Poincaré and log-Sobolev inequalities for mixtures

André Schlichting ⬤0000-0003-4140-491X

Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Templergraben 55, 52056 Aachen, Germany
schlichting@igpm.rwth-aachen.de

Abstract: This work studies mixtures of probability measures on $\mathbb{R}^n$ and gives bounds on the Poincaré and the log-Sobolev constants of two-component mixtures provided that each component satisfies the functional inequality, and both components are close in the χ^2-distance. The estimation of those constants for a mixture can be far more subtle than it is for its parts. Even mixing Gaussian measures may produce a measure with a Hamiltonian potential possessing multiple wells leading to metastability and large constants in Sobolev type inequalities. In particular, the Poincaré constant stays bounded in the mixture parameter whereas the log-Sobolev may blow up as the mixture ratio goes to 0 or 1. This observation generalizes the one by Chafaï and Malrieu to the multidimensional case. The behavior is shown for a class of examples to be not only a mere artifact of the method.

Keywords: Poincaré inequality; log-Sobolev inequality; relative entropy; Fisher information; Dirichlet form; mixture; finite Gaussian mixtures.

---

1. Introduction

A mixture of two probability measures μ_0 and μ_1 on $\mathbb{R}^n$ is for some parameter $p \in [0, 1]$ the probability measure μ_p defined by

$$\mu_p := p\mu_0 + (1 - p)\mu_1 . \tag{1}$$

Hereby, both measures μ_0 and μ_1 are assumed to be absolutely continuous with respect to the Lebesgue measure and their supports are nested, i.e. $\operatorname{supp} \mu_0 \subseteq \operatorname{supp} \mu_1$ or $\operatorname{supp} \mu_1 \subseteq \operatorname{supp} \mu_0$. Under these assumptions at least one measure is absolutely continuous to the other one

$$\mu_0 \ll \mu_1 \qquad \text{or} \qquad \mu_1 \ll \mu_0 \, ,$$

which implies that at least one of the measures has a density with respect to the other one

$$\mathrm{d}\mu_0 = \frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_1} \, \mathrm{d}\mu_1 \qquad \text{or} \qquad \mathrm{d}\mu_1 = \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0} \, \mathrm{d}\mu_0 \, .$$

This work establishes criteria to check in a simple way under which a mixture of measures satisfies a Poincaré or log-Sobolev inequality provided that each of the components satisfies one.

Definition 1 (PI(ϱ) and LSI(α))**.** *A probability measure μ on $\mathbb{R}^n$ satisfies the* Poincaré inequality *with constant $\varrho > 0$, if for all functions $f : \mathbb{R}^n \to \mathbb{R}$*

$$\operatorname{Var}_\mu[f] := \int \left| f - \int f \, \mathrm{d}\mu \right|^2 \mathrm{d}\mu \leq \frac{1}{\varrho} \int |\nabla f|^2 \, \mathrm{d}\mu \, . \qquad \text{PI}(\varrho)$$

A probability measure μ satisfies the log-Sobolev inequality *with constant $\alpha > 0$, if for all functions $f : \mathbb{R}^n \to \mathbb{R}^+$ holds*

$$\operatorname{Ent}_\mu[f] := \int f \log f \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu \, \log\left(\int f \, \mathrm{d}\mu \right) \leq \frac{1}{\alpha} \int \frac{|\nabla f|^2}{2f} \, \mathrm{d}\mu \, . \qquad \text{LSI}(\alpha)$$

By the change of variable $f \mapsto f^2$ the log-Sobolev inequality $\mathrm{LSI}(\alpha)$ is equivalent to

$$\mathrm{Ent}_\mu\left[f^2\right] \leq \frac{2}{\alpha} \int |\nabla f|^2 \, \mathrm{d}\mu. \tag{2}$$

The question of how ϱ_p and α_p in $\mathrm{PI}(\varrho_p)$ and $\mathrm{LSI}(\alpha_p)$ depend for a mixture μ_p on the parameter $p \in [0,1]$ was first studied by Chafaï and Malrieu [1] for measures on $\mathbb{R}^n$. The aim is to deduce simple criteria under which the measure μ_p (1) satisfies $\mathrm{PI}(\varrho_p)$ and $\mathrm{LSI}(\alpha_p)$ knowing that μ_0 and μ_1 satisfy $\mathrm{PI}(\varrho_0)$, $\mathrm{PI}(\varrho_1)$ and $\mathrm{LSI}(\alpha_0)$, $\mathrm{LSI}(\alpha_1)$, respectively. The approach by Chafaï and Malrieu [1] is based on a functional depending on the distribution function of the measures μ_0 and μ_1, which then lead to bounds on the Poincaré and log-Sobolev constant of the mixture in one dimension.

This work generalizes part of the results from Chafaï and Malrieu [1] to the multidimensional case by a simple argument. The estimates on the Poincaré and log-Sobolev constants hold for the case, when the χ^2-distance of μ_0 and μ_1 is bounded (see (5) for its definition). For this to be true, at least one of the measures μ_0 and μ_1 needs to be absolutely continuous to the other, which is also a necessary condition for the mixture having connected support. The resulting bound is optimal in the scaling behavior of the mixture parameter $p \to 0, 1$, i.e. a logarithmic blow-up behavior in p for the log-Sobolev constant, whereas the Poincaré constant stays bounded. This different behavior of the Poincaré and log-Sobolev constant was also observed in the setting of metastability in [2, Remark 2.20].

Let us first introduce the principle for the Poincaré inequality in Section 2 and then for the log-Sobolev inequality in Section 3. Then, the procedure is illustrated on specific examples of mixtures in Section 4.

2. Poincaré inequality

To keep the presentation concise, the following notation for the mean of a function $f : \mathbb{R}^n \to \mathbb{R}$ with respect to a measure μ is introduced

$$\mathrm{E}_\mu[f] := \int f \, \mathrm{d}\mu.$$

In this way, the variance in $\mathrm{PI}(\varrho)$ and relative entropy in $\mathrm{LSI}(\alpha)$ become

$$\mathrm{Var}_\mu[f] = \mathrm{E}_\mu\left[\left(f - \mathrm{E}_\mu[f]\right)^2\right] = \mathrm{E}_\mu[f] - \left(\mathrm{E}_\mu[f]\right)^2 \quad \text{and} \quad \mathrm{Ent}_\mu[f] = \mathrm{E}_\mu[f \log f] - \mathrm{E}_\mu[f] \log\left(\mathrm{E}_\mu[f]\right).$$

Likewise, the covariance of two functions $f, g : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\mathrm{Cov}_\mu[f, g] = \mathrm{E}_\mu\left[\left(f - \mathrm{E}_\mu[f]\right)\left(g - \mathrm{E}_\mu[g]\right)\right] = \mathrm{E}_\mu[f \, g] - \mathrm{E}_\mu[f] \, \mathrm{E}_\mu[g].$$

The Cauchy-Schwarz inequality for the covariance takes now the form

$$\mathrm{Cov}_\mu[f, g] \leq \mathrm{Var}_\mu[f] \, \mathrm{Var}_\mu[g].$$

The argument is based on an easy but powerful observation for measures μ_0 and μ_1 with joint support.

