ANNALES
DE LA FACULTÉ
DES SCIENCES
DE
TOULOUSE
Mathématiques

Bo'az Klartag and Eli Putterman

Spectral monotonicity under Gaussian convolution

Tome XXXII, nº 5 (2023), p. 939–967.

https://doi.org/10.5802/afst.1759

Publication membre du centre Mersenne pour l'édition scientifique ouverte
http://www.centre-mersenne.org/
e-ISSN : 2258-7519

Annales de la faculté des sciences de Toulouse Volume XXXII, n° 5, 2023
pp. 939-967

Spectral monotonicity under Gaussian convolution [(*)]

Bo'az Klartag [(1)] and Eli Putterman [(2)]

Abstract. — We show that the Poincaré constant of a log-concave measure in Euclidean space is monotone increasing along the heat flow. In fact, the entire spectrum of the associated Laplace operator is monotone decreasing. Two proofs of these results are given. The first proof analyzes a curvature term of a certain time-dependent diffusion, and the second proof constructs a contracting transport map following the approach of Kim and Milman.

Résumé. — Nous montrons que la constante de Poincaré d'une mesure log-concave dans l'espace euclidien est croissante le long du flot de la chaleur. En fait, le spectre entier de l'opérateur de Laplace associé est décroissant. Deux preuves de ces résultats sont données. La première preuve analyse un terme de courbure d'une certaine diffusion dépendant du temps, et la seconde preuve construit une application de transport contractante en suivant l'approche de Kim et Milman.

1. Introduction

The Poincaré constant $C_P(\mu)$ of a Borel probability measure μ on $\mathbb{R}^n$ is the smallest constant $C \geqslant 0$ such that for any locally-Lipschitz function $f \in L^2(\mu)$,

$$\mathrm{Var}_\mu(f) \leqslant C \cdot \int_{\mathbb{R}^n} |\nabla f|^2 \mathrm{d}\mu$$

where $\mathrm{Var}_\mu(f) = \int_{\mathbb{R}^n} f^2 \mathrm{d}\mu - \left(\int_{\mathbb{R}^n} f \mathrm{d}\mu\right)^2$ and $|\cdot|$ is the Euclidean norm. The Poincaré constant governs the rate of convergence to equilibrium of the Langevin dynamics in velocity space [33].

(*) Reçu le 20 juillet 2021, accepté le 3 décembre 2021.

(1) Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel — boaz.klartag@weizmann.ac.il

(2) School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel — putterman@mail.tau.ac.il

Supported by a grant from the Israel Science Foundation (ISF), number 765/19.

Article proposé par Vincent Guedj.

Suppose that μ admits a smooth, positive density ρ in $\mathbb{R}^n$. The Laplace operator associated with μ, defined a priori on smooth, compactly supported functions $u : \mathbb{R}^n \to \mathbb{R}$, is given by

$$Lu = L_\mu u = \Delta u + \nabla(\log \rho) \cdot \nabla u. \tag{1.1}$$

It satisfies

$$\int_{\mathbb{R}^n} (Lu) v \mathrm{d}\mu = -\int_{\mathbb{R}^n} \langle \nabla u, \nabla v \rangle \mathrm{d}\mu$$

for any two smooth functions $u, v : \mathbb{R}^n \to \mathbb{R}$, one of which is compactly supported. The operator L_μ is essentially self-adjoint in $L^2(\mu)$, negative semi-definite, with a simple eigenvalue at 0 corresponding to the constant eigenfunction (see [2, Corollary 3.2.2]). The Poincaré constant is given by

$$C_P(\mu) = 1/\lambda_1^{(\mu)},$$

where $\lambda_1^{(\mu)}$ is the *spectral gap* of L, the infimum over all positive $\lambda > 0$ that belong to the spectrum of $-L$. Under mild regularity assumptions the spectrum of L is discrete (e.g., when ρ is C^2 and $\Delta(\sqrt{\rho})/\sqrt{\rho}$ tends to infinity at infinity [2, Corollary 4.10.9], or when ρ is log-concave and $|\log \rho(x)|/|x|$ tends to infinity at infinity, as shown in Appendix A below). In this case we write

$$0 = \lambda_0^{(\mu)} < \lambda_1^{(\mu)} \leqslant \lambda_2^{(\mu)} \leqslant \lambda_3^{(\mu)} \leqslant \dots$$

for the eigenvalues of $-L$, repeated according to their multiplicity.

A non-negative function ρ on $\mathbb{R}^n$ is log-concave if $K = \{x \in \mathbb{R}^n ; \rho(x) > 0\}$ is convex, and $\log \rho$ is concave in K. An absolutely continuous probability measure on $\mathbb{R}^n$ is called log-concave if it has a log-concave density. An arbitrary probability measure on $\mathbb{R}^n$ is called log-concave if it is the pushforward of some absolutely continuous log-concave probability measure on $\mathbb{R}^k$ under an injective affine map. An example of a log-concave probability measure is γ_s, the Gaussian probability measure on $\mathbb{R}^n$ of mean zero and covariance $s \cdot \mathrm{Id}$. In a minor abuse of notation, we use γ_s to denote also its density function $\gamma_s(x) = (2\pi s)^{-n/2} \exp(-|x|^2/(2s))$. Another example of a log-concave probability measure is the uniform probability measure on any convex body in $\mathbb{R}^n$. The convolution of two log-concave probability measures is again log-concave, as follows from the Prékopa–Leindler inequality [8, Theorem 1.2.3] or from the earlier work by Davidovič, Korenbljum and Hacet [17].

The Poincaré constant is a particularly useful invariant in the class of log-concave probability measures. For example, when μ is absolutely-continuous and log-concave, its Poincaré constant is determined, up to a multiplicative universal constant, by the isoperimetric constant

$$h(\mu) = \inf_{A \subseteq \mathbb{R}^n} \frac{\int_{\partial A} \rho}{\min\{\mu(A), 1 - \mu(A)\}},$$

where the infimum runs over all open sets $A \subseteq \mathbb{R}^n$ with smooth boundary. Indeed, the Cheeger [13] and Buser–Ledoux [9, 28] inequalities state that for any absolutely-continuous, log-concave probability measure μ on $\mathbb{R}^n$,

$$\frac{1}{9} \leqslant C_P(\mu) \cdot h^2(\mu) \leqslant 4.$$

A well-known conjecture by Kannan, Lovász and Simonovits (KLS) states that the Poincaré constant of a log-concave probability measure is equivalent, up to a multiplicative universal constant, to the operator norm of the covariance matrix of μ. See the recent papers by Chen [14] and by Klartag and Lehec [25] for more background and for the best known results towards this conjecture.

Abbreviate $\gamma = \gamma_1$, the standard Gaussian measure in $\mathbb{R}^n$, whose Poincaré constant is $C_P(\gamma) = 1$ (e.g., [2, Proposition 4.1.1]). It was proven by Cattiaux and Guillin [12, Theorem 9.4.3] that when μ is a log-concave probability measure,

$$C_P(\mu) \leqslant C_P(\mu * \gamma) + 1, \tag{1.2}$$

where $\mu * \gamma$ is the convolution of μ and γ. The reverse inequality $C_P(\mu) \geqslant C_P(\mu * \gamma) - 1$ is much easier to obtain and does not require log-concavity (see, e.g., [3, Proposition 1] or [15]). Our main result in this paper is an improvement upon (1.2):

Theorem 1.1. — *Let μ be a log-concave probability measure on $\mathbb{R}^n$. Then,*

$$C_p(\mu) \leqslant C_P(\mu * \gamma). \tag{1.3}$$

Moreover, assuming that μ admits a density that is smooth and positive in $\mathbb{R}^n$ and that L_μ has a discrete spectrum tending to infinity, we have

$$\lambda_k^{(\mu*\gamma)} \leqslant \lambda_k^{(\mu)} \qquad (k = 1, 2, \ldots)$$

We remark that the reverse inequality $1/\lambda_k^{(\mu*\gamma)} \leqslant 1/\lambda_k^{(\mu)} + 1$ is known to hold without any log-concavity assumptions, see e.g. [15]. Two proofs of Theorem 1.1 are presented here. One of these proofs utilizes a method from Kim and Milman [23] to construct a contraction transporting $\mu * \gamma$ to μ. Recall that a map $T : \mathbb{R}^n \to \mathbb{R}^n$ is a contraction if $|T(x) - T(y)| \leqslant |x - y|$ for all $x, y \in \mathbb{R}^n$.

Theorem 1.2. — *Let μ be a log-concave probability measure on $\mathbb{R}^n$. Then there exists a contraction $T : \mathbb{R}^n \to \mathbb{R}^n$ that pushes forward $\mu * \gamma$ to μ.*

This result is reminiscent of Caffarelli's theorem [10], which states that there is a contraction pushing forward γ to μ in the case where the density of μ with respect to the measure γ is log-concave. As is well-known, Theorem 1.2 implies that the Poincaré constant of μ is not larger than that of $\mu * \gamma$. Moreover, as explained e.g. in Ledoux [27, Proposition 1.2], it follows from Theorem 1.2 that when μ is an absolutely-continuous, log-concave probability measure on $\mathbb{R}^n$,

$$h(\mu) \geqslant h(\mu * \gamma). \tag{1.4}$$

There is also a corresponding inequality between the log-Sobolev constants of μ and $\mu * \gamma$, or any other quantity involving a Rayleigh-type quotient, see Caffarelli [10, Corollary 8]. We explain the proof of Theorem 1.2 and its implications in Section 3.

