HAL open science

Intrinsic dimensional functional inequalities on model spaces

Alexandros Eskenazis, Yair Shenfeld

To cite this version:

Alexandros Eskenazis, Yair Shenfeld. Intrinsic dimensional functional inequalities on model spaces. 2023. hal-04272898

HAL Id: hal-04272898

https://hal.science/hal-04272898v1

Preprint submitted on 6 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INTRINSIC DIMENSIONAL FUNCTIONAL INEQUALITIES ON MODEL SPACES

ALEXANDROS ESKENAZIS AND YAIR SHENFELD

Abstract. We initiate a systematic study of *intrinsic* dimensional versions of classical functional inequalities which capture refined properties of the underlying objects. We focus on model spaces: Euclidean space, Hamming cube, and manifolds of constant curvature. In the latter settings, our intrinsic dimensional functional inequalities improve on a series of known results and lead to new Hamilton-type matrix inequalities. Our proofs rely on scaling, tensorization, and stochastic methods.

2020 Mathematics Subject Classification. Primary: 39B62; Secondary: 26D10, 58J65, 60J60.
Key words. Logarithmic Sobolev inequalities, Gagliardo–Nirenberg inequalities, scaling, Föllmer process, Hamilton-type matrix inequalities.

Contents

1. Introduction 1

Part 1. Euclidean and product spaces: scaling and tensorization 8
2. Logarithmic Sobolev inequalities in Euclidean spaces and Cramér–Rao bounds 8
3. Gagliardo–Nirenberg–Sobolev inequalities 12
4. Beckner inequalities 13
5. q-logarithmic Sobolev inequalities 14
6. Beyond linear rescalings 16
7. Tensorization of nonlinear logarithmic Sobolev inequalities in product spaces 17

Part 2. Space forms: stochastic methods 18
8. Preliminaries 18
9. Intrinsic dimensional logarithmic Sobolev inequality in space forms 21
10. Nonpositively curved space forms 28
References 35

1. Introduction

This work focuses on the development of *intrinsic* dimensional versions of classical functional inequalities. In order to explain the meaning of "intrinsic" in this context it is best to start with an important example. The logarithmic Sobolev inequality in Gauss space [61, 36] asserts that for every nice-enough absolutely continuous probability measure μ on $\mathbb{R}^n$,

$$\mathsf{H}(\mu\|\gamma_n) \leq \frac{1}{2}\mathsf{I}(\mu\|\gamma_n), \tag{1}$$

where γ_n is the standard Gaussian measure on $\mathbb{R}^n$. Here,

$$\mathsf{H}(\mu\|\nu) \stackrel{\text{def}}{=} \int \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)\mathrm{d}\mu \tag{2}$$

This material is based upon work supported by the NSF grant DMS-1929284 while A. E. was in residence at ICERM for the Harmonic Analysis and Convexity program. This material is based upon work supported by the National Science Foundation under Award Number 2002022.

is the *relative entropy* of μ with respect to ν and

$$\mathsf{I}(\mu\|\nu) \stackrel{\text{def}}{=} \int \left|\nabla \log \frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right|^2 \mathrm{d}\mu = \int \frac{|\nabla(\mathrm{d}\mu/\mathrm{d}\nu)|^2}{\mathrm{d}\mu/\mathrm{d}\nu}\,\mathrm{d}\nu \tag{3}$$

is the *relative Fisher information* of μ with respect to ν, provided that $\nu \ll \mu$.

Gross' motivation for (1) was to find a substitute for the Euclidean Sobolev inequalities holds in infinite-dimensional spaces (which was needed in constructive quantum field theories). Sobolev inequalities have the feature that the dimension n of the ambient space $\mathbb{R}^n$ appears explicitly in the constants of the inequalities, which leads to triviality upon taking the limit $n \to \infty$. In contrast, the constant 1/2 appearing in (1) is *dimension-free*, leading to (1) being well-defined in infinite dimensions. On the other hand, as was already observed by Stam [61], (1) can in fact be improved if the dimension n is taken into account. To see this improvement we first apply a standard change of measure (see [66]) which shows that (1) is equivalent to

$$\mathsf{H}(\mu\|\lambda_n) - \mathsf{H}(\gamma_n\|\lambda_n) \le \frac{\mathsf{I}(\mu\|\lambda_n) - n}{2}, \tag{4}$$

where λ_n is the Lebesgue measure on $\mathbb{R}^n$. The *dimensional* log-Sobolev inequality of [20],

$$\mathsf{H}(\mu\|\lambda_n) - \mathsf{H}(\gamma_n\|\lambda_n) \le \frac{n}{2}\log\left(\frac{\mathsf{I}(\mu\|\lambda_n)}{n}\right), \tag{5}$$

improves upon (4) as can be seen from the inequality $\log s \le s-1$ for $s \in (0,\infty)$. It is clear that when the Fisher information is large, (5) provides an exponential refinement over (4). Despite this quantitative improvement, (5) suffers from a lack of sensitivity to the *intrinsic* dimension of μ. To see this, suppose that μ is of the form $\mathrm{d}\mu(x_1,\ldots,x_n) = \mathrm{d}\tilde{\mu}(x_1,\ldots,x_k)\,\mathrm{d}\gamma_{n-k}(x_{k+1},\ldots,x_n)$, where $k < n$ and $\tilde{\mu}$ is an absolutely continuous probability measure on $\mathbb{R}^k$. Then (5) rephrased in terms of $\tilde{\mu}$ asserts that

$$\mathsf{H}(\tilde{\mu}\|\lambda_k) - \mathsf{H}(\gamma_k\|\lambda_k) \le \frac{n}{2}\log\left(1 + \frac{\mathsf{I}(\tilde{\mu}\|\lambda_k) - k}{n}\right), \tag{6}$$

which deteriorates to (4) as the *ambient* dimension n increases, despite the fact that the intrinsic dimension k of μ is fixed. In other words, (5) is insensitive to the structure of μ. In [27, p. 12], Dembo showed that (5) can be further improved to an inequality which captures the intrinsic dimension of μ:

$$\mathsf{H}(\mu\|\lambda_n) - \mathsf{H}(\gamma_n\|\lambda_n) \le \frac{1}{2}\log\det \mathcal{I}(\mu\|\lambda_n), \tag{7}$$

where

$$\mathcal{I}(\mu\|\nu) \stackrel{\text{def}}{=} \int \frac{(\nabla(\mathrm{d}\mu/\mathrm{d}\nu))^{\otimes 2}}{\mathrm{d}\mu/\mathrm{d}\nu}\,\mathrm{d}\nu \tag{8}$$

is the *relative Fisher information matrix* of μ with respect to ν. Observe that

$$\mathsf{I}(\mu\|\nu) = \operatorname{tr}\mathcal{I}(\mu\|\nu), \tag{9}$$

and thus (7) improves on (5) by the elementary inequality $\log\det C \le n\log\frac{\operatorname{tr} C}{n}$ which holds for every $n \times n$ positive semidefinite matrix C. In particular, both sides of (7) behave additively with respect to product measures: Plugging in $\mathrm{d}\mu = \mathrm{d}\tilde{\mu}\,\mathrm{d}\gamma_{n-k}$ into (7) yields

$$\mathsf{H}(\tilde{\mu}\|\lambda_k) - \mathsf{H}(\gamma_k\|\lambda_k) \le \frac{1}{2}\log\det \mathcal{I}(\tilde{\mu}\|\lambda_k) \tag{10}$$

which captures correctly the intrinsic dimension of μ. More generally, by considering the eigenvalues of the Fisher information matrix, (7) can quantify the extent to which μ degenerates along each eigenvector direction.

The goal of this work is to initiate a systematic study of intrinsic dimensional versions of classical functional inequalities. We focus on some important model spaces: Euclidean space, Hamming cube, and space forms (manifolds of constant sectional curvature). These model spaces have historically played a crucial role in the development of functional inequalities

and their study has been the impetus leading to fruitful generalizations and abstractions; see the monograph [9]. In view of the richness of the subject, our intrinsic dimensional functional inequalities on these spaces improve on multiple classical inequalities from the literature. The tools required to establish intrinsic dimensional functional inequalities in each of the models spaces will depend on the unique characteristics of the space itself: **scaling** (Euclidean space), **tensorization** (Hamming cube), and **stochastic methods** (space forms). In the rest of the introduction we will review each of these methods and present examples of the intrinsic dimensional functional inequalities which follow. We defer the statements of many of our results to the main body of the paper; see the following brief summary:

Part 1. Euclidean and product spaces: scaling and tensorization

- Logarithmic Sobolev inequalities for homogeneous measures (Section 2.2).
- Bayesian Cramér–Rao bounds (Section 2.3).
- Gagliardo–Nirenberg–Sobolev inequalities (Section 3).
- Beckner inequalities (Section 4).
- q-logarithmic Sobolev inequalities (Section 5).
- Nonlinear logarithmic Sobolev inequalities in product spaces (Section 7).

Part 2. Space forms: stochastic methods

- Local logarithmic Sobolev inequalities on space forms (Section 9).
- Local logarithmic Sobolev inequalities and Hamilton's matrix inequalities on nonpositively curved space forms (Section 10).

1.1. **Euclidean spaces: scaling.** Most classical functional inequalities on $\mathbb{R}^n$ are coordinate-free results phrased in a coordinate-dependent way. As such, they can often be substantially refined when expressed in a suitable basis. Concretely, the correct basis is found by performing a change of variables of the form $x \mapsto Ax$ and then optimizing over a prescribed class of symmetries $A \in \mathsf{G} \subseteq \mathsf{GL}_n$. Let us remark that explicit improvements of this form can be obtained only when it is possible to solve these optimization problems, which is not always the case. These improvements are moreover motivated by the study of equality cases. When a functional inequality has a non-constant function $h : \mathbb{R}^n \to \mathbb{R}$ as an equality case, then the refined inequality obtained in the manner described above would be saturated by all functions of the form $h_A(x) = h(Ax)$, where $A \in \mathsf{G}$. This principle has already been applied by Dembo [27] in the case of the Gaussian logarithmic Sobolev inequality (see also [30, 14] and Section 2 below). In the first part of the paper we shall present more applications of this idea to other important functional inequalities in Euclidean space and further consequences.

1.1.1. *Beckner inequalities.* In [13], Beckner proved that any smooth function $u \in C_0^\infty(\mathbb{R}^n)$ satisfies the estimates

$$\forall\ p \in [1,2), \qquad \|u\|^2_{L_2(\gamma_n)} - \|u\|^2_{L_p(\gamma_n)} \le (2-p)\|\nabla u\|^2_{L_2(\gamma_n)}. \tag{11}$$

This family of inequalities interpolates between the Gaussian Poincaré inequality (corresponding to $p = 1$) and Gross' logarithmic Sobolev inequality [36] which arises as a limit when $p \to 2^-$. We refer to the influential work of Latała and Oleszkiewicz [42] as well as [9, Section 7.6] for examples of Beckner-type inequalities satisfied by non-Gaussian measures.

In [28, Corollary 4], Dolbeault and Toscani proposed a *dimensional* refinement of Beckner's inequality (11) for functions satisfying a second moment normalization condition. More specifically, they showed that if a function $u \in C_0^\infty(\mathbb{R}^n)$ satisfies the normalization condition

$$\int_{\mathbb{R}^n} |x|^2 u(x)^2 \,\mathrm{d}\gamma_n(x) = n\|u\|^2_{L_2(\gamma_n)}, \tag{12}$$

then

$$\forall p \in [1,2), \qquad \|u\|^2_{L_2(\gamma_n)} \varphi_{p,n}\left(1 - \frac{\|u\|^2_{L_p(\gamma_n)}}{\|u\|^2_{L_2(\gamma_n)}}\right) \le \|\nabla u\|^2_{L_2(\gamma_n)}, \tag{13}$$

where the function $\varphi_{p,n}$ is given by

$$\forall\ s \in (0,1), \qquad \varphi_{p,n}(s) \stackrel{\text{def}}{=} \frac{n}{4}\Big((1-s)^{-\frac{2p}{n(2-p)}} - 1\Big). \tag{14}$$

Observe that (13) improves upon (11) up to the value of the implicit constant as

$$\forall\ s \in (0,1), \qquad \varphi_{p,n}(s) \geq \frac{p}{2(2-p)} \log\Big(\frac{1}{1-s}\Big) \geq \frac{p}{2(2-p)} \cdot s. \tag{15}$$

The improvement (13) becomes particularly substantial when $\|u\|_{L_p(\gamma_n)} \ll \|u\|_{L_2(\gamma_n)}$.

In the spirit of the matricial refinement (7) over the dimensional logarithmic Sobolev inequality (5), we present the following refinement of (13) for functions whose second moment *matrix* is appropriately normalized.

Theorem 1. *Fix $n \in \mathbb{N}$ and let $u \in C_0^\infty(\mathbb{R}^n)$ be such that*

$$\forall\ i,j \in \{1,\ldots,n\}, \qquad \int_{\mathbb{R}^n} x_i x_j u(x)^2 \,\mathrm{d}\gamma_n(x) = \delta_{ij}\|u\|^2_{L_2(\gamma_n)}, \tag{16}$$

where δ_{ij} is the Kronecker delta. Then, we have

$$\forall\ p \in [1,2), \qquad \frac{\|u\|^2_{L_2(\gamma_n)} - \|u\|^2_{L_p(\gamma_n)}}{\|u\|^2_{L_2(\gamma_n)}} \leq 1 - \left[\det\left(\frac{4}{\|u\|^2_{L_2(\gamma_n)}} \int_{\mathbb{R}^n} (\nabla u)^{\otimes 2} \,\mathrm{d}\gamma_n + \mathrm{Id}_n\right)\right]^{-\frac{2-p}{2p}}. \tag{17}$$

Applying the inequality $\det C \leq \left(\frac{\mathrm{tr}C}{n}\right)^n$ and rearranging, we see that (17) strengthens (13).

1.1.2. *Gagliardo–Nirenberg–Sobolev inequalities.* Fix $n \in \mathbb{N}$. The Gagliardo–Nirenberg inequality [34, 53] asserts that for every $p,q,r,s \in [1,\infty)$ and $\theta \in [0,1]$ satisfying the constraint

$$\frac{1}{p} = \frac{\theta}{q} + \Big(\frac{1}{r} - \frac{1}{n}\Big)(1-\theta), \tag{18}$$

there exists a universal (optimal) constant $C^{p,q,r,s} > 0$ such that every $u \in C_0^\infty(\mathbb{R}^n)$ satisfies

$$\|u\|_{L_p(\mathbb{R}^n)} \leq C^{p,q,r,s} \|u\|^\theta_{L_q(\mathbb{R}^n)} \|\nabla u\|^{1-\theta}_{L_r(\mathbb{R}^n;\ell_s^n)}, \tag{19}$$

where we use the standard notation

$$\|\nabla u\|_{L_r(\mathbb{R}^n;\ell_s^n)} = \left(\int_{\mathbb{R}^n} \Big(\sum_{i=1}^n |\partial_i u(x)|^s\Big)^{r/s} \mathrm{d}x\right)^{1/r}. \tag{20}$$

In the special case $r \in (1,n)$ and $\theta = 0$, inequality (19) boils down to the classical Sobolev inequality [59, 60]. The endpoint case $r = 1$ and $\theta = 0$ was due to [34, 53] and the corresponding optimal constant for $s = 2$ was found by Federer, Fleming and Rishel [31, 32]. The optimal constant in the range $r \in (1,n)$ and $\theta = 0$ for $s = 2$ was discovered by Aubin and Talenti [2, 62]. The logarithmic Sobolev inequality (4) can be obtained as an endpoint case of the Gagliardo–Nirenberg–Sobolev inequality (19) with the optimal constant when $s = 2$ (see [26, Section 1]). Finally, the optimal constant $C^{p,q,r,s}$ for general parameters was found by Cordero-Erausquin, Nazaret and Villani in [22, Section 3]. In this paper, we present a refined inequality for $r = s$.

Theorem 2. *Let $p,q,r \in [1,\infty)$, $\theta \in [0,1]$ and $C^{p,q,r,r} > 0$ be such that* (19) *is satisfied for all functions $u \in C_0^\infty(\mathbb{R}^n)$ with $r = s$ under the constraint* (18)*. Then, for every $u \in C_0^\infty(\mathbb{R}^n)$, we have*

$$\|u\|_{L_p(\mathbb{R}^n)} \leq C^{p,q,r,r} n^{\frac{1-\theta}{r}} \|u\|^\theta_{L_q(\mathbb{R}^n)} \Big(\prod_{j=1}^n \|\partial_j u\|_{L_r(\mathbb{R}^n)}\Big)^{\frac{1-\theta}{n}}. \tag{21}$$

The inequality (21) improves on (19) by the arithmetic mean-geometric mean inequality so Theorem 2 asserts that Euclidean Gagliardo–Nirenberg–Sobolev inequalities, that is, inequalities of the form (19) with the choice of parameter $r = 2$, self-improve via scaling. In particular, (21) captures the fact (absent from (19)) that $\partial_i u \equiv 0$ on $\mathbb{R}^n$ implies that $u \equiv 0$ under any L_s-integrability assumption for u

1.2. **Product spaces: tensorization.** If (Ω,π) is a probability space, then for a measurable function $f:\Omega\to\mathbb{R}_+$ we shall denote its entropy with respect to π by

$$\mathrm{Ent}_\pi[f]\stackrel{\mathrm{def}}{=}\int_\Omega f\log f\,\mathrm{d}\pi-\Big(\int_\Omega f\,\mathrm{d}\pi\Big)\log\Big(\int_\Omega f\,\mathrm{d}\pi\Big). \tag{22}$$

The usefulness of logarithmic Sobolev inequalities in probability and geometry stems largely from the fact that entropy satisfies a simple yet powerful tensorization principle, rendering them dimension-free estimates [43]. In the interesting work [55], Polyanskiy and Samorodnitsky introduced a family of more general inequalities for Markov semigroups called *nonlinear* logarithmic Sobolev inequalities (see also [36, 66, 23, 24, 20, 50, 57] for previous occurrences of such estimates in the literature and applications). Let $\{P_t\}_{t\geq 0}$ be a Markov semigroup acting on measurable functions $f:\Omega\to\mathbb{R}$ with stationary measure π. Following [55], we say that $\{P_t\}_{t\geq 0}$ satisfies the (p,Φ)–LSI, where $p\geq 1$ and $\Phi:\mathbb{R}_+\to\mathbb{R}_+$ is a concave, continuous function with $\Phi(0)=0$, if for every measurable function $f:\Omega\to\mathbb{R}_+$, we have

$$\mathrm{Ent}_\pi[f^p]\leq\mathbb{E}_\pi[f^p]\Phi\left(\frac{\mathcal{E}(f,f^{p-1})}{\mathbb{E}_\pi[f^p]}\right), \tag{23}$$

where $\mathcal{E}(\cdot,\cdot)$ is the Dirichlet form corresponding to $\{P_t\}_{t\geq 0}$. As usual, the term $\mathcal{E}(f,f^{p-1})$ is interpreted as $\mathcal{E}(f,\log f)$ in the endpoint case $p=1$.

In [55, Theorem 1], the authors proved a dimensional tensorization property for nonlinear log-Sobolev inequalities asserting that if $\{P_t\}_{t\geq 0}$ satisfies the (p,Φ)–LSI, then for any $n\geq 1$, the product semigroup $\{P_t^{\otimes n}\}_{t\geq 0}$ with stationary measure π^n satisfies the $\big(p,n\Phi(\frac{1}{n}\cdot)\big)$–LSI:

$$\mathrm{Ent}_{\pi^n}[f^p]\leq n\mathbb{E}_{\pi^n}[f^p]\Phi\left(\frac{\mathcal{E}(f,f^{p-1})}{n\mathbb{E}_{\pi^n}[f^p]}\right). \tag{24}$$

By considering functions f of the form $f(x_1,\ldots,x_n)=\tilde{f}(x_1,\ldots,x_k)$, for $k<n$, we see that (24) suffers from the problem of incorporating the ambient dimension n into the constant, thus ignoring the structure of f. In the Euclidean setting, we overcame this issue by finding the correct basis via an optimization procedure over the cone of positive semidefinite matrices. In contrast, such an approach is not suitable on the Hamming cube due to its discrete nature. Our solution to this problem is to refine tensorization instead of scaling. Indeed, as a consequence of a more general tensorization principle (see Theorem 18 below), we shall prove the following stronger nonlinear logarithmic Sobolev inequality for product spaces.

Theorem 3. *Let $(\Omega,\pi,\{P_t\}_{t\geq 0})$ be a stationary Markov semigroup satisfying the (p,Φ)–LSI for some $p\geq 1$ and some concave, continuous function $\Phi:\mathbb{R}_+\to\mathbb{R}_+$ with $\Phi(0)=0$. Then, for any $n\geq 1$, every measurable function $f:\Omega^n\to\mathbb{R}_+$ satisfies*

$$\mathrm{Ent}_{\pi^n}[f^p]\leq\mathbb{E}_{\pi^n}[f^p]\sum_{i=1}^n\Phi\left(\frac{\mathbb{E}_{\pi^n}\big[\mathcal{E}_i(f,f^{p-1})\big]}{\mathbb{E}_{\pi^n}[f^p]}\right), \tag{25}$$

where $\mathcal{E}_i(\cdot,\cdot)$ is the Dirichlet form associated with the i-th component of the semigroup $\{P_t^{\otimes n}\}_{t\geq 0}$.

It follows readily from Jensen's inequality that

$$\sum_{i=1}^n\Phi\left(\frac{\mathbb{E}_{\pi^n}\big[\mathcal{E}_i(f,f^{p-1})\big]}{\mathbb{E}_{\pi^n}[f^p]}\right)\leq n\Phi\left(\frac{\mathcal{E}(f,f^{p-1})}{n\mathbb{E}_{\pi^n}[f^p]}\right), \tag{26}$$

where $\mathcal{E}(\cdot,\cdot)$ is the Dirichlet form associated to $\{P_t^{\otimes n}\}_{t\geq 0}$ and thus (25) indeed strengthens (24). Moreover, in [55, Theorems 4 and 6], the authors found the *optimal* functions Φ_p such that the (p,Φ_p)-LSI is satisfied on the one-dimensional Hamming cube $\{0,1\}$ equipped with the uniform measure. Tensorizing their result via Theorem 3, one deduces an improved nonlinear logarithmic Sobolev inequality on the Hamming cube $\{0,1\}^n$.