Lemma 1 (Mean-difference as covariance). *If* $\mathrm{supp}\,\mu_0 = \mathrm{supp}\,\mu_1$, *then for any* $\vartheta \in [0,1]$ *and any function* $f : \mathbb{R}^n \to \mathbb{R}$ *holds*

$$\mathrm{E}_{\mu_0}[f] - \mathrm{E}_{\mu_1}[f] = -\vartheta \, \mathrm{Cov}_{\mu_0}\left[f, \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right] + (1 - \vartheta) \, \mathrm{Cov}_{\mu_1}\left[f, \frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_1}\right]. \tag{3}$$

Proof. The change of measure formula yields that the covariances above are just the difference of the expectation on the right-hand side

$$\mathrm{Cov}_{\mu_0}\left[f, \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right] = \mathrm{E}_{\mu_0}\left[f\,\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right] - \mathrm{E}_{\mu_0}[f]\,\mathrm{E}_{\mu_0}\left[\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right] = \mathrm{E}_{\mu_1}[f] - \mathrm{E}_{\mu_0}[f]$$

and likewise for $\mathrm{Cov}_{\mu_1}\left[f, \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right]$. $\square$

The subsequent strategy is based on (3) by using a Cauchy-Schwarz inequality to arrive at the product of two variances. Then, $\mathrm{PI}(\varrho_0)$ or $\mathrm{PI}(\varrho_1)$ can be applied and the parameter ϑ leaves freedom to optimize the resulting expression. This allows to prove the following theorem, which is the generalization of [1, Theorem 4.4] to the multidimensional case for the Poincaré inequality provided μ_0 and μ_1 are absolutely continuous to each other.

Theorem 1 (PI for absolutely continuous mixtures). *Let μ_0 and μ_1 satisfy $\mathrm{PI}(\varrho_0)$ and $\mathrm{PI}(\varrho_1)$, respectively, and let both measures be absolutely continuous to each other. Then for all $p \in (0,1)$ and $q = 1 - p$ the mixture measure $\mu_p = p\,\mu_0 + q\,\mu_1$ satisfies $\mathrm{PI}(\varrho_p)$ with*

$$\frac{1}{\varrho_p} \leq \begin{cases} \dfrac{1}{\varrho_0}, & \text{if } \dfrac{\varrho_1}{\varrho_0} \geq 1 + p\chi_1 \\[2mm] \dfrac{1}{\varrho_1}, & \text{if } \dfrac{\varrho_0}{\varrho_1} \geq 1 + q\chi_0 \\[2mm] \dfrac{p\,\chi_1 + p\,q\,\chi_0\,\chi_1 + q\,\chi_0}{\varrho_0\,p\,\chi_1 + \varrho_1\,q\,\chi_0}, & \text{else.} \end{cases} \tag{4}$$

where

$$\chi_0 := \mathrm{Var}_{\mu_0}\left[\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right] \qquad \text{and} \qquad \chi_1 := \mathrm{Var}_{\mu_1}\left[\frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_1}\right]. \tag{5}$$

Proof. The variance of f with respect to μ_p is decomposed to

$$\mathrm{Var}_{\mu_p}[f] = p\,\mathrm{Var}_{\mu_0}[f] + q\,\mathrm{Var}_{\mu_1}[f] + p\,q\left(\mathrm{E}_{\mu_0}[f] - \mathrm{E}_{\mu_1}[f]\right)^2.$$

Hereby, the first two terms are just the expectation of the conditional variances. The third term is the variance of a Bernoulli random variable. Now, the mean-difference is rewritten by Lemma 1 and the square is estimated with the Young inequality introducing an additional parameter $\eta > 0$

$$(a+b)^2 \leq (1+\eta)\,a^2 + (1+\eta^{-1})\,b^2.$$

Then, the Cauchy-Schwarz inequality is applied to the covariances to obtain

$$\begin{aligned}
\mathrm{Var}_{\mu}[f] \leq{}& p\,\mathrm{Var}_{\mu_0}[f] + q\,\mathrm{Var}_{\mu_1}[f] + \\
&+ p\,q\left((1+\eta)\,\vartheta^2\,\mathrm{Cov}_{\mu_0}^2\left[f, \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\right] + \left(1+\eta^{-1}\right)(1-\vartheta)^2\,\mathrm{Cov}_{\mu_1}\left[f, \frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_1}\right]\right) \\
\leq{}& \left(1 + (1+\eta)\,\vartheta^2\,q\,\chi_0\right)p\,\mathrm{Var}_{\mu_0}[f] + \left(1 + \left(1+\eta^{-1}\right)(1-\vartheta)^2\,p\,\chi_1\right)q\,\mathrm{Var}_{\mu_1}[f] \\
\leq{}& \frac{1+(1+\eta)\,\vartheta^2\,q\,\chi_0}{\varrho_0}\int|\nabla f|^2\,p\,\mathrm{d}\mu_0 + \frac{1+\left(1+\eta^{-1}\right)(1-\vartheta)^2\,p\,\chi_1}{\varrho_1}\int|\nabla f|^2\,q\,\mathrm{d}\mu_1 \\
\leq{}& \max\left\{\frac{1+(1+\eta)\,\vartheta^2\,q\,\chi_0}{\varrho_0}, \frac{1+\left(1+\eta^{-1}\right)(1-\vartheta)^2\,p\,\chi_1}{\varrho_1}\right\}\int|\nabla f|^2\,\mathrm{d}\mu.
\end{aligned} \tag{6}$$

The resulting maximum is now minimized in $\eta > 0$ and $\vartheta \in [0,1]$. To do so without loss of generality $\varrho_0 \geq \varrho_1$ is assumed. The other case can always be obtained by interchanging the roles of μ_0 and μ_1. If $\varrho_0 > \varrho_1$, then $\vartheta = 1$ and $\eta \to 0$ is optimal as long as

$$\frac{1 + q\chi_0}{\varrho_0} \leq \frac{1}{\varrho_1}.$$

This corresponds to the second case in (4). By symmetry the first case follows if $\varrho_1 \geq \varrho_0$.

Now, in the case $\varrho_0 \geq \varrho_1$ and $\varrho_0 \leq (1 + q\chi_0)\varrho_1$ there exists by monotonicity for every $\vartheta \in (0,1)$ a unique $\eta_* = \eta_*(\vartheta) > 0$ such that both terms in the max of the right-hand side in (6) are equal and hence the max is minimal. Since $q\chi_0 > 0$ and $p\chi_1 > 0$, the sum of the coefficients in front is then given by $h(\vartheta) := (1 + \eta)\vartheta^2 + (1 + \frac{1}{\eta})(1 - \vartheta)^2$ in ϑ as a function of η. The minimization of h in $\vartheta \in (0,1)$ leads to $\vartheta_* = \frac{1}{1+\eta}$ and it holds

$$h(\vartheta^*) = \frac{1}{1+\eta} + \frac{\eta}{1+\eta} = 1.$$

Hence, in this caes the parameter $s = (1 + \eta_*)\vartheta_*^2 = \frac{1}{1+\eta_*} \in (0,1)$ and $(1 + \eta_*^{-1})(1 - \vartheta_*)^2 = \frac{\eta_*}{1+\eta_*} = 1 - s$. Thus, the problem can be rephrased: Find $s_* \in (0,1)$ which solves

$$\frac{1 + s\,q\chi_0}{\varrho_0} = \frac{1 + (1-s)\,p\chi_1}{\varrho_1}.$$

The solution s_* is given by

$$s_* = \frac{(1 + p\chi_1)\varrho_0 - \varrho_1}{\varrho_0\,p\chi_1 + \varrho_1\,q\chi_0}.$$

For this value of s_* the value of the max in (6) is given by

$$\frac{1 + s_*\,q\chi_0}{\varrho_0} = \frac{p\chi_1 + \frac{\varrho_1}{\varrho_0}\,q\chi_0 + (1 + p\,{,}\chi_1)\,q\chi_0 - \frac{\varrho_1}{\varrho_0}\,q\chi_0}{\varrho_0\,p\chi_1 + \varrho_1\,q\chi_0} = \frac{p\chi_1 + p\,q\chi_0\chi_1 + q\chi_0}{\varrho_0\,p\chi_1 + \varrho_1\,q\chi_0}. \qquad \square$$