We continue with a discussion of an additional proof of Theorem 1.1, which was chronologically the first proof that we found. For $s > 0$ denote

$$\mu_s = \mu * \gamma_s,$$

the evolution of the measure μ under the heat flow. The log-concavity of μ implies that μ_s is log-concave as well. We will show that $C_P(\mu_s)$ is nondecreasing in s. For a function $f : \mathbb{R}^n \to \mathbb{R}$ we consider its evolution under the heat semigroup

$$P_s(f) = f * \gamma_s \qquad (s > 0)$$

whenever the integrals defining the convolution converge absolutely. Setting $P_0 = \mathrm{Id}$ we obtain the heat semigroup $(P_s)_{s \geqslant 0}$, which satisfies the Euclidean heat equation

$$\frac{\partial}{\partial s} P_s f = \frac{\Delta P_s f}{2} \qquad (s > 0). \tag{1.5}$$

The operator $P_s : L^2(\mu_s) \to L^2(\mu)$ is a contraction operator with $P_s(1) = 1$ since

$$\|P_s(f)\|^2_{L^2(\mu)} = \int_{\mathbb{R}^n} P_s(f)^2 \rho \leqslant \int_{\mathbb{R}^n} P_s(f^2) \rho = \int_{\mathbb{R}^n} f^2 P_s(\rho) = \|f\|^2_{L^2(\mu_s)}.$$

Write $\rho_s = \rho * \gamma_s$ for the density of the probability measure μ_s, which is a smooth positive function in $\mathbb{R}^n$. The adjoint operator $Q_s = P_s^* : L^2(\mu) \to L^2(\mu_s)$ is defined by $Q_0 \varphi = \varphi$ and

$$Q_s \varphi = \frac{P_s(\varphi \rho)}{\rho_s} \qquad (s > 0), \tag{1.6}$$

and again it is a contraction operator with $Q_s(1) = 1$. It follows from (1.5) and (1.6) that the evolution equation for Q_s is the parabolic equation

$$\frac{\partial}{\partial s} Q_s \varphi = \frac{\Delta Q_s \varphi}{2} + \nabla \log \rho_s \cdot \nabla Q_s \varphi. \tag{1.7}$$

We are thus led to define the "box operator"

$$\square_s u = \frac{\Delta u}{2} + \nabla(\log \rho_s) \cdot \nabla u. \tag{1.8}$$

This operator resembles the Laplace operator $L_s := L_{\mu_s}$. Indeed, we have

$$L_s = \square_s + \frac{\Delta}{2}. \tag{1.9}$$

The $\square_s$ operator obeys a Bochner-type formula, which is unsurprising as $\square_s$ equals half of the Laplace operator associated with the log-concave probability measure whose density is proportional to ρ_s^2. Indeed, we compute that for smooth $u, v : \mathbb{R}^n \to \mathbb{R}$,

$$\Gamma_2(u,v) := \square_s \langle \nabla u, \nabla v\rangle - \langle \nabla \square_s u, \nabla v\rangle - \langle \nabla u, \nabla \square_s v\rangle \tag{1.10}$$

satisfies

$$\Gamma_2(u,u) = \|\nabla^2 u\|_{HS}^2 - 2\nabla^2(\log \rho_s)\nabla u \cdot \nabla u, \tag{1.11}$$

where $\|\nabla^2 u\|_{HS}$ is the Hilbert–Schmidt norm of the Hessian matrix $\nabla^2 u$. The expression in (1.11) is similar to the Bochner-type formula of the operator L_s, the main difference being the factor 2 in front of the second summand in (1.11), which is the "curvature term." Moreover, setting $\Gamma_0(u,v) = uv$ and $\Gamma_1(u,v) = \nabla u \cdot \nabla v$, we have

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_{\mathbb{R}^n} \Gamma_i(Q_s\varphi, Q_s\varphi)\mathrm{d}\mu_s = -\int_{\mathbb{R}^n} \Gamma_{i+1}(Q_s\varphi, Q_s\varphi)\mathrm{d}\mu_s \qquad (i = 0, 1),$$

under some regularity assumptions to be explained below. It follows that the Rayleigh quotient

$$\frac{\int_{\mathbb{R}^n} |\nabla Q_s\varphi|^2 \mathrm{d}\mu_s}{\int_{\mathbb{R}^n} (Q_s\varphi)^2 \mathrm{d}\mu_s}$$

is non-increasing in $s \in (0, \infty)$. This fact, formulated as Theorem 2.4 below, implies Theorem 1.1. More details, explanations and rigourous proofs are provided in Section 2.

In Section 4 we discuss conceptual aspects of the evolution $(Q_s\varphi)_{s\geqslant 0}$, and explain how it is equivalent to *Eldan's stochastic localization* [19, 30] and *Föllmer's drift* [32]. We also provide a Bayesian interpretation of this evolution, and explore various connections between these points of view.

2. A dynamic variant of Γ-calculus

In this section we prove Theorem 1.1. Consider the linear differential operator $\square_s$ defined by formula (1.8) above. Similarly to the formalism from [2], for smooth functions $u, v : \mathbb{R}^n \to \mathbb{R}$ we define $\Gamma_0(u,v) = uv$, and for $i \geqslant 0$ and $s > 0$,

$$\Gamma_{i+1}(u,v) = \Box_s \Gamma_i(u,v) - \Gamma_i(u, \Box_s v) - \Gamma_i(\Box_s u, v) - \frac{\mathrm{d}\Gamma_i}{\mathrm{d}s}(u,v). \tag{2.1}$$

Thus $\Gamma_1(u,v) = \nabla u \cdot \nabla v$ and $\Gamma_2(u,v)$ coincides with definition (1.10) above. The rationale for definition (2.1) is that, under regularity assumptions stated below,

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_{\mathbb{R}^n} \Gamma_i(\varphi_s, \varphi_s)\mathrm{d}\mu_s = -\int_{\mathbb{R}^n} \Gamma_{i+1}(\varphi_s,\varphi_s)\mathrm{d}\mu_s. \tag{2.2}$$

where $\varphi_s = Q_s\varphi$. If we were allowed to ignore all regularity issues, (2.2) could be proven as follows: differentiating under the integral sign and applying (1.5) and (1.7),

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s}\int_{\mathbb{R}^n} \Gamma_i(\varphi_s,\varphi_s)\rho_s \\
&= 2\int_{\mathbb{R}^n} \Gamma_i\left(\Box_s\varphi_s, \varphi_s\right)\rho_s + \int_{\mathbb{R}^n} \Gamma_i\left(\varphi_s,\varphi_s\right)\frac{\Delta\rho_s}{2} + \int_{\mathbb{R}^n} \frac{\mathrm{d}\Gamma_i}{\mathrm{d}s}(\varphi_s,\varphi_s)\rho_s \\
&= \int_{\mathbb{R}^n}\left[2\Gamma_i\left(\Box_s\varphi_s,\varphi_s\right) + \frac{\Delta\Gamma_i\left(\varphi_s,\varphi_s\right)}{2}\right]\mathrm{d}\mu_s + \int_{\mathbb{R}^n}\frac{\mathrm{d}\Gamma_i}{\mathrm{d}s}(\varphi_s,\varphi_s)\mathrm{d}\mu_s.
\end{aligned} \tag{2.3}$$

Next we use (1.9) and the fact that $\int_{\mathbb{R}^n}(L_s u)\mathrm{d}\mu_s = 0$ under regularity assumptions (e.g., when u is smooth and compactly supported). This yields

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s}\int_{\mathbb{R}^n} \Gamma_i(\varphi_s,\varphi_s)\rho_s \\
&\quad= \int_{\mathbb{R}^n}\left[2\Gamma_i\left(\Box_s\varphi_s,\varphi_s\right) - \Box_s\Gamma_i\left(\varphi_s,\varphi_s\right)\right]\mathrm{d}\mu_s + \int_{\mathbb{R}^n}\frac{\mathrm{d}\Gamma_i}{\mathrm{d}s}(\varphi_s,\varphi_s)\mathrm{d}\mu_s \\
&\quad= -\int_{\mathbb{R}^n}\Gamma_{i+1}(\varphi_s,\varphi_s)\mathrm{d}\mu_s.
\end{aligned} \tag{2.4}$$

This would be a rigorous proof for (2.2) had we worked in the context of a compact Riemannian manifold (which also has a heat kernel $P_s : L^2(\mu) \to L^2(\mu_s)$ and a corresponding adjoint $Q_s = P_s^*$). However, in this paper we are interested in the non-compact situation of $\mathbb{R}^n$, since we rely on the fact that the heat flow preserves curvature conditions such as log-concavity, which is currently known to hold only for a Euclidean space [26]. Nevertheless, the operators $\Box_s$ and $(Q_s)_{s\geqslant 0}$ seem rather natural also in the Riemannian setting. We remark that our formalism is rather similar to the inhomogeneous Bakry-Émery method recently introduced by Bodineau and Bauerschmidt [5] in a different context.

Our first task in this section is to rigorously justify (2.2) for a fairly large class of functions φ. To do this, we shall express Q_s explicitly as an integral operator.

Recall that we work with an absolutely-continuous, log-concave probability measure μ on $\mathbb{R}^n$ having density ρ. As before, for $s > 0$ we write $\mu_s = \mu * \gamma_s$ and $\rho_s = P_s\rho = \rho * \gamma_s$, while the operator Q_s is defined via formula (1.6). For $s > 0$ and $y \in \mathbb{R}^n$ we define the probability density

$$p_{s,y}(x) = e^{\frac{\langle y,x\rangle}{s} - \frac{|x|^2}{2s}} \frac{\rho(x)}{Z_{s,y}} \tag{2.5}$$

where

$$Z_{s,y} = \int_{\mathbb{R}^n} e^{\frac{\langle y,x\rangle}{s} - \frac{|x|^2}{2s}} \rho(x)\mathrm{d}x = \frac{\rho_s(y)}{\gamma_s(y)} \tag{2.6}$$

is a normalizing constant (the "partition function"). In the next lemma we express the value of $Q_s\varphi$ at the point y as the average of φ with respect to the density $p_{s,y}$.