1.3. **Space forms: stochastic methods.** In order to explain our intrinsic dimensional functional inequalities on space forms we first recall the notion of *local* logarithmic Sobolev inequalities. Starting with the Euclidean setting, fix $T \geq 0$, $x \in \mathbb{R}^n$, and let $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda_n} = \frac{f P_T \delta_x}{P_T f(x)}$ where δ_x is the Dirac mass at x, $f : \mathbb{R}^n \to \mathbb{R}$ is a nonnegative function, and $\{P_t\}_{t\geq 0}$ is the Euclidean heat semigroup given by $P_t h(x) := \int h(x + \sqrt{t}z) d\gamma_n(z)$. Plugging μ into (5) yields (after integration by parts and using the explicit form of $P_T \delta_x$),

$$P_T(f \log f)(x) - P_T f(x) \log P_T f(x) \leq \frac{T}{2} P_T \Delta f(x) + \frac{n}{2} P_T f(x) \log\left(1 - \frac{T}{n} \frac{P_T(f \Delta \log f)(x)}{P_T f(x)}\right). \tag{27}$$

The inequality (27) is the *local* dimensional logarithmic Sobolev inequality on $\mathbb{R}^n$ [10]. While (27) provides an *upper* bound on the (local) entropy, the *reverse* local dimensional logarithmic Sobolev inequality [10] provides a *lower* bound,

$$P_T(f \log f)(x) - P_T f(x) \log P_T f(x) \geq \frac{T}{2} P_T \Delta f(x) - \frac{n}{2} P_T f(x) \log\left(1 + \frac{T}{n} \Delta \log P_T f(x)\right). \tag{28}$$

Analogously, we can use (7), instead of (5), to get the local intrinsic dimensional logarithmic Sobolev inequality on $\mathbb{R}^n$,

$$P_T(f \log f)(x) - P_T f(x) \log P_T f(x) \leq \frac{T}{2} P_T \Delta f(x) + \frac{1}{2} P_T f(x) \log\det\left(\mathsf{Id}_n - T \frac{P_T(f \nabla^2 \log f(x))}{P_T f(x)}\right), \tag{29}$$

which improves on (27). As for a *reverse* local intrinsic dimensional logarithmic Sobolev inequality in $\mathbb{R}^n$, we will establish below (Theorem 32) that

$$P_T(f \log f)(x) - P_T f(x) \log P_T f(x) \geq \frac{T}{2} \Delta P_T f(x) - \frac{1}{2} P_T f(x) \log\det\left(\mathsf{Id}_n + T \nabla^2 \log P_T f(x)\right), \tag{30}$$

which improves on (28).

Turning to the manifold setting, local dimensional logarithmic Sobolev inequalities exist on manifolds in forms which account for both the dimension of the manifold as well as the Ricci curvature [6]. In light of the existence of the local intrinsic dimensional logarithmic Sobolev inequalities on Euclidean spaces (29) and (30), we wish to understand whether such inequalities can also exist on manifolds. Upon closer inspection, however, it is clear that inequalities such as (29) and (30) cannot hold if the only curvature information given pertains to the Ricci tensor. On a conceptual level, the difference between the dimensional and intrinsic dimensional inequalities is that the former provide information about the *trace* of the Fisher information matrix, while the latter provide information about the *full spectrum*. Hence, while information on the trace of the Riemann tensor, i.e., Ricci curvature, suffices to yield a dimensional inequality, information on the full Riemann tensor, i.e., sectional curvature, should be required to give an intrinsic dimensional inequality.

A concrete manifestation of this intuition is exhibited by the inequalities of Li–Yau and Hamilton [47, 37]. As was realized in [10], the reverse local dimensional logarithmic Sobolev inequality (28) implicitly implies the Li–Yau inequality on $\mathbb{R}^n$,

$$\forall x \in \mathbb{R}^n, \qquad -\Delta \log P_T f(x) \leq \frac{n}{T}, \tag{31}$$

since the argument in the log term of (28) must be nonnegative. Analogously, the reverse local intrinsic dimensional logarithmic Sobolev inequality (30) implies Hamilton's inequality,

$$\forall x \in \mathbb{R}^n, \qquad -\nabla^2 \log P_T f(x) \preceq \frac{1}{T} \mathsf{Id}_n, \tag{32}$$

where $\preceq$ is the order of positive semidefinite matrices. In the manifold setting, the Li–Yau inequality, which is a statement about the *trace* of the Hessian of $\log P_T f$, holds under a nonnegativity assumption on the *trace* of the Riemann tensor, namely the Ricci tensor [47, 68]. Indeed, Bakry and Ledoux [10] (see also the follow-up work [6]) established (reverse) local dimensional logarithmic Sobolev inequalities on manifolds with nonnegative *Ricci* curvature

which imply the Li–Yau inequality. In contrast, Hamilton's inequality, which is a statement about the *Hessian* of $\log P_T f$, requires the manifold to have nonnegative sectional curvature (and also to be Einstein), which is an assumption on the *full spectrum* of the Riemann tensor [37]. It follows that if local intrinsic dimensional logarithmic Sobolev inequalities were to hold, then information about the sectional curvature should be provided.

In this work we establish local intrinsic dimensional logarithmic Sobolev inequalities as well as Hamilton-type matrix inequalities for space forms: Euclidean spaces, spheres, and hyperbolic spaces. In addition to serving as the model spaces for functional inequalities on manifolds, these spaces are the simplest non-trivial examples of manifolds where we could hope for local intrinsic dimensional logarithmic Sobolev inequalities to hold. The methods of scaling and tensorization which worked, respectively, for Euclidean spaces and product spaces no longer apply on curved spaces as they lack product and homogeneity structures. Hence, we take a different route and build on the stochastic approach of Lehec [44, 45] and Eldan, Lehec, and Shenfeld [30] towards logarithmic Sobolev inequalities. We start by stating our local intrinsic dimensional logarithmic Sobolev on space forms while deferring precise definitions to Part 2.

Theorem 4. *Let* (M, g) *be an n-dimensional Riemannian manifold with constant sectional curvature* $\kappa \in \mathbb{R} \setminus \{0\}$ *with the associated heat semigroup* $\{P_t\}_{t\geq 0}$. *Fix* $T > 0$, $x \in \mathsf{M}$, *a smooth positive function* $f : \mathsf{M} \to \mathbb{R}$ *with* $\int_M f \, \mathrm{d}P_T\delta_x = 1$, *and let* μ *be the probability measure with* $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x} = f$. *Define the 2-tensor* $C(t) = \frac{e^{n\kappa t}}{n\kappa}A + tB$ *for* $t \in \mathbb{R}$ *where* A, B *are the 2-tensors given by*

$$\begin{cases} A = -e^{-n\kappa T}\Big(P_T\nabla^2 f(x) - \frac{1}{n}P_T\Delta f(x)\cdot \mathrm{g}\Big) \\ B = \Big(\frac{(n-1)\kappa}{2} - \frac{\Delta P_T f(x)}{n}\Big)\cdot \mathrm{g}. \end{cases} \tag{33}$$

Then, we have the local intrinsic dimensional logarithmic Sobolev inequality

$$\begin{aligned} &P_T(f\log f)(x) - P_T f(x)\log P_T f(x) \\ &\leq \frac{1}{2}\int_0^T \mathrm{tr}\Big[e^{C(t)-C(T)}\Big(\mathrm{g} + \mathbb{E}_\mu\big(\nabla\log f\big)^{\otimes 2}\int_t^T e^{2C(s)-2C(T)}\,\mathrm{d}s\Big)^{-1}\mathbb{E}_\mu\big(\nabla\log f\big)^{\otimes 2}e^{C(t)-C(T)}\Big]\,\mathrm{d}t, \end{aligned} \tag{34}$$

and the reverse local intrinsic dimensional logarithmic Sobolev inequality

$$\begin{aligned} &P_T(f\log f)(x) - P_T f(x)\log P_T f(x) \\ &\geq \frac{1}{2}\int_0^T \mathrm{tr}\Big[e^{C(t)-C(0)}\Big(\mathrm{g} - \big(\nabla\log P_T f(x)\big)^{\otimes 2}\int_0^t e^{2C(s)-2C(0)}\,\mathrm{d}s\Big)^{-1}\big(\nabla\log P_T f(x)\big)^{\otimes 2}e^{C(t)-C(0)}\Big]\,\mathrm{d}t. \end{aligned} \tag{35}$$

As will become clear from the proof of Theorem 4, the theorem is not optimal and follows from a more powerful "master " matrix differential inequality (section 10.3). There are other inequalities which can be deduced from the master matrix differential inequality, specifically in space forms with nonpositive sectional curvature. In particular, we prove Hamilton-type matrix inequalities for the heat equation:

Theorem 5. *Let* (M, g) *be an n-dimensional Riemannian manifold with constant nonpositive sectional curvature* $\kappa \leq 0$. *Let* $\{P_t\}_{t\geq 0}$ *be the associated heat semigroup and let* $f : \mathsf{M} \to \mathbb{R}$ *be a positive function. Then, for every* $x \in \mathsf{M}$ *and every* $T \geq 0$,

$$\text{if, either } \kappa = 0, \text{ or } \kappa < 0 \text{ and } \frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} = 1, \quad \text{then} \quad -\nabla^2\log P_T f(x) \preceq \frac{1}{T}\mathrm{Id}_n \quad \forall x \in \mathsf{M}. \tag{36}$$

Further,

$$\begin{aligned} &\text{if} \quad \frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} > 1, \\ &\text{then} \quad -\nabla^2 \log P_T f(x) \le \frac{n\kappa}{2}\left\{\sqrt{\frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} - 1}\cot\left(\frac{n\kappa T}{2}\sqrt{\frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} - 1}\right) - 1\right\}\mathsf{Id}_n. \end{aligned} \tag{37}$$

In flat space, where $\kappa = 0$, Theorem 5 reduces to (32), namely, Hamilton's matrix inequality [37, Corollary 4.4]. In hyperbolic spaces, Theorem 5 is completely new. The constraint $\frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} > 1$ is natural. Indeed, Theorem 5 is a matrix version of the improved Li-Yau inequality of Bakry, Bolley, and Gentil—see Remark 34.

Going beyond matrix inequalities, we can use our master matrix differential inequality to obtain another form of local intrinsic dimensional logarithmic Sobolev inequalities.

Theorem 6. *Let* (M, g) *be the n-dimensional hyperbolic space with sectional curvature* $\kappa < 0$ *with the associated heat semigroup* $\{P_t\}_{t\ge 0}$*. Fix* $T > 0$*,* $x \in \mathsf{M}$*, a smooth positive function* $f : \mathsf{M} \to \mathbb{R}$ *with* $\int_M f \, \mathrm{d}P_T\delta_x = 1$*, and let* μ *be the probability measure with* $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x} = f$*. Then, with*

$$\lambda \stackrel{\text{def}}{=} \frac{n^2\kappa^2}{4}\left\{\frac{4}{n^2\kappa}\Delta P_T f(x) - 1\right\}, \quad \alpha_i \stackrel{\text{def}}{=} \begin{cases} \arctan\left(\frac{1}{\sqrt{\lambda}}\left(\sigma_i + \frac{n\kappa}{2}\right)\right) & \text{if } \lambda < 0, \\ -\frac{2}{2\sigma_i + n\kappa} & \text{if } \lambda = 0, \\ \operatorname{arctanh}\left(-\frac{1}{\sqrt{-\lambda}}\left(\sigma_i + \frac{n\kappa}{2}\right)\right) & \text{if } \lambda > 0, \end{cases} \tag{38}$$

we have the local intrinsic dimensional logarithmic Sobolev inequality

$$P_T(f\log f)(x) - P_T f(x)\log P_T f(x) \le -\frac{P_T\Delta f(x)}{2} - \frac{n^2\kappa T}{2} - \frac{1}{2}\begin{cases} \sum_{i=1}^n \log\left(\frac{\cos(\alpha_i)}{\cos(\sqrt{\lambda}T + \alpha_i)}\right) & \text{if } \lambda > 0 \\ \sum_{i=1}^n \log\left(\frac{\alpha_i}{T + \alpha_i}\right) & \text{if } \lambda = 0 \\ \sum_{i=1}^n \log\left(\frac{\cosh(\alpha_i)}{\cosh(\sqrt{-\lambda}T + \alpha_i)}\right) & \text{if } \lambda < 0 \end{cases} \tag{39}$$

where $\{\sigma_i\}_{i=1}^n$ *are the eigenvalues of* $\mathbb{E}_\mu[-\nabla^2 \log f]$*, and the reverse local intrinsic dimensional logarithmic Sobolev inequality*

$$P_T(f\log f)(x) - P_T f(x)\log P_T f(x) \ge -\frac{P_T\Delta f(x)}{2} - \frac{n^2\kappa T}{2} + \frac{1}{2}\begin{cases} \sum_{i=1}^n \log\left(\frac{\cos(\alpha_i)}{\cos(\sqrt{\lambda}T + \alpha_i)}\right) & \text{if } \lambda > 0 \\ \sum_{i=1}^n \log\left(\frac{\alpha_i}{T + \alpha_i}\right) & \text{if } \lambda = 0 \\ \sum_{i=1}^n \log\left(\frac{\cosh(\alpha_i)}{\cosh(\sqrt{-\lambda}T + \alpha_i)}\right) & \text{if } \lambda < 0 \end{cases} \tag{40}$$

where $\{\sigma_i\}_{i=1}^n$ *are the eigenvalues of* $-\nabla^2 \log P_T f(x)$*.*

Acknowledgements. We are grateful to Dario Cordero-Erausquin, Max Fathi, Nathael Gozlan, and Yury Polyanskiy for useful pointers to the literature and to Georgios Moschidis for many helpful discussions.

Part 1. Euclidean and product spaces: scaling and tensorization

2. Logarithmic Sobolev inequalities in Euclidean spaces and Cramér–Rao bounds

In this section we discuss strengthenings of logarithmic Sobolev inequalities for measures on Euclidean spaces by means of scaling. In addition, we derive an application of these inequalities to Bayesian Cramér–Rao bounds.

2.1. **Warm-up: Gross' inequality.** The Euclidean reformulation (4) of the logarithmic Sobolev inequality in Gauss space [36] asserts that if $f : \mathbb{R}^n \to \mathbb{R}_+$ is a probability density, then

$$\int_{\mathbb{R}^n} f(x)\log f(x)\,\mathrm{d}x - \mathsf{H}(\gamma_n\|\lambda_n) \le \frac{1}{2}\left(\int_{\mathbb{R}^n} \frac{|\nabla f(x)|^2}{f(x)}\,\mathrm{d}x - n\right). \tag{41}$$

Fix such a density f and consider the reparametrized density $f_A : \mathbb{R}^n \to \mathbb{R}_+$ which is given by $f_A(x) = (\det A)\cdot f(Ax)$, where $A \in \mathsf{GL}_n$ is a positive definite matrix. Applying (41) for f_A we get

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\log f(x)\,\mathrm{d}x + \log\det A - \mathsf{H}(\gamma_n\|\lambda_n) &= \int_{\mathbb{R}^n} f_A(x)\log f_A(x)\,\mathrm{d}x - \mathsf{H}(\gamma_n\|\lambda_n)\\ &\le \frac{1}{2}\left(\int_{\mathbb{R}^n}\frac{|\nabla f_A(x)|^2}{f_A(x)}\,\mathrm{d}x - n\right) = \frac{1}{2}\left(\int_{\mathbb{R}^n}\frac{|A\cdot\nabla f(x)|^2}{f(x)}\,\mathrm{d}x - n\right),\end{aligned} \tag{42}$$

which after rearranging becomes

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\log f(x)\,\mathrm{d}x - \mathsf{H}(\gamma_n\|\lambda_n) &\le \frac{1}{2}\left(\int_{\mathbb{R}^n}\frac{|A\cdot\nabla f(x)|^2}{f(x)}\,\mathrm{d}x - \log\det A^2 - n\right)\\ &= \frac{1}{2}\Big(\mathrm{tr}\Big(A^2\cdot\mathcal{I}(\mu\|\lambda_n)\Big) - \log\det A^2 - n\Big).\end{aligned} \tag{43}$$

For the optimal choice of matrix $A = \mathcal{I}(\mu\|\lambda_n)^{-1/2}$, (43) readily becomes Dembo's inequality (7). Observe that in this argument we made critical use of the change of variables formula for the Lebesgue measure, i.e., that $\lambda_n(AK) = (\det A)\cdot\lambda_n(K)$ for any Borel $K \subset \mathbb{R}^n$ and $A \in \mathsf{GL}_n$. While Lebesgue is the only measure on Euclidean space satisfying such an invariance property under all linear transformations, in the next section we shall observe that a weaker self-improvement can be deduced for measures which behave well under diagonal linear maps.

2.2. **Logarithmic Sobolev inequalities for homogeneous measures.** Let $p_1,\ldots,p_n \ge 0$. An absolutely continuous measure ρ on $\mathbb{R}^n$ with density $w : \mathbb{R}^n \to \mathbb{R}_+$ is called $(p_1,\ldots,p_n)$-homogeneous if for every $t_1,\ldots,t_n > 0$,

$$\forall\ x = (x_1,\ldots,x_n) \in \mathbb{R}^n, \qquad w(t_1x_1,\ldots,t_nx_n) = t_1^{p_1}\cdots t_n^{p_n} w(x_1,\ldots,x_n). \tag{44}$$

Theorem 7. *Fix $c_1, c_2 > 0$, $n \in \mathbb{N}$, $p_1,\ldots,p_n \ge 0$ and let ρ be a $(p_1,\ldots,p_n)$-homogeneous measure such that for any Borel probability measure μ on $\mathbb{R}^n$,*

$$\mathsf{H}(\mu\|\rho) \le c_1\mathsf{I}(\mu\|\rho) + c_2. \tag{45}$$

Then, for any Borel probability measure μ on $\mathbb{R}^n$ with positive differentiable density f, we have

$$\mathsf{H}(\mu\|\rho) \le \frac{1}{2}\sum_{k=1}^n (1+p_k)\log\left(\frac{2ec_1}{1+p_k}\int_{\mathbb{R}^n}\frac{(\partial_k f(y))^2}{f(y)}\,\mathrm{d}\rho(y)\right) + c_2. \tag{46}$$

The existence of homogeneous measures ρ satisfying inequalities of the form (45), as well as more general *entropy-energy inequalities* follows, for instance, from [9, Proposition 7.3.1].

Proof of Theorem 7. Let $f = \frac{\mathrm{d}\mu}{\mathrm{d}\rho}$ be an arbitrary positive function with ρ-integral equal to 1 and fix $t_1,\ldots,t_n > 0$. The measure μ_t with density $x \mapsto f_t(x) = t_1^{1+p_1}\cdots t_n^{1+p_n} f(t_1x_1,\ldots,t_nx_n)$ with respect to ρ is a probability measure, as

$$\prod_{i=1}^n t_i^{1+p_i}\int_{\mathbb{R}^n} f(t_1x_1,\ldots,t_nx_n)w(x)\,\mathrm{d}x = \prod_{i=1}^n t_i^{p_i}\int_{\mathbb{R}^n} f(y)w(t_1^{-1}y_1,\ldots,t_n^{-1}y_n)\,\mathrm{d}y = 1, \tag{47}$$

where we made the change of variables $(y_1,\dots,y_n) = (t_1x_1,\dots,t_nx_n)$. We have,

$$\mathsf{H}(\mu_t\|\rho) = \prod_{i=1}^n t_i^{1+p_i} \int_{\mathbb{R}^n} f(t_1x_1,\dots,t_nx_n)\left\{\log f(t_1x_1,\dots,t_nx_n) + \sum_{k=1}^n (1+p_k)\log t_k\right\} w(x)\,\mathrm{d}x$$
$$= \prod_{i=1}^n t_i^{p_i} \int_{\mathbb{R}^n} f(y)\left\{\log f(y) + \sum_{k=1}^n (1+p_k)\log t_k\right\} w(t_1^{-1}y_1,\dots,t_n^{-1}y_n)\,\mathrm{d}y = \mathsf{H}(\mu\|\rho) + \sum_{k=1}^n (1+p_k)\log t_k.$$

Similarly, assuming in addition that f is differentiable, for every $k\in\{1,\dots,n\}$ we have $\partial_k f_t(x) = t_k\prod_i t_i^{1+p_i}\partial_k f(t_1x_1,\dots,t_nx_n)$ and thus

$$\begin{aligned}\mathsf{I}(\mu_t\|\rho) &= \sum_{k=1}^n t_k^2\cdot\prod_{i=1}^n t_i^{1+p_i}\int_{\mathbb{R}^n}\frac{(\partial_k f(t_1x_1,\dots,t_nx_n))^2}{f(t_1x_1,\dots,t_nx_n)}w(x)\,\mathrm{d}x\\ &= \sum_{k=1}^n t_k^2\cdot\prod_{i=1}^n t_i^{p_i}\int_{\mathbb{R}^n}\frac{(\partial_k f(y))^2}{f(y)}w(t_1^{-1}y_1,\dots,t_n^{-1}y_n)\,\mathrm{d}y = \sum_{k=1}^n t_k^2\int_{\mathbb{R}^n}\frac{(\partial_k f(y))^2}{f(y)}\,\mathrm{d}\rho(y).\end{aligned} \tag{48}$$

Therefore, applying (45) for μ_t and reorganizing the terms, we deduce that

$$\mathsf{H}(\mu\|\rho) \le \inf_{t_1,\dots,t_n\ge 0}\left\{c_1\sum_{k=1}^n t_k^2\int_{\mathbb{R}^n}\frac{(\partial_k f(y))^2}{f(y)}\,\mathrm{d}\rho(y) - \sum_{k=1}^n(1+p_k)\log t_k + c_2\right\}.$$

It is now elementary to check that the above infimum is attained when

$$\forall\ k\in\{1,\dots,n\},\qquad t_k^2 = \frac{1+p_k}{2c_1}\Big(\int_{\mathbb{R}^n}\frac{(\partial_k f(y))^2}{f(y)}\,\mathrm{d}\rho(y)\Big)^{-1} \tag{49}$$

and plugging this choice of parameters completes the proof. □

Specifically for Lebesgue measure, Theorem 7 implies that if μ has differentiable density f,

$$\mathsf{H}(\mu\|\lambda_n) - \mathsf{H}(\gamma_n\|\lambda_n) \le \frac{1}{2}\sum_{k=1}^n \log\Big(\int_{\mathbb{R}^n}\frac{(\partial_k f(y))^2}{f(y)}\,\mathrm{d}y\Big), \tag{50}$$

which is weaker than Dembo's inequality (7) in view of the elementary estimate $\det C \le \prod_s C_{ss}$ which holds for all positive semidefinite matrices C. On the other hand, (50) combined with Jensen's inequality implies (5). We refer to [7, 10, 5, 6] for further dimensional logarithmic Sobolev inequalities and applications to Li–Yau-type estimates [47], hypercontractivity [51, 19, 52, 12, 4] and heat kernel estimates [4, 7].

2.3. **A Bayesian Cramér–Rao bound.** In [1], Aras, Lee, Pananjady and Courtade observed that logarithmic Sobolev inequalities formally imply Bayesian Cramér–Rao bounds, thus extending some results of Efroimovich [29] for Gaussian measures. In this section, we investigate similar applications of intrinsic dimensional log-Sobolev inequalities in the spirit of (46) and (7).

