Elementary proof of logarithmic Sobolev inequalities for Gaussian convolutions on $\mathbb{R}$


<http://ambp.cedram.org/item?id=AMBP_2016__23_1_129_0>
Elementary proof of logarithmic Sobolev inequalities for Gaussian convolutions on $\mathbb{R}$

David Zimmermann

Abstract

In a 2013 paper, the author showed that the convolution of a compactly supported measure on the real line with a Gaussian measure satisfies a logarithmic Sobolev inequality (LSI). In a 2014 paper, the author gave bounds for the optimal constants in these LSIs. In this paper, we give a simpler, elementary proof of this result.

1. Introduction

A probability measure $\mu$ on $\mathbb{R}^n$ is said to satisfy a logarithmic Sobolev inequality (LSI) with constant $c \in \mathbb{R}$ if

$$\text{Ent}_\mu(f^2) \leq c \mathcal{E}(f, f)$$

for all locally Lipschitz functions $f : \mathbb{R}^n \to \mathbb{R}_+$, where $\text{Ent}_\mu$, called the entropy functional, is defined as

$$\text{Ent}_\mu(f) := \int f \log \frac{f}{\int f \, d\mu} \, d\mu$$

Keywords: Logarithmic Sobolev inequality, convolutions.
and $\mathcal{E}(f,f)$, the energy of $f$, is defined as

$$\mathcal{E}(f,f) := \int |\nabla f|^2 d\mu,$$

with $|\nabla f|$ defined as

$$|\nabla f|(x) := \lim_{y \to x} \sup \frac{|f(x) - f(y)|}{|x - y|},$$

so that $|\nabla f|$ is defined everywhere and coincides with the usual notion of gradient where $f$ is differentiable. The smallest $c$ for which a LSI with constant $c$ holds is called the optimal log-Sobolev constant for $\mu$.

LSIs are a useful tool that have been applied in various areas of mathematics, such as geometry [1, 2, 5, 8, 9, 10, 14], probability [6, 11, 12, 13, 16], optimal transport [15, 17], and statistical physics [19, 20, 21]. In [23], the present author showed that the convolution of a compactly supported measure on $\mathbb{R}$ with a Gaussian measure satisfies a LSI, and an application of this fact to random matrix theory was given; that result, however, did not provide any quantitative information about the optimal log-Sobolev constants. In [22, Thms. 2 and 3], bounds for the optimal constants in these LSIs were given (stated as Theorem 1.1 below), and the results were extended to $\mathbb{R}^n$. (See [18] for statements about LSIs for convolutions with more general measures).

**Theorem 1.1.** Let $\mu$ be a probability measure on $\mathbb{R}$ whose support is contained in an interval of length $2R$, and let $\gamma_{\delta}$ be the centered Gaussian of variance $\delta > 0$, i.e., $d\gamma_{\delta}(t) = (2\pi\delta)^{-1/2} \exp(-\frac{t^2}{2\delta}) dt$. Then for some absolute constants $K_i$, the optimal log-Sobolev constant $c(\delta)$ for $\mu \ast \gamma_{\delta}$ satisfies

$$c(\delta) \leq K_1 \frac{\delta^{3/2} R}{4R^2 + \delta} \exp\left(\frac{2R^2}{\delta}\right) + K_2 (\sqrt{\delta} + 2R)^2.$$

In particular, if $\delta \leq R^2$, then

$$c(\delta) \leq K_3 \frac{\delta^{3/2}}{R} \exp\left(\frac{2R^2}{\delta}\right).$$

The $K_i$ can be taken in the above inequalities to be $K_1 = 6905, K_2 = 4989, K_3 = 7803$.

Theorem 1.1 was proved in [22] using the following theorem due to Bobkov and Götze [4, p. 25, Thm 5.3]:

$$130$$
**Theorem 1.2** (Bobkov, Götze). Let \( \mu \) be a Borel probability measure on \( \mathbb{R} \) with distribution function \( F(x) = \mu((\infty, x]) \). Let \( p \) be the density of the absolutely continuous part of \( \mu \) with respect to Lebesgue measure, and let \( m \) be a median of \( \mu \). Let

\[
D_0 = \sup_{x < m} \left( F(x) \cdot \log \frac{1}{F(x)} \cdot \int_x^m \frac{1}{p(t)} dt \right),
\]

\[
D_1 = \sup_{x > m} \left( (1 - F(x)) \cdot \log \frac{1}{1 - F(x)} \cdot \int_m^x \frac{1}{p(t)} dt \right),
\]

defining \( D_0 \) and \( D_1 \) to be zero if \( \mu((\infty, m]) = 0 \) or \( \mu((m, \infty)) = 0 \), respectively, and using the convention \( 0 \cdot \infty = 0 \). Then the optimal log Sobolev constant \( c \) for \( \mu \) satisfies

\[
\frac{1}{150} (D_0 + D_1) \leq c \leq 468(D_0 + D_1).
\]