Lemma 2.1. — *Let $y \in \mathbb{R}^n, s > 0$ and suppose that $\varphi \in L^1(\mu)$, or more generally, that $\varphi : \mathbb{R}^n \to \mathbb{R}$ is such that $\varphi(x)e^{-t|x|^2} \in L^1(\mu)$ for some $t \in (0, s)$. Then*

$$Q_s\varphi(y) = \int_{\mathbb{R}^n} \varphi \cdot p_{s,y}. \tag{2.7}$$

Furthermore, $Q_s\varphi(y)$ is a smooth function of $y \in \mathbb{R}^n$ and $s > 0$ which satisfies

$$\nabla Q_s\varphi = \frac{Q_s(x\varphi) - Q_s(\varphi)Q_s(x)}{s}, \tag{2.8}$$

$$\partial_s Q_s\varphi(y) = -\frac{Q_s(f_y\varphi) - Q_s(\varphi)Q_s(f_y)}{s^2} = \Box_s Q_s\varphi(y), \tag{2.9}$$

where $f_y(x) = \langle x, y\rangle - |x|^2/2$.

Proof. — According to (2.5) and (2.6),

$$\begin{aligned}
\rho_s(y) \cdot &\int_{\mathbb{R}^n} \varphi \cdot p_{s,y} \\
&= \int_{\mathbb{R}^n} \varphi(x) \cdot e^{\langle y,x\rangle/s - |x|^2/(2s)}\rho(x)\mathrm{d}x \cdot \gamma_s(y) \\
&= \int_{\mathbb{R}^n} \varphi(x)\rho(x) \cdot \gamma_s(y - x)\mathrm{d}x = [(\varphi\rho) * \gamma_s](y) = P_s(\varphi\rho).
\end{aligned} \tag{2.10}$$

Now (2.7) follows from (1.6) and (2.10). The smoothness of $Q_s\varphi$ and equations (2.8) and (2.9) follow by differentiating (2.6) and (2.7) under the integral sign. This is legitimate, since any partial derivative in the (s, y)-variables of the function $p_{s,y}(x)\varphi(x)$ is seen to be bounded by an integrable function, and the bound is locally uniform in s and y. $\square$

By a multi-index (k, α) we mean a non-negative integer k and a vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers. For a multi-index (k, α) and for a smooth function $f(s, y)$ we abbreviate

$$\partial_s^k \partial_y^\alpha f = \left(\frac{\partial}{\partial s}\right)^k \left(\frac{\partial}{\partial y_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial y_n}\right)^{\alpha_n} f(s, y) \qquad (s > 0, y \in \mathbb{R}^n).$$

We denote $|\alpha| = \sum_i |\alpha_i|$. We say that a measurable function $\varphi : \mathbb{R}^n \to \mathbb{R}$ has *subexponential decay relative to* ρ if there exists $a, C > 0$ such that $|\varphi(x)| \leqslant \frac{C}{\sqrt{\rho(x)}} e^{-a|x|}$ for all $x \in \mathbb{R}^n$ for which $\rho(x) > 0$. Since ρ is an integrable, log-concave function, there exist $\alpha, \beta > 0$ such that

$$\rho(x) \leqslant \alpha e^{-\beta|x|} \qquad \text{for all } x \in \mathbb{R}^n. \tag{2.11}$$

(see e.g., [8, Lemma 2.2.1]). In particular, a function φ that grows at most polynomially at infinity, has subexponential decay relative to ρ.

Lemma 2.2. — *Fix $s > 0$ and let (k, α) be a multi-index. Then,*

(i) *The function $\partial_s^k \partial_y^\alpha \log \rho_s(y)$ grows at most polynomially at infinity in $y \in \mathbb{R}^n$.*
(ii) *Let φ have subexponential decay relative to ρ. Then the function $\partial_s^k \partial_y^\alpha Q_s \varphi(y)$ has subexponential decay relative to ρ_s.*

Moreover, if s varies in an interval $[s_0, s_1]$ with $s_0 > 0$, then the implied constants in these two assertions may be chosen not to depend on s.

Proof. — Using the heat equation, we can replace time derivatives of ρ_s by space derivatives, so in (i) we only need consider space derivatives of $\log \rho_s$. By differentiating (2.6) with respect to y we see that

$$\nabla_y \log \rho_s(y) = -\frac{y}{s} + \nabla_y \log Z_{s,y} = \frac{Q_s(x)(y) - y}{s}.$$

Repeated differentiations show that conclusion (i) would follow once we prove the following claim: for any $d > 0$, the function $Q_s(|x|^d)(y)$ grows at most polynomially at infinity as a function of $y \in \mathbb{R}^n$, with the implied constants not depending on $s \in [s_0, s_1]$.

Let us prove this claim. It follows from (2.11) that $C_d := \int_{\mathbb{R}^n} (1 + |x|)^d \rho(x) \mathrm{d}x < \infty$. Therefore, for any $y \in \mathbb{R}^n$ and $s > 0$,

$$\frac{\int_{\mathbb{R}^n} |x - y|^d e^{-|x-y|^2/(2s)} \rho(x) \mathrm{d}x}{\int_{\mathbb{R}^n} e^{-|x-y|^2/(2s)} \rho(x) \mathrm{d}x} \leqslant \frac{\int_{\mathbb{R}^n} |x - y|^d \rho(x) \mathrm{d}x}{\int_{\mathbb{R}^n} \rho(x) \mathrm{d}x} \leqslant C_d (1 + |y|)^d,$$

where the first inequality follows from the fact that $|x - y|^d$ is increasing in $|x - y|$ while $e^{-|x-y|^2/(2s)}$ is decreasing in $|x - y|$, and the second inequality uses $|x-y| \leqslant (1+|x|)(1+|y|)$. Consequently, as $|x|^d \leqslant 2^d(|y|^d + |x-y|^d)$,

$$Q_s(|x|^d)(y) = \int_{\mathbb{R}^n} |x|^d p_{s,y}(x)\mathrm{d}x = \frac{\int_{\mathbb{R}^n} |x|^d e^{-|x-y|^2/(2s)}\rho(x)\mathrm{d}x}{\int_{\mathbb{R}^n} e^{-|x-y|^2/(2s)}\rho(x)\mathrm{d}x} \leqslant \widetilde{C}_d(1+|y|)^d,$$

for some coefficient $\widetilde{C}_d$ depending only on ρ and on d. This shows that $Q_s(|x|^d)(y)$ grows at most polynomially, from which (i) follows.

We move on to the proof of (ii). Let $a, C > 0$ be such that $\sqrt{\rho(x)}\cdot|\varphi(x)| \leqslant Ce^{-a|x|}$ for all x. From (1.6),

$$|\varphi_s| \leqslant \frac{C}{P_s(\rho)} P_s(\sqrt{\rho(x)}e^{-a|x|}) \leqslant \frac{C}{P_s(\rho)} \cdot \sqrt{P_s(\rho)P_s(e^{-2a|x|})} = \frac{C}{\sqrt{\rho_s}} P_s(e^{-2a|x|})^{\frac{1}{2}} \tag{2.12}$$

where we have used the Cauchy–Schwarz inequality for P_s. In order to conclude that φ_s has subexponential decay relative to ρ_s, it remains only to note the following: since P_s is convolution with a Gaussian of covariance $s \cdot \mathrm{Id}$, there exists $\widetilde{C} = \widetilde{C}_{a,s,n} > 0$ such that $P_s(e^{-2a|x|})(y) \leqslant \widetilde{C}e^{-2a|y|}$ for all $y \in \mathbb{R}^n$. (The constant $\widetilde{C} = \sup_{s\in[s_0,s_1]} \int_{\mathbb{R}^n} e^{2a|x|}\gamma_s(x)\mathrm{d}x$ works for all $s \in [s_0, s_1]$.)

We still need to bound the partial derivatives of $\varphi_s(y)$ with respect to the s-variable and y-variables. The first-order derivatives are given by formulas (2.8) and (2.9), and higher-order derivatives may be computed by repeated applications of these two formulas. Thus $\partial_s^k \partial_y^\alpha Q_s\varphi(y)$ can be expressed as a sum with a fixed number of summands. Each of these summands is a product of a term of the form $\frac{1}{s^m}Q_s(f\varphi)$, where f is a polynomial of degree bounded by $2k + |\alpha|$, and terms of the form $Q_s(p)$ with p a polynomial in the space variables. For any such p, the function $Q_s(p)$ grows at most polynomially because $Q_s(|x|^d)$ does for all d. In addition, $f\varphi$ has subexponential decay relative to ρ, so by the previous part of the proof, $Q_s(f\varphi)$ has subexponential decay relative to ρ_s. Consequently, each of the summands in $\partial_s^k \partial_y^\alpha Q_s\varphi(y)$ has subexponential decay relative to ρ_s, so $\partial_s^k \partial_y^\alpha Q_s\varphi(y)$ does as well. □

Recall the definition (2.1) of $\Gamma_i(u, v)$. In the next proposition we rigorously justify the computations in (2.3) and (2.4). We discuss only the case $i = 0, 1$; while the extension to higher-order carrés des champs presents no particular difficulty, it has been omitted as it is unnecessary for our purposes.

