“More than fifteen proofs of the logarithmic Sobolev inequality”

Chapter 5 of the monograph *Analysis and geometry of Markov diffusion operators* is devoted to logarithmic Sobolev inequalities. It is mentioned on page 275, after C. Villani, *Optimal transport, old and new*, Springer (2009), page 559, that: “At present, there are more than fifteen known proofs of the logarithmic Sobolev inequality”\(^1\). It is the purpose of this note to try to identify these proofs (as of 2015)\(^2\).

Only the logarithmic Sobolev inequality for the Gaussian measure is considered. A main source of references are the surveys [8, 7, 9].

Let \( \gamma = \gamma_n \) be the standard Gaussian probability distribution on \( \mathbb{R}^n \), with density
\[
\frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2} |x|^2}, \quad x \in \mathbb{R}^n,
\]
with respect to the Lebesgue measure. The logarithmic Sobolev inequality (for \( \gamma \)) expresses that, for every smooth function \( f : \mathbb{R}^n \to \mathbb{R} \) with \( \int_{\mathbb{R}^n} f^2 \, d\gamma = 1 \),
\[
\int_{\mathbb{R}^n} f^2 \log f^2 \, d\gamma \leq 2 \int_{\mathbb{R}^n} |\nabla f|^2 \, d\gamma. \tag{1}
\]
(By homogeneity, for every smooth, square integrable, \( f : \mathbb{R}^n \to \mathbb{R} \),
\[
\int_{\mathbb{R}^n} f^2 \log f^2 \, d\gamma - \int_{\mathbb{R}^n} f^2 \, d\gamma \log \left( \int_{\mathbb{R}^n} f^2 \, d\gamma \right) \leq 2 \int_{\mathbb{R}^n} |\nabla f|^2 \, d\gamma.
\]
It is a simple matter to check that the inequality is sharp on the exponential functions \( f(x) = e^{\langle x, a \rangle - |a|^2}, \quad x \in \mathbb{R}^n \), where \( a \in \mathbb{R}^n \).

\(^1\)The number 15 seems to be taken from the survey [8].

\(^2\)Any relevant informations, references, omitted or new proofs, are most welcome (and will be incorporated).
This inequality can take various equivalent forms. For example, for any (smooth strictly positive) probability density $f$ with respect to $\gamma$ (i.e. $\int_{\mathbb{R}^n} f \, d\gamma = 1$),

$$\int_{\mathbb{R}^n} f \log f \, d\gamma \leq \frac{1}{2} \int_{\mathbb{R}^n} |\nabla f|^2 f \, d\gamma. \quad (2)$$

The right-hand side is the so-called Fisher information $I(\mu \mid \gamma)$ of the probability measure $d\mu = f \, d\gamma$ with respect to $\gamma$, so that the latter can also be formulated as

$$H(\mu \mid \gamma) \leq \frac{1}{2} I(\mu \mid \gamma) \quad (3)$$

where $H(\mu \mid \gamma) = \int_{\mathbb{R}^n} f \log f \, d\gamma$ is the relative entropy of $\mu$ with respect to $\gamma$.

It is an important feature that the inequality and the constants do not depend on the dimension of the underlying state space. By affine transformations, the logarithmic Sobolev inequality may be formulated for arbitrary Gaussian measures. Due to its dimension-free character, infinite dimensional Gaussian measures may also be considered.

After preliminary steps and formulations by


in a (one-dimensional) information theoretic framework (see 16.), and


in a mathematical physics context,

logarithmic Sobolev inequalities have been introduced and emphasized by L. Gross in the fundamental contribution


where, in particular, (1) is established.

A further main aspect of this work by L. Gross is the equivalence (in a general Dirichlet form framework) of logarithmic Sobolev inequalities with hypercontractivity properties. For the Gaussian measure $\gamma$, let

$$P_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}} y) \, d\gamma(y), \quad t > 0, \ x \in \mathbb{R}^n, \quad (4)$$
be the so-called Ornstein-Uhlenbeck semigroup (acting on suitable functions \( f : \mathbb{R}^n \rightarrow \mathbb{R} \)). By Jensen’s inequality, the operators \( P_t \) are contractions in all \( L^p(\gamma) \), \( p \geq 1 \), spaces. The logarithmic Sobolev inequality (1), holding for all smooth functions \( f \), is equivalent to the hypercontractivity property

\[ \| P_t f \|_q \leq \| f \|_p \]  