**Remark 1.3.** In \( \mathbb{R}^n \), the analogue of Theorem 1.1 holds for measures supported in a ball of radius \( R \), with optimal log-Sobolev constant \( c(\delta) \) bounded by

\[
c(\delta) \leq K R^2 \exp \left( 20n + \frac{5R^2}{\delta} \right)
\]

for some absolute constant \( K \) and for \( \delta \leq R^2 \). This was proved in [22] using a theorem due to Cattiaux, Guillin, and Wu [7, Thm. 1.2] that gives satisfaction of a LSI under a Lyapunov condition.

The goal of the present paper is to provide an elementary proof of Theorem 1.1. The result proved is the following:

**Theorem 1.4.** Let \( \mu \) be a probability measure on \( \mathbb{R} \) whose support is contained in an interval of length \( 2R \), and let \( \gamma_\delta \) be the centered Gaussian of variance \( \delta > 0 \), i.e., \( d\gamma_\delta(t) = (2\pi\delta)^{-1/2}\exp(-\frac{t^2}{2\delta})dt \). Then the optimal log-Sobolev constant \( c(\delta) \) for \( \mu \ast \gamma_\delta \) satisfies

\[
c(\delta) \leq \max \left( 2\delta \exp \left( \frac{4R^2}{\delta} + \frac{4R}{\sqrt{\delta}} + \frac{1}{4} \right), 2\delta \exp \left( \frac{12R^2}{\delta} \right) \right).
\]

In particular, if \( \delta \leq 3R^2 \), we have

\[
c(\delta) \leq 2\delta \exp \left( \frac{12R^2}{\delta} \right).
\]
Remark 1.5. The bound in Theorem 1.4 is worse than the bound in Theorem 1.1 for small $\delta$, but still has an order of magnitude that is exponential in $R^2/\delta$. (It is shown in [22, Example 21] that one cannot do better than exponential in $R^2/\delta$ for small $\delta$.)

Remark 1.6. In fact, the proof of Theorem 1.4 yields a slightly stronger result. The proof is based upon showing that the convolution of the compactly supported measure with the Gaussian is the push-forward of the Gaussian under a Lipschitz map. This fact, together with the Gaussian isoperimetric inequality, yields the isoperimetric inequality for the convolution measure, which implies the logarithmic Sobolev inequality; see [3], in which Bakry and Ledoux show that a probability measure on $\mathbb{R}$ satisfies this isoperimetric inequality if and only if the measure is a Lipschitz push-forward of the Gaussian.)

The author would like to thank his Ph.D. advisor, Todd Kemp, for his valuable insights and discussions regarding this topic. The author would also like to thank the anonymous referee for his or her corrections and suggestions (especially Remark 1.6) that have improved the overall presentation and exposition of this paper.

2. Proof of Theorem 1.4

The proof of Theorem 1.4 is based on two facts: first, the Gaussian measure $\gamma_1$ of unit variance satisfies a LSI with constant $2$. Second, Lipschitz functions preserve LSIs. We give a precise statement of this second fact below.

Proposition 2.1. Let $\mu$ be a measure on $\mathbb{R}$ that satisfies a LSI with constant $c$, and let $T: \mathbb{R}^n \to \mathbb{R}^n$ be Lipschitz. Then the push-forward measure $T_* \mu$ also satisfies a LSI with constant $c \|T\|_{\text{Lip}}^2$.

Proof. Let $g: \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz. Then $g \circ T$ is locally Lipschitz, so by the LSI for $\mu$,

$$\int (g \circ T)^2 \log \frac{(g \circ T)^2}{\int (g \circ T)^2 \, d\mu} \, d\mu \leq c \int |\nabla (g \circ T)|^2 \, d\mu.$$  \hspace{1cm} (2.1)

But since $T$ is Lipschitz,

$$|\nabla (g \circ T)| \leq (|\nabla g| \circ T) \|T\|_{\text{Lip}}.$$
So by a change of variables, (2.1) simply becomes
\[ \int g^2 \log g^2 dT* \mu \leq c \|T\|^2_{\text{Lip}} \int |\nabla g|^2 dT* \mu. \]
as desired. \qed

We now prove Theorem 1.4.