Proposition 2.3. — *Fix $s > 0$ and $i = 0, 1$. Suppose that φ has subexponential decay relative to ρ and $\varphi_s = Q_s(\varphi)$. Then equation (2.2) holds for φ_s.*

Proof. — Recall that $\Gamma_0(u, v) = uv$, $\Gamma_1(u, v) = \nabla u \cdot \nabla v$ and that $\varphi_s(y)$ is smooth in (s, y). Therefore,

$$\partial_s \left[\Gamma_i(\varphi_s, \varphi_s)\rho_s\right] = 2\Gamma_i(\partial_s\varphi_s, \varphi_s)\rho_s + \Gamma_i(\varphi_s, \varphi_s)\rho_s \cdot \partial_s(\log \rho_s). \tag{2.13}$$

According to Lemma 2.2 we may bound the expression in (2.13) by the integrable function $Ce^{-a|y|}$ for some $C, a > 0$, and the bound is locally uniform in s. This justifies interchanging differentiation and integration to obtain

$$\begin{aligned}\partial_s \int_{\mathbb{R}^n} \Gamma_i(\varphi_s, \varphi_s)\rho_s &= \int_{\mathbb{R}^n} \partial_s \left[\Gamma_i(\varphi_s, \varphi_s)\rho_s\right] \\ &= 2\int_{\mathbb{R}^n} \Gamma_i\left(\Box_s\varphi_s, \varphi_s\right)\rho_s + \int_{\mathbb{R}^n} \Gamma_i\left(\varphi_s, \varphi_s\right)\frac{\Delta\rho_s}{2},\end{aligned}$$

where we have used Lemma 2.1 and the heat equation (1.5). Next we need to carry out the integrations by parts of (2.3) and (2.4) and show that no boundary terms arise. When integrating the term $\Gamma_i(\varphi_s, \varphi_s)\Delta\rho_s/2$ by parts twice, we encounter the boundary integrands $\Gamma_i(\varphi_s, \varphi_s)\nabla\rho_s$ and $\nabla\Gamma_i(\varphi_s, \varphi_s)\cdot \rho_s$. Both of these decay exponentially at infinity, so the integration by parts over $\mathbb{R}^n$ introduces no boundary terms, verifying (2.3). In (2.4), we use the integration by parts formula

$$\int_{\mathbb{R}^n} L_s\Gamma_i(\varphi_s, \varphi_s) \cdot \rho_s = \int_{\mathbb{R}^n} \operatorname{div}(\rho_s \nabla\Gamma_i(\varphi_s, \varphi_s)) = 0,$$

which is again justified by the exponential decay of $\nabla\Gamma_i(\varphi_s, \varphi_s)\cdot\rho_s$ at infinity. This completes the proof of (2.2). □

We write $H^1(\mu)$ for the space of all functions in $L^2(\mu)$ whose weak derivatives belong to $L^2(\mu)$, equipped with the norm

$$\|f\|_{H^1(\mu)} = \sqrt{\int_{\mathbb{R}^n} f^2 \mathrm{d}\mu + \int_{\mathbb{R}^n} |\nabla f|^2 \mathrm{d}\mu}.$$

See e.g. the appendix of [4] and the references therein for information about weak derivatives, the Sobolev space $H^1(\mu)$, and for a proof of the fact that the space of smooth, compactly supported functions in $\mathbb{R}^n$ is dense in $H^1(\mu)$.

Theorem 2.4. — *Let μ be an absolutely-continuous, log-concave probability measure on $\mathbb{R}^n$ and let $0 \not\equiv \varphi \in H^1(\mu)$. Then with $\varphi_s = Q_s\varphi$, the Rayleigh quotient*

$$R_\varphi(s) = \frac{\int_{\mathbb{R}^n} |\nabla\varphi_s|^2 \mathrm{d}\mu_s}{\int_{\mathbb{R}^n} \varphi_s^2 \mathrm{d}\mu_s} \tag{2.14}$$

is non-increasing in $s \in [0, \infty)$. *Consequently, the function* $\log \|\varphi_s\|_{L^2(\mu_s)}$ *is convex in* $s \in [0, \infty)$.

For the proof of Theorem 2.4 we require the following technical lemma:

Lemma 2.5.

(i) *For any fixed* $s \geqslant 0$, *the quantities* $R_\varphi(s)$ *and* $\|\varphi_s\|_{L^2(\mu_s)}$ *depend continuously on* $\varphi \in H^1(\mu) \setminus \{0\}$.

(ii) *For any* $0 \not\equiv \varphi \in H^1(\mu)$,

$$R_\varphi(0) = \lim_{s \to 0^+} R_\varphi(s) \quad \textit{and} \quad \|\varphi\|_{L^2(\mu)} = \lim_{s \to 0^+} \|\varphi_s\|_{L^2(\mu_s)}. \tag{2.15}$$

Proof. — We first prove part (i). Let $\varphi \in L^2(\mu) \setminus \{0\}$. By Cauchy–Schwarz, we have the pointwise bound

$$\varphi_s^2 \rho_s = \frac{P_s^2(\rho\varphi)}{P_s(\rho)} \leqslant \frac{P_s(\varphi^2\rho)P_s(\rho)}{P_s(\rho)} = P_s(\varphi^2\rho). \tag{2.16}$$

Moreover, the log-concavity of μ implies that whenever $\varphi \in H^1(\mu)$,

$$|\nabla\varphi_s|^2 \rho_s \leqslant P_s(|\nabla\varphi|^2\rho). \tag{2.17}$$

Indeed, the probability density $p_{s,y}$ from (2.5) is "more log-concave than γ_s," in the sense that $p_{s,y}/\gamma_s$ is log-concave. The Brascamp–Lieb inequality (see e.g., [2, §4.9]) thus implies that the Poincaré constant of the probability density $p_{s,y}$ is at most s. That is, letting X be a random vector with density $p_{s,y}$ and f a weakly differentiable function with $\mathbb{E}|f(X)|^2 < \infty$ and $\mathbb{E}|\nabla f(X)|^2 < \infty$,

$$\operatorname{Var} f(X) \leqslant s \cdot \mathbb{E}|\nabla f(X)|^2. \tag{2.18}$$

Hence, by Lemma 2.1, for any $\theta \in S^{n-1} = \{x \in \mathbb{R}^n\,;\, |x| = 1\}$,

$$\begin{aligned}\nabla\varphi_s \cdot \theta &= \frac{\mathbb{E}(X\cdot\theta)\varphi(X) - \mathbb{E}(X\cdot\theta)\mathbb{E}\varphi(X)}{s} \\ &\leqslant \frac{\sqrt{\operatorname{Var}(X\cdot\theta)\operatorname{Var}(\varphi(X))}}{s} \leqslant \sqrt{\mathbb{E}|\nabla\varphi(X)|^2},\end{aligned}$$

which implies (2.17) since $\mathbb{E}|\nabla\varphi(X)|^2 = Q_s(|\nabla\varphi|^2) = P_s(|\nabla\varphi|^2\rho)/\rho_s$. By integrating over $\mathbb{R}^n$, the inequalities (2.16) and (2.17) imply that

$$\begin{aligned}&\|\varphi_s\|_{L^2(\mu_s)} \leqslant \|\varphi\|_{L^2(\mu)} && \text{for } \varphi \in L^2(\mu),\\ &\|\nabla\varphi_s\|_{L^2(\mu_s)} = \sqrt{\int_{\mathbb{R}^n} |\nabla\varphi_s|^2 \mathrm{d}\mu_s} \leqslant \|\varphi\|_{H^1(\mu)} && \text{for } \varphi \in H^1(\mu).\end{aligned}$$

Consequently, the functional $\varphi \mapsto \|\varphi_s\|_{L^2(\mu_s)}$ is 1-Lipschitz in $L^2(\mu)$, while the functional $\varphi \mapsto \|\nabla\varphi_s\|_{L^2(\mu_s)}$ is 1-Lipschitz in $H^1(\mu)$. In particular, for any fixed $s \geqslant 0$, the quantities $R_\varphi(s)$ and $\|\varphi_s\|_{L^2(\mu_s)}$ depend continuously on $\varphi \in H^1(\mu) \setminus \{0\}$, proving (i).

For part (ii), note that $\varphi \mapsto R_\varphi(s)$ is locally uniformly continuous in $H^1(\mu)\setminus\{0\}$, being the quotient of two positive, 1-Lipschitz functions. Hence, it suffices to prove (2.15) for φ in a dense subset of $H^1(\mu)\setminus\{0\}$. We may thus assume that φ is smooth and compactly supported. We claim that for almost every $y \in \mathbb{R}^n$,

$$\varphi_s^2(y)\rho_s(y) \xrightarrow{s\to 0} \varphi^2(y)\rho(y) \quad \text{and} \quad |\nabla\varphi_s(y)|^2\rho_s(y) \xrightarrow{s\to\infty} |\nabla\varphi(y)|^2\rho(y). \tag{2.19}$$

Let $K = \{x \in \mathbb{R}^n\,;\, \rho(x) > 0\}$. In proving (2.19), we may thus assume that $y \notin \partial K$, since the boundary of the convex set K has Lebesgue measure zero. If $y \notin \overline{K}$ then ρ vanishes in a neighborhood of y, hence

$$P_s(\varphi^2\rho)(y) \xrightarrow{s\to 0} 0 \qquad \text{and} \qquad P_s(|\nabla\varphi|^2\rho)(y) \xrightarrow{s\to 0} 0,$$

and (2.19) follows from the bounds (2.16) and (2.17). As ρ is log-concave, it is locally Lipschitz on K, so by the Rademacher theorem, ρ is differentiable almost everywhere in the interior of K. It thus suffices to prove (2.19) for $y \in K$ such that ρ is differentiable at y. Differentiating φ_s yields

$$\nabla\varphi_s = \nabla\left(\frac{P_s(\varphi\rho)}{P_s(\rho)}\right) = \frac{\nabla P_s(\varphi\rho)}{P_s(\rho)} - \frac{P_s(\varphi\rho)\nabla\rho_s}{\rho_s^2} \qquad (s > 0). \tag{2.20}$$

It is a property of the heat semigroup that if f is a bounded measurable function differentiable at a point $y \in \mathbb{R}^n$, then $\nabla P_s(f)(y) \to \nabla f(y)$ as $s \to 0$; this is easily shown by writing $\nabla P_s f = f * \nabla\gamma_s$ and approximating f by its first-order Taylor polynomial. Applying this to the functions ρ and $\varphi\rho$ which are bounded in $\mathbb{R}^n$ and differentiable at y, we obtain $\nabla\rho_s(y) \to \nabla\rho(y)$ and $\nabla P_s(\varphi\rho)(y) \to \nabla(\varphi\rho)(y)$ as $s \to 0$. Moreover, $\varphi_s(y) \to \varphi(y), \rho_s(y) \to \rho(y)$ because ρ and $\varphi\rho$ are continuous at y and bounded in $\mathbb{R}^n$. It follows that $\nabla\varphi_s(y) \to \nabla\varphi(y)$, completing the proof of (2.19).