Remark 1. *The constants χ_0 and χ_1 can be rewritten if μ_0 and μ_1 are mutual absolutely continuous as*

$$\chi_0 = \int\left(\frac{d\mu_1}{d\mu_0}\right)^2 d\mu_0 - 1 = \int\frac{d\mu_1}{d\mu_0}\,d\mu_1 - 1 \quad \textit{and} \quad \chi_1 = \int\left(\frac{d\mu_0}{d\mu_1}\right)^2 d\mu_1 - 1 = \int\frac{d\mu_0}{d\mu_1}\,d\mu_0 - 1.$$

This quantity is also known as χ^2-distance on the space of probability measures (cf. [3]). The χ^2-distance is a rather weak distance and therefore bounds many other probability distances. Among them is also the relative entropy. Indeed, by the concavity of the logarithm and the Jensen inequality follows

$$\mathrm{Ent}_{\mu_0}\left[\frac{d\mu_1}{d\mu_0}\right] = \int\log\frac{\mu_1}{\mu_0}\,d\mu_1 \leq \log\left(\int\frac{\mu_1}{\mu_0}\,d\mu_1\right) = \log(1 + \chi_0) \leq \chi_0,$$

Remark 2. *The proof of Theorem 1 shows that the expression for $\frac{1}{\varrho}$ in the last case of (4) can be bounded above and below by*

$$\max\left\{\frac{1}{\varrho_0}, \frac{1}{\varrho_1}\right\} \leq \frac{p\chi_1 + p\,q\chi_0\chi_1 + q\chi_0}{\varrho_0\,p\chi_1 + \varrho_1\,q\chi_0} \leq \max\left\{\frac{1 + q\chi_0}{\varrho_0}, \frac{1 + p\chi_1}{\varrho_1}\right\}. \tag{7}$$

In the case, where $\chi_0 = \chi_1 = \chi$, the formula for ϱ_p (4) simplifies to

$$\frac{1}{\varrho_p} \leq \frac{1 + p\,q\chi}{p\,\varrho_0 + q\,\varrho_1}. \tag{8}$$

Corollary 1. *Let $\mu_0 \ll \mu_1$ and μ_0, μ_1 satisfy $\mathrm{PI}(\varrho_0)$, $\mathrm{PI}(\varrho_1)$, respectively. Then for all $p \in [0,1]$ with $q = 1 - p$ the mixture measure $\mu_p = p\,\mu_0 + q\,\mu_1$ satisfies $\mathrm{PI}(\varrho_p)$ with*

$$\frac{1}{\varrho_p} = \max\left\{ \frac{1}{\varrho_0}, \frac{1 + p\,\chi_1}{\varrho_1} \right\}.$$

Likewise, if $\mu_1 \ll \mu_0$, then it holds

$$\frac{1}{\varrho_p} = \max\left\{ \frac{1}{\varrho_1}, \frac{1 + q\,\chi_0}{\varrho_0} \right\}.$$

Proof. The proof is a simple consequence of Lemma 1 with $\vartheta = 0$ and a similar line of estimates as in (6). $\square$

3. Log-Sobolev inequality

In this section a criterion for $\mathrm{LSI}(\alpha)$ is established. It will be convenient, to establish it in the form (2). For a function $g : \mathbb{R}^n \to \mathbb{R}^+$ and two probability measures μ_0 and μ_1 the averaged function $\bar{g} : \{0,1\} \to \mathbb{R}^+$ is defined by

$$\bar{g}(0) := \mathrm{E}_{\mu_0}[g] \quad \text{and} \quad \bar{g}(1) := \mathrm{E}_{\mu_1}[g] .$$

Moreover, the mixture of two Dirac measures δ_0 and δ_1 is by slight abuse of notation denoted by $\delta_p := p\,\delta_0 + q\,\delta_1$ for $p \in [0,1]$ and $q = 1 - p$. Then, the entropy of the mixture $\mu_p = p\,\mu_0 + q\,\mu_1$ is given by

$$\mathrm{Ent}_{\mu_p}[f^2] = p\,\mathrm{Ent}_{\mu_0}[f^2] + q\,\mathrm{Ent}_{\mu_1}[f^2] + \mathrm{Ent}_{\delta_p}\left[\overline{f^2} \right] . \tag{9}$$

The following discrete log-Sobolev inequality for a Bernoulli random variable is used to estimate the entropy of the averaged function $\bar{f}$. The optimal log-Sobolev constant was found by Higuchi and Yoshida [4] and Diaconis and Saloff-Coste [5, Theorem A.2.] at the same time.

Lemma 2 (Optimal log-Sobolev inequality for Bernoulli measures). *A Bernoulli measure μ_p on $\{0,1\}$, i.e. a mixture of two Dirac measures $\delta_p = p\,\delta_0 + q\,\delta_1$ with $p \in [0,1]$ and $q = 1 - p$ satisfies the discrete log-Sobolev inequality*

$$\mathrm{Ent}_{\delta_p}[g] \leq \frac{p\,q}{\Lambda(p,q)} \left(g(0) - g(1) \right)^2 \quad \text{for all } g : \{0,1\} \to \mathbb{R}^+ ,$$

where $\Lambda : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is the logarithmic mean *defined by*

$$\Lambda(p,q) := \frac{p - q}{\log p - \log q}, \quad \text{for } p \neq q \quad \text{and} \quad \Lambda(p,p) := \lim_{q \to p} \Lambda(p,q) = p .$$

The above result allows to estimate the coarse-grained entropy in (9).

Lemma 3 (Estimate of the coarse-grained entropy). *Let $\overline{f^2} : \{0,1\} \to \mathbb{R}^+$ be given by $\overline{f^2}(i) := \mathrm{E}_{\mu_i}[f^2]$ for $i \in \{0,1\}$. Then for all $p \in [0,1]$ and $q = 1 - p$ holds*

$$\mathrm{Ent}_{\delta_p}\left[\overline{f^2} \right] \leq \frac{p\,q}{\Lambda(p,q)} \left(\mathrm{Var}_{\mu_0}[f] + \mathrm{Var}_{\mu_1}[f] + \left(\mathrm{E}_{\mu_0}[f] - \mathrm{E}_{\mu_1}[f] \right)^2 \right) . \tag{10}$$

Proof. Lemma 2 applied to $\mathrm{Ent}_{\delta_p}(\overline{f^2})$ yields

$$\mathrm{Ent}_{\bar{\mu}}(\overline{f^2})) \leq \frac{p\,q}{\Lambda(p,q)} \left(\sqrt{\overline{f^2}(0)} - \sqrt{\overline{f^2}(1)} \right)^2 . \tag{11}$$

The square-root-mean-difference on the right-hand side of (11) can be estimated by using the fact that the function $(a, b) \mapsto (\sqrt{a} - \sqrt{b})^2$ is jointly convex on $\mathbb{R}^+ \times \mathbb{R}^+$. Indeed, by introducing the functions $f_0, f_1 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^+$ defined by $f_0(x, y) = f(x)$ and $f_1(x, y) = f(y)$, an application of the Jensen inequality yields the estimate

$$
\begin{aligned}
\left(\sqrt{\mathrm{E}_{\mu_0}[f^2]} - \sqrt{\mathrm{E}_{\mu_1}[f^2]} \right)^2 &= \left(\sqrt{\mathrm{E}_{\mu_0 \times \mu_1}[f_0^2]} - \sqrt{\mathrm{E}_{\mu_0 \times \mu_1}[f_1^2]} \right)^2 \\
&\leq \mathrm{E}_{\mu_0 \times \mu_1} \left[(f_0 - f_1)^2 \right] \\
&\leq \mathrm{E}_{\mu_0}[f^2] - 2\,\mathrm{E}_{\mu_0}[f]\,\mathrm{E}_{\mu_1}[f] + \mathrm{E}_{\mu_1}[f^2] \\
&= \mathrm{Var}_{\mu_0}[f] + \mathrm{Var}_{\mu_1}[f] + \left(\mathrm{E}_{\mu_0}[f] - \mathrm{E}_{\mu_1}[f] \right)^2 .
\end{aligned}
\tag{12}
$$