**Proof of Theorem 1.4.** In light of Proposition 2.1, we will establish the theorem by showing that $\mu * \gamma_\delta$ is the push-forward of $\gamma_1$ under a Lipschitz map. By translation invariance of LSI, we can assume that $\text{supp}(\mu) \subseteq [-R, R]$. We will also first assume that $\delta = 1$ (the general case will be handled at the end of the proof by a scaling argument).

Let $F$ and $G$ be the cumulative distribution functions of $\gamma_1$ and $\mu * \gamma_1$, i.e.,
\[ F(x) = \int_{-\infty}^{x} p(t) \, dt, \quad G(x) = \int_{-\infty}^{x} q(t) \, dt, \]
where
\[ p(t) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{t^2}{2} \right) \quad \text{and} \quad q(t) = \int_{-R}^{R} p(t - s) \, d\mu(s). \]

Notice that $q$ is smooth and strictly positive, so that $G^{-1} \circ F$ is well-defined and smooth. It is readily seen that $(G^{-1} \circ F)(\gamma_1) = \mu * \gamma_1$, so to establish the theorem we simply need to bound the derivative of $G^{-1} \circ F$.

Now
\[ (G^{-1} \circ F)'(x) = \frac{1}{G'(G^{-1} \circ F)(x))} \cdot F'(x) = \frac{p(x)}{q((G^{-1} \circ F)(x))}. \]

We will bound the above derivative in cases – when $x \geq 2R$, when $-2R \leq x \leq 2R$, and when $x \leq -2R$.

We first consider the case $x \geq 2R$. Define
\[ \Lambda(x) = \int_{-R}^{R} e^{xs} d\mu(s), \quad K(x) = \frac{\log \Lambda(x) + R}{x}. \]

Note $\Lambda$ and $K$ are smooth for $x \neq 0$.

**Lemma 2.2.** For $x \geq 2R$, 
\[ \exp \left( -2R^2 - 2R - \frac{1}{8} \right) p(x) \leq q(x + K(x)) \leq e^{-R} p(x). \]
Proof. By definition of \( q, p, \Lambda, \) and \( K, \)

\[
q(x + K(x)) = \int_{-R}^{R} p(x + K(x) - s) \, d\mu(s) \\
= p(x) \cdot e^{-xK(x)} \int_{-R}^{R} \exp \left( -\frac{(K(x) - s)^2}{2} \right) \cdot e^{xs} \, d\mu(s) \\
= \frac{e^{-R} p(x)}{\Lambda(x)} \int_{-R}^{R} \exp \left( -\frac{(K(x) - s)^2}{2} \right) \cdot e^{xs} \, d\mu(s) \\
\leq \frac{e^{-R} p(x)}{\Lambda(x)} \int_{-R}^{R} e^{xs} \, d\mu(s) \\
= e^{-R} p(x).
\]

To get the other inequality, first note that \( e^{-Rx} \leq \Lambda(x) \leq e^{Rx}. \) (These are just the maximum and minimum values in the integrand defining \( \Lambda. \)) This implies that \(-R + R/x \leq K(x) \leq R + R/x, \) so for \(-R \leq s \leq R\) and \(x \geq 2R,\) we have

\[
-2R - \frac{R}{x} \leq -2R + \frac{R}{x} \leq K(x) - s \leq 2R + \frac{R}{x}
\]

so that

\[
\exp \left( -\frac{(K(x) - s)^2}{2} \right) \geq \exp \left( -\frac{(2R + R/x)^2}{2} \right) \geq \exp \left( -\frac{(2R + R/(2R))^2}{2} \right) = \exp \left( -2R^2 - R - \frac{1}{8} \right).
\]

Therefore

\[
q(x + K(x)) = \frac{e^{-R} p(x)}{\Lambda(x)} \int_{-R}^{R} \exp \left( -\frac{(K(x) - s)^2}{2} \right) \cdot e^{xs} \, d\mu(s) \\
\geq \exp \left( -2R^2 - 2R - \frac{1}{8} \right) p(x). \quad \Box
\]