Finally, the functions $|\nabla\varphi|^2\rho$ and $\varphi^2\rho$ are bounded and compactly supported. Hence, for $s \in (0,1]$ the Gaussian convolutions $P_s(|\nabla\varphi|^2\rho)$ and $P_s(\varphi^2\rho)$ are bounded by $Ce^{-|y|^2/2}$ in $\mathbb{R}^n$ for some constant C that does not depend on s. From (2.16), (2.17), the dominated convergence theorem, and (2.19), we obtain

$$\int_{\mathbb{R}^n} \varphi_s^2\rho_s \xrightarrow{s\to 0} \int_{\mathbb{R}^n} \varphi^2\rho \qquad \text{and} \qquad \int_{\mathbb{R}^n} |\nabla\varphi_s|^2\rho_s \xrightarrow{s\to 0} |\nabla\varphi|^2\rho,$$

completing the proof of (2.15). □

Remark 2.6. — Inequality (2.17) states that $|\nabla\varphi_s|^2 \leqslant Q_s(|\nabla\varphi|^2)$. Using the interpretation in [37, §3.1] and arguing as in [37, §3.2], one may prove the stronger gradient bound

$$|\nabla\varphi_s| \leqslant Q_s(|\nabla\varphi|),$$

which we do not need here.

It follows from (2.1) and a straightforward computation that

$$\Gamma_2(u,u) = \Box_s |\nabla u|^2 - 2\langle \nabla \Box_s u, \nabla u\rangle \\ = \|\nabla^2 u\|_{HS}^2 - 2\langle \nabla^2(\log\rho_s)\nabla u, \nabla u\rangle. \quad (2.21)$$

On the other hand, the Bochner formula for the differential operator $L_s = L_{\mu_s}$ states that for any smooth, compactly supported function $u : \mathbb{R}^n \to \mathbb{R}$,

$$\int_{\mathbb{R}^n} (-L_s)^2 u \cdot u \mathrm{d}\mu_s = \int_{\mathbb{R}^n} (L_s u)^2 \mathrm{d}\mu_s \\ = \int_{\mathbb{R}^n} \left[\|\nabla^2 u\|_{HS}^2 - \langle \nabla^2(\log\rho_s)\nabla u, \nabla u\rangle\right] \mathrm{d}\mu_s. \quad (2.22)$$

See [2, §1.16.1] for a proof of (2.22). Formula (2.22) remains valid when u and its partial derivatives are smooth functions with subexponential decay relative to ρ_s, since the integration by parts yield no boundary terms as in the proof of Proposition 2.3. Thanks to Lemma 2.2, we know that formula (2.22) is valid for $u = Q_s\varphi$ whenever φ has subexponential decay relative to ρ.

The integrand on the right-hand side of (2.22) is almost identical to the expression in (2.21), the only difference is the coefficient 2 in front of the second summand.

Proof of Theorem 2.4. — When u is a smooth function such that u and its partial derivatives have subexponential decay relative to ρ_s, we write for $i = 1, 2$,

$$\|u\|_{\dot{H}^i(\mu_s)} = \sqrt{\int_{\mathbb{R}^n} (-L_s)^i u \cdot u \mathrm{d}\mu}.$$

Thus $\|u\|_{\dot{H}^1(\mu_s)}^2 = \int_{\mathbb{R}^n} |\nabla u|^2 \mathrm{d}\mu_s$. The operator L_s is initially defined by the formula $L_s u = \Delta u + \nabla \log\rho_s \cdot \nabla u$ assuming u and its partial derivatives have subexponential decay relative to ρ_s. This operator is essentially self-adjoint and negative semi-definite in $L^2(\mu_s)$ (e.g., [2, Corollary 3.2.2]). Hence, by the spectral theorem and the Cauchy–Schwarz inequality,

$$\|u\|_{\dot{H}^1(\mu_s)}^2 \leqslant \|u\|_{\dot{H}^2(\mu_s)} \cdot \|u\|_{L^2(\mu_s)}. \quad (2.23)$$

Consider first the case where $0 \not\equiv \varphi \in H^1(\mu)$ has subexponential decay relative to ρ and $s > 0$. Thanks to Proposition 2.3 we may apply (2.2) and compute that

$$\frac{\mathrm{d}}{\mathrm{d}s} R_\varphi(s) = \frac{\|\varphi_s\|_{\dot{H}^1(\mu_s)}^4 - \int_{\mathbb{R}^n} \Gamma_2(\varphi_s,\varphi_s)\mathrm{d}\mu_s \cdot \|\varphi_s\|_{L^2(\mu_s)}^2}{\|\varphi_s\|_{L^2(\mu_s)}^4}.$$

Hence, from (2.21) and (2.22),

$$\frac{\mathrm{d}}{\mathrm{d}s}R_\varphi(s) = \frac{\|\varphi_s\|^4_{\dot{H}^1(\mu_s)} - \|\varphi_s\|^2_{\dot{H}^2(\mu_s)}\|\varphi_s\|^2_{L^2(\mu_s)}}{\|\varphi_s\|^4_{L^2(\mu_s)}} + \frac{\int_{\mathbb{R}^n}\langle(\nabla^2\log\rho_s)\nabla\varphi_s,\nabla\varphi_s\rangle\mathrm{d}\mu_s}{\|\varphi_s\|^2_{L^2(\mu_s)}}.$$

By log-concavity $\nabla^2\log\rho_s \leqslant 0$. Hence we conclude from (2.23) that

$$\frac{\mathrm{d}}{\mathrm{d}s}R_\varphi(s) \leqslant 0.$$

Therefore $R_\varphi(s)$ is non-increasing in $s\in(0,\infty)$. It follows from (2.2) that

$$\partial_s \log\|\varphi\|_{L^2(\mu_s)} = -R_\varphi(s),$$

and consequently $\log\|\varphi\|_{L^2(\mu_s)}$ is convex in $s\in(0,\infty)$. Lemma 2.5 now implies that $R_\varphi(s)$ is decreasing in $s\in[0,\infty)$ and $\log\|\varphi_s\|_{L^2(\mu_s)}$ is convex in $s\in[0,\infty)$.

Finally, compactly supported smooth functions, which certainly have subexponential decay relative to ρ, are dense in $H^1(\mu)$. The Rayleigh quotient and $\|\varphi_s\|_{L^2(\mu_s)}$ are continuous on $H^1(\mu)\setminus\{0\}$ by Lemma 2.5, hence we obtain that $R_\varphi(s)$ is non-increasing and $\log\|\varphi_s\|_{L^2(\mu_s)}$ is convex in $s\in[0,\infty)$ for any $0\not\equiv\varphi\in H^1(\mu)$. $\square$

Using the min-max characterization of eigenvalues, we derive our main result as a corollary to Theorem 2.4.

Proof of Theorem 1.1. — We may set $s=1$, since $\mu_s=\mu*\gamma$ for $s=1$. We may assume that μ is absolutely continuous, as otherwise we may pass to a lower dimension thanks to the well-known fact that the Poincaré constant of a Cartesian product of two measures is the maximum of the Poincaré constants of the factors. The Poincaré constant of μ, which is finite and positive (see [6]), satisfies

$$\frac{1}{C_P(\mu)} = \inf\left\{R_\varphi(0)\,;\,0\not\equiv\varphi\in H^1(\mu),\int_{\mathbb{R}^n}\varphi\,\mathrm{d}\mu=0\right\},$$

and similarly for μ_s. For any $\varepsilon>0$ there exists $0\not\equiv\varphi\in H^1(\mu)$ with $\int\varphi\mathrm{d}\mu=0$ such that $R_\varphi(0)<C_P(\mu)^{-1}+\varepsilon$. Since $\int\varphi_s\mathrm{d}\mu_s=\int\varphi\mathrm{d}\mu=0$, we deduce from Theorem 2.4 that,

$$\frac{1}{C_P(\mu_s)} \leqslant R_\varphi(s)\leqslant R_\varphi(0) < \frac{1}{C_P(\mu)}+\varepsilon.$$

As $\varepsilon>0$ was arbitrary, inequality (1.3) is proven.