Now, a combination (11) and (12) gives (10). $\square$

The decomposition (9) together with (10) yields that a mixture $\mu_p = p\,\mu_0 + q\,\mu_1$ for $p \in [0, 1]$ and $q = 1 - p$ satisfies

$$
\begin{aligned}
\mathrm{Ent}_{\mu_p}\left[f^2 \right] &\leq p\,\mathrm{Ent}_{\mu_0}\left[f^2 \right] + q\,\mathrm{Ent}_{\mu_1}\left[f^2 \right] \\
&\quad + \frac{p\,q}{\Lambda(p, q)}\,\mathrm{Var}_{\mu_0}[f] + \mathrm{Var}_{\mu_1}[f] + \left(\mathrm{E}_{\mu_0}[f] - \mathrm{E}_{\mu_1}[f] \right)^2 .
\end{aligned}
\tag{13}
$$

The right-hand side of (13) consists of quantities, which can be estimated under the assumption that μ_0 and μ_1 satisfy $\mathrm{LSI}(\alpha_0)$ and $\mathrm{LSI}(\alpha_1)$. The following theorem provides an extension of the result [1, Theorem 4.4] to the multidimensional case for the log-Sobolev inequality.

Theorem 2 (LSI for absolutely continuous mixtures). *Let μ_0 and μ_1 satisfy $\mathrm{LSI}(\alpha_0)$ and $\mathrm{LSI}(\alpha_1)$, respectively, and let both measures be absolutely continuous to each other. Then for all $p \in (0, 1)$ and $q = 1 - p$ the mixture measure $\mu_p = p\,\mu_0 + q\,\mu_1$ satisfies $\mathrm{LSI}(\alpha_p)$ with*

$$
\frac{1}{\alpha_p} \leq \begin{cases} \dfrac{1 + q\,\lambda_p}{\alpha_0} , & \text{if } \dfrac{\alpha_1}{\alpha_0} \geq 1 + p\,\lambda_p(1 + \chi_1(1 + q\,\lambda_p)) , \\[2ex] \dfrac{1 + p\,\lambda_p}{\alpha_1} , & \text{if } \dfrac{\alpha_0}{\alpha_1} \geq 1 + q\,\lambda_p(1 + \chi_0(1 + p\,\lambda_p)) , \\[2ex] \dfrac{p(1 + q\,\lambda_p)\chi_1 + p\,q\,\lambda_p\,\chi_0\,\chi_1 + q(1 + p\,\lambda_p)\,\chi_0}{\alpha_0\,p\,\chi_1 + \alpha_1\,q\,\chi_0} , & \text{else} . \end{cases}
\tag{14}
$$

Hereby, χ_0 and χ_1 are given in (5) and λ_p is used for the inverse logarithmic mean

$$
\lambda_p := \frac{1}{\Lambda(p, q)} = \frac{\log p - \log q}{p - q} , \qquad \text{for } p \neq \frac{1}{2} , \qquad \text{and} \qquad \lambda_{1/2} = 2 .
$$

Proof. The starting point is the splitting obtained from (13). The variances and mean-difference in (13) can be estimated in the same way as in the proof (6) of Theorem 1. Additionally, the fact [6] that $\mathrm{LSI}(\alpha)$ implies $\mathrm{PI}(\alpha)$ is used to derive for any $\eta > 0$ and any $\vartheta \in (0, 1)$

$$
\begin{aligned}
\mathrm{Ent}_{\mu_p}[f^2] &\leq \frac{1}{\alpha_0}\left(1 + q\,\lambda_p\left(1 + (1 + \eta)\,\vartheta^2\,\chi_0 \right) \right) \int |\nabla f|^2\,p\,\mathrm{d}\mu_0 \\
&\quad + \frac{1}{\alpha_1}\left(1 + p\,\lambda_p\left(1 + (1 + \eta^{-1})\,(1 - \vartheta)^2\,\chi_1 \right) \right) \int |\nabla f|^2\,q\,\mathrm{d}\mu_1 \\
&\leq \max\left\{ \frac{1 + q\,\lambda_p\left(1 + (1 + \eta)\,\vartheta^2\,\chi_0 \right)}{\alpha_0} , \frac{1 + p\,\lambda_p\left(1 + (1 + \eta^{-1})\,(1 - \vartheta)^2\,\chi_1 \right)}{\alpha_1} \right\} \int |\nabla f|^2\,\mathrm{d}\mu_p .
\end{aligned}
\tag{15}
$$

By introducing reduced log-Sobolev constants

$$\tilde{\alpha}_0 := \frac{\alpha_0}{1 + q\lambda_p} \qquad \text{and} \qquad \tilde{\alpha}_1 := \frac{\alpha_1}{1 + p\lambda_p}, \tag{16}$$

as well as defining the constants $\tilde{\chi}_0$ and $\tilde{\chi}_1$ by

$$\tilde{\chi}_0 := \frac{\chi_0 \lambda_p}{1 + q\lambda_p} \qquad \text{and} \qquad \tilde{\chi}_1 = \frac{\chi_1 \lambda_p}{1 + p\lambda_p}, \tag{17}$$

the bound (15) takes the form

$$\text{Ent}_{\mu_p}(f^2) \le \max\left\{ \frac{1 + (1 + \eta)\vartheta^2 \tilde{\chi}_0}{\tilde{\alpha}_0}, \frac{1 + (1 + \frac{1}{\eta})(1 - \vartheta)^2 \tilde{\chi}_1}{\tilde{\alpha}_1} \right\} \int |\nabla f|^2 \, d\mu_p. \tag{18}$$

The estimate (18) has the same structure as the estimate (6), where $\tilde{\alpha}_0$, $\tilde{\alpha}_i$ play the role of ϱ_0, ϱ_1 and $\tilde{\chi}_0$, $\tilde{\chi}_1$ the roles of χ_0, χ_1. Hence, the optimization procedure from the proof of Theorem 1 applies to this case and the last step consists of translating the constants $\tilde{\alpha}_0$, $\tilde{\alpha}_1$ and $\tilde{\chi}_0$, $\tilde{\chi}_1$ back to the original ones. $\square$

Remark 3. *Let the bound for $\frac{1}{\alpha_p}$ in the last case of (14) be denoted by $\frac{1}{A_p}$. Then the proof shows that it can be bounded above and below in the same way as in (7) in terms of the reduced constants (16) and (17)*

$$\max\left\{ \frac{1 + q\lambda_p}{\alpha_0}, \frac{1 + p\lambda_p}{\alpha_1} \right\} \le \frac{1}{A_p} \le \max\left\{ \frac{1 + q\lambda_p(1 + \chi_0)}{\alpha_0}, \frac{1 + p\lambda_p(1 + \chi_1)}{\alpha_1} \right\}.$$

In the case $\chi_0 = \chi_1 = \chi$, it holds the simplified bound

$$\frac{1}{\alpha_p} \le \frac{1 + \lambda_p + pq\lambda_p\chi}{p\alpha_0 + q\alpha_1}. \tag{19}$$

The inverse logarithmic mean $\lambda_p = \frac{1}{\Lambda(p,q)}$ blows up logarithmically for $p \to \{0,1\}$. Hence, even in the case $\chi = 0$, the bound (19) diverges logarithmically. This logarithmic divergence looks at first sight artificial, especially in comparison to (8) showing that the Poincaré constant is bounded. However, the next section with examples shows, that this blow-ups may actually occure. Hence, the bound in (14) is actually optimal on this level of generality.