Following [1], we work in the setting of parametric statistics. Let $\{\mu_\theta\}_{\theta\in\mathbb{R}^n}$ be a family of probability measures on a measurable space $(\Omega,\mathcal{F})$. Assume moreover that there exists a dominating σ-finite measure λ on Ω such that μ_θ has a positive density with respect to λ,

$$\forall\ \theta\in\mathbb{R}^n,\qquad \mathrm{d}\mu_\theta(x) = f(x;\theta)\,\mathrm{d}\lambda(x). \tag{51}$$

We shall assume throughout that each function $\theta\mapsto f(x;\theta)$ is smooth and that

$$\int_\Omega \nabla_\theta f(x;\theta)\,\mathrm{d}\lambda(x) = 0 \tag{52}$$

for almost every $\theta\in\mathbb{R}^n$. The Fisher information of the parametric family $\{\mu_\theta\}_{\theta\in\mathbb{R}^n}$ is

$$\forall\ \theta\in\mathbb{R}^n,\qquad \mathcal{J}(\theta) \stackrel{\mathrm{def}}{=} \int_\Omega\frac{|\nabla_\theta f(x;\theta)|^2}{f(x;\theta)}\,\mathrm{d}\lambda(x). \tag{53}$$

Finally, if π is a probability measure on $\mathbb{R}^n$, we denote the mutual information of π with the family $\{\mu_\theta\}_{\theta\in\mathbb{R}^n}$ by

$$I\Big(\pi;\{\mu_\theta\}\Big) \stackrel{\text{def}}{=} \int_{\mathbb{R}^n}\int_\Omega f(x;\theta)\log\left(\frac{f(x;\theta)}{\int_{\mathbb{R}^n} f(x,\phi)\,\mathrm{d}\pi(\phi)}\right)\mathrm{d}\lambda(x)\,\mathrm{d}\pi(\theta). \tag{54}$$

The main result of [1, Theorem 1] specified to the standard Gaussian measure γ_n asserts that for every absolutely continuous probability measure π on $\mathbb{R}^n$,

$$I\Big(\pi;\{\mu_\theta\}\Big) + \mathsf{H}(\pi\|\gamma_n) \le \frac{1}{2}\Big(\mathsf{I}(\pi\|\gamma_n) + \int_{\mathbb{R}^n} \mathcal{J}(\theta)\,\mathrm{d}\pi(\theta)\Big). \tag{55}$$

Inequality (55) implies the Gaussian logarithmic Sobolev inequality (1) since choosing $\mu_\theta = \lambda$ independently of θ, the terms $I(\pi;\{\mu_\theta\})$ and $\mathcal{J}(\theta)$ both vanish. We present inequalities in the spirit of (55) for homogeneous measures satisfying a log-Sobolev inequality of the form (45).

Theorem 8. *Fix $c_1, c_2 > 0$, $n \in \mathbb{N}$, $p_1,\ldots,p_n \ge 0$ and let ρ be a $(p_1,\ldots,p_n)$-homogeneous measure such that for any Borel measure μ on $\mathbb{R}^n$,*

$$\mathsf{H}(\mu\|\rho) \le c_1\mathsf{I}(\mu\|\rho) + c_2. \tag{56}$$

Then, for every parametric family $\{\mu_\theta\}_{\theta\in\mathbb{R}^n}$ and every absolutely continuous measure π on $\mathbb{R}^n$ whose density with respect to ρ is $h:\mathbb{R}^n\to\mathbb{R}_+$, we have

$$\begin{aligned} &I\Big(\pi;\{\mu_\theta\}\Big) + \mathsf{H}(\pi\|\rho) \\ &\le \frac{1}{2}\sum_{k=1}^n (1+p_k)\log\left(\frac{2ec_1}{1+p_k}\left(\int_{\mathbb{R}^n}\frac{(\partial_k h(\theta))^2}{h(\theta)}\,\mathrm{d}\rho(\theta) + \int_{\mathbb{R}^n}\int_\Omega \frac{(\partial_{\theta_k} f(x;\theta))^2}{f(x;\theta)}\,\mathrm{d}\lambda(x)\,\mathrm{d}\pi(\theta)\right)\right) + c_2. \end{aligned} \tag{57}$$

Observe that the terms inside the logarithm on the right-hand side are the k-th component of the Fisher informations $\mathsf{I}(\pi|\rho)$ and $\mathcal{J}(\theta)$ respectively, in analogy with Theorem 7.

Proof of Theorem 8. Consider the function $f:\Omega\to\mathbb{R}_+$ given by

$$\forall\ x\in\Omega, \qquad f(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} f(x;\theta)\,\mathrm{d}\pi(\theta) \tag{58}$$

and observe that

$$\int_\Omega f(x)\,\mathrm{d}\lambda(x) = \int_\Omega\int_{\mathbb{R}^n} f(x;\theta)\,\mathrm{d}\pi(\theta)\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}^n}\int_\Omega \mathrm{d}\mu_\theta(x)\,\mathrm{d}\pi(\theta) = 1. \tag{59}$$

Moreover, for $x\in\Omega$, consider the function $h_x:\mathbb{R}^n\to\mathbb{R}_+$ given by

$$\forall\ \theta\in\mathbb{R}^n, \qquad h_x(\theta) \stackrel{\text{def}}{=} \frac{h(\theta)f(x;\theta)}{f(x)} \tag{60}$$

and notice that the measure ν_x on $\mathbb{R}^n$ with $\mathrm{d}\nu_x(\theta) = h_x(\theta)\,\mathrm{d}\rho(\theta)$ is a probability measure since

$$\nu_x(\mathbb{R}^n) = \int_{\mathbb{R}^n}\frac{h(\theta)f(x;\theta)}{f(x)}\,\mathrm{d}\rho(\theta) = \int_{\mathbb{R}^n}\frac{f(x;\theta)}{f(x)}\,\mathrm{d}\pi(\theta) \stackrel{(58)}{=} 1. \tag{61}$$

By Theorem 7 and the assumption on ρ, for every $x\in\Omega$ we have

$$\mathsf{H}(\nu_x\|\rho) \le \frac{1}{2}\sum_{k=1}^n (1+p_k)\log\left(\frac{2ec_1}{1+p_k}\int_{\mathbb{R}^n}\frac{(\partial_k h_x(\theta))^2}{h_x(\theta)}\,\mathrm{d}\rho(\theta)\right) + c_2. \tag{62}$$

Integrating this inequality with respect to the probability measure $f(x)\,\mathrm{d}\lambda(x)$, we get

$$\begin{aligned} \int_\Omega \mathsf{H}(\nu_x\|\rho)f(x)\,\mathrm{d}\lambda(x) &\le \frac{1}{2}\sum_{k=1}^n (1+p_k)\int_\Omega \log\left(\frac{2ec_1}{1+p_k}\int_{\mathbb{R}^n}\frac{(\partial_k h_x(\theta))^2}{h_x(\theta)}\,\mathrm{d}\rho(\theta)\right)f(x)\,\mathrm{d}\lambda(x) + c_2 \\ &\le \frac{1}{2}\sum_{k=1}^n (1+p_k)\log\left(\frac{2ec_1}{1+p_k}\int_\Omega\int_{\mathbb{R}^n}\frac{(\partial_k h_x(\theta))^2}{h_x(\theta)}f(x)\,\mathrm{d}\rho(\theta)\,\mathrm{d}\lambda(x)\right) + c_2, \end{aligned} \tag{63}$$

where the last line follows from Jensen's inequality. Moreover, by definition we have

$$\begin{aligned}\int_\Omega \mathsf{H}(\nu_x\|\rho) f(x)\,\mathrm{d}\lambda(x) &= \int_\Omega \int_{\mathbb{R}^n} h(\theta) f(x;\theta) \log\Big(\frac{h(\theta) f(x;\theta)}{f(x)}\Big)\,\mathrm{d}\rho(\theta)\,\mathrm{d}\lambda(x) \\ &= \int_{\mathbb{R}^n} h(\theta)\log h(\theta)\,\mathrm{d}\rho(\theta) + \int_{\mathbb{R}^n}\int_\Omega f(x;\theta)\log\Big(\frac{f(x;\theta)}{f(x)}\Big)\,\mathrm{d}\lambda(x)\,\mathrm{d}\pi(\theta) \\ &= \mathsf{H}(\pi\|\rho) + I\Big(\pi;\{\mu_\theta\}\Big).\end{aligned} \tag{64}$$

Similarly, computing the integral on the right-hand side of (63), gives

$$\begin{aligned}\int_\Omega\int_{\mathbb{R}^n} \frac{(\partial_k h_x(\theta))^2}{h_x(\theta)} f(x)\,\mathrm{d}\rho(\theta)\,\mathrm{d}\lambda(x) &= \int_\Omega\int_{\mathbb{R}^n} \frac{\Big(f(x;\theta)\partial_k h(\theta) + h(\theta)\partial_{\theta_k} f(x;\theta)\Big)^2}{h(\theta) f(x;\theta)}\,\mathrm{d}\rho(\theta)\,\mathrm{d}\lambda(x) \\ &= \int_{\mathbb{R}^n} \frac{(\partial_k h(\theta))^2}{h(\theta)} \int_\Omega f(x;\theta)\,\mathrm{d}\lambda(x)\,\mathrm{d}\rho(\theta) + 2\int_{\mathbb{R}^n} \partial_k h(\theta) \int_\Omega \partial_{\theta_k} f(x;\theta)\,\mathrm{d}\lambda(x)\,\mathrm{d}\rho(\theta) \\ &\qquad + \int_{\mathbb{R}^n} h(\theta) \int_\Omega \frac{(\partial_{\theta_k} f(x;\theta))^2}{f(x;\theta)}\,\mathrm{d}\lambda(x)\,\mathrm{d}\rho(\theta) \\ &\overset{(52)}{=} \int_{\mathbb{R}^n} \frac{(\partial_k h(\theta))^2}{h(\theta)}\,\mathrm{d}\rho(\theta) + \int_{\mathbb{R}^n}\int_\Omega \frac{(\partial_{\theta_k} f(x;\theta))^2}{f(x;\theta)}\,\mathrm{d}\lambda(x)\,\mathrm{d}\pi(\theta).\end{aligned}$$

Combining everything, we deduce the desired inequality. □

Remark 9. *In the case of the Gaussian measure $\rho = \gamma_n$, we have at our disposal the intrinsic dimensional logarithmic Sobolev inequality* (7)*. Repeating the same proof mutatis mutandis while replacing* (46) *with* (7)*, we conclude that for any probability measure π on $\mathbb{R}^n$ whose density with respect to γ_n is $h : \mathbb{R}^n \to \mathbb{R}_+$, and for every parametric family $\{\mu_\theta\}_{\theta\in\mathbb{R}^n}$, we have*

$$I\Big(\pi;\{\mu_\theta\}\Big) + \mathsf{H}(\pi\|\gamma_n) \le \frac{\mathrm{tr}\mathsf{M}_{2,\pi} - n}{2} + \frac{1}{2}\log\det\Big(2\mathsf{Id}_n + \mathcal{I}(\pi\|\gamma_n) + \int_{\mathbb{R}^n}\int_\Omega \frac{(\nabla_\theta f(x;\theta))^{\otimes 2}}{f(x;\theta)}\,\mathrm{d}\lambda(x)\,\mathrm{d}\pi(\theta) - \mathsf{M}_{2,\pi}\Big),$$

where $\mathsf{M}_{2,\pi} = \int \theta^{\otimes 2}\,\mathrm{d}\pi(\theta)$. This recovers a result of Efroimovich [29, Theorem 5]*. Combining the inequalities $\log\det C \le n\log\frac{\mathrm{tr}C}{n}$ and $\log y \le y - 1$, which hold for all $y > 0$ and all $n\times n$ positive definite matrices C, we see that Efroimovich's inequality is a strengthening of* (55)*.*

3. Gagliardo–Nirenberg–Sobolev inequalities

In this section we shall prove Theorem 2:

Theorem 10. *Let $p, q, r \in [1,\infty)$, $\theta \in [0,1]$ and $C^{p,q,r,r} > 0$ be such that*

$$\|u\|_{L_p(\mathbb{R}^n)} \le C^{p,q,r,r} \|u\|^{\theta}_{L_q(\mathbb{R}^n)} \|\nabla u\|^{1-\theta}_{L_r(\mathbb{R}^n;\ell_r^n)} \tag{65}$$

is satisfied for all functions $u \in C_0^\infty(\mathbb{R}^n)$ under the constraint

$$\frac{1}{p} = \frac{\theta}{q} + \Big(\frac{1}{r} - \frac{1}{n}\Big)(1-\theta). \tag{66}$$

Then, for every $u \in C_0^\infty(\mathbb{R}^n)$, we have

$$\|u\|_{L_p(\mathbb{R}^n)} \le C^{p,q,r,r} n^{\frac{1-\theta}{r}} \|u\|^{\theta}_{L_q(\mathbb{R}^n)} \Big(\prod_{j=1}^n \|\partial_j u\|_{L_r(\mathbb{R}^n)}\Big)^{\frac{1-\theta}{n}}. \tag{67}$$

Proof. Fix $t = (t_1,\ldots,t_n) \in \mathbb{R}^n_+$ and consider the function $u_t \in C_0^\infty(\mathbb{R}^n)$ given by

$$\forall\ x = (x_1,\ldots,x_n) \in \mathbb{R}^n, \qquad u_t(x) \overset{\mathrm{def}}{=} u(t_1x_1,\ldots,t_nx_n). \tag{68}$$

Then, for $s \ge 1$ we have

$$\|u_t\|_{L_s(\mathbb{R}^n)} = \Big(\int_{\mathbb{R}^n} |u(t_1x_1,\ldots,t_nx_n)|^s\,\mathrm{d}x\Big)^{1/s} = \prod_{j=1}^n t_j^{-1/s} \|u\|_{L_s(\mathbb{R}^n)} \tag{69}$$

and

$$\|\nabla u_t\|^r_{L_r(\mathbb{R}^n;\ell_r^n)} = \sum_{i=1}^n \int_{\mathbb{R}^n} \|\partial_i u_t\|^r_{L_r(\mathbb{R}^n)} = \sum_{i=1}^n t_i^r \prod_{j=1}^n t_j^{-1} \|\partial_i u\|^r_{L_r(\mathbb{R}^n)}. \tag{70}$$

Therefore, applying (65) to u_t and rearranging, we deduce that

$$\begin{aligned} \|u\|_{L_p(\mathbb{R}^n)} &\leq C^{p,q,r,r} \Big(\prod_{j=1}^n t_j\Big)^{\frac{1}{p}-\frac{\theta}{q}-\frac{1-\theta}{r}} \|u\|^\theta_{L_q(\mathbb{R}^n)} \Big(\sum_{i=1}^n t_i^r \|\partial_i u\|^r_{L_r(\mathbb{R}^n)}\Big)^{\frac{1-\theta}{r}} \\ &\overset{(66)}{=} C^{p,q,r,r} \Big(\prod_{j=1}^n t_j\Big)^{-\frac{1-\theta}{n}} \|u\|^\theta_{L_q(\mathbb{R}^n)} \Big(\sum_{i=1}^n t_i^r \|\partial_i u\|^r_{L_r(\mathbb{R}^n)}\Big)^{\frac{1-\theta}{r}} \end{aligned} \tag{71}$$

for every $t_1, \ldots, t_n > 0$. Choosing

$$t_i = \|\partial_i u\|^{-1}_{L_r(\mathbb{R}^n)} \tag{72}$$

gives the desired inequality (67). □

4. Beckner inequalities

In this section we shall prove Theorem 1:

Theorem 11. *Fix $n \in \mathbb{N}$ and let $u \in C_0^\infty(\mathbb{R}^n)$ be such that*

$$\forall\ i,j \in \{1,\ldots,n\}, \qquad \int_{\mathbb{R}^n} x_i x_j u(x)^2 \,\mathrm{d}\gamma_n(x) = \delta_{ij} \|u\|^2_{L_2(\gamma_n)}, \tag{73}$$

where δ_{ij} is the Kronecker delta. Then, we have

$$\forall\ p \in [1,2), \qquad \frac{\|u\|^2_{L_2(\gamma_n)} - \|u\|^2_{L_p(\gamma_n)}}{\|u\|^2_{L_2(\gamma_n)}} \leq 1 - \det\left(\frac{4}{\|u\|^2_{L_2(\gamma_n)}} \int_{\mathbb{R}^n} (\nabla u)^{\otimes 2} \,\mathrm{d}\gamma_n + \mathrm{Id}_n\right)^{-\frac{2-p}{2p}}. \tag{74}$$

For the proof of Theorem 11 we shall use the intrinsic dimensional logarithmic Sobolev inequality (7) which takes the following simple form for appropriately normalized functions in Gauss space.

Lemma 12. *Let $u \in C_0^\infty(\mathbb{R}^n)$ be such that $\|u\|_{L_2(\gamma_n)} = 1$ and*

$$\forall\ i,j \in \{1,\ldots,n\}, \qquad \int_{\mathbb{R}^n} x_i x_j u(x)^2 \,\mathrm{d}\gamma_n(x) = \delta_{ij}. \tag{75}$$

Then, we have

$$\mathrm{Ent}_{\gamma_n}[u^2] \leq \frac{1}{2} \log\det\left(4 \int_{\mathbb{R}^n} (\nabla u)^{\otimes 2} \,\mathrm{d}\gamma_n + \mathrm{Id}_n\right). \tag{76}$$

Proof. Let $u \in C_0^\infty(\mathbb{R}^n)$ satisfy the assumptions of the lemma and define $f(x) = u(x)^2 \frac{\exp(-|x|^2/2)}{(2\pi)^{n/2}}$ which is the density of a probability measure μ on $\mathbb{R}^n$. Then, we have

$$\mathsf{H}(\mu\|\lambda_n) - \mathsf{H}(\gamma_n\|\lambda_n) = \mathrm{Ent}_{\gamma_n}[u^2] - \frac{1}{2}\int_{\mathbb{R}^n} |x|^2 u(x)^2 \,\mathrm{d}\gamma_n(x) - \frac{n}{2}\log 2\pi + \frac{n}{2}\log 2\pi e \overset{(75)}{=} \mathrm{Ent}_{\gamma_n}[u^2]. \tag{77}$$

On the other hand, for $k \in \{1,\ldots,n\}$, we compute

$$\partial_k f(x) = \Big(2u(x)\partial_k u(x) - x_k u(x)^2\Big) \frac{e^{-|x|^2/2}}{(2\pi)^{n/2}} \tag{78}$$

and thus for $i,j \in \{1,\ldots,n\}$, we get

$$\mathcal{I}(\mu\|\lambda_n)_{ij} = 4\int_{\mathbb{R}^n} \partial_i u \partial_j u \,\mathrm{d}\gamma_n - 2\int_{\mathbb{R}^n} \Big(x_j \partial_i u(x) + x_i \partial_j u(x)\Big) u(x) \,\mathrm{d}\gamma_n(x) + \int_{\mathbb{R}^n} x_i x_j u(x)^2 \,\mathrm{d}\gamma_n(x). \tag{79}$$

For $i \neq j$, integration by parts gives

$$2\int_{\mathbb{R}^n} x_j \partial_i u(x) u(x)\, \mathrm{d}\gamma_n(x) = \int_{\mathbb{R}^n} \partial_i\Big(x_j u(x)^2\Big)\, \mathrm{d}\gamma_n(x) = \int_{\mathbb{R}^n} x_i x_j u(x)^2\, \mathrm{d}\gamma_n(x) \overset{(75)}{=} 0, \tag{80}$$

whereas for $i = j$, again by integration by parts,

$$2\int_{\mathbb{R}^n} x_i \partial_i u(x) u(x)\, \mathrm{d}\gamma_n(x) = \int_{\mathbb{R}^n} \partial_i\Big(x_i u(x)^2\Big)\, \mathrm{d}\gamma_n(x) - \int_{\mathbb{R}^n} u(x)^2\, \mathrm{d}\gamma_n(x) = \int_{\mathbb{R}^n} x_i^2 u(x)^2\, \mathrm{d}\gamma_n(x) - 1 \overset{(75)}{=} 0.$$

Plugging the above in (79) and using (75) again for the last term, we deduce that

$$\mathcal{I}(\mu\|\lambda_n) = 4\int_{\mathbb{R}^n} (\nabla u)^{\otimes 2}\, \mathrm{d}\gamma_n + \mathrm{Id}_n \tag{81}$$

and the conclusion of the lemma follows from (7). □

Equipped with Lemma 12, we proceed to the proof of Theorem 1.

Proof of Theorem 11. Assume, without loss of generality, that $\|u\|_{L_2(\gamma_n)} = 1$. Combining a lemma of Dolbeault and Toscani [28, Lemma 5] (see also [42]) with Lemma 12, we get that

$$\frac{1}{\|u\|^2_{L_p(\gamma_n)}} = \frac{\|u\|^2_{L_2(\gamma_n)}}{\|u\|^2_{L_p(\gamma_n)}} \leq \exp\left(\frac{2-p}{p}\mathrm{Ent}_{\gamma_n}[u^2]\right) \overset{(76)}{\leq} \det\left(4\int_{\mathbb{R}^n} (\nabla u)^{\otimes 2}\, \mathrm{d}\gamma_n + \mathrm{Id}_n\right)^{\frac{2-p}{2p}}. \tag{82}$$

Therefore,

$$1 - \|u\|^2_{L_p(\gamma_n)} \leq 1 - \det\left(4\int_{\mathbb{R}^n} (\nabla u)^{\otimes 2}\, \mathrm{d}\gamma_n + \mathrm{Id}_n\right)^{-\frac{2-p}{2p}}, \tag{83}$$

which is the desired estimate under the normalization $\|u\|_{L_2(\gamma_n)} = 1$. □

5. q-logarithmic Sobolev inequalities

Following Bobkov and Zegarlinski [18] (see also [11]) we say that a probability measure μ on the real line satisfies the q-logarithmic Sobolev inequality with constant $C > 0$ if for any $f \in C_0^\infty(\mathbb{R})$ we have

$$\mathrm{Ent}_\mu[|f|^q] \leq C\int_{\mathbb{R}} |f'(x)|^q\, \mathrm{d}\mu(x). \tag{84}$$

Standard tensorization principles show that if (84) holds, then for any $f \in C_0^\infty(\mathbb{R}^n)$,

$$\mathrm{Ent}_{\mu^n}[|f|^q] \leq C\sum_{i=1}^n \int_{\mathbb{R}^n} |\partial_i f(x)|^q\, \mathrm{d}\mu^n(x), \tag{85}$$

where $\mu^n = \mu \otimes \mu \otimes \cdots \otimes \mu$ is the product measure of i.i.d. coordinates distributed like μ. In particular, it has been established in [18, Corollary 5.6] (see also [16, Section 5]) that the measure μ_p with density $\frac{1}{Z_p}e^{-|x|^p}$, where $p > 2$, satisfies the q-logarithmic Sobolev inequality for $q = \frac{p}{p-1}$ with some constant $C_q > 1$. In order to investigate scale-invariant refinements of (85) for this family of measures in the spirit of (50), we first need to formulate them as Euclidean inequalities.