Next, assume that L_μ has discrete spectrum, and let $k\geqslant 1$. There exists a $(k+1)$-dimensional subspace $E\subseteq H^1(\mu)$ such that $R_\varphi(0)\leqslant\lambda_k^{(\mu)}$ for any

$0 \not\equiv \varphi \in E$. For $s > 0$ the linear operator Q_s defined in (1.6) is one-to-one in $L^1(\mu)$. (Indeed, given $P_s(\varphi\rho)$ we may recover the Fourier transform of $\varphi\rho \in L^1(\mathbb{R}^n)$ which determines $\varphi \in L^1(\mu)$.) Hence

$$E_s = \{Q_s\varphi\,;\, \varphi \in E\}$$

is a $(k+1)$-dimensional subspace, and $R_\varphi(s) \leqslant R_\varphi(0) \leqslant \lambda_k^{(\mu)}$ for all $\varphi \in E$. In other words, there exists a $(k+1)$-dimensional subspace $E_s \subseteq H^1(\mu_s)$ on which the Rayleigh quotient is at most $\lambda_k^{(\mu)}$. By the min-max characterization of eigenvalues,

$$\lambda_k^{(\mu_s)} \leqslant \lambda_k^{(\mu)},$$

completing the proof. □

The proof of Theorem 1.1 clearly shows that $C_P(\mu * \gamma_s) \geqslant C_P(\mu)$ for all $s > 0$, so by the semigroup property $s \mapsto C_P(\mu * \gamma_s)$ is non-decreasing in $s \in [0, \infty)$.

Remark 2.7. — Let μ be a log-concave probability measure in $\mathbb{R}^n$ with density $\rho = e^{-W}$, where W is a smooth function such that

$$\lim_{x\to\infty} \frac{|\nabla W(x)|^2}{2} - \Delta W(x) = \infty.$$

In this case, we have the strict inequality

$$\lambda_k^{(\mu*\gamma)} < \lambda_k^{(\mu)} \qquad (k = 1, 2, \ldots) \tag{2.24}$$

In order to prove (2.24), we first observe that $\nabla^2 \log \rho_s(y) < 0$ for all $y \in \mathbb{R}^n$ as follows from the equality case of the Brascamp–Lieb inequality or from [16, 18]. Arguing as in the proof of Theorem 2.4 and using the fact that $\nabla\varphi_s \not\equiv 0$ as φ_s is non-constant, we conclude that $\mathrm{d}R_\varphi(s)/\mathrm{d}s < 0$ whenever $0 \not\equiv \varphi \in H^1(\mu)$ has subexponential decay relative to ρ.

Therefore (2.24) would follow from Theorem 2.4, as in the proof of Theorem 1.1 above, had we known that any eigenfunction φ of L_μ has subexponential decay relative to ρ.

Indeed, let $A : L^2(\mathbb{R}^n) \xrightarrow{\sim} L^2(\mu)$ be the isometry given by $A(g) = e^{\frac{W}{2}}g$. It is well-known and easy to verify that $A^{-1}L_\mu A$ is the Schrödinger operator

$$-\Delta + \frac{|\nabla W|^2}{4} - \frac{\Delta W}{2},$$

which is of the form $-\Delta + V$ with $V \geqslant 0$ and $V \to \infty$ as $x \to \infty$. By results on the decay of eigenfunctions of Schrödinger operators [36, Theorem XIII.70], the function $A^{-1}\varphi$ has subexponential decay at infinity, and hence φ has subexponential decay relative to ρ.

3. A contraction transporting $\mu * \gamma$ to μ

In this section we prove Theorem 1.2 using the arguments of Kim and Milman [23]. To begin with, we work with a log-concave probability measure μ with a smooth, strictly positive density ρ on $\mathbb{R}^n$. We furthermore make the regularity assumption that there exists $\varepsilon > 0$ such that

$$-\nabla^2 \log \rho(y) \leqslant \frac{1}{\varepsilon} \cdot \mathrm{Id} \qquad (y \in \mathbb{R}^n). \tag{3.1}$$

We shall later remove these assumptions on ρ. As above, for $s \geqslant 0$ we write $\mu_s = \mu * \gamma_s$ and ρ_s is the density of μ_s. Thus ρ_s is smooth, positive and log-concave in $\mathbb{R}^n$. For $s \geqslant 0$ consider the *advection field*

$$W_s(y) = -\frac{1}{2}\nabla \log \rho_s(y). \tag{3.2}$$

The "physical" interpretation of this vector field is as follows. One of the derivations of the heat equation is based on Fourier's law, according to which the flux of heat across a tiny surface in a short time interval is proportional to the temperature gradient across the surface. If we think of the heat as carried by a fluid of particles with density $\rho(x,t)$, this means that the current of heat is proportional to $-\nabla\rho$ (we take $\frac{1}{2}$ to be the constant of proportionality); since the current of heat is simply ρv, where $v(x,t)$ is the bulk velocity of the fluid, we obtain $v = -\frac{1}{2}\frac{\nabla\rho}{\rho}$, which is (3.2). For more details, see [38, §5.4].

With this point of view, the trajectory of a particle located at time $s = 0$ at the point $y \in \mathbb{R}^n$ is the curve $s \mapsto T_s(y)$ where

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}s} T_s(y) = W_s(T_s(y)), \quad s \geqslant 0 \\ \quad T_0(y) = y \end{cases} \tag{3.3}$$

Lemma 3.1. — *Under the above assumptions on ρ, the ordinary differential equation* (3.3) *determines the family of maps $(T_s : \mathbb{R}^n \to \mathbb{R}^n)_{s\geqslant 0}$. These maps are all diffeomorphisms, and $T_s(y)$ is smooth in $(s,y) \in [0,\infty) \times \mathbb{R}^n$.*

Proof. — Since $\rho_s = \rho * \gamma_s$, the function $\rho_s(y)$ is smooth and positive in $(s,y) \in [0,\infty) \times \mathbb{R}^n$. Therefore $W_s(y)$ is smooth in $(s,y) \in [0,\infty) \times \mathbb{R}^n$ as well. It remains to show that W_s is $1/(2\varepsilon)$-Lipschitz on $\mathbb{R}^n$ for any $s \geqslant 0$. Once this is shown, the standard theory of ordinary differential equations implies the existence and uniqueness of solutions to (3.3) and their smooth dependence on initial conditions (e.g., [22, Chapter V]). The fact that the T_s are diffeomorphisms follows from the theory of flows of time-dependent vector fields on manifolds (e.g., [29, Chapter 17]).

We need to compute the derivative of W_s. As in the beginning of the proof of Lemma 2.2 above, by differentiating (2.6) we see that for any $s > 0$

and $y \in \mathbb{R}^n$,

$$DW_s(y) = -\frac{1}{2}\nabla^2 \log \rho_s(y) = \frac{s \cdot \mathrm{Id} - \mathrm{Cov}(p_{s,y})}{2s^2} \tag{3.4}$$

where $\mathrm{Cov}(p_{s,y}) \in \mathbb{R}^{n\times n}$ is the covariance matrix of the probability density $p_{s,y}$. Since ρ_s is log-concave, the differential DW_s is a symmetric positive semidefinite matrix. From (2.5) and the regularity assumption (3.1) we see that for $s \geqslant 0$ and $x \in \mathbb{R}^n$,

$$-\nabla^2 \log p_{s,y}(x) \leqslant \left(\frac{1}{\varepsilon} + \frac{1}{s}\right) \cdot \mathrm{Id} \tag{3.5}$$

in the sense of symmetric matrices. It is well-known (see [7, Theorem 5.4]) that (3.5) implies that

$$\left(\frac{1}{\varepsilon} + \frac{1}{s}\right)^{-1} \cdot \mathrm{Id} \leqslant \mathrm{Cov}(p_{s,y}). \tag{3.6}$$

From (3.4) and (3.6) we deduce the pointwise bound

$$\|DW_s(y)\|_{op} \leqslant \frac{1}{2(s+\varepsilon)} \qquad (y \in \mathbb{R}^n),$$

where $\|\cdot\|_{op}$ is the operator norm. This bound clearly applies also for $s = 0$. Therefore $W_s : \mathbb{R}^n \to \mathbb{R}^n$ is $1/(2\varepsilon)$-Lipschitz for any $s \geqslant 0$, completing the proof. □

As explained in Kim and Milman [23], the diffeomorphism T_s is an expansion, i.e., $|T_s(x) - T_s(y)| \geqslant |x - y|$ for all x, y and s. In order to prove this, we show that everywhere in $\mathbb{R}^n$,

$$(DT_s)^*(DT_s) \geqslant \mathrm{Id}\,. \tag{3.7}$$

Inequality (3.7) is certainly true when $s = 0$, while the fact that DW_s is positive semidefinite implies that

$$\frac{\partial}{\partial s}(DT_s)^*(DT_s) = 2(DT_s)^*(DW_s)DT_s \geqslant 0.$$

Therefore (3.7) holds true. This implies that $\left\|D\left(T_s^{-1}\right)\right\|_{op} \leqslant 1$, hence T_s^{-1} is a contraction and T_s is an expansion. Next, from (3.2) and the heat equation $\partial\rho_s/\partial s = \Delta\rho_s/2$ we obtain the linear transport equation (also known as the continuity equation),

$$\frac{\partial \rho_s}{\partial s} + \mathrm{div}(\rho_s W_s) = 0 \qquad (s \geqslant 0, y \in \mathbb{R}^n).$$

The continuity equation implies that ρ_s is the density of the pushforward of μ under the diffeomorphism T_s (see e.g. [38, Theorem 5.34]). Consequently, the map T_s^{-1} is a contraction that pushes forward μ_s to μ.

Proof of Theorem 1.2. — Set $s = 1$ so that $\mu_s = \mu * \gamma$. We have just established the existence of a contraction transporting μ_s to μ under the additional requirement that μ admits a smooth, positive density satisfying the regularity assumption (3.1).