An analogue statement as Corollary 1 for the Poincaré constant is obtained for the lob-Sobolev constant, which proof follows along the same lines and is omitted.

Corollary 2. *Let $\mu_0 \ll \mu_1$ and μ_0, μ_1 satisfy $\text{LSI}(\alpha_0)$ and $\text{LSI}(\alpha_1)$, respectively. Then for any $p \in (0,1)$ and $p = 1 - q$ the mixture measure $\mu_p = p\mu_0 + q\mu_1$ satisfies $\text{LSI}(\alpha_p)$ with*

$$\frac{1}{\alpha_p} \le \max\left\{ \frac{1 + q\lambda_p}{\alpha_0}, \frac{1 + p\lambda_p(1 + \chi_1)}{\alpha_1} \right\}$$

Likewise, if $\mu_1 \ll \mu_0$, then it holds

$$\frac{1}{\alpha_p} \le \max\left\{ \frac{1 + p\lambda_p}{\alpha_1}, \frac{1 + q\lambda_p(1 + \chi_0)}{\alpha_0} \right\}$$

4. Examples

The results of Theorem 1 and Theorem 2 are illustrated for some specific examples and also compared to the results [1, Section 4.5], which however are restricted to one-dimensional

measures. Although the criterion of Theorem 1 and Theorem 2 can only give upper bounds for the multidimensional case, when at least one of the mixture component is absolutely continuous to the other, it is still possible to obtain the optimal results in terms of scaling in the mixture parameter $p \to \{0,1\}$.

4.1. Mixture of two Gaussian measures with equal covariance matrix

Let us consider the mixtures of two Gaussians $\mu_0 := \mathcal{N}(0, \Sigma)$ and $\mu_1 := \mathcal{N}(y, \Sigma)$, for some $y \in \mathbb{R}^n$ and a strictly positive definite covariance matrix $\Sigma \geq \sigma \, \mathrm{Id}$ in the sense of quadratic forms for some $\sigma > 0$. Then, μ_0 and μ_1 satisfy $\mathrm{PI}(\sigma^{-1})$ and $\mathrm{LSI}(\sigma^{-1})$ by the Bakry-Émery criterion (Theorem A1), i.e. $\varrho_0 = \alpha_0 = \varrho_1 = \alpha_1 = \sigma^{-1}$. Furthermore, the χ^2-distance between μ_0 and μ_1 can be explicitly calculated as a Gaussian integral (see also [7])

$$
\begin{aligned}
\chi_0 = \chi_1 &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} \int \exp\left(-x \cdot \Sigma^{-1} x + \tfrac{1}{2}(x-y) \cdot \Sigma^{-1}(x-y)\right) \mathrm{d}x - 1 \\
&= \exp\left(y \cdot \Sigma^{-1} y\right) \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} \int \exp\left(-\tfrac{1}{2}(x+y)\Sigma^{-1}(x+y)\right) \mathrm{d}x - 1 \leq e^{|y|^2/\sigma} - 1.
\end{aligned}
$$

Then the bound from Theorem 1 in the form (8) yields

$$
\frac{1}{\varrho_p} \leq \left(1 + pq\left(e^{|y|^2/\sigma} - 1\right)\right)\sigma. \tag{20}
$$

Likewise, the log-Sobolev constant follows from Theorem 2 in the form (19) leads to

$$
\frac{1}{\alpha_p} \leq \left(1 + pq\lambda_p\left(e^{|y|^2/\sigma} + 1\right)\right)\sigma.
$$

By noting that $pq \leq pq\lambda_p \leq \frac{1}{4}$, both constants stay uniformly bounded in p. The large exponential factor in the distance $e^{|y|^2/\sigma}$ cannot be avoided on this level of generality since the mixed measure μ_p has a bimodal structure leading to metastable effects [2, Remark 2.20].

The result [1, Corollary 4.7] deduced the following bound on $\frac{1}{\varrho_p}$ for the mixture of two one-dimensional standard Gaussians $\sigma = 1$ in (20)

$$
\frac{1}{\varrho_p} \leq 1 + pq\,|y|^2\left(\Phi(|y|)\,e^{|y|^2} + \frac{|y|}{\sqrt{2\pi}}\,e^{|y|^2/2} + \tfrac{1}{2}\right), \tag{21}
$$

where $\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-y^2/2} \, \mathrm{d}y$. The elementary inequalities $e^{a^2} - 1 \leq a^2 e^{a^2}$ and $\Phi(a) \geq 1 + \frac{a}{\sqrt{2\pi}} e^{-a^2/2}$ for all $a \in \mathbb{R}$ show that the bound (20) is better than the bound (21) for all parameter values $p \in [0,1]$ and $|y| \geq 0$.

Hence, this example shows, that for mixtures with components that are absolutely continuous to each other as well as whose tail behavior is controlled in terms of the χ^2-distance, Theorem 1 and Theorem 2 even improve the bound of [1] and generalize it to the multidimensional case.

4.2. Mixture of a Gaussian and sub-Gaussian measure

Let us consider $\mu_1 = \mathcal{N}(0, \Sigma)$ where $\Sigma \geq \sigma \, \mathrm{Id}$ is strictly positive definite. In addition, let the density of μ_0 with respect to μ_1 be bounded uniformly by some $\kappa \geq 1$, that is the relative density satisfies $\mathrm{d}\mu_0/\mathrm{d}\mu_1 \leq \kappa$ almost everywhere on $\mathbb{R}^n$. By the Bakry-Émery criterion (Theorem A1), it holds $\varrho_1 = \alpha_1 = \frac{1}{\sigma}$. Further, an upper bound for χ_1 is obtained by the assumption on the bound on the relative density

$$
\chi_1 = \mathrm{Var}_{\mu_1}\left[\frac{\mu_0}{\mu_1}\right] = \int \left(\frac{\mu_0}{\mu_1}\right)^2 \mathrm{d}\mu_1 - 1 \leq \kappa^2 - 1.
$$

Provided that μ_0 satisfies $\mathrm{PI}(\varrho_0)$, the Poincaré constant of the mixture $\mu_p = p\,\mu_0 + q\,\mu_1$ satisfies by Corollary 1 the estimate

$$\frac{1}{\varrho_p} \le \max\left\{\frac{1}{\varrho_0}, (1 + p(\kappa^2 - 1))\sigma\right\}.$$

Similarly, Corollary 2 provides whenever μ_0 satisfies $\mathrm{LSI}(\alpha_0)$ the following bound for the log-Sobolev constant of the mixture measure μ_p

$$\frac{1}{\alpha_p} \le \max\left\{\frac{1 + q\lambda_p}{\alpha_0}, (1 + p\lambda_p\kappa^2)\sigma\right\}.$$

In this case, the logarithmic blow-up of the log-Sobolev constant cannot be ruled out for $p \to \{0, 1\}$, without any further information on μ_0.