(5)

holding for all (measurable) \( f : \mathbb{R}^n \rightarrow \mathbb{R} \), \( 1 < p < q < \infty \), \( e^{2t} \geq \frac{q-1}{p-1} \) (where \( \| \cdot \|_p \) is the \( L^p(\gamma) \) norm). The proof of the equivalence (put forward by L. Gross, and already apparent in the aforementioned work of P. Federbush) amounts to take the time derivative of \( \| P_t f \|_{q(t)} \) where \( q(t) = 1 + e^{2t}(p - 1) \), which yields the logarithmic Sobolev inequality up to a power-type change of function.

The hypercontractive bound (5) was actually put forward first, before the logarithmic Sobolev inequality, by E. Nelson in


The proof of (5) therein is a reduction to a problem of multipliers for Hermite series, analyzed in the broader context of continuous Fock spaces. Several related papers in the mathematical physics literature appeared in the period (cf. \([8, 7]\)).

For the discussion, it is significant to emphasize the corresponding logarithmic Sobolev inequality and hypercontractivity property on the discrete cube \( \{-1, +1\}^n \), and more specifically on the two-point space \( \{-1, +1\} \), equipped with the uniform probability measure, where they amount to two-point inequalities. That is, for the logarithmic Sobolev inequality,

\[ \frac{1}{2} [a^2 \log a^2 + b^2 \log b^2] - \frac{1}{2} (a^2 + b^2) \log \left( \frac{1}{2} (a^2 + b^2) \right) \leq (a - b)^2 \]  

(6)

for every \( a, b \in \mathbb{R} \), and, for hypercontractivity,

\[ \left( \frac{1}{2} [a + e^{-t}b]^q + |a - e^{-t}b|^q \right)^{1/q} \leq \left( \frac{1}{2} [a + b]^p + |a - b|^p \right)^{1/p} \]  

(7)

for every \( a, b \in \mathbb{R} \), \( 1 < p < q < \infty \), \( e^{2t} \geq \frac{q-1}{p-1} \). Of course, according to Gross’ principle, these two families of inequalities are equivalent ((6) being the infinitesimal version of (7)). The point is that actually the two-point inequalities tensorize to the discrete cube \( \{-1, +1\}^n \), and then, after a suitable rescaling, yield the Gaussian inequalities via the classical central limit theorem, which is the original proof emphasized by L. Gross in his 1975 article (see e.g. \([1]\) for a detailed discussion). This reasoning has been very fruitful in the origin and proofs of the logarithmic Sobolev inequality and hypercontractivity.
From the same tensorization property, it is enough to prove the logarithmic Sobolev
inequality (1) and hypercontractivity (5) in dimension one, and many of the proofs presented
below are restricted to this case. Besides, the lists only emphasize proofs yielding optimal
constants. In this regard, it may be mentioned that a defective term in the logarithmic
Sobolev inequality may be accepted, provided it is dimension-free. More precisely, suppose
that it may be established that for every \( \varepsilon > 0 \), there is a constant \( C(\varepsilon) > 0 \) such that, for
every \( n \geq 1 \) and all smooth functions \( f : \mathbb{R}^n \to \mathbb{R} \) such that \( \int_{\mathbb{R}^n} f^2 d\gamma = 1 \),
\[
\int_{\mathbb{R}^n} f^2 \log f^2 \, d\gamma \leq (2 + \varepsilon) \int_{\mathbb{R}^n} |\nabla f|^2 \, d\gamma + C(\varepsilon),
\]
then (1) holds true. Indeed, given \( g : \mathbb{R} \to \mathbb{R} \) with \( \int_{\mathbb{R}} g^2 d\gamma = 1 \), application of the latter to \( f = g^{\otimes n} \) yields
\[
\int_{\mathbb{R}} g^2 \log g^2 \, d\gamma \leq (2 + \varepsilon) \int_{\mathbb{R}} g^2 \, d\gamma + \frac{C(\varepsilon)}{n}
\]
from which the claim follows after letting \( n \) tend to infinity and \( \varepsilon \) to zero.