Consider now the case where μ is an arbitrary absolutely-continuous, log-concave probability measure in $\mathbb{R}^n$. For any $\varepsilon > 0$, the measure $\mu_\varepsilon = \mu * \gamma_\varepsilon$ has a smooth, positive, log-concave density satisfying the regularity assumption (3.1), as follows from the computation in (3.4) above. Hence there exists a contraction transporting $\mu_\varepsilon * \gamma$ to μ_ε. By [23, Lemma 3.3], in order to show that there exists a contraction from $\mu * \gamma$ to μ, it suffices to show that

$$\mu_\varepsilon \xrightarrow{\varepsilon \to 0^+} \mu \tag{3.8}$$

in the total variation metric, and that $\mu_\varepsilon * \gamma \longrightarrow \mu * \gamma$ as $\varepsilon \to 0$ in the weak topology. Since $\mu_\varepsilon * \gamma = (\mu * \gamma)_\varepsilon$ and since convergence in total variation implies convergence in the weak topology, it suffices to prove (3.8). Thus we need to show that

$$\int_{\mathbb{R}^n} |\rho_\varepsilon(x) - \rho(x)| \mathrm{d}x \xrightarrow{\varepsilon \to 0^+} 0. \tag{3.9}$$

Arguing as in (2.11) and the paragraph following (2.12) above, we know that there exist $a, b > 0$ such that $\rho_\varepsilon(x) \leqslant a e^{-b|x|}$ for all $x \in \mathbb{R}^n$ and $0 \leqslant \varepsilon \leqslant 1$, with $\rho_0 = \rho$. Since ρ is continuous almost everywhere in $\mathbb{R}^n$, the integrand in (3.9) converges to zero almost everywhere, and (3.9) follows from the dominated convergence theorem.

Thus the conclusion of the theorem is valid when μ is an absolutely continuous, log-concave probability measure. Finally, if μ is not absolutely continuous, then we may project to a lower dimension using an orthogonal projection, which is a contraction, and reduce matters to the absolutely continuous case. □

Theorem 1.2 implies that for $\nu = \mu * \gamma$ and $0 \not\equiv \varphi \in H^1(\nu)$ we have the following inequality between Rayleigh quotients:

$$\frac{\int_{\mathbb{R}^n} |\nabla(\varphi \circ T)|^2 \mathrm{d}\nu}{\int_{\mathbb{R}^n} (\varphi \circ T)^2 \mathrm{d}\nu} \leqslant \frac{\int_{\mathbb{R}^n} |(\nabla \varphi) \circ T|^2 \mathrm{d}\nu}{\int_{\mathbb{R}^n} (\varphi \circ T)^2 \mathrm{d}\nu} = \frac{\int_{\mathbb{R}^n} |\nabla \varphi|^2 \mathrm{d}\mu}{\int_{\mathbb{R}^n} \varphi^2 \mathrm{d}\mu}. \tag{3.10}$$

We may now repeat the proof of Theorem 1.1 from Section 2, with the linear map $\varphi \mapsto \varphi \circ T$ playing the role of the linear map $\varphi \mapsto Q_1 \varphi$. This yields another proof of Theorem 1.1, relying on (3.10) in place of Theorem 2.4.

4. A Bayesian interpretation of Eldan's stochastic localization

Eldan's stochastic localization technique was introduced by Eldan in [19] and developed since then by several authors in different settings [14, 20, 24, 30]. The method has turned out to be useful in particular for the study of log-concave measures, culminating thus far in the breakthrough result of Chen [14] showing that the isotropic constant grows more slowly than any power of the dimension. In this section, we give a "Bayesian" interpretation of Eldan's stochastic localization relating it to the heat flow and to the operator Q_s introduced above, as well as to the Föllmer drift in the theory of Wiener space. It was this line of development which led us to the results announced in the introduction; however, this section may be read independently.

We refer to [35] for background on stochastic processes. Let μ be an absolutely continuous probability measure on $\mathbb{R}^n$ with density p_0 and with finite second moments. Let $(W_t)_{t\geqslant 0}$ be a standard Brownian motion on $\mathbb{R}^n$ with $W_0 = 0$.

The stochastic localization process, in the version introduced by [30], is a density-valued stochastic process p_t driven by W_t, defined as follows: for every $x \in \mathbb{R}^n$, the process $(p_t(x))_{t\geqslant 0}$ is the solution to the stochastic differential equation

$$\mathrm{d}p_t(x) = p_t(x)\langle x - a_t, \mathrm{d}W_t\rangle \tag{4.1}$$

with initial condition p_0, where $a_t = \int_{\mathbb{R}^n} x \cdot p_t(x)\,\mathrm{d}x$ is the barycenter of p_t. As this equation has no drift term, $p_t(x)$ is a martingale, and p_t is almost surely a probability density. In particular, $\mathbb{E}[p_t(x)] = p_0(x)$, and for any test function φ, we have $\mathbb{E}_{X\sim p_0}[\varphi(X)] = \mathbb{E}[\mathbb{E}_{X\sim p_t}[\varphi(X)]]$.

The process $(p_t)_{t\geqslant 0}$ has another description, as a stochastic "tilt" of p_0. In this section, for $t \geqslant 0$ and $\theta \in \mathbb{R}^n$ let $p_{t,\theta}$ denote the probability density given by

$$p_{t,\theta}(x) = \frac{1}{Z(t,\theta)} e^{\langle \theta, x\rangle - \frac{t|x|^2}{2}} p_0(x), \tag{4.2}$$

where $Z(t,\theta) = \int_{\mathbb{R}^n} e^{\langle \theta, x\rangle - \frac{t|x|^2}{2}} p_0(x)\,\mathrm{d}x$ is a normalization constant. Let $a(t,\theta)$ denote the barycenter of $p_{t,\theta}$, and define the stochastic process θ_t via the differential equation

$$\mathrm{d}\theta_t = a(t,\theta_t)\,\mathrm{d}t + \mathrm{d}W_t, \qquad \theta_0 = 0. \tag{4.3}$$

It turns out that when θ_t and p_t are driven by the same Brownian motion, p_t is precisely equal to p_{t,θ_t}. For proofs of these and other formulas relating to the stochastic localization process, and for the application to the KLS conjecture, see [30, 31] or [14].

The Bayesian interpretation of the Eldan process is quite simple: let X be a random vector distributed according to μ, independent of the Brownian motion $(W_t)_{t\geqslant 0}$. Denote

$$\widetilde{\theta}_t = tX + W_t \qquad (t \geqslant 0). \tag{4.4}$$

Our main observations are the following two claims:

(i) The process $(\widetilde{\theta}_t)_{t\geqslant 0}$ coincides in law with the process $(\theta_t)_{t\geqslant 0}$ which solves (4.3) above.
(ii) For any fixed $t > 0$ and $\theta \in \mathbb{R}^n$, the probability density $p_{t,\theta}$ on $\mathbb{R}^n$ is precisely the conditional probability distribution of X given that $\widetilde{\theta}_t = \theta$.

Thus, when we observe the tilt process $(\theta_t)_{t\geqslant 0}$, we actually see a Brownian motion with a constant drift X which is unknown, but whose prior distribution is known to us. Moreover, the posterior probability density for the unknown drift X given the observation of the process $(\theta_s)_{0\leqslant s\leqslant t}$ until time t depends only on the state of the process at time t, and is equal to p_{t,θ_t}.

In the following proposition we prove these two claims. For $T > 0$ let $\mathcal{V}_T = C_0([0,T],\mathbb{R}^n)$ be the Wiener space of $\mathbb{R}^n$-valued continuous functions $(W_t)_{0\leqslant t\leqslant T}$ with $W_0 = 0$. Slightly abusing notation, we write γ_T for the Wiener probability measure on $\mathcal{V}_T$ and $\{\mathcal{F}_t\}_{0\leqslant t\leqslant T}$ for the natural filtration, i.e., $\mathcal{F}_t$ is the σ-algebra generated by $(W_s)_{0\leqslant s\leqslant t}$.

Proposition 4.1. — *Let μ be a probability measure on $\mathbb{R}^n$ which is absolutely continuous with respect to the Lebesgue measure λ. Fix $T > 0$, and consider the space $\Omega = \mathbb{R}^n \times \mathcal{V}_T$ and the transformation $\tau : \Omega \to \Omega$ given by*

$$\tau(x,(W_t)_{0\leqslant t\leqslant T}) = (x,(W_t + tx)_{0\leqslant t\leqslant T}).$$

Write $\nu = \tau_*(\mu \otimes \gamma_T)$. Then,

(i) *The stochastic process $(\widetilde{\theta}_t)_{t\geqslant 0}$ described in (4.4) coincides in law with the Itô process $(\theta_t)_{t\geqslant 0}$ defined as the solution to the stochastic differential equation (4.3).*
(ii) *The measure ν is absolutely continuous with respect to $\lambda \otimes \gamma_T$ on Ω with density*

$$\frac{\mathrm{d}\nu}{\mathrm{d}(\lambda \otimes \gamma_T)}(x,\widetilde{\theta}) = p_0(x)e^{\langle \widetilde{\theta}_T, x\rangle - \frac{T|x|^2}{2}},$$

for $x \in \mathbb{R}^n$ and $\widetilde{\theta} = (\widetilde{\theta}_t)_{0\leqslant t\leqslant T} \in \mathcal{V}_T$. Consequently, when $(X,(\widetilde{\theta}_t)_{0\leqslant t\leqslant T})$ is the stochastic process described in (4.4), the conditional distribution of X with respect to $\widetilde{\theta} = (\widetilde{\theta}_t)_{0\leqslant t\leqslant T}$ is given by *the probability density*

$$q_T(x|\widetilde{\theta}) = \frac{p_0(x)e^{\langle \tilde{\theta}_T, x\rangle - \frac{T|x|^2}{2}}}{\int_{\mathbb{R}^n} p_0(y)e^{\langle \tilde{\theta}_T, y\rangle - \frac{T|y|^2}{2}}\,\mathrm{d}y} = p_{T,\tilde{\theta}_T}(x) \qquad (x \in \mathbb{R}^n). \tag{4.5}$$