4.3. Mixture of two centered Gaussians with different variance

For $\mu_0 = \mathcal{N}(0, \mathrm{Id})$ and $\mu_1 = \mathcal{N}(0, \sigma\,\mathrm{Id})$, the Bakry-Émery criterion (Theorem A1) implies $\varrho_0 = \alpha_0 = 1$ and $\varrho_1 = \alpha_1 = \sigma^{-1}$. The calculation of the χ^2-distance can be done using the spherical symmetry and is reduced to the one dimensional integral

$$\chi_0 = \int \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0}\,\mathrm{d}\mu_1 - 1 = \frac{\mathcal{H}^{n-1}(\partial B_1)}{(2\pi)^{\frac{n}{2}}\sigma^n}\int_{\mathbb{R}^+} r^{n-1}e^{-\left(\frac{1}{\sigma} - \frac{1}{2}\right)r^2}\,\mathrm{d}r - 1\,.$$

Hereby, $\mathcal{H}^{n-1}(S^{n-1})$ denotes the $n-1$-dimensional Hausdorff measure of the sphere $\partial B_1 = \{x \in \mathbb{R}^n : |x| = 1\}$. The integral does only exist for $\sigma < 2$. In this case, it can be evaluated and simplified. The bound for the constant χ_1 follows by duality under the substitution $\sigma \mapsto \sigma^{-1}$ and is given by

$$\chi_0 = \begin{cases} \frac{1}{(\sigma(2-\sigma))^{\frac{n}{2}}} - 1 & ,\sigma < 2 \\ +\infty & ,\sigma \ge 2 \end{cases} \quad \text{and} \quad \chi_1 = \begin{cases} \frac{1}{\left(\sigma^{-1}(2-\sigma^{-1})\right)^{\frac{n}{2}}} - 1 & ,\sigma > \frac{1}{2} \\ +\infty & ,\sigma \le \frac{1}{2} \end{cases}. \tag{22}$$

If $\sigma \le 1/2$, that is for $\chi_1 = \infty$, the bound given in Corollary 1 yields

$$\frac{1}{\varrho_p} \le \max\{\sigma, 1 + q\chi_0\} = \max\left\{\sigma, (1-q) + \frac{q}{(\sigma(2-\sigma))^{\frac{n}{2}}}\right\} = p + \frac{q}{(\sigma(2-\sigma))^{\frac{n}{2}}}$$

Similarly, if $\sigma \ge 2$, that is for $\chi_0 = \infty$, the bound becomes

$$\frac{1}{\varrho_p} \le \max\{1, (1 + p\chi_1)\sigma\} \le \sigma\left(q + \frac{p}{(\sigma^{-1}(2 - \sigma^{-1}))^{\frac{n}{2}}}\right).$$

In the case $\frac{1}{2} < \sigma < 2$, the interpolation bound (4) of Theorem 1 could be applied. However, the scaling behavior for the Poincaré constant can already be observed with the estimate (7) in Remark 2, where again thanks to the symmetry $\sigma \mapsto \frac{1}{\sigma}$ it holds

$$\frac{1}{\varrho_p} \le \begin{cases} p + \frac{q}{(\sigma(2-\sigma))^{\frac{n}{2}}}\,, & \text{for } \sigma \le 1\,, \\[2mm] \sigma\left(q + \frac{p}{(\sigma^{-1}(2-\sigma^{-1}))^{\frac{n}{2}}}\right), & \text{for } \sigma \ge 1\,. \end{cases} \tag{23}$$

Hence, the Poincaré constant stays bounded for the full range of parameter $p \in [0, 1]$ and $\sigma > 0$.

In the case for the log-Sobolev constant, the bound from Corollary 2 gives

$$\frac{1}{\alpha_p} \leq \begin{cases} 1 + \frac{q\lambda_p}{(\sigma(2-\sigma))^{\frac{n}{2}}}, & \sigma \leq 1 \\ \sigma\left(1 + \frac{p\lambda_p}{\left(\sigma^{-1}(2-\sigma^{-1})\right)^{\frac{n}{2}}}\right), & \sigma \geq 1 \end{cases}. \tag{24}$$

The bound (24) blows up logarithmically for $p \to \{0,1\}$ in general. However, the special case $\sigma = 1$, although trivially, allows for the combined bound $\frac{1}{\alpha_p} \leq 1 + \min\{p,q\}\lambda_p$, which stays bounded. This behavior can be extended to the range $\sigma \in (\frac{1}{2}, 2)$ thanks to (22) and the interpolation bound of Theorem 2.

The result (23) can be compared with the one of [1, Section 4.5.2.], which states that for some $C > 0$, all $\sigma > 1$ and $p \in (0, 1/2)$ it holds

$$\frac{1}{\varrho_{p,\mathrm{CM}}} \leq \sigma + C p^{\frac{1}{\sigma-1}}. \tag{25}$$

In general, depending on the constant C the bound (23) is better for σ small, whereas the scaling in σ is better for (25), namely linear instead of $\sigma^{\frac{3}{2}}$ as in (20).

4.4. Mixture of uniform and Gaussian measure

Let $\mu_0 = \mathcal{N}(0, \mathrm{Id})$ and $\mu_1 = \frac{1}{\mathcal{H}^n(B_1)}\mathbb{1}_{B_1}$ with B_1 the unit ball around zero. Then it holds $\varrho_0 = 1$ by the Bakry-Émery criterion (Theorem A1) and $\varrho_1 \geq \frac{\pi^2}{\mathrm{diam}(B_1)^2} = \frac{\pi^2}{4}$ by the result of [8]. Furthermore, since $\mu_1 \ll \mu_0$ the χ^2-distance between μ_0 and μ_1 becomes thanks to the spherical symmetry

$$\chi_0 + 1 = \int \left(\frac{\mu_1}{\mu_0}\right)^2 \mathrm{d}\mu_0 = \frac{(2\pi)^{\frac{n}{2}}}{\mathcal{H}^n(B_1)^2} \int_{B_1} e^{|x|/2} \,\mathrm{d}x = \frac{(2\pi)^{\frac{n}{2}} \mathcal{H}^{n-1}(\partial B_1)}{\mathcal{H}^n(B_1)^2} \int_0^1 r^{n-1} e^{r^2/2} \,\mathrm{d}r. \tag{26}$$

The volume $\mathcal{H}^n(B_1)$ and the surface area $\mathcal{H}^{n-1}(\partial B_1)$ of the n-sphere satisfy the following relations

$$\frac{\mathcal{H}^{n-1}(\partial B_1)}{\mathcal{H}^n(B_1)} = n \quad \text{and} \quad \frac{(2\pi)^{\frac{n}{2}}}{\mathcal{H}^n(B_1)} = 2^{\frac{n}{2}}\Gamma\left(\frac{n}{2} + 1\right) =: g_n. \tag{27}$$

The integral on the right-hand side in (26) can be bounded below by $\frac{1}{n}$ and above by $\frac{\sqrt{e}}{n}$, which alltogether yields

$$g_n \leq \chi_0 + 1 \leq \sqrt{e}g_n.$$

Corollary 1 implies that the Poincaré constant of the mixture $\mu_p = p\,\mu_0 + q\,\mu_1$ satisfies

$$\frac{1}{\varrho_p} \leq \max\left\{\frac{1}{\varrho_1}, 1 + q\chi_0\right\} \leq p + q\sqrt{e}\,g_n, \tag{28}$$

where the last inequality follows from $\frac{4}{\pi^2} \leq p + q\sqrt{e}\,g_n$ for $n \geq 1$ and all $p \in [0, 1]$.

The estimate of the log-Sobolev constant uses that $\alpha_0 = 1$ by the Bakry-Émery criterion (Theorem A1) and $\alpha_1 \geq \frac{2}{e}$ from (A1). Then, Corollary 2 yields the bound

$$\frac{1}{\alpha_p} \leq \max\left\{\frac{1 + p\lambda_p}{\alpha_1}, \frac{1 + q\lambda_p(1 + \chi_0)}{\alpha_0}\right\} \leq \max\left\{\frac{(1 + p\lambda_p)e}{2}, 1 + q\lambda_p\sqrt{e}g_n\right\}. \tag{29}$$

There is a logarithmically blow-up of the bound for $p \to \{0, 1\}$.