Theorem 13. *For any $q \in (1,2)$, there exists a constant $\tilde{C}_q > 0$ such that for any $n \in \mathbb{N}$ and any probability measure μ on $\mathbb{R}^n$ with positive differentiable density g,*

$$\mathrm{H}(\mu\|\lambda_n) \leq \tilde{C}_q \sum_{i=1}^n \int_{\mathbb{R}^n} \left(\left|\frac{\partial_i g(x)}{g(x)}\right|^q + |x_i|^{\frac{q}{q-1}}\right) \mathrm{d}\mu(x). \tag{86}$$

Proof. For $p = \frac{q}{q-1} > 2$ consider the probability measure $\mathrm{d}\mu_p(x) = \frac{e^{-|x|^p}}{Z_p}$ on $\mathbb{R}$, where the normalizing constant is $Z_p = 2\Gamma(1+1/p) > 2$. Let μ be a probability measure on $\mathbb{R}^n$ with differentiable density $g : \mathbb{R}^n \to \mathbb{R}_+$ and consider the function $f : \mathbb{R}^n \to \mathbb{R}_+$ given by

$$\forall\, x \in \mathbb{R}^n, \qquad f(x) = Z_p^{n/q} g(x)^{1/q} e^{\|x\|_p^p/q}, \tag{87}$$

which satisfies $\int_{\mathbb{R}^n} f(x)^q \,\mathrm{d}\mu_p^n(x) = 1$. Therefore, the q-logarithmic Sobolev inequality for μ_p^n applied to the function f implies that

$$\frac{1}{Z_p^n}\int_{\mathbb{R}^n} f(x)^q e^{-\|x\|_p^p} \log f(x)^q \,\mathrm{d}x = \mathrm{Ent}_{\mu_p^n}[f^q] \le \frac{C_q}{Z_p^n}\sum_{i=1}^n \int_{\mathbb{R}^n} |\partial_i f(x)|^q e^{-\|x\|_p^p} \,\mathrm{d}x. \tag{88}$$

Observe that

$$\begin{aligned}\frac{1}{Z_p^n}\int_{\mathbb{R}^n} f(x)^q e^{-\|x\|_p^p} \log f(x)^q \,\mathrm{d}x &= \int_{\mathbb{R}^n} g(x)\log\Big(Z_p^n g(x) e^{\|x\|_p^p}\Big)\,\mathrm{d}x \\ &= \mathsf{H}(\mu\|\lambda_n) + \int_{\mathbb{R}^n} \|x\|_p^p \,\mathrm{d}\mu(x) + \log Z_p^n,\end{aligned} \tag{89}$$

and for $i \in \{1,\dots,n\}$,

$$\frac{1}{Z_p^n}\int_{\mathbb{R}^n} |\partial_i f(x)|^q e^{-\|x\|_p^p} \,\mathrm{d}x = \frac{1}{q^q}\int_{\mathbb{R}^n} \Big|\frac{\partial_i g(x)}{g(x)} + p\,\mathrm{sign}(x_i)|x_i|^{p-1}\Big|^q \,\mathrm{d}\mu(x). \tag{90}$$

Therefore, rearranging (88) we deduce that

$$\begin{aligned}\mathsf{H}(\mu\|\lambda_n) &\le \frac{C_q}{q^q}\sum_{i=1}^n \int_{\mathbb{R}^n} \Big|\frac{\partial_i g(x)}{g(x)} + p\,\mathrm{sign}(x_i)|x_i|^{p-1}\Big|^q \,\mathrm{d}\mu(x) - \sum_{i=1}^n \int_{\mathbb{R}^n} |x_i|^p \,\mathrm{d}\mu(x) - \log Z_p^n \\ &\le \tilde{C}_q \sum_{i=1}^n \int_{\mathbb{R}^n} \left(\Big|\frac{\partial_i g(x)}{g(x)}\Big|^q + |x_i|^p\right)\mathrm{d}\mu(x)\end{aligned} \tag{91}$$

for some different constant $\tilde{C}_q > 0$ and the proof is complete. □

This Euclidean weakening of the q-logarithmic Sobolev inequality (85) for μ_p^n makes it amenable to refinements via scaling.

Theorem 14. *For any $q \in (1,2)$ and $p = \frac{q}{q-1}$, there exists a constant $\tilde{C}_q > 0$ such that for any $n \in \mathbb{N}$ and any probability measure μ on $\mathbb{R}^n$,*

$$\mathsf{H}(\mu\|\lambda_n) \le \sum_{i=1}^n \inf_{t_i>0}\Big\{\tilde{C}_q t_i^q \int_{\mathbb{R}^n} \Big|\frac{\partial_i g(x)}{g(x)}\Big|^q \,\mathrm{d}\mu(x) + \frac{\tilde{C}_q}{t_i^p}\int_{\mathbb{R}^n} |x_i|^p \,\mathrm{d}\mu(x) - \log t_i\Big\}. \tag{92}$$

Proof. Fix $t_1,\dots,t_n > 0$ and consider the probability measure μ_t whose density is given by $x \mapsto g_t(x) \stackrel{\text{def}}{=} t_1\cdots t_n g(t_1x_1,\dots,t_nx_n)$. Then, we have

$$\mathsf{H}(\mu_t\|\lambda_n) = \mathsf{H}(\mu\|\lambda_n) + \sum_{i=1}^n \log t_i \tag{93}$$

and for every $i \in \{1,\dots,n\}$,

$$\int_{\mathbb{R}^n} \left(\Big|\frac{\partial_i g_t(x)}{g_t(x)}\Big|^q + |x_i|^p\right)\mathrm{d}\mu_t(x) = \int_{\mathbb{R}^n}\left(t_i^q\Big|\frac{\partial_i g(x)}{g(x)}\Big|^q + \frac{|x_i|^p}{t_i^p}\right)\mathrm{d}\mu(x). \tag{94}$$

Therefore, applying (86) to μ_t and rearranging, we deduce that

$$\mathsf{H}(\mu\|\lambda_n) \le \sum_{i=1}^n \Big\{\tilde{C}_q t_i^q \int_{\mathbb{R}^n} \Big|\frac{\partial_i g(x)}{g(x)}\Big|^q \,\mathrm{d}\mu(x) + \frac{\tilde{C}_q}{t_i^p}\int_{\mathbb{R}^n} |x_i|^p \,\mathrm{d}\mu(x) - \log t_i\Big\} \tag{95}$$

and taking an infimum over $t_1,\dots,t_n > 0$ completes the proof. □

6. Beyond linear rescalings

The simple idea of the previous sections can be summarized as follows. Let

$$\mathcal{K}(f) \leq \mathcal{L}(f) \tag{96}$$

be a functional inequality valid for regular enough functions f on $\mathbb{R}^n$ and fix a subgroup of symmetries $\mathsf{G} \subseteq \mathsf{GL}_n$. For a fixed $f : \mathbb{R}^n \to \mathbb{R}$ for which inequality (96) is valid and $A \in \mathsf{G}$, consider the function $f_A : \mathbb{R}^n \to \mathbb{R}$ given by $f_A(x) = f(Ax)$. *If* (96) applied to f_A can be rearranged to an upper bound for $\mathcal{K}(f)$ of the form

$$\mathcal{K}(f) \leq \widetilde{L}(f, A), \tag{97}$$

then taking an infimum over $A \in \mathsf{G}$ yields a stronger inequality as (96) just amounts to the choice $A = \mathsf{Id}_n$. Observe that enhancing inequalities in this way, always produces a larger family of extremals. For instance, (4) becomes an equality only when μ is a translate of γ_n, (5) becomes an equality when μ is a Gaussian measure with covariance matrix of the form $\sigma \mathsf{Id}_n$, where $\sigma > 0$, and (7) becomes an equality for any Gaussian measure on $\mathbb{R}^n$.

In this section, we will discuss the possibility of refining functional inequalities by using changes of variables via *nonlinear* maps and we shall illustrate this in the case of the logarithmic Sobolev inequality (4). Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a smooth diffeomorphism and for a measure μ on $\mathbb{R}^n$ with a differentiable density $f : \mathbb{R}^n \to \mathbb{R}_+$ consider the measure μ_T whose density is given by $f_T(x) = (f \circ T)(x)|\det \mathsf{D}T(x)|$, where $x \in \mathbb{R}^n$ and $\mathsf{D}T \in \mathsf{M}_n(\mathbb{R})$ is the differential of T. We need the following computations for the relative entropy and Fisher information of μ_T.

Lemma 15. *In the setting above,*

$$\mathsf{H}(\mu_T \| \lambda_n) = \mathsf{H}(\mu \| \lambda_n) + \int_{\mathbb{R}^n} \log|\det \mathsf{D}T(T^{-1}(x))| \, \mathrm{d}\mu(x) \tag{98}$$

and

$$\mathsf{I}(\mu_T \| \lambda_n) = \int_{\mathbb{R}^n} \left| \frac{\mathsf{D}T(T^{-1}(x)) \cdot \nabla f(x)}{f(x)} + \nabla \log|\det \mathsf{D}T(T^{-1}(x))| \right|^2 \mathrm{d}\mu(x). \tag{99}$$

The proof is a straightforward computation using a change of variables and is thus omitted. These formulas along with the fact that any absolutely continuous measure can be transported to γ_n give rise to the following variational formula for relative entropy on $\mathbb{R}^n$.

Theorem 16. *Let μ be an absolutely continuous measure on $\mathbb{R}^n$. Then,*

$$\mathsf{H}(\mu \| \lambda_n) - \mathsf{H}(\gamma_n \| \lambda_n) = \min_{T \in \mathrm{Diff}(\mathbb{R}^n)} \psi(T) \tag{100}$$

with equality if T is a transport map from μ to γ_n, where

$$\psi(T) \stackrel{\text{def}}{=} \frac{1}{2} \int_{\mathbb{R}^n} \left| \frac{\mathsf{D}T(T^{-1}(x)) \cdot \nabla f(x)}{f(x)} + \nabla \log|\det \mathsf{D}T(T^{-1}(x))| \right|^2 \mathrm{d}\mu(x) - \int_{\mathbb{R}^n} \log|\det \mathsf{D}T(T^{-1}(x))| \, \mathrm{d}\mu(x).$$

Proof. Applying the logarithmic Sobolev inequality (4) to μ_T and using Lemma 15, we get

$$\begin{aligned} \mathsf{H}(\mu \| \lambda_n) + \int_{\mathbb{R}^n} \log|\det \mathsf{D}T(T^{-1}(x))| \, \mathrm{d}\mu(x) - \mathsf{H}(\gamma_n \| \lambda_n) \\ \leq \frac{1}{2} \int_{\mathbb{R}^n} \left| \frac{\mathsf{D}T(T^{-1}(x)) \cdot \nabla f(x)}{f(x)} + \nabla \log|\det \mathsf{D}T(T^{-1}(x))| \right|^2 \mathrm{d}\mu(x), \end{aligned} \tag{101}$$

with equality only if $\mu_T = \gamma_n$. The existence of a map T transporting μ to γ_n is a classical fact in optimal transport going back to at least [56, 41] (see also [64]). □

We are not aware of a proof of (100) which does not rely on the logarithmic Sobolev inequality (4). It remains very interesting to understand whether (100) can lead to stability estimates for (4), or even (7), in the spirit of [30, Theorem 3].

Formula (100) becomes more tractable when specified to specific kinds of diffeomorphisms. For instance, when T is a product map of the form $T(x) = (\tau_1(x_1), \ldots, \tau_n(x_n))$, we get

$$\mathsf{H}(\mu\|\lambda_n) - \mathsf{H}(\gamma_n\|\lambda_n) \leq \inf_{\tau_1,\ldots,\tau_n \in \mathrm{Diff}(\mathbb{R})} \sum_{i=1}^n \int_{\mathbb{R}^n} \left\{ \frac{1}{2} \left| \frac{(\tau_i' \circ \tau_i^{-1})\partial_i f}{f} + (\log|\tau_i' \circ \tau_i^{-1}|)' \right|^2 - \log|\tau_i' \circ \tau_i^{-1}| \right\} \mathrm{d}\mu. \tag{102}$$

A similar simplified formula can be derived if T is a rotationally invariant map of the form $T(x) = \sigma(|x|)x$. The equality cases of Theorem 16 show that if μ is a product measure or a rotationally invariant measure, then the inequalities obtained by optimizing over the corresponding class of nonlinear transformations become equalities. For the case of a general probability measure μ, we pose the following question.

Question 17. *Let μ be an arbitrary absolutely continuous probability measure on $\mathbb{R}^n$. For which collection of diffeomorphisms $\tau_1, \ldots, \tau_n \in \mathrm{Diff}(\mathbb{R})$ is the infimum* (102) *attained?*

A similar question can be asked for the optimal rotationally invariant change of variables.

We have not investigated whether nonlinear changes of variables may give rise to variational formulas à la (100) when applied to other estimates like the Gagliardo–Nirenberg–Sobolev inequality (65) or Beckner's inequality (11).

7. Tensorization of nonlinear logarithmic Sobolev inequalities in product spaces

Let I be a countable set, $\{(\mathbb{X}_i, \mu_i)\}_{i\in I}$ a family of probability spaces where $\mathbb{X}_i$ is countable and denote their product space by $(\mathbb{X}, \mu) = (\prod_{i\in I} \mathbb{X}_i, \otimes_{i\in I}\mu_i)$. For a point $x = (x_i)_{i\in I} \in \mathbb{X}$ and $i \in I$, we shall denote by $x_{\sim i}$ the point $(x_j)_{j\neq i} \in \prod_{j\neq i} \mathbb{X}_j$ and by $\mu_{\sim i} \stackrel{\mathrm{def}}{=} \otimes_{j\neq i}\mu_j$. Moreover, for a point $z \in \prod_{j\neq i} \mathbb{X}_j$ and a function $f : \mathbb{X} \to \mathbb{R}$, we shall denote by $f_z : \mathbb{X}_i \to \mathbb{R}$ the restriction of f given by

$$\forall\, y \in \mathbb{X}_i, \qquad f_z(y) \stackrel{\mathrm{def}}{=} f(z, y). \tag{103}$$

For each $i \in I$, let B_i be a functional acting on measurable functions $g : \prod_{j\in J}\mathbb{X}_j \to \mathbb{R}$ for any $J \subseteq I$. We shall say that the family of functionals $\{B_i\}_{i\in I}$ *disintegrates* if it satisfies the identities

$$\forall\, i \in I, \qquad \int B_i\Big(f_{x_{\sim i}}\Big)\, \mathrm{d}\mu_{\sim i}(x_{\sim i}) = B_i(f). \tag{104}$$

Our main tensorization principle for nonlinear entropy inequalities is the following.

Theorem 18. *Fix a countable set I and two collections of functionals $\{Q_i\}_{i\in I}$, $\{M_i\}_{i\in I}$ which disintegrate in the above sense. Let $\Phi : \mathbb{R} \to \mathbb{R}$ be a concave function and suppose that, for any $i \in I$, every function $f_i : \mathbb{X}_i \to \mathbb{R}_+$ satisfies the inequality*

$$\mathrm{Ent}_{\mu_i}[f_i] \leq Q_i(f_i) + \mathbb{E}_{\mu_i}[f_i]\Phi\left(\frac{M_i(f_i) - Q_i(f_i)}{\mathbb{E}_{\mu_i}[f_i]}\right). \tag{105}$$

Then, every function $f : \mathbb{X} \to \mathbb{R}_+$ satisfies

$$\mathrm{Ent}_\mu[f] \leq \sum_{i\in I} Q_i(f) + \mathbb{E}_\mu[f] \sum_{i\in I} \Phi\left(\frac{M_i(f) - Q_i(f)}{\mathbb{E}_\mu[f]}\right). \tag{106}$$

Proof. Combining the subadditivity of entropy and the assumptions of the theorem (including the disintegration of $\{Q_i\}_{i\in I}$) we get that, for every $f : \mathbb{X} \to \mathbb{R}_+$,

$$\begin{aligned}
\mathrm{Ent}_\mu[f] &\leq \sum_{i\in I} \int \mathrm{Ent}_{\mu_i}[f_{x_{\sim i}}]\, \mathrm{d}\mu_{\sim i}(x_{\sim i}) \\
&\stackrel{(105)}{\leq} \sum_{i\in I} \int \left[Q_i(f_{x_{\sim i}}) + \mathbb{E}_{\mu_i}[f_{x_{\sim i}}]\Phi\left(\frac{M_i(f_{x_{\sim i}}) - Q_i(f_{x_{\sim i}})}{\mathbb{E}_{\mu_i}[f_{x_{\sim i}}]}\right)\right] \mathrm{d}\mu_{\sim i}(x_{\sim i}) \\
&= \sum_{i\in I} Q_i(f) + \mathbb{E}_\mu[f] \sum_{i\in I} \int \Phi\left(\frac{M_i(f_{x_{\sim i}}) - Q_i(f_{x_{\sim i}})}{\mathbb{E}_{\mu_i}[f_{x_{\sim i}}]}\right) \frac{\mathbb{E}_{\mu_i}[f_{x_{\sim i}}]}{\mathbb{E}_\mu[f]}\, \mathrm{d}\mu_{\sim i}(x_{\sim i}).
\end{aligned} \tag{107}$$

Since $\int \mathbb{E}_{\mu_i}[f_{x_{\sim i}}]\,\mathrm{d}\mu_{\sim i}(x_{\sim i}) = \mathbb{E}_\mu[f]$, and $\mathbb{E}_{\mu_i}[f_{x_{\sim i}}] \geq 0$, it follows that $\frac{\mathbb{E}_{\mu_i}[f_{x_{\sim i}}]}{\mathbb{E}_\mu[f]}\,\mathrm{d}\mu_{\sim i}(x_{\sim i})$ defines a probability measure on $\prod_{j\neq i}\mathbb{X}_j$. Hence, by Jensen's inequality and disintegration, we get

$$\begin{aligned}\mathrm{Ent}_\mu[f] &\leq \sum_{i\in I} Q_i(f) + \mathbb{E}_\mu[f]\sum_{i\in I}\Phi\left(\int \frac{M_i(f_{x_{\sim i}}) - Q_i(f_{x_{\sim i}})}{\mathbb{E}_{\mu_i}[f_{x_{\sim i}}]}\cdot\frac{\mathbb{E}_{\mu_i}[f_{x_{\sim i}}]}{\mathbb{E}_\mu[f]}\,\mathrm{d}\mu_{\sim i}(x_{\sim i})\right)\\ &= \sum_{i\in I} Q_i(f) + \mathbb{E}_\mu[f]\sum_{i\in I}\Phi\left(\frac{M_i(f) - Q_i(f)}{\mathbb{E}_\mu[f]}\right).\end{aligned} \tag{108}$$

This completes the proof of the theorem. □

Remark 19. *While Theorem 18 is stated in a general form which contains the disintegrating additive errors $\{Q_i\}_{i\in I}$, in its main application (Theorem 3) which refines the result of* [55]*, these are assumed to be vanishing. We chose to include the deficits in the general formulation above as such terms often appear in modified logarithmic Sobolev-type inequalities, especially in discrete settings (see, for instance,* [15, 67, 17, 39]*).*

Proof of Theorem 3. The conclusion (25) directly follows from Theorem 18 with $Q_i(f) = 0$ and $M_i(f) = \mathbb{E}_{\pi^n}[\mathcal{E}_i(f^{1/p}, f^{1-1/p})]$ since these functionals disintegrate. □

Remark 20. *A different refinement of the log-Sobolev inequality on the discrete cube in terms of the logarithmic Laplace transform of the underlying measure can be found in* [3, Equation (12)]

Part 2. Space forms: stochastic methods

8. Preliminaries

In this section we will introduce the necessary prerequisites from stochastic calculus on manifolds required to prove Theorem 4. We will be following the standard notation of [38, 54].

8.1. **The frame bundle.** Let (M, g) be a complete n-dimensional Riemannian manifold. The *orthonormal frame bundle* $\mathcal{O}(\mathsf{M})$ of M is the set of all pairs of the form $(x, \mathbf{u})$, where $x \in \mathsf{M}$ and $\mathbf{u} : \mathbb{R}^n \to T_x\mathsf{M}$ is a Euclidean isometry. We shall denote by $\pi : \mathcal{O}(\mathsf{M}) \to \mathsf{M}$ the natural projection given by $\pi(x, \mathbf{u}) = x$. Any scalar-valued function $f : \mathsf{M} \to \mathbb{R}$ admits a natural *lift* $\mathbf{f} : \mathcal{O}(\mathsf{M}) \to \mathbb{R}$ given by

$$\forall\ (x, \mathbf{u}) \in \mathcal{O}(\mathsf{M}), \qquad \mathbf{f}(x, \mathbf{u}) = f(x). \tag{109}$$

Abusing notation, we shall often identify the pair $(x, \mathbf{u}) \in \mathcal{O}(\mathsf{M})$ with the isomorphism $\mathbf{u}$. The frame bundle $\mathcal{O}(\mathsf{M})$ is equipped with Bochner's horizontal Laplacian

$$\Delta_{\mathcal{O}(\mathsf{M})} \stackrel{\text{def}}{=} \sum_{i=1}^n H_i^2, \tag{110}$$

and can be verified (see [38, Proposition 3.1.2]) that the lift $\mathbf{f}$ of any function $f : \mathsf{M} \to \mathbb{R}$ satisfies

$$\forall\ \mathbf{u} \in \mathcal{O}(\mathsf{M}), \qquad \Delta_{\mathcal{O}(\mathsf{M})}\mathbf{f}(\mathbf{u}) = \Delta f(\pi\mathbf{u}), \tag{111}$$

where Δ is the Laplace–Beltrami operator of (M, g).

A curve $\{\mathbf{u}_t\}_{t\in[0,1]}$ in $\mathcal{O}(\mathsf{M})$ is called *horizontal* if for every $a \in \mathbb{R}^n$, the vector field $\{\mathbf{u}_t a\}_{t\in[0,1]}$ is parallel along the curve $\{\pi\mathbf{u}_t\}_{t\in[0,1]}$ in M. A tangent vector $\mathbf{X} \in T_{\mathbf{u}}\mathcal{O}(\mathsf{M})$ is called horizontal if it is the tangent vector of a horizontal curve passing from $\mathbf{u}$. For any vector $X \in T_{\pi\mathbf{u}}\mathsf{M}$ there exists a unique horizontal vector $\mathbf{X} \in T_{\mathbf{u}}\mathcal{O}(\mathsf{M})$ such that $\pi_*\mathbf{X} = X$; we say that $\mathbf{X}$ is the horizontal lift of X at $\mathbf{u}$. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$. The i-th fundamental *horizontal vector field* H_i evaluated at a point $\mathbf{u} \in \mathcal{O}(\mathsf{M})$ is the horizontal lift of the vector $\mathbf{u}e_i \in T_{\pi\mathbf{u}}\mathsf{M}$. Thus, for any $i \in \{1, \ldots, n\}$, the lift $\mathbf{f}$ of a function $f : \mathsf{M} \to \mathbb{R}$ satisfies

$$\forall\ \mathbf{u} \in \mathcal{O}(\mathsf{M}), \qquad \mathsf{H}_i\mathbf{f}(\mathbf{u}) = \nabla_{\mathbf{u}e_i} f(\pi\mathbf{u}). \tag{112}$$

A vector field on $\mathcal{O}(\mathsf{M})$ is called *horizontal* if it lies in the span of $\{\mathsf{H}_1,\dots,\mathsf{H}_n\}$. We denote by $\langle\cdot,\cdot\rangle_{\mathrm{hor}}$ the natural inner product on the space of horizontal vector fields on $\mathcal{O}(\mathsf{M})$ given by

$$\left\langle \sum_{i=1}^n Z_i\mathsf{H}_i, \sum_{i=1}^n W_i\mathsf{H}_i \right\rangle_{\mathrm{hor}} = \sum_{i=1}^n Z_i W_i. \tag{113}$$

Moreover, we shall denote by $\nabla_{\mathrm{hor}}\mathbf{f} = (\mathsf{H}_1\mathbf{f},\dots\mathsf{H}_n\mathbf{f}) \in \mathbb{R}^n$ the *horizontal gradient* of a given function $\mathbf{f}:\mathcal{O}(\mathsf{M})\to\mathbb{R}$. In this terminology, the horizontal Laplacian takes the form $\Delta_{\mathcal{O}(\mathsf{M})}\mathbf{f} = \sum_i \mathsf{H}_i^2\mathbf{f}$.

We record for future reference the following very useful expression for the action of the commutator of $\Delta_{\mathcal{O}(\mathsf{M})}$ with H_i on lifted functions.

Lemma 21. *If $f:(\mathsf{M},\mathsf{g})\to\mathbb{R}$ is a smooth function, then for any $i\in\{1,\dots,n\}$, its lift $\mathbf{f}$ satisfies*

$$\forall\ \mathbf{u}\in\mathcal{O}(\mathsf{M}),\qquad \Delta_{\mathcal{O}(\mathsf{M})}\mathsf{H}_i\mathbf{f}(\mathbf{u}) - \mathsf{H}_i\Delta_{\mathcal{O}(\mathsf{M})}\mathbf{f}(\mathbf{u}) = \mathrm{Ric}(\nabla f,\mathbf{u}e_i)(\pi\mathbf{u}), \tag{114}$$

where $\mathrm{Ric}(\cdot,\cdot)$ *is the Ricci tensor on* M.