Lemma 2.3. \( K'(x) \leq R \) for \( x > 0. \)
Proof. Recall that $e^{-Rx} \leq \Lambda(x)$. (Again, $e^{-Rx}$ is the minimum value in the integrand defining $\Lambda$). We therefore have

$$K'(x) = \frac{\Lambda'(x)}{x\Lambda(x)} - \frac{\log \Lambda(x)}{x^2} - \frac{R}{x^2} - \frac{\int_{-R}^{R} s e^{sx} d\mu(s)}{x\Lambda(x)} - \frac{\log \Lambda(x)}{x^2} - \frac{R}{x^2}$$

$$\leq \frac{R \int_{-R}^{R} e^{sx} d\mu(s)}{x\Lambda(x)} + \frac{Rx}{x^2} - \frac{R}{x^2}$$

$$= \frac{2R}{x} - \frac{R}{x^2}.$$

By elementary calculus, the above has a maximum value of $R$. □

Lemma 2.4. For $x \geq 2R$,

$$x - R \leq (G^{-1} \circ F)(x) \leq x + K(x).$$

Proof. Since $G$ and $G^{-1}$ are increasing, the lemma is equivalent to

$$G(x - R) \leq F(x) \leq G(x + K(x)).$$

The first inequality follows from the definition of $G$ and the Fubini-Tonelli Theorem:

$$G(x - R) = \int_{-\infty}^{x-R} q(t) \, dt = \int_{-\infty}^{x-R} \int_{-R}^{R} p(t-s) \, d\mu(s) \, dt$$

$$= \int_{-R}^{R} \int_{-\infty}^{x-R} p(t-s) \, dt \, d\mu(s)$$

$$= \int_{-R}^{R} \int_{-\infty}^{x-R-s} p(u) \, du \, d\mu(s)$$

where $u = t - s$

$$\leq \int_{-R}^{R} \int_{-\infty}^{x} p(u) \, du \, d\mu(s)$$

$$= F(x).$$
To establish the other inequality, we use Lemmas 2.2 and 2.3:

\[ 1 - G(x + K(x)) = \int_{x+K(x)}^{\infty} q(t) \, dt = \int_{x}^{\infty} q(u + K(u))(1 + K'(u)) \, du \]

where \( t = u + K(u) \)

\[ \leq \int_{x}^{\infty} p(u)e^{-R}(1 + R) \, du \]

by Lemmas 2.2 and 2.3

\[ \leq \int_{x}^{\infty} p(u) \, du \]

since \( e^R \geq 1 + R \)

\[ = 1 - F(x), \]

so that \( F(x) \leq G(x + K(x)) \), as desired. \( \square \)

We are almost ready to bound \( (G^{-1} \circ F)'(x) \) for \( x \geq 2R \). The last observation to make is that \( q \) is decreasing on \([R, \infty)\) since

\[ q'(t) = \int_{-R}^{R} p'(t - s) \, d\mu(s) = \int_{-R}^{R} -(t - s)p(t - s) \, d\mu(s) \leq 0 \quad \text{for } t \geq R. \]

So for \( x \geq 2R \) we have, by Lemma 2.4,

\[ q((G^{-1} \circ F)(x)) \geq q(x + K(x)). \]

Combining this with Lemma 2.2, we get

\[ (G^{-1} \circ F)'(x) = \frac{p(x)}{q((G^{-1} \circ F)(x))} \leq \frac{p(x)}{q(x + K(x))} \leq \exp \left( 2R^2 + 2R + \frac{1}{8} \right) \]

for \( x \geq 2R \).

In the case where \(-2R \leq x \leq 2R\), first note that for all \( x \),

\[ x - R \leq (G^{-1} \circ F)(x) \leq x + R; \]

136
the first inequality above was done in Lemma 2.4, and the second inequality is proven in the same way. So
\[
\sup_{-2R \leq x \leq 2R} (G^{-1} \circ F)'(x) = \sup_{-2R \leq x \leq 2R} \frac{p(x)}{q((G^{-1} \circ F)(x))} \leq \sup_{-2R \leq x \leq 2R \atop -R \leq y \leq R} \frac{p(x)}{q(x + y)} \leq \left( \inf_{-2R \leq x \leq 2R \atop -R \leq y \leq R} \frac{q(x+y)}{p(x)} \right)^{-1}.
\]