The blow-up for $p \to 1$ is artificial, which can be shown by a combination Bakry-Émery criterion and the Holley-Stroock perturbation principle. To do so, the Hamiltonian of μ_p is decomposed into a convex function and some error term

$$
\begin{aligned}
H_p(x) := -\log \mu_p(x) &= -\log\left(\frac{p}{(2\pi)^{\frac{n}{2}}} e^{-\frac{|x|^2}{2}} + \frac{1-p}{\mathcal{H}^n(B_1)} \mathbb{1}_{B_1(0)}(x) \right) \\
&= -\log\left(e^{-\frac{|x|^2}{2}+\frac{1}{2}} + \frac{1-p}{p} \frac{(2\pi)^{\frac{n}{2}}}{\mathcal{H}^n(B_1)} \sqrt{e} \mathbb{1}_{B_1(0)}(x) \right) + C_{p,n} \\
&= \frac{|x|^2 - 1}{2} - \psi_p(x) + \tilde{C}_{p,n},
\end{aligned}
$$
(30)

where

$$
\psi_p(x) := \left(\log\left(e^{-\frac{|x|^2}{2}+\frac{1}{2}} + \frac{1-p}{p} \frac{(2\pi)^{\frac{n}{2}}}{\mathcal{H}^n(B_1)} \sqrt{e} \right) + \frac{|x|^2 - 1}{2} \right) \mathbb{1}_{B_1(0)}(x).
$$

The function ψ_p is radially monotone towards the boundary of B_1, which yields for $|x| \to 1$ the bound

$$
0 \le \psi_p(x) \le \log\left(1 + \frac{1-p}{p} \frac{(2\pi)^{\frac{n}{2}}}{\mathcal{H}^n(B_1)} \sqrt{e} \right).
$$
(31)

From (30) the Hamiltonian H_p is compared with the convex potential $\frac{|x|^2-1}{2}$ with the bound (31) on the perturbation ψ_p. This together yields by the Bakry-Émergy criterion (Theorem A1) and the Holley-Stroock perturbation principle (Theorem A2) the μ_p satisfies PI($\tilde{\varrho}_p$) and LSI($\tilde{\alpha}_p$) with

$$
\frac{1}{\tilde{\varrho}_p} \le \frac{1}{\tilde{\alpha}_p} \le 1 + \frac{1-p}{p} \sqrt{e}\, g_n,
$$
(32)

where g_n is the same constant as in (27). This bound only blows up for $p \to 0$. But the blow-up is like $\frac{1}{p}$. Furthermore, the bound on the Poincaré constant is worse than the one from (28). Therefore, both approaches need to be combined.

The combination of the bounds obtained in (29) and (32) results in the improved bound

$$
\frac{1}{\alpha} \le C_n(1 + q\lambda_p g_n), \quad \text{with } C_n \text{ some universal constant,}
$$
(33)

which only logarithmically blows up for $p \to 0$.

This example shows that the Poincaré constant and log-Sobolev constant may have different scaling behavior for $p \to 0$. Indeed, [1] show for this specific mixture in the one-dimensional case that the log-Sobolev constant can be bounded below by

$$
C|\log p| \le \frac{1}{\alpha},
$$

for p small enough and a constant C independent of p. In one dimension, lower bounds are accessible via the functional introduced by Bobkov-Götze [9]. Hence the bound (33) is optimal in the one-dimensional case, which strongly indicates also optimality for the higher dimension case in terms of scaling in the mixture ration p.

To conclude, the Bakry-Émery criterion in combination with the Holley-Stroock perturbation principle is effective for detecting blow-ups of the log-Sobolev constant for mixtures, but has, in general, the wrong scaling behavior in the mixing parameter p. On the other hand, the criterion presented in Theorem 2 provides the right scaling of the blow-up but may give artificial blow-ups, if the components of the mixture become singular in the sense of the χ^2-distance.

5. Conclusion and future work

Recently, the investigation of mixtures can be found in many different applications, and the main results of this work may be useful to the investigation of asymmetric Kalman filter estimates [10], the study of asymmetric mixtures in Marine Biology [11], Econometrics [12], Gradient-quadratic and fixed-point iteration algorithms [7] and estimates of multivariate Gaussian mixtures [13].

Theorem 1 and Theorem 2 provide a simple estimate of the Poincaré and log-Sobolev constants of a two-component mixture measure $\mu_p = p\,\mu_0 + q\,\mu_1$ if the χ^2-distance of μ_0 and μ_1 is bounded and each of the components satisfies a Poincaré or log-Sobolev inequality. Section 4 reviews several examples with the following findings:

- For mixtures with components that are mutually absolutely continuous and whose tail behavior is mutually controlled in terms of the χ^2-distance, Theorem 1 and Theorem 2 are very effective.
- If only one of the components is absolutely continuous to the other one with bounded density, then it is still possible to obtain a bound on the Poincaré and log-Sobolev constant. However, the log-Sobolev constant blows up logarithmically in the mixture parameter p approaching 0 or 1. It is shown for specific examples that this blow-up is at least for one limit $p \to 0$ or $p \to 1$ not artificial due to the applied method.
- A necessary condition for the finiteness of the χ^2-distance between two measures is that at least one of the measures μ_0 and μ_1 is absolutely continuous to the other one, which in particular provides a mixture with connected support. This condition is too strong since one can easily decompose a measure into a mixture, where the joint support of the components is a null set. In this case, the present approach would not be helpful, even though the mixture may still satisfy both functional inequalities.

Future work could overcome the limits of the present approach by revisiting the crucial ingredient for both the Poincaré and log-Sobolev inequality, which was the representation of the mean-difference in Lemma 1 regarding covariances. The formula (3) from Lemma 1 applies only in the case where both measures are mutually absolutely continuous. However, the idea of an interpolation bound can be generalized to suitable weighted Sobolev spaces. For this, since $\mu_0, \mu_1 \ll \mu_p$ for all $p \in (0,1)$, one can formally write and estimate

$$\mathrm{E}_{\mu_0}[f] - \mathrm{E}_{\mu_1}[f] = \mathrm{Cov}_{\mu_p}\left[f, \frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_p} - \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_p}\right] \leq \|f\|_{\dot{H}^1(\mu_p)} \left\|\frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_p} - \frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_p}\right\|_{\dot{H}^{-1}(\mu_p)}. \tag{34}$$

Hereby, $\dot{H}^1(\mu_p)$ is the homogeneous weighted $\dot{H}^1$ space with norm $\|f\|_{\dot{H}^1(\mu_p)}^2 := \int |\nabla f|^2 \,\mathrm{d}\mu_p$ and $\dot{H}^{-1}(\mu_p)$ is its dual space with norm

$$\|\omega\|_{\dot{H}^{-1}(\mu_p)}^2 := \sup_{f \in \dot{H}^1(\mu_p)} \left\{2\langle f, \omega\rangle_{\mu_p} - \|f\|_{\dot{H}^1(\mu_p)}^2\right\}.$$

The representation (34) is fruitful to many more applications in which the components of the mixture do not need to be absolutely continuous. Similar ideas for estimating mean-differences were successfully applied in the metastable setting [2,14], in which suitable bounds on the $\dot{H}^{-1}$-norm are obtained. In this regard the bound (34) promises many interesting new insights for future studies.