Proof. We shall follow the notation of [38, Section 5.5]. For $i,k\in\{1,\dots,n\}$, it follows from [38, Lemma 5.5.1] that the commutator $[\mathsf{H}_i,\mathsf{H}_k]$ is a vertical vector field and thus $[\mathsf{H}_i,\mathsf{H}_k]\mathbf{f}=0$, i.e.

$$\mathsf{H}_k\mathsf{H}_i\mathbf{f} = \mathsf{H}_i\mathsf{H}_k\mathbf{f}. \tag{115}$$

Therefore, we have

$$\mathsf{H}_k^2\mathsf{H}_i\mathbf{f} = \mathsf{H}_k\mathsf{H}_i\mathsf{H}_k\mathbf{f} = [\mathsf{H}_k,\mathsf{H}_i]\mathsf{H}_k\mathbf{f} + \mathsf{H}_i\mathsf{H}_k^2\mathbf{f}. \tag{116}$$

Substituting the expression of [38, Lemma 5.5.1] for $[\mathsf{H}_k,\mathsf{H}_i]$, we get

$$[\mathsf{H}_k,\mathsf{H}_i]\mathsf{H}_k\mathbf{f} = -\sum_{a,b}\Omega_{ki}^{ab}V_{ab}\mathsf{H}_k\mathbf{f} = -\sum_{a,b}\Omega_{ki}^{ab}[V_{ab},\mathsf{H}_k]\mathbf{f}, \tag{117}$$

where in the last identity we used that $V_{ab}\mathbf{f}=0$. Again, by [38, Lemma 5.5.1], if we denote by $A_{ab}^{k\ell}$ the number $\frac{1}{2}$ for $(a,b)=(k,\ell)$ and $-\frac{1}{2}$ for $(a,b)=(\ell,k)$, and zero otherwise, we obtain

$$-\sum_{a,b}\Omega_{ki}^{ab}[V_{ab},\mathsf{H}_k]\mathbf{f} = -\sum_{a,b,\ell}\Omega_{ki}^{ab}A_{ab}^{k\ell}\mathsf{H}_\ell\mathbf{f} = \frac{1}{2}\sum_{\ell}\left\{\Omega_{ki}^{\ell k} - \Omega_{ki}^{k\ell}\right\}\mathsf{H}_\ell\mathbf{f} = \sum_{\ell}\Omega_{ki}^{\ell k}\mathsf{H}_\ell\mathbf{f}, \tag{118}$$

where the antisymmetry of Ω on the top indices follows from its definition in [38, p. 153] as it is an $\mathfrak{o}(d)$-valued tensor. Combining (116), (117), (118) and summing over k, we deduce that

$$\Delta_{\mathcal{O}(\mathsf{M})}\mathsf{H}_i\mathbf{f}(\mathbf{u}) - \mathsf{H}_i\Delta_{\mathcal{O}(\mathsf{M})}\mathbf{f}(\mathbf{u}) = \sum_{k,\ell}\Omega_{ki}^{\ell k}\mathsf{H}_\ell\mathbf{f}(\mathbf{u}) = \sum_{k,\ell}\Omega_{ki}^{\ell k}\nabla_{\mathbf{u}e_\ell}f(\pi\mathbf{u}). \tag{119}$$

Now, observe that by the definition of Ω in terms of the Riemann tensor R of M in [38, p. 149],

$$\Omega_{ki}^{\ell k} = \mathsf{g}\Big(\mathsf{R}(\mathbf{u}e_k,\mathbf{u}e_i)\mathbf{u}e_k,\mathbf{u}e_\ell\Big) \tag{120}$$

and the conclusion follows from the definition of Ricci curvature. □

8.2. **Brownian motion on manifolds.** Let $W_t = (W_t^1,\dots,W_t^n)$ be a standard Brownian motion on $\mathbb{R}^n$ and (M,g) be a complete n-dimensional Riemannian manifold. We consider the following stochastic differential equation on the frame bundle $\mathcal{O}(\mathsf{M})$,

$$\mathrm{d}\Phi_t = \sum_{i=1}^n \mathsf{H}_i(\Phi_t)\circ\mathrm{d}W_t^i, \tag{121}$$

where the shorthand notation $\circ$ refers to the Stratonovitch integral. In Itô terms, the above SDE asserts that for every smooth $g:\mathcal{O}(\mathsf{M})\to\mathbb{R}$, we have

$$\mathrm{d}g(\Phi_t) = \sum_{i=1}^n \mathsf{H}_i g(\Phi_t)\,\mathrm{d}W_t^i + \frac{1}{2}\Delta_{\mathcal{O}(\mathsf{M})}g(\Phi_t)\,\mathrm{d}t. \tag{122}$$

For any initial condition $\Phi_0 = \mathbf{u}\in\mathcal{O}(\mathsf{M})$, this equation has a strong solution which does not blow up in finite time if the Ricci curvature of M is bounded from below by any constant $\kappa\in\mathbb{R}$ (see [38, Theorem 4.2.4] and [63] for a sufficient and almost necessary condition for stochastic

completeness). We denote by $B_t = \pi\Phi_t$, where $t \geq 0$, the *Brownian motion* on M whose starting point is $x = \pi\mathbf{u} \in \mathsf{M}$. Applying (122), we deduce that for any smooth function $f : \mathsf{M} \to \mathbb{R}$, the Brownian motion $\{B_t\}_{t\geq 0}$ satisfies the SDE

$$\mathrm{d}f(B_t) = \sum_{i=1}^{n} \Phi_t^i f(B_t)\,\mathrm{d}W_t^i + \frac{1}{2}\Delta f(B_t)\,\mathrm{d}t. \tag{123}$$

8.3. **The Föllmer process and Lehec's formula.** In this section we introduce an analogue of the classical Föllmer process [58, 33, 46] on Riemannian manifolds (see also [38, Section 5.4]). We then present a result of Lehec [45] who used this process to give a stochastic proof of the dimensional logarithmic Sobolev inequality for manifolds with Ricci curvature bounded below (see [10, 6] for more general statements proven via semigroup arguments).

Let $W_t = (W_t^1, \ldots, W_t^n)$ be a standard Brownian motion on $\mathbb{R}^n$ and (M, g) be a complete n-dimensional Riemannian manifold whose Ricci curvature is bounded from below. We shall denote by $\mathrm{d}x$ the volume measure on M and by $\{P_t\}_{t\geq 0}$ the heat semigroup on M. Recall that for a smooth function $h : \mathsf{M} \to \mathbb{R}$, the action of the heat flow $\{P_t\}_{t\geq 0}$ on g is characterized by the ordinary differential equation

$$\forall\ t > 0,\ x \in \mathsf{M}, \qquad \frac{\partial P_t h}{\partial t}(x) = \frac{1}{2}\Delta P_t h(x) \tag{124}$$

with initial condition $P_0 h = h$ on M. We recall that the heat semigroup and the Laplacian commute: $\Delta P_t h = P_t \Delta h$, and we write $P_t \nabla^2 f(x)$ for the 2-tensor on $T_x\mathsf{M}$ identified with the symmetric matrix $(P_t \nabla^2 f(\Phi_0 e_i, \Phi_0 e_j)(x))_{i,j=1}^n$. Note that P_t and ∇^2 *do not* commute (cf. Theorem 24).

For a positive function $f : \mathsf{M} \to \mathbb{R}_+$ and $T > 0$, we consider the following system of stochastic differential equations with respect to $(\Psi_t, X_t) \in \mathcal{O}(\mathsf{M}) \times \mathsf{M}$

$$\begin{cases} \mathrm{d}\Psi_t = \sum_{i=1}^n \mathsf{H}_i(\Psi_t) \circ \Big(\mathrm{d}W_t^i + \Psi_t^{-1} \nabla \log P_{T-t} f(X_t)\,\mathrm{d}t \Big) \\ X_t = \pi\Psi_t \end{cases}, \tag{125}$$

where the notation $\circ$ again refers to the Stratonovitch integral. It is known (see [45, Theorem 7]) that if f is a smooth-enough positive function, then for any initial condition $\Psi_0 = \mathbf{u} \in \mathcal{O}(\mathsf{M})$, the system (125) has a strong solution on $[0, T]$. In [45, Theorem 7], Lehec proved the manifold version of an important representation formula for relative entropy in terms of the *Föllmer process* X_t, first proven in their earlier work [44].

Theorem 22 (Lehec)**.** *Let* (M, g) *be a complete n-dimensional Riemannian manifold whose Ricci curvature is bounded from below and fix a smooth enough positive density function* $f : \mathsf{M} \to \mathbb{R}_+$ *and* $T > 0$*. If* $\{X_t\}_{t\in[0,T]}$ *is a solution of* (125) *with initial condition* $\Psi_0 = \mathbf{u}$ *and* $\pi\mathbf{u} = x$*, then the relative entropy of the measure* μ *with density* $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x} = f$ *is*

$$\mathsf{H}(\mu \| P_T\delta_x) = \frac{1}{2}\mathbb{E}\left[\int_0^T \left|\nabla \log P_{T-t} f(X_t)\right|^2 \mathrm{d}t\right]. \tag{126}$$

where $|v| \stackrel{\text{def}}{=} \mathsf{g}_x(v, v)$ *for* $x \in \mathsf{M}$ *and* $v \in T_x\mathsf{M}$.

It is worth pointing out that, in view of the decay and regularity of the heat kernel on space forms (see, e.g., [21, Chapter 6] and [48, 25, 35]), it suffices to assume that the functions for which we wish to prove the logarithmic Sobolev inequalities of Theorem 4 are Lipschitz and bounded away from 0. Therefore, the regularity conditions required for the function f in Lehec's theorem will always be tacitly assumed to hold.

We record for future reference the following computations (see also [38, Equations (5.5.2) – (5.5.4)]) on the SDE satisfied by partial derivatives of the logarithm of the heat kernel.

Lemma 23. *Let* (M, g) *be a complete n-dimensional Riemannian manifold and fix a smooth enough positive density function* $f : \mathsf{M} \to \mathbb{R}_+$ *and* $T > 0$*. Denote by* $F_t : \mathsf{M} \to \mathbb{R}$ *the function given by*

$$\forall\ x \in \mathsf{M}, \qquad F_t(x) = \log P_{T-t} f(x) \tag{127}$$

and by $\mathbf{F}_t$ *the lift of* F_t *onto* $\mathcal{O}(\mathsf{M})$. *If* $\{X_t\}_{t\in[0,T]}$ *is a solution of* (125) *and* $\{W_t\}_{t\geq 0}$ *is a standard Brownian motion on* $\mathbb{R}^n$, *then for every* $i \in \{1,\dots,n\}$ *we have*

$$\mathrm{d}\mathsf{H}_i\mathbf{F}_t(\Psi_t) = \Big\langle \nabla^{\mathrm{hor}}\mathsf{H}_i\mathbf{F}_t(\Psi_t), \mathrm{d}W_t\Big\rangle + \frac{1}{2}\mathrm{Ric}\Big(\nabla F_t, \Psi_t e_i\Big)(X_t)\,\mathrm{d}t. \tag{128}$$

Proof. Using Itô's formula and (125), we get (omitting the dependence on Ψ_t on the right-hand side of (129))

$$\mathrm{d}\mathsf{H}_i\mathbf{F}_t(\Psi_t) = \Big\langle \nabla^{\mathrm{hor}}\mathsf{H}_i\mathbf{F}_t, \mathrm{d}W_t\Big\rangle + \left\{\frac{\partial \mathsf{H}_i\mathbf{F}_t}{\partial t} + \frac{1}{2}\Delta_{\mathcal{O}(\mathsf{M})}\mathsf{H}_i\mathbf{F}_t + \langle \nabla^{\mathrm{hor}}\mathsf{H}_i\mathbf{F}_t, \nabla^{\mathrm{hor}}\mathbf{F}_t\rangle\right\}\mathrm{d}t. \tag{129}$$

Observe that the function F_t satisfies the equation

$$\frac{\partial \mathbf{F}_t}{\partial t} = -\frac{1}{2}\Delta_{\mathcal{O}(\mathsf{M})}\mathbf{F}_t - \frac{1}{2}\left|\nabla^{\mathrm{hor}}\mathbf{F}_t\right|^2, \tag{130}$$

which, after applying H_i on both sides, gives

$$\frac{\partial \mathsf{H}_i\mathbf{F}_t}{\partial t} = \mathsf{H}_i\frac{\partial \mathbf{F}_t}{\partial t} = -\frac{1}{2}\mathsf{H}_i\Delta_{\mathcal{O}(\mathsf{M})}\mathbf{F}_t - \frac{1}{2}\mathsf{H}_i\left|\nabla^{\mathrm{hor}}\mathbf{F}_t\right|^2. \tag{131}$$

Moreover, we have

$$\mathsf{H}_i\left|\nabla^{\mathrm{hor}}\mathbf{F}_t\right|^2 = \mathsf{H}_i\langle \nabla^{\mathrm{hor}}\mathbf{F}_t, \nabla^{\mathrm{hor}}\mathbf{F}_t\rangle = 2\langle \mathsf{H}_i\nabla^{\mathrm{hor}}\mathbf{F}_t, \nabla^{\mathrm{hor}}\mathbf{F}_t\rangle = 2\langle \nabla^{\mathrm{hor}}\mathsf{H}_i\mathbf{F}_t, \nabla^{\mathrm{hor}}\mathbf{F}_t\rangle, \tag{132}$$

where in the last identity we use that $[\mathsf{H}_i, \mathsf{H}_k]\mathbf{h} = 0$ for any lifted function $\mathbf{h}$ on $\mathcal{O}(\mathsf{M})$ [38, Lemma 5.5.1]. Substituting (131) and (132) in (129), we finally obtain

$$\mathrm{d}\mathsf{H}_i\mathbf{F}_t(\Psi_t) = \Big\langle \nabla^{\mathrm{hor}}\mathsf{H}_i\mathbf{F}_t(\Psi_t), \mathrm{d}W_t\Big\rangle + \frac{1}{2}[\Delta_{\mathcal{O}(\mathsf{M})}, \mathsf{H}_i]\mathbf{F}_t(\Psi_t)\,\mathrm{d}t \tag{133}$$

and the desired identity follows immediately from Lemma 21. □

8.4. **The heat flow on space forms.** The classical Bochner formula (see, e.g., [65]) implies that if (M, g) is a Riemannian manifold with constant Ricci curvature $\mathrm{Ric} = \kappa \in \mathbb{R}$, then

$$\forall\ t \geq 0, \qquad \nabla P_t f = e^{-\kappa t/2} P_t \nabla f \tag{134}$$

for every smooth function $f : \mathsf{M} \to \mathbb{R}$. In [65], Wang investigated commutation relations of this form for second order derivatives instead of the gradient ∇. We shall use the following result.

Theorem 24 (Wang). *A Riemannian manifold* (M, g) *of dimension* n *has constant sectional curvature* $\kappa \in \mathbb{R}$ *if and only if the Hessian tensor of every smooth function* $f : \mathsf{M} \to \mathbb{R}$ *satisfies*

$$\forall\ r \geq 0, \qquad \nabla^2 P_r f = e^{-n\kappa r} P_t \nabla^2 f + \frac{1 - e^{-n\kappa r}}{n} P_r \Delta f \cdot \mathsf{g}. \tag{135}$$

9. Intrinsic dimensional logarithmic Sobolev inequality in space forms

Having explained the necessary background we can now present Theorem 4. We first recall that when $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x} = f$, we have

$$\mathsf{H}(\mu \| P_T\delta_x) = P_T(f\log f)(x) - P_T f(x)\log P_T f(x).$$

Theorem 25. *Let* (M, g) *be an* n*-dimensional Riemannian manifold with constant sectional curvature* $\kappa \in \mathbb{R}\setminus\{0\}$ *with the associated heat semigroup* $\{P_t\}_{t\geq 0}$. *Fix* $T > 0$, $x \in \mathsf{M}$, *a smooth positive function* $f : \mathsf{M} \to \mathbb{R}$ *with* $\int_M f\,\mathrm{d}P_T\delta_x = 1$, *and let* μ *be the probability measure with* $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x} = f$. *Define the 2-tensor* $C(t) = \frac{e^{n\kappa t}}{n\kappa}A + tB$ *for* $t \in \mathbb{R}$, *where* A, B *are 2-tensors given by*

$$\begin{cases} A = -e^{-n\kappa T}\Big(P_T\nabla^2 f(x) - \frac{1}{n}P_T\Delta f(x)\cdot \mathsf{g}\Big) \\ B = \left(\frac{(n-1)\kappa}{2} - \frac{\Delta P_T f(x)}{n}\right)\cdot \mathsf{g}, \end{cases} \tag{136}$$

and let $\mathbb{E}_\mu\big(\nabla \log f\big)^{\otimes 2} \stackrel{\text{def}}{=} \mathbb{E}\Big[\big(\nabla^{\text{hor}} \log \mathbf{f}(\Psi_T)\big)^{\otimes 2}\Big]$. *Then, we have the local intrinsic dimensional logarithmic Sobolev inequality*

$$\begin{aligned}&\mathsf{H}(\mu\|P_T\delta_x)\\ &\le \frac{1}{2}\int_0^T \operatorname{tr}\Big[e^{C(t)-C(T)}\Big(\mathsf{g}+\mathbb{E}_\mu\big(\nabla\log f\big)^{\otimes 2}\int_t^T e^{2C(s)-2C(T)}\,\mathrm{d}s\Big)^{-1}\mathbb{E}_\mu\big(\nabla\log f\big)^{\otimes 2}e^{C(t)-C(T)}\Big]\,\mathrm{d}t,\end{aligned} \tag{137}$$

and the reverse local intrinsic dimensional logarithmic Sobolev inequality

$$\begin{aligned}&\mathsf{H}(\mu\|P_T\delta_x)\\ &\ge \frac{1}{2}\int_0^T \operatorname{tr}\Big[e^{C(t)-C(0)}\Big(\mathsf{g}-\big(\nabla\log P_T f(x)\big)^{\otimes 2}\int_0^t e^{2C(s)-C(0)}\,\mathrm{d}s\Big)^{-1}\big(\nabla\log P_T f(x)\big)^{\otimes 2}e^{C(t)-C(0)}\Big]\,\mathrm{d}t.\end{aligned} \tag{138}$$

The proof of Theorem 25 (see also the stronger Theorem 30) is modeled after the stochastic proof by Eldan, Lehec, and Shenfeld [30] of the intrinsic dimensional logarithmic Sobolev inequality in flat space (7) (and a weaker reverse inequality [30, Theorem 3]). A basic ingredient of this approach is deriving a stochastic differential equation for the *tensor* whose trace is the term $\big|\nabla \log P_{T-t} f(X_t)\big|^2$ in (126). This is the content of the next lemma for which we establish the following notation. Let $\{B_t\}_{t\ge 0}$ be a Brownian motion on M with $B_0 = x$. As before, we denote by F_t the function $\log P_{T-t} f$ and by $\mathbf{F_t}$ its horizontal lift on $\mathcal{O}(\mathsf{M})$. Moreover, we shall denote by G_t the function $\exp F_t = P_{T-t} f$ and by $\mathbf{G}_t = \exp \mathbf{F}_t$ its lift. Consider the random matrices $\mathsf{Q}(t), \mathsf{P}(t) \in \mathbb{M}_n(\mathbb{R})$ (the space of $n\times n$ square matrices over $\mathbb{R}$) given by

$$\begin{aligned}&\mathsf{Q}_{ij}(t) \stackrel{\text{def}}{=} \mathsf{H}_i\mathsf{H}_j\mathbf{F_t}(\Psi_t) = \mathsf{H}_j\mathsf{H}_i\mathbf{F_t}(\Psi_t) = \mathsf{Q}_{ji}(t),\\ &\mathsf{P}(t) \stackrel{\text{def}}{=} \mathsf{Q}(t)^2.\end{aligned} \tag{139}$$

We can now derive the aforementioned stochastic differential equation.

Lemma 26. *Let* (M,g) *be a Riemannian manifold. In the terminology above, for every* $i,j\in\{1,\dots,n\}$, *there exists a martingale* $\{M_{ij}(t)\}_{t\in[0,T]}$ *such that for* $t\in[0,T]$, *we have*

$$\mathsf{H}_i\mathbf{F_t}(\Psi_t)\cdot\mathsf{H}_j\mathbf{F}_t(\Psi_t) = M_{ij}(t) + \frac{1}{2}\int_0^t \mathrm{Ric}\Big(\nabla F_s(X_s), \mathsf{H}_i\mathbf{F}_s(\Psi_s)\cdot\Psi_s e_j + \mathsf{H}_j\mathbf{F}_s(\Psi_s)\cdot\Psi_s e_i\Big)\,\mathrm{d}s + \int_0^t \mathsf{P}_{ij}(s)\,\mathrm{d}s.$$

Proof. Observe that by the chain rule, we have (omitting the dependence on Ψ_t on the right-hand side below)

$$\mathsf{Q}_{ij}(t) = \mathsf{H}_i\mathsf{H}_j\mathbf{F}_t = \mathsf{H}_i\mathsf{H}_j\log\mathbf{G}_t = \frac{\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t}{\mathbf{G}_t} - \mathsf{H}_i\mathbf{F}_t\cdot\mathsf{H}_j\mathbf{F}_t \tag{140}$$

and by the definition and symmetry of the matrix $\mathsf{Q}(t)$,

$$\mathsf{P}_{ij}(t) = \sum_{k=1}^n \mathsf{Q}_{ik}(t)\mathsf{Q}_{kj}(t) = \sum_{k=1}^n \mathsf{H}_k\mathsf{H}_i\mathbf{F}_t\cdot\mathsf{H}_k\mathsf{H}_j\mathbf{F}_t = \Big\langle\nabla^{\text{hor}}\mathsf{H}_i\mathbf{F}_t, \nabla^{\text{hor}}\mathsf{H}_j\mathbf{F}_t\Big\rangle. \tag{141}$$

Combining Itô's product rule with Lemma 23, we get that for $i,j\in\{1,\dots,n\}$,

$$\begin{aligned}\mathrm{d}\Big\{\mathsf{H}_i\mathbf{F_t}(\Psi_t)\cdot\mathsf{H}_j\mathbf{F}_t(\Psi_t)\Big\} &= \mathsf{H}_i\mathbf{F}_t(\Psi_t)\,\mathrm{d}\mathsf{H}_j\mathbf{F}_t(\Psi_t) + \mathsf{H}_j\mathbf{F}_t(\Psi_t)\,\mathrm{d}\mathsf{H}_i\mathbf{F}_t(\Psi_t) + \mathrm{d}\mathsf{H}_j\mathbf{F}_t(\Psi_t)\cdot\mathrm{d}\mathsf{H}_i\mathbf{F}_t(\Psi_t)\\ &= \Big\{\frac{1}{2}\mathrm{Ric}\Big(\nabla F_t, \mathsf{H}_i\mathbf{F}_t\cdot\Psi_t e_j + \mathsf{H}_j\mathbf{F}_t\cdot\Psi_t e_i\Big) + \Big\langle\nabla^{\text{hor}}\mathsf{H}_i\mathbf{F}_t, \nabla^{\text{hor}}\mathsf{H}_j\mathbf{F}_t\Big\rangle\Big\}\,\mathrm{d}t\\ &\qquad + \Big\langle\mathsf{H}_i\mathbf{F}_t\nabla^{\text{hor}}\mathsf{H}_j\mathbf{F}_t + \mathsf{H}_j\mathbf{F}_t\nabla^{\text{hor}}\mathsf{H}_i\mathbf{F}_t, \mathrm{d}W_t\Big\rangle,\end{aligned} \tag{142}$$

where in the right-hand side we again omitted the dependence on Ψ_t and X_t. Denoting the term in the last line by $\mathrm{d}\mathsf{M}_{ij}(t)$, it is clear that $\{M_{ij}(t)\}_{t\in[0,T]}$ is a martingale and (142) becomes

$$\mathsf{H}_i\mathbf{F_t}(\Psi_t)\cdot\mathsf{H}_j\mathbf{F}_t(\Psi_t) = \mathsf{M}_{ij}(t) + \frac{1}{2}\int_0^t \mathrm{Ric}\Big(\nabla F_s(X_s), \mathsf{H}_i\mathbf{F}_s(\Psi_s)\cdot\Psi_s e_j + \mathsf{H}_j\mathbf{F}_s(\Psi_s)\cdot\Psi_s e_i\Big)\,\mathrm{d}s + \int_0^t \mathsf{P}_{ij}(s)\,\mathrm{d}s,$$

where we also used (141). This is the desired identity. □

The stochastic differential equation of Lemma 26 will allow us to derive a differential equation for

$$\forall\ t \in (0,T), \qquad v_{ij}(t) \stackrel{\text{def}}{=} \mathbb{E}\Big[\mathsf{H}_i\mathbf{F}_t(\Psi_t)\cdot\mathsf{H}_j\mathbf{F}_t(\Psi_t)\Big], \qquad i,j\in\{1,\dots,n\}; \tag{143}$$

note that with this notation, (126) reads $\mathsf{H}(\mu\|P_T\delta_x) = \frac{1}{2}\int_0^T \mathrm{tr}\Big[v(t)\Big]\,\mathrm{d}t$. We will then turn the differential equation into a differential *inequality* from which Theorem 30 and Theorem 25 shall follow. To derive the differential equation for $v(t)$ we start by defining

$$m(t) \stackrel{\text{def}}{=} \mathbb{E}[-\mathsf{Q}(t)] \quad \text{and} \quad n(t) \stackrel{\text{def}}{=} \mathbb{E}[\mathsf{P}(t)]. \tag{144}$$

Assuming that the underlying manifold M is Einstein and taking expectations, we deduce the following differential equation for $v(t)$.