Logarithmic Sobolev inequalities may be considered in a wide generality (cf. e.g. [2]).
Only the Gaussian logarithmic Sobolev inequality and hypercontractivity are considered in
this note. The Gaussian logarithmic Sobolev inequality is sometimes considered, and proved,
in its Euclidean (dimensional) form, with respect to the Lebesgue measure (obtained after a
simple, equivalent, change of function – seemingly first emphasized in [17]),
\[
\int_{\mathbb{R}^n} f^2 \log f^2 \, dx \leq \frac{n}{2} \log \left( \frac{2}{n \pi e} \int_{\mathbb{R}^n} |\nabla f|^2 \, dx \right)
\]  
(8)
for every smooth function \( f : \mathbb{R}^n \to \mathbb{R} \) such that \( \int_{\mathbb{R}^n} f^2 \, dx = 1 \).

The list of proofs is divided into proofs of hypercontractivity (historically first) and
proofs of the logarithmic Sobolev inequality. For each of them, a two-line description of
the corresponding underlying argument is presented. Some works address at the same time
hypercontractivity and logarithmic Sobolev inequality. While only the Gaussian logarithmic
Sobolev inequality is classified, it should be mentioned that several proofs apply similarly
to the family of log-concave probability measures \( d\mu = e^{-V} \, dx \) on \( \mathbb{R}^n \) where \( V : \mathbb{R}^n \to \mathbb{R} \)
is a potential which is more convex than the quadratic one in the sense that \( V(x) - \frac{c}{2} |x|^2 \),
\( x \in \mathbb{R}^n \), is convex for some \( c > 0 \) (as well as even more general settings cf. [2]).

Only the first references on a given type of proof are emphasized, hopefully not omitting
any and addressing proper acknowledgement. It is a bit difficult to settle whether there are
more than fifteen different proofs, since a certain amount of overlaps and similar arguments
in different contexts may be detected.

4
1 Hypercontractivity


Proofs of the two-point inequality (7).


Proof of hypercontractivity from a general integral inequality of Hölder type for functions composed along directions in \( \mathbb{R}^n \), established by symmetrization techniques\(^3\).


Martingale representation of the conditional expectation with respect to Brownian motion and use of stochastic calculus.


Proof of hypercontractivity via the study of Gaussian extremizers of operators given by Gaussian kernels.


Heat flow proof of both hypercontractivity and the logarithmic Sobolev inequality along the line of the Bakry-Émery proof of logarithmic Sobolev inequality (10.).

\(^3\)Numerous alternate proofs of this inequality, known as Brascamp-Lieb inequality, are available nowadays.
2 Logarithmic Sobolev inequality


Simple use of Jensen’s inequality and the fact that the lowest eigenfunction for a Sturm-Liouville boundary value problem with Dirichlet boundary conditions is positive.


Solving an optimization problem on the line by calculus of variations.


Analytic investigation of multipliers along ultraspherical polynomials yielding at the extremal cases the two-point inequality and the Gaussian logarithmic Sobolev inequality.


Consequence of a Gaussian Polyá-Szegő energy inequality established by Gaussian symmetrization.


Study of the second derivative of entropy along the Ornstein-Uhlenbeck semigroup, outlined, in a simplified form, in Section 3 (a scheme of proof valid in a general Markov triple context cf. [2]).


Deduction of the (Euclidean) logarithmic Sobolev inequality from the classical Euclidean isoperimetric inequality by the co-area formula.

—

Further, improved works in this direction include [6, 14].

Use of non-linear equation solved by extremal functions under the curvature condition fulfilled by Gauss space (in a compact setting though).


Deduction of the (Gaussian) logarithmic Sobolev inequality from the Gaussian isoperimetric inequality by the co-area formula.


Proof by duality and general rearrangement inequalities, of both hypercontractivity and logarithmic Sobolev inequality, exploiting competition between rearrangement on the sphere and Gauss space and the familiar Riesz rearrangement (partly inspired by E. Lieb’s proof [13] of the sharp Hardy-Littlewood-Sobolev inequalities).


Proof of the Euclidean logarithmic Sobolev inequality via the Fokker-Planck equation, similar in nature to the Bakry-Émery argument (10.), leading to a deficit term identifying the extremal functions$^5$.

---

$^5$see also [11] about the latter point.