Appendix Bakry-Émery criterion and Holley-Stroock perturbation principle

Two classical conditions for Poincaré and log-Sobolev inequalities are stated in this part of the appendix. The *Bakry-Émery criterion* relates the convexity of the Hamiltonian of a measure and positive curvature of the underlying space to constants for the Poincaré and log-Sobolev inequalities. Although the result is classical for the case of $\mathbb{R}^n$, the result for general convex domain was established in [15, Theorem 2.1].

Theorem A1 (Bakry-Émery criterion [16, Proposition 3, Corollaire 2], [15, Theorem 2.1]). *Let $\Omega \subset \mathbb{R}^n$ be convex and let $H : \Omega \to \mathbb{R}$ be a Hamiltonian with Gibbs measure $\mu(dx) = Z_\mu^{-1} e^{-H(x)} \mathbb{1}_\Omega(x)\,dx$ and assume that $\nabla^2 H(x) \geq \kappa > 0$ for all $x \in \operatorname{supp} \mu$. Then μ satisfies $\mathrm{PI}(\varrho)$ and $\mathrm{LSI}(\alpha)$ with*

$$\varrho \geq \kappa \qquad and \qquad \alpha \geq \kappa.$$

The second condition is the *Holley-Stroock perturbation principle*, which allows to show Poincaré and log-Sobolev inequalities for a very large class of measures.

Theorem A2 (Holley-Stroock perturbation principle [17, p. 1184]). *Let $\Omega \subset \mathbb{R}^n$ and $H : \Omega \to \mathbb{R}$ and $\psi : \Omega \to \mathbb{R}^n$ be a bounded function. Let μ and $\tilde{\mu}$ be the Gibbs measures with Hamiltonian H and $H + \psi$, respectively*

$$\mu(dx) = \frac{1}{Z_\mu} e^{-H(x)} \mathbb{1}_\Omega(x)\,dx \qquad and \qquad \tilde{\mu}(dx) = \frac{1}{Z_{\tilde{\mu}}} e^{-H(x) - \psi(x)} \mathbb{1}_\Omega(x)\,dx.$$

Then, if μ satisfies $\mathrm{PI}(\varrho)$ and $\mathrm{LSI}(\alpha)$, then $\tilde{\mu}$ satisfy $\mathrm{PI}(\tilde{\varrho})$ and $\mathrm{LSI}(\tilde{\alpha})$, respectively. Hereby the constants satisfy

$$\tilde{\varrho} \geq e^{-\operatorname{osc}_\Omega \psi} \varrho \qquad and \qquad \tilde{\alpha} \geq e^{-\operatorname{osc}_\Omega \psi} \alpha,$$

where $\operatorname{osc}_\Omega \psi := \sup_\Omega \psi - \inf_\Omega \psi$.

Proofs relying on semigroup theory of Theorem A1 and Theorem A2 can be found in the exposition by Ledoux [6, Corollary 1.4, Corollary 1.6 and Lemma 1.2].

Example A1 (Uniform measure on the ball). *The measure $\mu_1 = \frac{1}{\mathcal{H}^n(B_1)} \mathbb{1}_{B_1}$, with B_1 is the unit ball around zero, satisfies $\mathrm{LSI}(\alpha_1)$ with*

$$\alpha_1 \geq \frac{2}{e}. \tag{A1}$$

The proof compares the measure μ_1 with a family of measures

$$\nu_\sigma(dx) = \frac{1}{Z_\sigma} \exp\left(-\sigma|x|^2 + \frac{\sigma}{2}\right) \mathbb{1}(x)\,dx \qquad for\ \sigma > 0.$$

Then, it holds that ν_σ satisfies $\mathrm{LSI}(2\sigma)$ by the Bakry-Émery criterion (Theorem A1). Moreover, it holds that $\operatorname{osc}_{x \in B_1} |-\sigma|x|^2 + \sigma/2| = \frac{\sigma}{2}$ and hence μ_1 satisfies $\mathrm{LSI}(2\sigma e^{-\sigma})$ by the Holley-Stroock perturbation principle (Theorem A2) for all $\sigma > 0$. Optimizing the expression $2\sigma e^{-\sigma}$ in σ gives the bound (A1).

References

1. Chafaï, D.; Malrieu, F. On fine properties of mixtures with respect to concentration of measure and Sobolev type inequalities. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **2010**, *46*, 72–96. doi:10.1214/08-AIHP309.

2. Menz, G.; Schlichting, A. Poincaré and logarithmic Sobolev inequalities by decomposition of the energy landscape. *Annals of Probability* **2014**, *42*, 1809–1884. doi:10.1214/14-AOP908.

3. Gibbs, A.L.; Su, F.E. On Choosing and Bounding Probability Metrics. *International Statistical Review / Revue Internationale de Statistique* **2002**, *70*, 419. doi:10.2307/1403865.

4. Higuchi, Y.; Yoshida, N. Analytic conditions and phase transition for Ising models., 1995.

5. Diaconis, P.; Saloff-Coste, L. Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability* **1996**, *6*, 695–750. doi:10.1214/aoap/1034968224.

6. Ledoux, M. Logarithmic sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités XXXV*; Number x, 1999; pp. 167–194. doi:10.1007/978-3-540-44671-2.

7. Carreira-Perpinan, M.A. Mode-finding for mixtures of Gaussian distributions. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1318–1323. doi:10.1.1.31.6631.

8. Payne, L.E.; Weinberger, H.F. An optimal Poincaré inequality for convex domains. *Archive for Rational Mechanics and Analysis* **1960**, *5*, 286–292. doi:10.1007/BF00252910.

9. Bobkov, S.G.; Götze, F. Exponential Integrability and Transportation Cost Related to Logarithmic Sobolev Inequalities. *Journal of Functional Analysis* **1999**, *163*, 1–28. doi:10.1006/jfan.1998.3326.

10. Nurminen, H.; Ardeshiri, T.; Piche, R.; Gustafsson, F. Skew-t Filter and Smoother With Improved Covariance Matrix Approximation. *IEEE Trans. Signal Process.* **2018**, *66*, 5618–5633. doi:10.1109/TSP.2018.2865434.

11. Contreras-Reyes, J.; López Quintero, F.; Yáñez, A. Towards Age Determination of Southern King Crab (Lithodes santolla) Off Southern Chile Using Flexible Mixture Modeling. *J. Mar. Sci. Eng.* **2018**, *6*, 157. doi:10.3390/jmse6040157.

12. Tasche, D. Exact Fit of Simple Finite Mixture Models. *J. Risk Financ. Manag.* **2014**, *7*, 150–164. doi:10.3390/jrfm7040150.

13. McLachlan, G.; Peel, D. *Finite Mixture Models*; Wiley Series in Probability and Statistics, John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2000. doi:10.1002/0471721182.

14. Schlichting, A.; Slowik, M. Poincaré and logarithmic Sobolev constants for metastable Markov chains via capacitary inequalities. *arXiv*, [1705.05135].

15. Kolesnikov, A.V.; Milman, E. Riemannian metrics on convex sets with applications to Poincaré and log-Sobolev inequalities. *Calculus of Variations and Partial Differential Equations* **2016**, *55*, 1–36. doi:10.1007/s00526-016-1018-3.

16. Bakry, D.; Émery, M. *Diffusions hypercontractives*; Séminaire de probabilités, XIX, Springer: Berlin, 1985; pp. 177–206.

17. Holley, R.; Stroock, D. Logarithmic Sobolev inequalities and stochastic Ising models. *Journal of Statistical Physics* **1987**, *46*, 1159–1194. doi:10.1007/BF01011161.

18. Schlichting, A. The Eyring-Kramers formula for Poincaré and logarithmic Sobolev inequalities. PhD thesis, Universität Leipzig, 2012.