Lemma 27. *Let* (M,g) *be an Einstein manifold with constant Ricci curvature* $\mathrm{Ric}=\rho$ *for some* $\rho\in\mathbb{R}$. *For every* $i,j\in\{1,\dots,n\}$ *and* $t\in(0,T)$, *we have*

$$\frac{\mathrm{d}v_{ij}(t)}{\mathrm{d}t} = n_{ij}(t) + \rho v_{ij}(t). \tag{145}$$

Proof. Since M has constant Ricci curvature ρ, we have

$$\mathrm{Ric}\Big(\nabla F_s(X_s), \mathsf{H}_i\mathbf{F}_s(\Psi_s)\cdot\Psi_s e_j\Big) = \rho\mathsf{H}_i\mathbf{F}_s(\Psi_s)\cdot\mathsf{g}(\nabla F_s(X_s),\Psi_s e_j) = \rho\mathsf{H}_i\mathbf{F}_s(\Psi_s)\cdot\mathsf{H}_j\mathbf{F}_s(\Psi_s). \tag{146}$$

Plugging this in the rightmost term of Lemma 26, we get that

$$\mathsf{H}_i\mathbf{F_t}(\Psi_t)\cdot\mathsf{H}_j\mathbf{F}_t(\Psi_t) = \mathsf{M}_{ij}(t) + \rho\int_0^t \mathsf{H}_i\mathbf{F}_s(\Psi_s)\cdot\mathsf{H}_j\mathbf{F}_s(\Psi_s)\,\mathrm{d}s + \int_0^t \mathsf{P}_{ij}(s)\,\mathrm{d}s. \tag{147}$$

The result follows after taking expectation (since $\mathbb{E}M_{ij}(t) = M_{ij}(0) = 0$) and differentiating. □

In order to turn (145) into a differential *inequality* we will use Jensen's inequality $n(t)\geq m(t)^2$ where we used $\mathsf{P}=\mathsf{Q}^2$. To use the latter inequality we need to better understand the term $m(t)$. On manifolds of constant curvature, $m(t)$ takes the following simple form.

Lemma 28. *Let* (M,g) *be an* n*-dimensional Riemannian manifold with constant sectional curvature* $\kappa\in\mathbb{R}$. *For every* $i,j\in\{1,\dots,n\}$ *and* $t\in(0,T)$, *we have*

$$m_{ij}(t) = v_{ij}(t) - e^{-n\kappa(T-t)}P_T\nabla^2 f(\Phi_0 e_i,\Phi_0 e_j)(x) - \frac{1-e^{-n\kappa(T-t)}}{n}P_T\Delta f(x)\cdot\delta_{ij}. \tag{148}$$

Proof. Taking expectations in (140), we obtain

$$m_{ij}(t) = \mathbb{E}[-\mathsf{Q}_{ij}(t)] \stackrel{(140)}{=} v_{ij}(t) - \mathbb{E}\left[\frac{\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t(\Psi_t)}{\mathbf{G}_t(\Psi_t)}\right]. \tag{149}$$

It follows from (125) and (121) that Ψ_t has law $f(B_T)$ with respect to Φ_t for every $t\in[0,T]$ (see also the proof of [45, Theorem 7] for an argument based on Girsanov's theorem). Therefore, by the tower property of conditional expectation, we have

$$\begin{aligned}\mathbb{E}\left[\frac{\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t(\Psi_t)}{\mathbf{G}_t(\Psi_t)}\right] &= \mathbb{E}\left[\frac{\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t(\Phi_t)}{\mathbf{G}_t(\Phi_t)}f(B_T)\right] = \mathbb{E}\left[\frac{\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t(\Phi_t)}{\mathbf{G}_t(\Phi_t)}\mathbb{E}\Big[f(B_T)\Big|\{\Phi_r\}_{r\leq t}\Big]\right]\\ &= \mathbb{E}\left[\frac{\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t(\Phi_t)}{\mathbf{G}_t(\Phi_t)}P_{T-t}f(B_t)\right] = \mathbb{E}\Big[\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t(\Phi_t)\Big].\end{aligned} \tag{150}$$

Recall that for any function $h:\mathsf{M}\to\mathbb{R}$ with horizontal lift $\mathbf{h}$, we have

$$\forall\ \mathbf{u}\in\mathcal{O}(\mathsf{M}), \qquad \mathsf{H}_i\mathsf{H}_j\mathbf{h}(\mathbf{u}) = \nabla^2 h(\mathbf{u}e_i,\mathbf{u}e_j)(\pi\mathbf{u}), \tag{151}$$

see, e.g., [38, Equation (2.2.3)]. Combining (151) with Theorem 24, we deduce that

$$\begin{aligned}\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t(\Phi_t) &\overset{(151)}{=} \nabla^2 P_{T-t} f(\Phi_t e_i, \Phi_t e_j)(B_t)\\ &\overset{(135)}{=} e^{-n\kappa(T-t)} P_{T-t}\nabla^2 f(\Phi_t e_i, \Phi_t e_j)(B_t) + \frac{1-e^{-n\kappa(T-t)}}{n} P_{T-t}\Delta f(B_t)\cdot \mathsf{g}(\Phi_t e_i, \Phi_t e_j) \qquad (152)\\ &= e^{-n\kappa(T-t)} P_{T-t}\nabla^2 f(\Phi_t e_i, \Phi_t e_j)(B_t) + \frac{1-e^{-n\kappa(T-t)}}{n} P_{T-t}\Delta f(B_t)\cdot \delta_{ij},\end{aligned}$$

where in the last equality we used that $\{\Phi_t e_1, \ldots, \Phi_t e_n\}$ is an orthonormal basis of $T_{B_t}\mathsf{M}$. Taking expectations on both sides, we get

$$\mathbb{E}\Big[\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t(\Phi_t)\Big] = e^{-n\kappa(T-t)}\mathbb{E}\Big[P_{T-t}\nabla^2 f(\Phi_t e_i, \Phi_t e_j)(B_t)\Big] + \frac{1-e^{-n\kappa(T-t)}}{n}\mathbb{E}\Big[P_{T-t}\Delta f(B_t)\Big]\cdot\delta_{ij}. \qquad (153)$$

By the definition [65, Equation (1.2)] of the action of $\{P_s\}_{s\geq 0}$ on tensors, we have

$$\mathbb{E}\Big[P_{T-t}\nabla^2 f(\Phi_t e_i, \Phi_t e_j)(B_t)\Big] = \mathbb{E}\Big[\nabla^2 f(\Phi_T e_i, \Phi_T e_j)(B_T)\Big] = P_T\nabla^2 f(\Phi_0 e_i, \Phi_0 e_j)(x),$$

where the last identity follows from the definition of stochastic parallel transport given by $\{\Phi_s\circ\Phi_0^{-1}\}_{s\geq 0}$ (see [38, Section 2.3]). Similarly, we have

$$\mathbb{E}\Big[P_{T-t}\Delta f(B_t)\Big] = \mathbb{E}\Big[\Delta f(B_T)\Big] = P_T\Delta f(x) \qquad (154)$$

and combining everything we deduce that

$$\mathbb{E}\Big[\mathsf{H}_i\mathsf{H}_j\mathbf{G}_t(\Phi_t)\Big] = e^{-n\kappa(T-t)}P_T\nabla^2 f(\Phi_0 e_i, \Phi_0 e_j)(x) + \frac{1-e^{-n\kappa(T-t)}}{n}P_T\Delta f(x)\cdot\delta_{ij}. \qquad (155)$$

Plugging (155) and (150) in (149) completes the proof. □

We are now ready to derive the differential inequality for $v(t)$. For simplicity, we shall denote by $c_T \overset{\text{def}}{=} P_T\Delta f(x)$ and by J_T the symmetric matrix with

$$(J_T)_{ij} \overset{\text{def}}{=} P_T\nabla^2 f(\Phi_0 e_i, \Phi_0 e_j)(x) - \frac{1}{n}P_T\Delta f(x)\cdot\delta_{ij}, \qquad (156)$$

which satisfies $\mathrm{tr}J_T = 0$. Combining all of the above, we get the following matrix inequality:

Proposition 29. *Let* (M,g) *be an n-dimensional Riemannian manifold with constant sectional curvature* $\kappa\in\mathbb{R}$. *For every* $t\in(0,T)$, *we have*

$$\begin{aligned}\frac{\mathrm{d}v(t)}{\mathrm{d}t} \geq v(t)^2 &- \Big(e^{-n\kappa(T-t)}J_T + \tfrac{c_T}{n}\cdot\mathsf{Id}_n\Big)v(t) - v(t)\Big(e^{-n\kappa(T-t)}J_T + \tfrac{c_T}{n}\cdot\mathsf{Id}_n\Big)\\ &+ \Big(e^{-n\kappa(T-t)}J_T + \tfrac{c_T}{n}\cdot\mathsf{Id}_n\Big)^2 + (n-1)\kappa v(t),\end{aligned} \qquad (157)$$

so in particular,

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} \geq v(t)^2 + \Big(\Big(\tfrac{(n-1)\kappa}{2} - \tfrac{c_T}{n}\Big)\cdot\mathsf{Id}_n - e^{-n\kappa(T-t)}J_T\Big)v(t) + v(t)\Big(\Big(\tfrac{(n-1)\kappa}{2} - \tfrac{c_T}{n}\Big)\cdot\mathsf{Id}_n - e^{-n\kappa(T-t)}J_T\Big), \qquad (158)$$

where $\geq$ *is the inequalities in the positive semidefinite ordering.*

Proof. Combining the matrix Jensen inequality

$$n(t) = \mathbb{E}\Big[Q(t)^2\Big] \geq \mathbb{E}\Big[-Q(t)\Big]^2 = m(t)^2 \qquad (159)$$

with (145), (148) and expanding, we get (157)

$$\begin{aligned}\frac{\mathrm{d}v(t)}{\mathrm{d}t} \geq v(t)^2 &- \Big(e^{-n\kappa(T-t)}J_T + \tfrac{c_T}{n}\cdot\mathsf{Id}_n\Big)v(t) - v(t)\Big(e^{-n\kappa(T-t)}J_T + \tfrac{c_T}{n}\cdot\mathsf{Id}_n\Big)\\ &+ \Big(e^{-n\kappa(T-t)}J_T + \tfrac{c_T}{n}\cdot\mathsf{Id}_n\Big)^2 + (n-1)\kappa v(t).\end{aligned} \qquad (160)$$

The inequality (158) follows since the squared matrix is positive semidefinite. □

Proposition 29 allows us to deduce the following local intrinsic dimensional logarithmic Sobolev inequalities which are, however, non-explicit.

Theorem 30. *Let* (M,g) *be an n-dimensional Riemannian manifold with constant sectional curvature* $\kappa\in\mathbb{R}$. *Fix* $T>0$, $x\in\mathsf{M}$, *a smooth positive function* $f:\mathsf{M}\to\mathbb{R}$ *with* $\int_M f\,\mathrm{d}P_T\delta_x=1$, *and let* μ *be the probability measure with* $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x}=f$. *Suppose there is a family of matrices* $U(t)\in\mathbb{M}_n(\mathbb{R})$ *for* $t\in[0,T]$ *which solves the equation*

$$\begin{aligned}\frac{\mathrm{d}U(t)}{\mathrm{d}t}=U(t)^2-\Big(e^{-n\kappa(T-t)}J_T+\tfrac{c_T}{n}\cdot\mathsf{Id}_n\Big)U(t)-U(t)\Big(e^{-n\kappa(T-t)}J_T+\tfrac{c_T}{n}\cdot\mathsf{Id}_n\Big)\\+\Big(e^{-n\kappa(T-t)}J_T+\tfrac{c_T}{n}\cdot\mathsf{Id}_n\Big)^2+(n-1)\kappa U(t),\end{aligned}\tag{161}$$

with either initial condition $U(0):=v(0)$ *or* $U(T):=v(T)$. *Then, we have the local intrinsic dimensional logarithmic Sobolev inequality*

$$\mathsf{H}(\mu\|P_T\delta_x)\le\frac{1}{2}\int_0^T\mathrm{tr}\Big[U(t)\Big]\,\mathrm{d}t,\quad U(T)=\mathbb{E}\Big(\nabla^{\mathrm{hor}}\log f(\Psi_T)\Big)^{\otimes2}.\tag{162}$$

and the reverse local intrinsic dimensional logarithmic Sobolev inequality

$$\mathsf{H}(\mu\|P_T\delta_x)\ge\frac{1}{2}\int_0^T\mathrm{tr}\Big[U(t)\Big]\,\mathrm{d}t,\quad U(0)=\Big(\nabla^{\mathrm{hor}}\log P_Tf(x)\Big)^{\otimes2}.\tag{163}$$

Proof. Lehec's formula (126) implies

$$\mathsf{H}(\mu\|P_T\delta_x)\overset{(126)}{=}\frac{1}{2}\mathbb{E}\Big[\int_0^T\big|\nabla F_t(X_t)\big|^2\,\mathrm{d}t\Big]=\frac{1}{2}\sum_{i=1}^n\int_0^T\mathbb{E}\Big[\mathsf{H}_i\mathbf{F}_t(\Psi_t)^2\Big]\,\mathrm{d}t=\frac{1}{2}\int_0^T\mathrm{tr}\Big[v(t)\Big]\,\mathrm{d}t.\tag{164}$$

For the reverse local intrinsic dimensional logarithmic Sobolev inequality, we note that $U(0)=v(0)$ so the result follows by (157) and standard comparison principles for matrix Ricatti equations, see [40]. For the local intrinsic dimensional logarithmic Sobolev inequality, we have $U(T)=v(T)$ and the conclusion follows by reversing time. □

Theorem 30 provides sharp results which are, however, not explicit since the solutions of (161) are complicated. They are expressed in terms of special functions, except in the flat space case where they simplify considerably– see Section 9.1. To avoid the complication of Theorem 30 we will use (158), rather than the stronger inequality (157), which will lead to explicit bounds, namely Theorem 25. To this end, we shall need the following technical lemma on matrix Bernoulli differential inequalities.

Lemma 31. *Fix* $T>\varepsilon>0$, $n\in\mathbb{N}$, $\gamma\in\mathbb{R}\setminus\{0\}$ *and let* $A,B\in\mathbb{M}_n(\mathbb{R})$ *be symmetric matrices with* $AB=BA$. *Consider* $C(t)\overset{\mathrm{def}}{=}\frac{e^{\gamma t}}{\gamma}A+tB$, *where* $t\in\mathbb{R}$. *For any positive definite matrix* $V_\varepsilon\in\mathbb{M}_n(\mathbb{R})$, *if a continuous function* $V:[\varepsilon,T]\to\mathbb{M}_n(\mathbb{R})$ *for which every* $V(t)$ *is a positive semi-definite matrix satisfies the ordinary differential inequality*

$$\forall\ t\in(\varepsilon,T),\qquad\frac{\mathrm{d}V(t)}{\mathrm{d}t}\ge V(t)^2+\Big(e^{\gamma t}A+B\Big)V(t)+V(t)\Big(e^{\gamma t}A+B\Big)\tag{165}$$

with boundary condition $V(\varepsilon)=V_\varepsilon$, *then it also satisfies the matrix inequalities*

$$\forall\ t\in[\varepsilon,T],\qquad V(t)\le e^{C(t)-C(T)}\Big(\mathsf{Id}_n+V(T)\int_t^T e^{2C(s)-2C(T)}\,\mathrm{d}s\Big)^{-1}V(T)e^{C(t)-C(T)}.\tag{166}$$

and

$$\forall\ t\in[\varepsilon,T],\qquad V(t)\ge e^{C(t)-C(\varepsilon)}\Big(\mathsf{Id}_n-V(\varepsilon)\int_\varepsilon^t e^{2C(s)-2C(\varepsilon)}\,\mathrm{d}s\Big)^{-1}V(\varepsilon)e^{C(t)-C(\varepsilon)}.\tag{167}$$

Moreover, the right-hand side of (167) *is positive definite for every* $t\in(\varepsilon,T)$.

Proof. Since A and B commute, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{C(t)} = \Big(e^{\gamma t}A + B\Big)e^{C(t)} = e^{C(t)}\Big(e^{\gamma t}A + B\Big). \tag{168}$$

As V_ε is positive definite, the same holds for $V(t)$ for t near ε so let $t_{\max} \in [\varepsilon, T]$ be the supremum over $t \in [\varepsilon, T]$ where $V(t)$ is positive definite. For $t \in (\varepsilon, t_{\max})$, multiplying (165) by $V(t)^{-1}$ on both sides, we deduce that

$$\frac{\mathrm{d}V(t)^{-1}}{\mathrm{d}t} \preceq -\mathsf{Id}_n - V(t)^{-1}C'(t) - C'(t)V(t)^{-1}, \tag{169}$$

where $C'(t) \stackrel{\text{def}}{=} \frac{\mathrm{d}C(t)}{\mathrm{d}t}$. Therefore, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[e^{C(t)}V(t)^{-1}e^{C(t)}\Big] \stackrel{(168)}{=} e^{C(t)}\Big(C'(t)V(t)^{-1} + \frac{\mathrm{d}V(t)^{-1}}{\mathrm{d}t} + V(t)^{-1}C'(t)\Big)e^{C(t)} \stackrel{(169)}{\preceq} -e^{2C(t)}, \tag{170}$$

where in the last inequality we used that $C(t)$ is symmetric. Integrating from ε to t, we get

$$e^{C(t)}V(t)^{-1}e^{C(t)} - e^{C(\varepsilon)}V(\varepsilon)^{-1}e^{C(\varepsilon)} \preceq -\int_\varepsilon^t e^{2C(s)}\,\mathrm{d}s \tag{171}$$

which can be rearranged to give, for every $t \in [\varepsilon, t_{\max})$,

$$V(t)^{-1} \preceq e^{-C(t)}\Big(e^{C(\varepsilon)}V(\varepsilon)^{-1}e^{C(\varepsilon)} - \int_\varepsilon^t e^{2C(s)}\,\mathrm{d}s\Big)e^{-C(t)}. \tag{172}$$

Since the right-hand side of (172) is finite for every $t \in [\varepsilon, T]$, we can take the limit $t \uparrow t_{\max}$ to conclude that $V(t_{\max})$ is positive definite, and hence $t_{\max} = T$. Since the function $A \mapsto A^{-1}$ is operator decreasing on positive definite matrices, this proves (167) after some simple algebraic manipulations. Moreover, as a consequence of (172), the right-hand side of (167) is indeed positive definite. Similarly, integrating (170) from t to T and rearranging gives

$$V(t)^{-1} \succeq e^{-C(t)}\Big(e^{C(T)}V(T)^{-1}e^{C(T)} + \int_t^T e^{2C(s)}\,\mathrm{d}s\Big)e^{-C(t)}. \tag{173}$$

However, since $V(t)^{-1}$ is positive definite for every $t \in [\varepsilon, T]$ this is equivalent to

$$V(t) \preceq e^{C(t)-C(T)}\Big(\mathsf{Id}_n + V(T)\int_t^T e^{2C(s)-2C(T)}\,\mathrm{d}s\Big)^{-1}V(T)e^{C(t)-C(T)}, \tag{174}$$

which concludes the proof of (166). □

Proof of Theorem 25. Fix $T > 0$, $\varepsilon > 0$, and $x \in \mathsf{M}$. Let $f : \mathsf{M} \to \mathbb{R}$ be a smooth positive function with $\int_{\mathsf{M}} f\,\mathrm{d}P_T\delta_x = 1$ and let μ be the probability measure on M with $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x} = f$. Without loss of generality, we can perturb f and assume that

$$v_\varepsilon \stackrel{\text{def}}{=} \mathbb{E}\Big[\Big(\nabla^{\text{hor}}\log P_{T-\varepsilon}\mathbf{f}(\Psi_\varepsilon)\Big)^{\otimes 2}\Big] \tag{175}$$

is a positive definite matrix. Following the terminology above, Lehec's formula (126) implies

$$\mathsf{H}(\mu\|P_T\delta_x) \stackrel{(126)}{=} \frac{1}{2}\mathbb{E}\Big[\int_0^T \big|\nabla F_t(X_t)\big|^2\,\mathrm{d}t\Big] = \frac{1}{2}\sum_{i=1}^n\int_0^T \mathbb{E}\Big[\mathsf{H}_i\mathbf{F}_t(\Psi_t)^2\Big]\,\mathrm{d}t = \frac{1}{2}\int_0^T \mathrm{tr}\Big[v(t)\Big]\,\mathrm{d}t. \tag{176}$$

Since $v(\varepsilon) = v_\varepsilon$ is a positive definite matrix, Proposition 29 and Lemma 31 give

$$\forall\ t \in [\varepsilon, T], \qquad v(t) \preceq e^{C(t)-C(T)}\Big(\mathsf{Id}_n + v_T\int_t^T e^{2C(s)-2C(T)}\,\mathrm{d}s\Big)^{-1}v_T e^{C(t)-C(T)} \tag{177}$$

where $C(t) = \frac{e^{\gamma t}}{\gamma} A + tB$, for the matrices

$$\begin{cases} A = -e^{-n\kappa T} J_T \\ B = \left(\frac{(n-1)\kappa}{2} - \frac{c_T}{n}\right) \cdot \mathsf{Id}_n. \\ \gamma = n\kappa \end{cases} \tag{178}$$

By the perturbation above, we have thus established the validity of (177) for an arbitrary smooth positive density f and for any $\varepsilon > 0$. Since $v_T = \mathbb{E}_\mu\big(\nabla \log f\big)^{\otimes 2}$, the logarithmic Sobolev inequality of Theorem 25 follows by combining (176) and (177) with $\varepsilon \to 0^+$. The reverse logarithmic Sobolev inequality follows by using (167) since $v_0 = \big(\nabla \log P_T f(x)\big)^{\otimes 2}$. □

9.1. **Intrinsic dimensional local logarithmic Sobolev inequalities in flat spaces.** Our next goal is to prove the intrinsic dimensional local logarithmic Sobolev inequalities in flat spaces, i.e., equations (29) and (30). In contrast to the proof of Theorem 25, which uses the weaker inequality (158), here we will use the stronger inequality (157) which in flat space has an explicit clear solution.