For convenience, let \( S = \{(x, y) : -2R \leq x \leq 2R, -R \leq y \leq R\} \). Now
\[
\inf_{(x,y) \in S} \frac{q(x+y)}{p(x)} = \inf_{(x,y) \in S} \frac{1}{p(x)} \int_{-R}^{R} p(x+y-s) \, d\mu(s).
\]

Since \( p \) has no local minima, the minimum value of the above integrand occurs at either \( s = R \) or \( s = -R \). Without loss of generality, we assume the minimum is achieved at \( s = R \) (otherwise, we can replace \((x, y)\) with \((-x, -y)\) by symmetry of \( S \) and \( p \)). So
\[
\inf_{(x,y) \in S} \frac{q(x+y)}{p(x)} \geq \inf_{(x,y) \in S} \frac{1}{p(x)} \cdot p(x + y + R).
\]

Elementary calculus shows that the above infimum is equal to \( e^{-6R^2} \) (achieved at \( x = 2R, y = R \)). Therefore
\[
\sup_{-2R \leq x \leq 2R} (G^{-1} \circ F)'(x) \leq \left( \inf_{(x,y) \in S} \frac{q(x+y)}{p(x)} \right)^{-1} \leq e^{6R^2}.
\]

The case \( x \leq -2R \) is dealt with in the same way as the case \( x \geq 2R \), the analogous statements being:
\[
\exp \left( -2R^2 - 2R - \frac{1}{8} \right) p(x) \leq q(x + K(x)) \leq e^{-R} p(x),
\]
\[
K'(x) \leq R,
\]
\[
x + K(x) \leq (G^{-1} \circ F)(x) \leq x + R,
\]
and \( q \) is increasing for \( x \leq -2R \). The upper bound for \( (G^{-1} \circ F)'(x) \) obtained in this case is the same as the one in the case \( x \geq 2R \).
We therefore have
\[ ||G^{-1} \circ F||_{\text{Lip}} \leq \max \left( \exp \left( 2R^2 + 2R + \frac{1}{8} \right), e^{6R^2} \right) \]
So by Proposition 2.1, \( \mu \ast \gamma_1 \) satisfies a LSI with constant \( c(1) \) satisfying
\[ c(1) \leq 2||G^{-1} \circ F||_{\text{Lip}}^2 \leq \max \left( 2 \exp \left( 4R^2 + 4R + \frac{1}{4} \right), 2e^{12R^2} \right). \]
This proves the theorem for the case \( \delta = 1 \).

To establish the theorem for a general \( \delta > 0 \), first observe that
\[ \mu \ast \gamma_\delta = (h_{\sqrt{\delta}})_* \left( ((h_{1/\sqrt{\delta}})_* \mu) \ast \gamma_1 \right), \]
where \( h_\lambda \) denotes the scaling map with factor \( \lambda \), i.e., \( h_\lambda(x) = \lambda x \). Now \( (h_{1/\sqrt{\delta}})_* \mu \) is supported in \([-R/\sqrt{\delta}, R/\sqrt{\delta}]\), so by the case \( \delta = 1 \) just proven, \( ((h_{1/\sqrt{\delta}})_* \mu) \ast \gamma_1 \) satisfies a LSI with constant
\[ \max \left( 2 \exp \left( 4(R/\sqrt{\delta})^2 + 4(R/\sqrt{\delta}) + \frac{1}{4} \right), 2e^{12(R/\sqrt{\delta})^2} \right). \]
Finally, since \( ||h_{\sqrt{\delta}}||_{\text{Lip}}^2 = \delta \), we have by Proposition 2.1,
\[ c(\delta) \leq \max \left( 2\delta \exp \left( \frac{4R^2}{\delta} + \frac{4R}{\sqrt{\delta}} + \frac{1}{4} \right), 2\delta \exp \left( \frac{12R^2}{\delta} \right) \right). \]
In particular, when \( \delta \leq 3R^2 \) (in fact when \( \delta \leq (160 - 64\sqrt{6})R^2 \approx 3.23R^2 \)), we have
\[ 2\delta \exp \left( \frac{4R^2}{\delta} + \frac{4R}{\sqrt{\delta}} + \frac{1}{4} \right) \leq 2\delta \exp \left( \frac{12R^2}{\delta} \right) \]
so the above bound on \( c(\delta) \) simplifies to
\[ c(\delta) \leq 2\delta \exp \left( \frac{12R^2}{\delta} \right). \]
\[ \square \]

References

log-Sobolev inequalities for convolutions


David Zimmermann
Department of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093, USA
dszimmer@ucsd.edu