The Euclidean logarithmic Sobolev inequality is developed as a geometric asymptotic estimate with respect to Lebesgue measure from the classical isoperimetric inequality in $\mathbb{R}^n$ in its functional form with sharp constants.

19. W. Beckner (as quoted in [12], p. 146).

Deduction from S. Bobkov’s functional form of the Gaussian isoperimetric inequality [5, 2].


Proof by stochastic calculus on Brownian paths, starting from the Clark-Ocone-Haussmann formula. Going back to unpublished works by B. Maurey and J. Neveu, and actually close in spirit to the Neveu proof of hypercontractivity.


Deduction from the functional form (Prékopa-Leindler) of the Brunn-Minkowski inequality at the edge of the parameters.


Proof of a strengthened form of the logarithmic Sobolev inequality linking entropy, Fisher information and Wasserstein distance (HWI inequality) by optimal transport and pde tools.


Proof of the Otto-Villani HWI inequality by means of the logarithmic form of Wang’s Harnack inequality [16], itself following from the Bakry-Émery scheme 10.


The Euclidean logarithmic Sobolev inequality appears as an extreme point in the proof of best constants in a family of Gagliardo-Nirenberg inequalities achieved by non-linear analysis.

Direct mass transportion methods (including the Otto-Villani HWI inequality).


Proof from a stochastic representation formula for Gaussian relative entropy in Wiener space, relying on stochastic calculus and Girsanov’s theorem.

### 3 The simplest proof?

In [8], L. Gross presents the 1978 proof by O. Rothaus as the simplest proof of the logarithmic Sobolev inequality. In 2010, he agreed that, probably, the simplest proof is the one by D. Bakry and M. Émery (1985) [10]. This section thus outlines this argument.

Let
\[ h_t(x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{1}{4t}||x||^2}, \quad t > 0, \quad x \in \mathbb{R}^n, \]
be the standard heat kernel, fundamental solution of the heat equation \( \partial_t h_t = \Delta h_t \). The convolution semigroup \( H_t f(x) = f * h_t(x), \ t > 0, \) solves \( \partial_t H_t f = \Delta H_t f = H_t \Delta f \) with initial data \( f \). At \( t = \frac{1}{2} \), \( h_t \) is just the standard Gaussian density so that \( H_{\frac{1}{2}} f(0) = \int_{\mathbb{R}^n} f \, d\gamma \) (while \( P_0 f = f \)).

For \( f > 0 \) smooth, \( t > 0, \) at any point, by the heat equation,
\[
H_t(f \log f) - H_t f \log H_t f
= \int_0^t \frac{d}{ds} H_s(H_{t-s} f \log H_{t-s} f) \, ds
= \int_0^t H_s \left( \Delta (H_{t-s} f \log H_{t-s} f) - \Delta H_{t-s} f \log H_{t-s} f - \Delta H_{t-s} f \right) \, ds
= \int_0^t H_s \left( \frac{\nabla H_{t-s} f}{H_{t-s} f} \right)^2 \, ds
\]
(since \( \Delta(g \log g) - \Delta g \log g - \Delta g = \frac{\nabla g^2}{g} \) for any smooth positive function \( g \)). Now, for any \( u > 0, \nabla H_u f = H_u (\nabla f) \), so that
\[
|\nabla H_u f|^2 \leq [H_u(|\nabla f|)]^2 \leq H_u \left( \frac{|\nabla f|^2}{f} \right) H_u f
\]
by the Cauchy-Schwarz inequality for the integral kernel \( H_u \). With \( u = t - s \), it follows that
\[
\frac{|\nabla H_{t-s} f|^2}{H_{t-s} f} \leq H_{t-s} \left( \frac{|\nabla f|^2}{f} \right),
\]
and hence

\[ H_t(f \log f) - H_t f \log H_t f \leq \int_0^t H_s \left( H_{t-s} \left( \frac{|\nabla f|^2}{f} \right) \right) ds = t H_t \left( \frac{|\nabla f|^2}{f} \right) \]

from the semigroup property. For \( t = \frac{1}{2} \), at \( x = 0 \) (for example), the latter amounts to (2). Thus, the logarithmic Sobolev inequality is (just) a consequence of the Cauchy-Schwarz inequality.

References