Theorem 32. *Fix $T > 0$ and $x \in \mathbb{R}^n$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth positive function with $\int_{\mathbb{R}^n} f \, \mathrm{d}P_T\delta_x = 1$ and let μ be the probability measure on $\mathbb{R}^n$ with $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x} = f$. Then, we have the local intrinsic dimensional logarithmic Sobolev inequality*

$$\mathsf{H}(\mu \| P_T\delta_x) \le \frac{T}{2} \Delta P_T f(x) + \frac{1}{2} P_T f(x) \log\det\left(\mathsf{Id}_n - T\frac{P_T(f\nabla^2 \log f(x))}{P_T f(x)}\right), \tag{179}$$

and the reverse local intrinsic dimensional logarithmic Sobolev inequality

$$\mathsf{H}(\mu \| P_T\delta_x) \ge \frac{T}{2} \Delta P_T f(x) - \frac{1}{2} P_T f(x) \log\det\big(\mathsf{Id}_n + T\nabla^2 \log P_T f(x)\big). \tag{180}$$

Proof. The inequality (179) follows by setting $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda_n} \stackrel{\text{def}}{=} \frac{f P_T\delta_x}{P_T f(x)}$ in (7). To prove (180), we may assume without loss of generality assume that $-\nabla^2 \log P_T f(x)$ is invertible. Set $U(0) \stackrel{\text{def}}{=} v(0) = (\nabla \log P_T f(x))^{\otimes 2}$ and use the normalization assumption $\int_{\mathbb{R}^n} f \, \mathrm{d}P_T\delta_x = P_T f(x) = 1$ to conclude that

$$U(0) - \nabla^2 P_T f(x) = (\nabla \log P_T f(x))^{\otimes 2} - \frac{\nabla^2 P_T f(x)}{P_T f(x)} = -\nabla^2 \log P_T f(x) \tag{181}$$

is invertible. In flat space, using $\kappa = 0$ and $J_T + \frac{c_T}{n} \cdot \mathsf{Id}_n = P_T \nabla^2 f(x)$, equation (161) becomes

$$\frac{\mathrm{d}U(t)}{\mathrm{d}t} = U(t)^2 - P_T\nabla^2 f(x) U(t) - U(t) P_T \nabla^2 f(x) + (P_T \nabla^2 f(x))^2. \tag{182}$$

The solution of (182) can be verified to be

$$\forall\ t \in (0,T), \qquad U(t) = \Big([U(0) - P_T\nabla^2 f(x)]^{-1} - t\Big)^{-1} + P_T \nabla^2 f(x), \tag{183}$$

where we used Hamilton's matrix inequality (32) (see also Theorem 33 below) to justify the invertibility of $[U(0) - P_T\nabla^2 f(x)]^{-1} - t$. Applying (163) of Theorem 30 yields

$$\mathsf{H}(\mu \| P_T\delta_x) \ge -\frac{1}{2} \int_0^T \mathrm{tr}\Big[\Big((\nabla^2 \log P_T f(x))^{-1} + t\Big)^{-1}\Big] \, \mathrm{d}t + \frac{T}{2} P_T \Delta f(x), \tag{184}$$

where again we used normalization assumption $P_T f(x) = 1$. To rewrite the right-hand side of (184) let $\{\lambda_i\}_{i=1}^n$ stand for the eigenvalues of $\nabla^2 \log P_T f(x)$ so

$$\begin{aligned} \int_0^T \mathrm{tr}\Big[\Big((\nabla^2 \log P_T f(x))^{-1} + t\Big)^{-1}\Big] \, \mathrm{d}t &= \sum_{i=1}^n \int_0^T (\lambda_i^{-1} + t)^{-1} \, \mathrm{d}t = \sum_{i=1}^n \log\left(\frac{\lambda_i^{-1} + T}{\lambda_i^{-1}}\right) \\ &= \log\det\big(\mathsf{Id}_n + T\nabla^2 \log P_T f(x)\big). \end{aligned} \tag{185}$$

It follows that (184) reads

$$\mathsf{H}(\mu\|P_T\delta_x) \geq \frac{T}{2}\Delta P_T f(x) - \frac{1}{2}\log\det\Big(\mathsf{Id}_n + T\nabla^2 \log P_T f(x)\Big), \tag{186}$$

so using again $P_T f(x) = 1$, (186) is equivalent to

$$\mathsf{H}(\mu\|P_T\delta_x) \geq \frac{T}{2}\Delta P_T f(x) - \frac{1}{2}P_T f(x)\log\det\Big(\mathsf{Id}_n + T\nabla^2 \log P_T f(x)\Big) \qquad \square$$

Semigroup vs. stochastic interpolation. The idea of writing the relative entropy as an integral of a *gradient term* goes back to the beginning of the Bakry–Émery theory of functional inequalities (see [8, 4] or [9, Section 5.5]). Such gradient terms often satisfy differential inequalities à la Proposition 29 which allow for the use of comparison principles in the spirit of Lemma 31. For instance, Lehec in [45] considered the scalar-valued function $\alpha : [0,T] \to \mathbb{R}$ given by

$$\forall\ t \in [0,T], \qquad \alpha(t) = \mathbb{E}\Big[|\nabla F_t(X_t)|^2\Big] \tag{187}$$

and showed that

$$\forall\ t \in (0,T), \qquad \alpha'(t) \geq \frac{1}{n}\Big(\alpha(t) - c_T\Big)^2 + (n-1)\kappa\alpha(t) \geq \frac{\alpha(t)}{n}\Big(\alpha(t) + n(n-1)\kappa - 2c_T\Big). \tag{188}$$

Applying a standard comparison principle to the latter inequality, he then derived a dimensional upper bound for the relative entropy, see [45, Equation (25)]. It is worth pointing out that (188) is also a consequence of (157) after taking traces and using the elementary inequality $\mathrm{tr}C^2 \geq \frac{(\mathrm{tr}C)^2}{n}$ which holds for all $n \times n$ positive semidefinite matrices C. A close inspection of the arguments of this section reveals that the logarithmic Sobolev inequality of Theorem 25 is a strengthening of Lehec's result for manifolds of constant sectional curvature.

By reasoning similar to (150), for every $i \in \{1,\ldots,n\}$, we have

$$\mathbb{E}\Big[(\mathsf{H}_i\mathbf{F_t}(\Psi_t))^2\Big] = \mathbb{E}\Big[\frac{(\mathsf{H}_i\mathbf{G}_t(\Psi_t))^2}{\mathbf{G}_t(\Psi_t)^2}\Big] = \mathbb{E}\Big[\frac{(\mathsf{H}_i\mathbf{G}_t(\Phi_t))^2}{(\mathbf{G}_t(\Phi_t))^2}f(B_T)\Big] = \mathbb{E}\Big[\frac{(\mathsf{H}_i\mathbf{G}_t(\Phi_t))^2}{\mathbf{G}_t(\Phi_t)}\Big] \tag{189}$$

and thus

$$\forall\ t \in (0,T), \qquad \alpha(t) = P_t\left[\frac{|\nabla P_{T-t}f|^2}{P_{T-t}f}\right](x). \tag{190}$$

This semigroup representation of $\alpha(t)$ was used by Bakry, Bolley and Gentil in [6, p. 405] to give an independent proof of inequality (188) for semigroups satisfying the curvature dimension condition $\mathsf{CD}(\rho,n)$, where $\rho = (n-1)\kappa$ (observe that $\alpha(t)$ is denoted by $\Lambda'(t)$ in their paper). Their main result [6, Theorem 2.2] improves upon [45, Equation (25)] as they did not disregard the nonnegative constant in the second inequality of (188) and thus get to apply a tighter comparison principle. One could implement a similar strategy in the matricial setting treated here, by replacing the matrix inequality (158) with the stronger inequality (157) and solving the corresponding ordinary differential equation. However, solutions of this equation appear to be non-explicit, so we chose the comparatively simpler presentation of Theorem 25 for clarity of the exposition. If one were to implement this reasoning, it is clear from the proofs of this section that the resulting inequality would improve upon [6, Theorem 2.2].

10. Nonpositively curved space forms

The goal of this section is to prove Theorem 5 (section 10.1) and Theorem 6 (section 10.2). We conclude with section 10.3 which discusses some of the ideas behind our proofs.

10.1. **Matrix inequalities.** The main result of this section is the following Hamilton-type matrix inequality, namely, Theorem 5.

Theorem 33. *Let* (M, g) *be an n-dimensional Riemannian manifold with constant nonpositive sectional curvature* $\kappa \leq 0$*. Then, for every* $T \geq 0$*,*

$$\text{if, either } \kappa = 0, \text{ or } \kappa < 0 \text{ and } \frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} = 1, \quad \text{then} \quad -\nabla^2 \log P_T f(x) \preceq \frac{1}{T}\mathsf{Id}_n \quad \forall x \in \mathsf{M}, \tag{191}$$

Further,

$$\begin{aligned} &\text{if} \quad \frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)}) > 1, \\ &\text{then} \quad -\nabla^2 \log P_T f(x) \preceq \frac{n\kappa}{2}\left\{\sqrt{\frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} - 1}\cot\left(\frac{n\kappa T}{2}\sqrt{\frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} - 1}\right) - 1\right\}\mathsf{Id}_n. \end{aligned} \tag{192}$$

Remark 34. *To put Theorem 5 in a larger context, and also to shed light on the conditions regarding* $\frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)}$*, let us recall the improved Li-Yau inequality of Bakry, Bolley, and Gentil* [6, Corollaries 2.3,2.4]*: Let* (M, g) *be an n-dimensional Riemannian manifold with lower bound* $(n-1)\kappa$ *on its Ricci curvature. Let* $\{P_t\}_{t\geq 0}$ *be the associated heat semigroup and let* $f : \mathsf{M} \to \mathbb{R}$ *be a positive function. Then, for every* $x \in \mathsf{M}$ *and every* $T \geq 0$*,*

$$\frac{4}{n(n-1)\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} < 1 + \frac{4\pi^2}{(n-1)^2\kappa^2 T^2} \tag{193}$$

and

$$-\Delta \log P_T f(x) < \tag{194}$$

$$\begin{cases} \frac{n(n-1)\kappa}{2}\left\{\sqrt{\frac{4}{n(n-1)\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} - 1}\coth\left(\frac{(n-1)\kappa T}{2}\sqrt{\frac{4}{n(n-1)\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} - 1}\right) - 1\right\} \text{if } \frac{4}{n(n-1)\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} \leq 1 \\ \frac{n(n-1)\kappa}{2}\left\{\sqrt{\frac{4}{n(n-1)\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} - 1}\cot\left(\frac{(n-1)\kappa T}{2}\sqrt{\frac{4}{n(n-1)\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} - 1}\right) - 1\right\} \text{if } 1 \leq \frac{4}{n(n-1)\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} < 1 + \frac{4\pi^2}{(n-1)^2\kappa^2 T^2}. \end{cases} \tag{195}$$

Hence, in the regime $1 + \frac{1}{n-1} \leq \frac{4}{n(n-1)\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} < 1 + \frac{4\pi^2}{(n-1)^2\kappa^2 T^2}$ *we are able to obtain in hyperbolic spaces a* matrix *version of the improved Li-Yau inequality.*

Proof of Theorem 33. We start by showing that

$$m(t) \stackrel{\text{def}}{=} \mathbb{E}[-\nabla^2 \log P_{T-t} f(X_t)] \tag{196}$$

satisfies the following differential inequality.

Lemma 35. *When* $\kappa \leq 0$*,*

$$\forall\ t \in [0, T], \qquad \frac{\mathrm{d}m(t)}{\mathrm{d}t} \geq m(t)^2 + n\kappa m(t) + \kappa \Delta P_T f(x) \cdot \mathrm{g}. \tag{197}$$

Proof. Define

$$u(t) \stackrel{\text{def}}{=} e^{-n\kappa(T-t)} J_T + \frac{c_T}{n} \cdot \mathrm{g} \tag{198}$$

where $(J_T)_{ij} \stackrel{\text{def}}{=} P_T \nabla^2 f(\Phi_0 e_i, \Phi_0 e_j)(x) - \frac{1}{n} P_T \Delta f(x) \cdot \delta_{ij}$ and $c_T \stackrel{\text{def}}{=} P_T \Delta f(x)$. By (148),

$$m(t) = v(t) - u(t), \tag{199}$$

so, by (157),

$$\frac{\mathrm{d}m(t)}{\mathrm{d}t} = \frac{\mathrm{d}v(t)}{\mathrm{d}t} - n\kappa e^{-n\kappa(T-t)} J_T = \frac{\mathrm{d}v(t)}{\mathrm{d}t} - n\kappa u(t) + \kappa c_T \cdot \mathrm{g} \tag{200}$$

$$\geq v(t)^2 - u(t)v(t) - v(t)u(t) + u(t)^2 + (n-1)\kappa v(t) - n\kappa u(t) + \kappa c_T \cdot \mathrm{g} \tag{201}$$

$$= m(t)^2 + n\kappa m(t) - \kappa v(t) + \kappa c_T \cdot \mathrm{g} \tag{202}$$

$$\geq m(t)^2 + n\kappa m(t) + \kappa c_T \cdot \mathrm{g} \tag{203}$$

where the last inequality uses that $\kappa \leq 0$ and that $v(t) \geq 0$ (since it is a nonnegative sum of rank-one matrices). □

The following technical lemma on matrix differential inequalities will allow us to further control the matrix $m(t)$.

Lemma 36. *Fix $T > 0$ and let $W(t)$ be a family of matrices for $t \in [0,T]$ satisfying the differential inequality*

$$\forall\ t \in [0,T], \qquad \frac{\mathrm{d}W(t)}{\mathrm{d}t} \geq W(t)^2 + \alpha W(t) + \beta \cdot \mathrm{Id}_n \tag{204}$$

for some constant $\alpha, \beta \in \mathbb{R}$. Fix $\theta \in S^{n-1}$ and let $\phi(t) \stackrel{\text{def}}{=} \langle W(t)\theta, \theta\rangle$ for $t \in [0,T]$. Then, fixing $\phi(0) = c$ we have

$$\forall\ t \in [0,T], \qquad \phi(t) \geq \xi_\lambda(t) - \frac{\alpha}{2} \tag{205}$$

where

$$\xi_\lambda(t) \stackrel{\text{def}}{=} \begin{cases} \sqrt{\lambda}\tan(\sqrt{\lambda}t + c_1) & \text{if } \lambda > 0 \\ -\frac{1}{t+c_2} & \text{if } \lambda = 0 \\ -\sqrt{-\lambda}\tanh(\sqrt{-\lambda}t + c_3) & \text{if } \lambda < 0, \end{cases} \tag{206}$$

with

$$\lambda \stackrel{\text{def}}{=} \beta - \frac{\alpha^2}{4}, \tag{207}$$

and

$$c_1 \stackrel{\text{def}}{=} \arctan\left(\frac{1}{\sqrt{\lambda}}\left(c + \frac{\alpha}{2}\right)\right), \quad c_2 \stackrel{\text{def}}{=} -\frac{2}{2c+\alpha}, \quad c_3 \stackrel{\text{def}}{=} \operatorname{arctanh}\left(-\frac{1}{\sqrt{-\lambda}}\left(c + \frac{\alpha}{2}\right)\right). \tag{208}$$

Proof. Since $W(t)$ satisfies $\frac{\mathrm{d}W(t)}{\mathrm{d}t} \geq W(t)^2 + \alpha W(t) + \beta \cdot \mathrm{Id}_n$ for all $t \in [0,T]$, we get that

$$\begin{aligned} \forall\ t \in [0,T], \qquad \frac{\mathrm{d}\phi(t)}{\mathrm{d}t} &= \left\langle \frac{\mathrm{d}W(t)}{\mathrm{d}t}\theta, \theta \right\rangle \geq \langle W(t)^2\theta, \theta\rangle + \alpha\langle W(t)\theta, \theta\rangle + \beta|\theta|^2 \\ &\geq \langle W(t)\theta, \theta\rangle^2 + \alpha\langle W(t)\theta, \theta\rangle + \beta|\theta|^2 = \phi(t)^2 + \alpha\phi(t) + \beta. \end{aligned} \tag{209}$$

Hence, ϕ satisfies the ordinary differential inequality

$$\forall\ t \in [0,T], \qquad \frac{\mathrm{d}\phi(t)}{\mathrm{d}t} \geq \phi(t)^2 + \alpha\phi(t) + \beta. \tag{210}$$

The solution of the ordinary differential equation

$$\forall\ t \in [0,T], \qquad \frac{\mathrm{d}\sigma(t)}{\mathrm{d}t} = \sigma(t)^2 + \alpha\sigma(t) + \beta, \quad \sigma(0) = c \tag{211}$$

is

$$\sigma(t) = \xi_\lambda(t) - \frac{\alpha}{2}$$

where

$$\xi_\lambda(t) \stackrel{\text{def}}{=} \begin{cases} \sqrt{\lambda}\tan(\sqrt{\lambda}t + c_1) & \text{if } \lambda > 0 \\ -\frac{1}{t+c_2} & \text{if } \lambda = 0 \\ -\sqrt{-\lambda}\tanh(\sqrt{-\lambda}t + c_3) & \text{if } \lambda < 0, \end{cases} \tag{212}$$

with

$$\lambda \stackrel{\text{def}}{=} \beta - \frac{\alpha^2}{4}, \tag{213}$$

and

$$c_1 \stackrel{\text{def}}{=} \arctan\left(\frac{1}{\sqrt{\lambda}}\left(c + \frac{\alpha}{2}\right)\right), \quad c_2 \stackrel{\text{def}}{=} -\frac{2}{2c+\alpha}, \quad c_3 \stackrel{\text{def}}{=} \operatorname{arctanh}\left(-\frac{1}{\sqrt{-\lambda}}\left(c + \frac{\alpha}{2}\right)\right). \tag{214}$$

Applying standard comparison theorems [49] we get that

$$\forall\ t \in [0,T], \qquad \phi(t) \geq \xi_\lambda(t) - \frac{\alpha}{2}$$

□

We are now ready for the proof of Theorem 33. Recall that Lemma 36 showed that the matrix $m(t)$ satisfies (204) with $\alpha = n\kappa$ and $\beta = \kappa\Delta P_T f(x)$. In the following we let $\phi(t) \stackrel{\text{def}}{=} \langle m(t)\theta, \theta\rangle$ for $\theta \in S^{n-1}$. Let us distinguish between the flat and negatively curved cases.

When $\kappa = 0$ we need to show

$$\phi(0) \leq \frac{1}{T} \tag{215}$$

since we can choose θ to be any normalized eigenvector of $-\nabla^2 \log P_T f(x)$. If $\phi(0) \leq 0$ then (215) is trivial so we may assume from now on that $\phi(0) > 0$. As $\kappa = 0$ we have $\alpha = \beta = 0$ so $\lambda = \beta - \frac{\alpha^2}{4} = 0$. Hence, applying (205) we see that

$$\forall\ t \in [0,T], \qquad \phi(t) \geq \frac{\phi(0)}{1 - \phi(0)t}. \tag{216}$$

In particular, (216) implies that the denominator $1 - \phi(0)t$ never vanishes since, otherwise, the right-hand side of (216) is $+\infty$ (as $\phi(0) > 0$) while the left-hand side is finite (as $\phi(t) < +\infty$). The non-vanishing of $1 - \phi(0)t$, together with $\phi(0) > 0$, implies that

$$0 < 1 - t\phi(0) \quad \Longleftrightarrow \quad \phi(0) \leq \frac{1}{t}. \tag{217}$$

Taking $t = T$ establishes (215).

