlan, C. Roberto and P.-M. Samson, Bounds on the deficit in the logarithmic Sobolev inequality, *J. Funct. Anal.* 267 (2014), no. 11, 4110–4138.

[10] L. Caffarelli, A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity, *Ann. of Math.* 131 (1990), 129–134.

[11] L. Caffarelli, The regularity of mappings with a convex potential, *J. Amer. Math. Soc.* 5 (1992), 99–104.

[12] E. Carlen, Some integral identities and inequalities for entire functions and their application to the coherent state transform, *J. Funct. Anal.* 97 (1991), 231–249.

[13] E. Carlen, Superadditivity of Fisher's information and logarithmic Sobolev inequalities, *J. Funct. Anal.* 101 (1991), 194–211.

[14] M. Christ, A sharpened Hausdorff-Young inequality, arXiv:1406.1210.

[15] A. Cianchi, N. Fusco, F. Maggi, and A. Pratelli, The sharp Sobolev inequality in quantitative form, *J. Eur. Math. Soc.* 11 (2009), 1105–1139.

[16] D. Cordero-Erausquin, Some applications of mass transport to Gaussian type inequalities, *Arch. Rational Mech. Anal.* 161 (2002), 257–269.

[17] G. De Philipis and A. Figalli, $W^{2,1}$ regularity of solutions to the Monge-Ampère equation, *Invent. Math.* 192 (2013), 55-69.

[18] R. Eldan, A two-sided estimate for the Gaussian noise stability deficit, *Invent. Math.*, (2015) 201: 561.

[19] A. Figalli and E. Indrei, A sharp stability result for the relative isoperimetric inequality inside convex cones, *J. Geom. Anal.* 23 (2013), 938–969.

[20] A. Figalli and D. Jerison, Quantitative stability for the Brunn-Minkowski inequality, *J. Eur. Math. Soc.*, to appear.

[21] A. Figalli, F. Maggi and A. Pratelli, A refined Brunn-Minkowski inequality for convex sets, *Ann. Inst. H. Poincaré Anal. Non Linéaire* 26 (2009), 2511–2519.
[22] A. Figalli, F. Maggi and A. Pratelli, A mass transportation approach to quantitative isoperimetric inequalities, *Invent. Math.* 182 (2010), 167–211.
[23] A. Figalli, F. Maggi and A. Pratelli, Sharp stability theorems for the anisotropic Sobolev and log-Sobolev inequalities on functions of bounded variation, *Adv. Math.* 242 (2013), 80–101.
[24] J. Fontbona, N. Gozlan and J.-F. Jabir, A variational approach to some transport inequalities, *preprint* (2015).
[25] N. Fusco, F. Maggi and A. Pratelli, The sharp quantitative isoperimetric inequality, *Ann. of Math.* 168 (2008), 941–980.
[26] A. Figalli and R. Neumayer, Gradient stability for the Sobolev inequality: the case $p \geq 2$, *preprint* (2015).
[27] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.* 97 (1975), 1061–1083.
[28] E. Indrei, A sharp lower bound on the polygonal isoperimetric deficit, *Proc. Amer. Math. Soc.*, DOI: 10.1090/proc/12947 (2015).
[29] E. Indrei and D. Marcon, A quantitative log-Sobolev inequality for a two parameter family of functions, *Int. Math. Res. Not. IMRN*, DOI: 10.1093/imrn/rnt138 (2013).
[30] E. Indrei and L. Nurbekyan, On the stability of the polygonal isoperimetric inequality, *Advances in Mathematics,* 276: 62-86, 2015.
[31] M. Ledoux, Logarithmic Sobolev inequalities for unbounded spin systems revisited, *Séminaire de Probabilités XXXV, Lecture Notes in Math* 1755 (2001), 167–194. Springer.
[32] J. Lehec, Representation formula for the entropy and functional inequalities, *Ann. IHP: Probab. Stat.* 43(3), 885–899 (2013).
[33] E. Lieb, Proof of an entropy conjecture of Wehrl, *Comm. Math. Phys.* 62 (1978), 35-41.
[34] E. Lieb, Thomas-Fermi and related theories of atoms and molecules, *Rev. Mod. Phys* 53 (1981), 603-641.
[35] R. J. McCann, Existence and uniqueness of monotone measure-preserving maps, *Duke Math. J.* 80 (1995), 309-323.
[36] C. Mooney, Partial regularity for singular solutions to the Monge-Ampere equation, *Comm. Pure Appl. Math.*, 68(2015), 1066-1084.
[37] F. Otto and C. Villani, Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* 173 (2000), 361–400.
[38] I. Segal, Mathematical characterization of the physical vacuum of the physical vacuum for a linear Bose-Einstein field, *Illinois J. Math.* 6 (1962), 500-523.
[39] I. Segal, Mathematical Problems in Relativistic Quantum Mechanics, *American Mathematical Society*, Providence, 1963.
[40] I. Segal, Construction of non-linear local quantum processes I, *Ann. of Math.* 92 (1970), 462-481.
[41] M. Talagrand, Transportation cost for Gaussian and other product measures, *Geom. Funct. Anal.* 6 (1996), 587–600.
[42] C. Villani, Topics in optimal transportation, *Graduate Studies in Mathematics 58, American Mathematical Society*, Providence, 2003.
[43] C. Villani, Optimal transport. Old and new, *Grundlehren der mathematischen Wissenschaften* 338, Springer, Berlin, 2009.
[44] A. Wehrl, On the relation between classical and quantum mechanical entropy, *Rep. Mat. Phys.* 16 (1979), 353-358.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: max.fathi@etu.upmc.fr
E-mail address: egi@cmu.edu
E-mail address: ledoux@math.univ-toulouse.fr