When $\kappa < 0$ we have $\alpha = \kappa c_T$ and $\beta = n\kappa$ so $\lambda = \kappa\Delta P_T f(x) - \frac{n^2\kappa^2}{4}$. If $\frac{4}{n^2 k}\Delta P_T f(x) = 1$, so that $\lambda = 0$, then the argument proceeds as in the case $\kappa = 0$. When $\frac{4}{n^2 k}\Delta P_T f(x) < 1$, so that $\lambda > 0$, applying (205) yields

$$\forall\ t \in [0,T], \qquad \phi(t) \geq \sqrt{\lambda}\tan(\sqrt{\lambda}t + c) - \frac{n\kappa}{2} \tag{218}$$

where

$$c \stackrel{\text{def}}{=} \arctan\left(\frac{1}{\sqrt{\lambda}}\left(\phi(0) + \frac{n\kappa}{2}\right)\right). \tag{219}$$

In particular, as $\phi(t)$ is finite, it follows that $\tan(\sqrt{\lambda}t + c) < +\infty$ for every $t \in [0,T]$. At $t = 0$, $c \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ by (219) and the range of arctan. It follows that, for every $t \in [0,T]$,

$$\frac{\pi}{2} > \sqrt{\lambda}t + c = \sqrt{\lambda}t + \arctan\left(\frac{1}{\sqrt{\lambda}}\left(\phi(0) + \frac{n\kappa}{2}\right)\right). \tag{220}$$

Plugging in $t = T$ into (220) and rearranging yields

$$\phi(0) < \sqrt{\lambda}\tan\left(\frac{\pi}{2} - \sqrt{\lambda}T\right) - \frac{n\kappa}{2} = \sqrt{\lambda}\cot\left(\sqrt{\lambda}T\right) - \frac{n\kappa}{2}. \tag{221}$$

Letting θ to be any normalized eigenvector of $-\nabla^2 \log P_T f(x)$ and recalling the definition of λ yields (192) upon rearrangement. $\square$

Remark 37. *We cannot address in Theorem 33 the regime* $\frac{4}{n^2\kappa}\frac{\Delta P_T f(x)}{P_T f(x)} < 1$ *since the* tanh *function, which will replace in the proof the* tan *function, is well-defined everywhere. It remains to be seen whether this is an artifact of the proof or an inherent obstacle.*

Remark 38. *The proof of Theorem 33 was obtained by using the inequality* (205) *subject to fixing the value of* $\phi(0)$*. Analogously, we can fix the value of* $\phi(T)$ *and use* (205) *(by reversing the time) to get different matrix inequalities. Indeed, our proof of Theorem 40 below makes use of the freedom to choose the initial (terminal) condition.*

10.2. **Intrinsic dimensional local logarithmic Sobolev inequalities in hyperbolic spaces.** In this section we prove local intrinsic dimensional logarithmic Sobolev inequalities for the hyperbolic space, namely, Theorem 6. The inequalities provided by Theorem 40 below will be obtained as a consequence of the differential inequality of Lemma 35 together with the following simple observation:

Lemma 39. *Let* $f : \mathsf{M} \to \mathbb{R}$ *be such that* $\int_M f \, \mathrm{d}P_T\delta_x = 1$ *and let* μ *be the probability measure with* $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x} = f$*. Then,*

$$P_T(f\log f)(x) - P_T f(x)\log P_T f(x) = H(\mu|P_T\delta_x) = -\frac{P_T\Delta f(x)}{2} + \frac{1}{2}\int_0^T \mathrm{tr}[m(t)]\,\mathrm{d}t. \tag{222}$$

Proof. By Theorem 22, (143), and (199),

$$H(\mu|P_T\delta_x) = \frac{1}{2}\int_0^T \mathrm{tr}[v(t)]\,\mathrm{d}t = \frac{1}{2}\int_0^T \mathrm{tr}[m(t)]\,\mathrm{d}t - \frac{1}{2}\int_0^T \mathrm{tr}[u(t)]\,\mathrm{d}t \tag{223}$$

where we recall (198),

$$u(t) \stackrel{\mathrm{def}}{=} e^{-n\kappa(T-t)}J_T + \frac{P_T\Delta f(x)}{n}\cdot \mathrm{g} \tag{224}$$

with $(J_T)_{ij} \stackrel{\mathrm{def}}{=} P_T\nabla^2 f(\Phi_0 e_i, \Phi_0 e_j)(x) - \frac{1}{n}P_T\Delta f(x)\cdot\delta_{ij}$. The proof is complete since $\mathrm{tr}[J_T] = 0$. $\square$

Theorem 40. *Let* (M, g) *be the n-dimensional hyperbolic space with sectional curvature* $\kappa < 0$ *with the associated heat semigroup* $\{P_t\}_{t\geq 0}$*. Fix* $T > 0$*,* $x \in \mathsf{M}$*, a smooth positive function* $f : \mathsf{M} \to \mathbb{R}$ *with* $\int_M f \, \mathrm{d}P_T\delta_x = 1$*, and let* μ *be the probability measure with* $\frac{\mathrm{d}\mu}{\mathrm{d}P_T\delta_x} = f$*. Then, with*

$$\lambda \stackrel{\mathrm{def}}{=} \frac{n^2\kappa^2}{4}\left\{\frac{4}{n^2\kappa}\Delta P_T f(x) - 1\right\}, \quad \alpha_i \stackrel{\mathrm{def}}{=} \begin{cases} \arctan\left(\frac{1}{\sqrt{\lambda}}\left(\sigma_i + \frac{n\kappa}{2}\right)\right) & \text{if } \lambda < 0, \\ -\frac{2}{2\sigma_i + n\kappa} & \text{if } \lambda = 0, \\ \mathrm{arctanh}\left(-\frac{1}{\sqrt{-\lambda}}\left(\sigma_i + \frac{n\kappa}{2}\right)\right) & \text{if } \lambda > 0, \end{cases} \tag{225}$$

we have the local intrinsic dimensional logarithmic Sobolev inequality

$$P_T(f\log f)(x) - P_T f(x)\log P_T f(x) \leq -\frac{P_T\Delta f(x)}{2} - \frac{n^2\kappa T}{2} - \frac{1}{2}\begin{cases} \sum_{i=1}^n \log\left(\frac{\cos(\alpha_i)}{\cos(\sqrt{\lambda}T + \alpha_i)}\right) & \text{if } \lambda > 0 \\ \sum_{i=1}^n \log\left(\frac{\alpha_i}{T + \alpha_i}\right) & \text{if } \lambda = 0 \\ \sum_{i=1}^n \log\left(\frac{\cosh(\alpha_i)}{\cosh(\sqrt{-\lambda}T + \alpha_i)}\right) & \text{if } \lambda < 0 \end{cases} \tag{226}$$

where $\{\sigma_i\}_{i=1}^n$ *are the eigenvalues of* $\mathbb{E}_\mu[-\nabla^2 \log f]$*, and the reverse local intrinsic dimensional logarithmic Sobolev inequality*

$$P_T(f\log f)(x) - P_T f(x)\log P_T f(x) \geq -\frac{P_T \Delta f(x)}{2} - \frac{n^2\kappa T}{2} + \frac{1}{2}\begin{cases} \sum_{i=1}^n \log\left(\frac{\cos(\alpha_i)}{\cos(\sqrt{\lambda}T+\alpha_i)}\right) & \text{if } \lambda > 0 \\ \sum_{i=1}^n \log\left(\frac{\alpha_i}{T+\alpha_i}\right) & \text{if } \lambda = 0 \\ \sum_{i=1}^n \log\left(\frac{\cosh(\alpha_i)}{\cosh(\sqrt{-\lambda}T+\alpha_i)}\right) & \text{if } \lambda < 0 \end{cases} \tag{227}$$

where $\{\sigma_i\}_{i=1}^n$ *are the eigenvalues of* $-\nabla^2 \log P_T f(x)$*.*

Proof. Given any basis $\{\theta_i\}_{i=1}^n$ of $\mathbb{R}^n$ we have

$$\forall\ t \in [0,T], \qquad \mathrm{tr}[m(t)] = \sum_{i=1}^n \phi_i(t) \tag{228}$$

where $\phi_i(t) \stackrel{\text{def}}{=} \langle m(t)\theta_i, \theta_i\rangle$ for $i = 1,\ldots,n$. It follows from Lemma 35 and Lemma 36 that

$$\forall\ t \in [0,T], \qquad \mathrm{tr}[m(t)] \geq \left(\sum_{i=1}^n \xi_{i,\lambda}(t)\right) - \frac{n^2\kappa}{2} \tag{229}$$

where

$$\xi_{i,\lambda}(t) \stackrel{\text{def}}{=} \begin{cases} \sqrt{\lambda}\tan(\sqrt{\lambda}t + c_{i,1}) & \text{if } \lambda > 0 \\ -\frac{1}{t+c_{i,2}} & \text{if } \lambda = 0 \\ -\sqrt{-\lambda}\tanh(\sqrt{-\lambda}t + c_{i,3}) & \text{if } \lambda < 0, \end{cases} \tag{230}$$

with

$$\lambda \stackrel{\text{def}}{=} \kappa \Delta P_T f(x) - \frac{n^2\kappa^2}{4}, \tag{231}$$

and

$$c_{i,1} \stackrel{\text{def}}{=} \arctan\left(\frac{1}{\sqrt{\lambda}}\left(\phi_i(0) + \frac{n\kappa}{2}\right)\right), \quad c_{i,2} \stackrel{\text{def}}{=} -\frac{2}{2\phi_i(0) + n\kappa}, \quad c_{i,3} \stackrel{\text{def}}{=} \operatorname{arctanh}\left(-\frac{1}{\sqrt{-\lambda}}\left(\phi_i(0) + \frac{n\kappa}{2}\right)\right). \tag{232}$$

It follows from Lemma 39 that

$$P_T(f\log f)(x) - P_T f(x)\log P_T f(x) \geq -\frac{P_T \Delta f(x)}{2} - \frac{n^2\kappa T}{2} + \frac{1}{2}\sum_{i=1}^n \int_0^T \xi_{i,\lambda}(t)\,\mathrm{d}t. \tag{233}$$

Hence, taking $\{\theta_i\}_{i=1}^n$ to be the eigenvectors of $m(0) = -\nabla^2 \log P_T f(x)$, and integrating $\{\xi_{i,\lambda}(t)\}$, yields (227).

To prove (226) we define $\tilde{\phi}_i(t) := \phi_i(T-t)$ which satisfies

$$\forall\ t \in [0,T], \qquad \tilde{\phi}(t) \leq \xi_{i,\lambda}(T-t) - \frac{\alpha}{2} \tag{234}$$

where now

$$c_{i,1} \stackrel{\text{def}}{=} \arctan\left(\frac{1}{\sqrt{\lambda}}\left(\phi_i(T) + \frac{n\kappa}{2}\right)\right), \quad c_{i,2} \stackrel{\text{def}}{=} -\frac{2}{2\phi_i(T) + n\kappa}, \quad c_{i,3} \stackrel{\text{def}}{=} \operatorname{arctanh}\left(-\frac{1}{\sqrt{-\lambda}}\left(\phi_i(T) + \frac{n\kappa}{2}\right)\right). \tag{235}$$

The proof now proceeds as in the proof of (227). □

10.3. **Discussion.** We conclude this section by discussing the roles of matrix differential inequalities in our proofs.

Matrix differential inequalities. The master matrix differential inequality (157), which is at the core of all of our proofs, can be expressed either in terms of $v(t)$,

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} \geq v(t)^2 - u(t)v(t) - v(t)u(t) + u(t)^2 + (n-1)\kappa v(t), \tag{236}$$

or in terms of $m(t)$,

$$\frac{\mathrm{d}m(t)}{\mathrm{d}t} \geq m(t)^2 + n\kappa m(t) - \kappa v(t) + \kappa c_T \cdot \mathrm{g}. \tag{237}$$

The inequalities (236) and (237) are equivalent and contain the same information. In particular, in *flat* space forms, where $\kappa = 0$, both inequalities are of the form $\frac{\mathrm{d}W(t)}{\mathrm{d}t} \geq W(t)^2$. In curved spaces, there are two different ways to proceed from (236) and (237):

(1) Omit the term $u(t)^2$ from (236) to get

$$v(t) \geq U(t) \tag{238}$$

where

$$\frac{\mathrm{d}U(t)}{\mathrm{d}t} = U(t)^2 - u(t)U(t) - U(t)u(t) + (n-1)\kappa U(t). \tag{239}$$

The point of omitting $u^2(t)$ is that equation (239) can be solved explicitly, in contrast to the equation resulting if we keep the $u^2(t)$ term.[1]

(2) Omit the term $-\kappa v(t)$ from (237), which can be done only in *negatively* curved space forms to get

$$m(t) \geq U(t) \tag{240}$$

where

$$\frac{\mathrm{d}U(t)}{\mathrm{d}t} = U(t)^2 + n\kappa U(t) + \kappa c_T \cdot \mathrm{g}. \tag{241}$$

Again, the point of omitting $-\kappa v(t)$ is so that (241) can be solved explicitly. Note that in flat spaces, there is no loss in omitting $-\kappa v(t)$.

Matrix vs. trace differential inequalities. The proofs of Theorem 40 and Theorem 25 proceed along similar but different lines. Both proofs start by establishing an inequality of the form

$$\frac{\mathrm{d}W(t)}{\mathrm{d}t} \geq \mathcal{F}(W(t)) \tag{242}$$

for some quadratic functional $\mathcal{F}$. The goal is to bound $\mathrm{tr}[W(t)]$ which can be achieve by two means. Letting $\{U(t)\}$ be the solution to

$$\frac{\mathrm{d}U(t)}{\mathrm{d}t} = \mathcal{F}(U(t)) \tag{243}$$

we could:

(1) Argue that $W(t) \geq U(t)$ and then take the trace on both sides to get

$$\mathrm{tr}[W(t)] \geq \mathrm{tr}[U(t)]. \tag{244}$$

This is the method used to prove Theorem 25 and Theorem 32 (with different functionals $\mathcal{F}$).

[1] If we take the trace in (236) then the equation with the $u(t)^2$ term can be solved explicitly—see the end of section 9. However, if we do so we would get the ambient dimension n rather than the intrinsic dimension.

(2) When $\mathcal{F}$ has scalar (rather than matrix) coefficients, it holds that

$$\langle \mathcal{F}(W(t))\theta, \theta\rangle \geq \mathcal{F}(\langle W(t)\theta, \theta\rangle) \tag{245}$$

with strict inequality unless θ is an eigenvector of $W(t)$. We can then $\{\theta_i\}$ to be any basis and let $\phi_{i,W}(t) := \langle W(t)\theta_i, \theta_i\rangle$, $\phi_{i,U}(t) := \langle U(t)\theta_i, \theta_i\rangle$ so

$$\frac{\mathrm{d}\phi_{i,W}(t)}{\mathrm{d}t} \geq \mathcal{F}(\phi_{i,W}(t)), \quad \frac{\mathrm{d}\phi_{i,U}(t)}{\mathrm{d}t} = \mathcal{F}(\phi_{i,U}(t)), \tag{246}$$

which shows $\phi_{i,W}(t) \geq \phi_{i,U}(t)$. Hence, for any basis $\{\theta_i\}$ we have

$$\mathrm{tr}[W(t)] = \sum_i \phi_{i,W}(t) \geq \sum_i \phi_{i,U}(t) = \mathrm{tr}[U(t)]. \tag{247}$$

This is the method used to prove Theorem 40.

While both methods lead to the inequality

$$\mathrm{tr}[W(t)] \geq \mathrm{tr}[U(t)], \tag{248}$$

the second method is weaker since the inequality $\frac{\mathrm{d}\phi_{i,W}(t)}{\mathrm{d}t} \geq \mathcal{F}(\phi_{i,W}(t))$ is weaker in principle than $\frac{\mathrm{d}W(t)}{\mathrm{d}t} \geq \mathcal{F}(W(t))$ unless θ_i is an eigenvector of $W(t)$. However, for the purpose of proving an inequality for the *trace*, there is no loss since the trace is invariant under rotations so for each t we can introduce a rotation $R(t)$ which takes $\{\theta_i\}$ to the eigenvectors of $W(t)$ or $U(t)$.

References

[1] Efe Aras, Kuan-Yun Lee, Ashwin Pananjady, and Thomas A. Courtade. A family of Bayesian Cramér-Rao bounds, and consequences for log-concave priors. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 2699–2703, 2019.

[2] Thierry Aubin. Problèmes isopérimétriques et espaces de Sobolev. *J. Differential Geometry*, 11(4):573–598, 1976.

[3] Fanny Augeri. A transportation approach to the mean-field approximation. *Probab. Theory Related Fields*, 180(1-2):1–32, 2021.

[4] Dominique Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994.

[5] Dominique Bakry, François Bolley, and Ivan Gentil. Dimension dependent hypercontractivity for Gaussian kernels. *Probab. Theory Related Fields*, 154(3-4):845–874, 2012.

[6] Dominique Bakry, François Bolley, and Ivan Gentil. The Li-Yau inequality and applications under a curvature-dimension condition. *Ann. Inst. Fourier (Grenoble)*, 67(1):397–421, 2017.

[7] Dominique Bakry, Didier Concordet, and Michel Ledoux. Optimal heat kernel bounds under logarithmic Sobolev inequalities. *ESAIM Probab. Statist.*, 1:391–407, 1995/97.

[8] Dominique Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[9] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[10] Dominique Bakry and Michel Ledoux. A logarithmic Sobolev form of the Li-Yau parabolic inequality. *Rev. Mat. Iberoam.*, 22(2):683–702, 2006.

[11] Franck Barthe and Cyril Roberto. Modified logarithmic Sobolev inequalities on $\mathbb{R}$. *Potential Anal.*, 29(2):167–193, 2008.

[12] William Beckner. Inequalities in Fourier analysis. *Ann. of Math. (2)*, 102(1):159–182, 1975.

[13] William Beckner. A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.*, 105(2):397–400, 1989.

[14] Neal Bez, Shohei Nakamura, and Hiroshi Tsuji. Stability of hypercontractivity, the logarithmic Sobolev inequality, and Talagrand's cost inequality. Preprint available at `https://arxiv.org/abs/2201.12478`, 2022.

[15] Sergey G. Bobkov and Michel Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.*, 156(2):347–365, 1998.

[16] Sergey G. Bobkov and Michel Ledoux. From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. *Geom. Funct. Anal.*, 10(5):1028–1052, 2000.

[17] Sergey G. Bobkov and Prasad Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.*, 19(2):289–336, 2006.

[18] Sergey G. Bobkov and Bogusław Zegarlinski. Entropy bounds and isoperimetry. *Mem. Amer. Math. Soc.*, 176(829):x+69, 2005.

[19] Aline Bonami. Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier (Grenoble)*, 20(fasc. 2):335–402 (1971), 1970.
[20] Eric A. Carlen. Superadditivity of Fisher's information and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 101(1):194–211, 1991.
[21] Isaac Chavel. *Eigenvalues in Riemannian geometry*, volume 115 of *Pure and Applied Mathematics*. Academic Press, Inc., Orlando, FL, 1984. Including a chapter by Burton Randol, With an appendix by Jozef Dodziuk.
[22] Dario Cordero-Erausquin, Bruno Nazaret, and Cédric Villani. A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities. *Adv. Math.*, 182(2):307–332, 2004.
[23] Max H. M. Costa and Thomas M. Cover. On the similarity of the entropy power inequality and the Brunn-Minkowski inequality. *IEEE Trans. Inform. Theory*, 30(6):837–839, 1984.
[24] Edward Brian Davies and Barry Simon. Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians. *J. Funct. Anal.*, 59(2):335–395, 1984.
[25] Amédée Debiard, Bernard Gaveau, and Edmond Mazet. Théorèmes de comparaison en géométrie riemannienne. *Publ. Res. Inst. Math. Sci.*, 12(2):391–425, 1976/77.
[26] Manuel Del Pino and Jean Dolbeault. The optimal Euclidean L^p-Sobolev logarithmic inequality. *J. Funct. Anal.*, 197(1):151–161, 2003.
[27] Amir Dembo. Information inequalities and uncertainty principles. Technical report no. 75, *Department of Statistics, Stanford University*, 1990.
[28] Jean Dolbeault and Giuseppe Toscani. Stability results for logarithmic Sobolev and Gagliardo-Nirenberg inequalities. *Int. Math. Res. Not. IMRN*, (2):473–498, 2016.
[29] S. Yu. Efroimovich. Information contained in a sequence of observations. *Problems in Information Transmission*, 15:24–39, 1980.
[30] Ronen Eldan, Joseph Lehec, and Yair Shenfeld. Stability of the logarithmic Sobolev inequality via the Föllmer process. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(3):2253–2269, 2020.
[31] Herbert Federer and Wendell H. Fleming. Normal and integral currents. *Ann. of Math. (2)*, 72:458–520, 1960.
[32] Wendell H. Fleming and Raymond Rishel. An integral formula for total gradient variation. *Arch. Math. (Basel)*, 11:218–222, 1960.
[33] Hans Föllmer. Time reversal on Wiener space. In *Stochastic processes—mathematics and physics (Bielefeld, 1984)*, volume 1158 of *Lecture Notes in Math.*, pages 119–129. Springer, Berlin, 1986.
[34] Emilio Gagliardo. Ulteriori proprietà di alcune classi di funzioni in più variabili. *Ricerche Mat.*, 8:24–51, 1959.
[35] Alexander Grigor'yan and Masakazu Noguchi. The heat kernel on hyperbolic space. *Bull. London Math. Soc.*, 30(6):643–650, 1998.
[36] Leonard Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.
[37] Richard S. Hamilton. A matrix Harnack estimate for the heat equation. *Comm. Anal. Geom.*, 1(1):113–126, 1993.
[38] Elton P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.
[39] Oliver Johnson. A discrete log-Sobolev inequality under a Bakry-Émery type condition. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(4):1952–1970, 2017.
[40] Robert A. Jones. Comparison theorems for matrix Riccati equations. *SIAM J. Appl. Math.*, 29:77–90, 1975.
[41] Herbert Knothe. Contributions to the theory of convex bodies. *Michigan Math. J.*, 4:39–52, 1957.
[42] Rafał Latała and Krzysztof Oleszkiewicz. Between Sobolev and Poincaré. In *Geometric aspects of functional analysis*, volume 1745 of *Lecture Notes in Math.*, pages 147–168. Springer, Berlin, 2000.
[43] Michel Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 120–216. Springer, Berlin, 1999.
[44] Joseph Lehec. Representation formula for the entropy and functional inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(3):885–899, 2013.
[45] Joseph Lehec. Borell's formula on a Riemannian manifold and applications. In *Convexity and concentration*, volume 161 of *IMA Vol. Math. Appl.*, pages 267–284. Springer, New York, 2017.
[46] Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014.
[47] Peter Li and Shing-Tung Yau. On the parabolic kernel of the Schrödinger operator. *Acta Math.*, 156(3-4):153–201, 1986.
[48] Henry P. McKean. An upper bound to the spectrum of Δ on a manifold of negative curvature. *J. Differential Geometry*, 4:359–366, 1970.
[49] M. Michel Petrovitch. Sur une manière d'étendre le théorème de la moyenne aux équations différentielles du premier ordre. *Math. Ann.*, 54(3):417–436, 1901.
[50] Laurent Miclo. Une majoration sous-exponentielle pour la convergence de l'entropie des chaînes de Markov à trou spectral. *Ann. Inst. H. Poincaré Probab. Statist.*, 35(3):261–311, 1999.
[51] Edward Nelson. A quartic interaction in two dimensions. In *Mathematical Theory of Elementary Particles (Proc. Conf., Dedham, Mass., 1965)*, pages 69–73. M.I.T. Press, Cambridge, Mass., 1966.
[52] Edward Nelson. The free Markoff field. *J. Functional Analysis*, 12:211–227, 1973.

[53] Louis Nirenberg. On elliptic partial differential equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, 13:115–162, 1959.
[54] Peter Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*. Springer, Cham, third edition, 2016.
[55] Yury Polyanskiy and Alex Samorodnitsky. Improved log-Sobolev inequalities, hypercontractivity and uncertainty principle on the hypercube. *J. Funct. Anal.*, 277(11):108280, 33, 2019.
[56] Murray Rosenblatt. Remarks on a multivariate transformation. *Ann. Math. Statistics*, 23:470–472, 1952.
[57] Alex Samorodnitsky. A modified logarithmic Sobolev inequality for the Hamming cube and some applications. Preprint available at `https://arxiv.org/abs/0807.1679`, 2008.
[58] Erwin Schrödinger. Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. *Ann. Inst. H. Poincaré*, 2(4):269–310, 1932.
[59] Sergeĭ L. Sobolev. On a theorem in functional analysis. 1938.
[60] Sergeĭ L. Sobolev. *Applications of functional analysis in mathematical physics*. Translations of Mathematical Monographs, Vol. 7. American Mathematical Society, Providence, R.I., 1963. Translated from the Russian by F. E. Browder.
[61] Aart Johannes Stam. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control*, 2:101–112, 1959.
[62] Giorgio Talenti. Best constant in Sobolev inequality. *Ann. Mat. Pura Appl. (4)*, 110:353–372, 1976.
[63] Nicholas Th. Varopoulos. Potential theory and diffusion on Riemannian manifolds. In *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*, Wadsworth Math. Ser., pages 821–837. Wadsworth, Belmont, CA, 1983.
[64] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
[65] Feng-Yu Wang. Identifying constant curvature manifolds, Einstein manifolds, and Ricci parallel manifolds. *J. Geom. Anal.*, 29(3):2374–2409, 2019.
[66] Fred B. Weissler. Logarithmic Sobolev inequalities for the heat-diffusion semigroup. *Trans. Amer. Math. Soc.*, 237:255–269, 1978.
[67] Liming Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields*, 118(3):427–438, 2000.
[68] Chengjie Yu and Feifei Zhao. Sharp Li-Yau-type gradient estimates on hyperbolic spaces. *J. Geom. Anal.*, 30(1):54–68, 2020.

(A. E.) CNRS, Institut de Mathématiques de Jussieu, Sorbonne Université, France and Trinity College, University of Cambridge, UK.
Email address: `alexandros.eskenazis@imj-prg.fr, ae466@cam.ac.uk`

(Y. S.) Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA
Email address: `shenfeld@mit.edu`