Proof. — We first prove (ii). For $x \in \mathbb{R}^n$, let $\tau_x : \mathcal{V}_T \to \mathcal{V}_T$ be defined by $\tau_x((W_t)_{t\leqslant T}) = (W_t + tx)_{t\leqslant T}$ so that $\tau(x,\omega) = (x, \tau_x(\omega))$. By Fubini's theorem,

$$\nu = \tau_*(\mu \otimes \gamma_T) = \int_{\mathbb{R}^n} (x, \tau_x)_* \gamma_T \,\mathrm{d}\mu(x). \tag{4.6}$$

This means that for any test function g,

$$\int_\Omega g\,\mathrm{d}\nu = \int_{\mathbb{R}^n} \left(\int_{\mathcal{V}_T} g(x,\widetilde{\theta})\,\mathrm{d}((\tau_x)_*\gamma_T)(\widetilde{\theta}) \right) \mathrm{d}\mu(x).$$

Since τ_x is just a translation in Wiener space by the deterministic function $f_x(t) = tx$, the Cameron–Martin theorem [11] yields that the density of $(\tau_x)_*\gamma_T$ with respect to γ_T at the point $(\widetilde{\theta}_t)_{0\leqslant t\leqslant T} \in \mathcal{V}_T$ equals

$$\begin{aligned} \frac{\mathrm{d}(\tau_x)_*\gamma_T}{\mathrm{d}\gamma_T}((\widetilde{\theta}_t)_{t\leqslant T}) &= \exp\left(\int_0^T \langle f_x'(t), \mathrm{d}\widetilde{\theta}_t\rangle - \frac{1}{2}\int_0^T |f_x'(t)|^2\,\mathrm{d}t \right) \\ &= \exp\left(\int_0^T x\,\mathrm{d}\widetilde{\theta}_t - \frac{1}{2}\int_0^T |x|^2\,\mathrm{d}t \right) = e^{\langle \tilde{\theta}_T, x\rangle - \frac{T|x|^2}{2}}. \end{aligned} \tag{4.7}$$

It follows from (4.6) and (4.7) that

$$\frac{\mathrm{d}\nu}{\mathrm{d}(\lambda\otimes\gamma_T)}(x,\widetilde{\theta}) = p_0(x)\cdot \frac{\mathrm{d}\nu}{\mathrm{d}(\mu\otimes\gamma_T)}(x,\widetilde{\theta}) = p_0(x)e^{\langle \tilde{\theta}_T, x\rangle - T\frac{|x|^2}{2}}. \tag{4.8}$$

The probability measure ν is the joint distribution of the stochastic process

$$(X, (\widetilde{\theta}_t)_{0\leqslant t\leqslant T})$$

described in (4.4). Therefore, when conditioning on the entire stochastic process $(\widetilde{\theta}_t)_{0\leqslant t\leqslant T}$, it follows from (4.8) that the probability density function of X is proportional to $x \mapsto p_0(x)e^{\langle \tilde{\theta}_T, x\rangle - T\frac{|x|^2}{2}}$ in $\mathbb{R}^n$. This completes the proof of (ii).

We move on to the proof of (i). We endow Ω with the probability measure $\mu\otimes\gamma_T$, and assume that $(X, (W_t)_{t\geqslant 0})$ is distributed according to this measure, while $\widetilde{\theta}_t = tX + W_t$. Thus,

$$\mathrm{d}\widetilde{\theta}_t = X\mathrm{d}t + \mathrm{d}W_t. \tag{4.9}$$

Write $\mathcal{N}_t$ for the σ-algebra generated by $(\widetilde{\theta}_s)_{0\leqslant s\leqslant t}$. Abbreviate $\mathbb{E}[X|\widetilde{\theta}] = \mathbb{E}[X|\mathcal{N}_t](\widetilde{\theta})$ for the conditional expectation of X with respect to $\mathcal{N}_t$, which is a function of $(\widetilde{\theta}_s)_{0\leqslant s\leqslant t}$. According to (4.9) and [35, Theorem 8.4.3], the process $(\widetilde{\theta}_t)_{0\leqslant t\leqslant T}$ coincides in law with the process $(\theta_t)_{0\leqslant t\leqslant T}$ defined by the initial condition $\theta_0 = \widetilde{\theta}_0 = 0$ and the stochastic differential equation

$$\mathrm{d}\theta_t = b(t,\theta_t)\,\mathrm{d}t + \mathrm{d}W_t,$$

if the function $b(t,x)$ defined for $0 < t \leqslant T$ and $x \in \mathbb{R}^n$ satisfies

$$b(t,\widetilde{\theta}_t) = \mathbb{E}[X|\widetilde{\theta}] \qquad \forall\, \widetilde{\theta} \in \mathcal{V}_t. \tag{4.10}$$

(To be precise, the statement in [35, Theorem 8.4.3] only treats time-independent diffusions, but the proofs generalize almost verbatim to the time-dependent case which we need.) The random variable $\mathbb{E}[X|\widetilde{\theta}]$, viewed as an $\mathcal{N}_t$-measurable function on Ω, is the conditional expectation of X given $\widetilde{\theta} = (\widetilde{\theta}_s)_{0\leqslant s\leqslant t}$. According to (ii), the conditional distribution of X given $\widetilde{\theta}$ is given by the probability density $q_t(x|\widetilde{\theta})$ from (4.5). Hence for any $0 < t < T$ and $\widetilde{\theta} \in \mathcal{V}_t$,

$$\mathbb{E}[X|\widetilde{\theta}] = \int_{\mathbb{R}^n} x \cdot q_t(x|\widetilde{\theta})\,\mathrm{d}x = \int_{\mathbb{R}^n} x \cdot p_{t,\tilde{\theta}_t}(x)\,\mathrm{d}x = a(t,\widetilde{\theta}_t).$$

We have thus verified condition (4.10) with $b(t,x) = a(t,x)$, completing the proof of (i). □

To reiterate, we have interpreted Eldan's stochastic localization for the measure μ as the following procedure: a value x is sampled from the distribution μ, and a Brownian motion with constant drift x, namely $\theta_t = tx + W_t$, is shown to an observer who knows the distribution μ, but not the value of x. From the observer's perspective, θ_t satisfies the stochastic differential equation (4.3), and at time t, the observer's posterior probability distribution for the hidden drift coefficient x is precisely p_{t,θ_t}. The fact that p_{t,θ_t} is a martingale now follows immediately from the law of total probability: for $s < t$,

$$\begin{aligned}
&\mathbb{E}[p_{t,\theta_t}(x)|\theta_s] \\
&\qquad = \mathbb{E}[p(X=x|\theta_t)|\theta_s] = \int_{\mathbb{R}^n} p(X=x|\theta_t=\theta)p(\theta_t=\theta|\theta_s)\,\mathrm{d}\theta \\
&\qquad = \int_{\mathbb{R}^n} p(X=x,\theta_t=\theta|\theta_s)\,\mathrm{d}\theta = p(X=x|\theta_s) = p_{s,\theta_s}(x).
\end{aligned} \tag{4.11}$$

In Bayesian terms, this simply means that if we continually obtain information about an unknown random variable X and update our posteriors for X accordingly, our expectation at time s for our estimate of X, or any function of X, at time t must coincide with our current estimate of X.

Remark 4.2.

(i) A curious property of Eldan's stochastic localization, in the Bayesian interpretation, is that the posterior distribution at time t depends only on θ_t: the full path $(\theta_s)_{0\leqslant s\leqslant t}$ contains no more information about X than θ_t alone. This is a limiting case of an amusing exercise in linear algebra and statistics which we now describe. Suppose that we are given N noisy observations of an unknown quantity X, of the form

$$X + Z_1, X + \frac{Z_1 + Z_2}{2}, \dots, X + \frac{Z_1 + \ldots + Z_N}{N} \tag{4.12}$$

where $Z_1, \dots, Z_N$ are independent, standard Gaussian random variables. Assume that the apriori distribution of X is known to us. What is the posterior distribution of X given the N observations in (4.12)? As it turns out, the posterior distribution depends only on the last of these N observations, for which the Gaussian noise is of the smallest variance. The first $N-1$ observations are completely useless in this context.

(ii) A suitably generalized version of this interpretation applies to the general stochastic localization process with a control matrix $C_t = C(t, \theta_t)$, as defined by [31]: a random variable X is drawn from μ as above, but instead of a Brownian motion with drift X, what the observer sees is an Itô process defined by the SDE $\mathrm{d}\theta_t = C(t,\theta_t) X \,\mathrm{d}t + C(t,\theta_t)^{\frac{1}{2}}\,\mathrm{d}W_t$. Again, p_t represents the observer's posterior distribution for X given the observation of θ_t up to time t. To prove that this description corresponds to the definition of the process in [31] one repeats the above argument using Girsanov's theorem, rather than the Cameron–Martin theorem. The "path-independence" property of the posteriors from the previous remark does not hold in this case.

4.1. Time inversion

Let us now explain the relationship between the tilt process in its Bayesian interpretation and our work in Section 2. A well-known identity for Brownian motion is the time-inversion property: suppose that $(W_t)_{t\in[0,\infty)}$ is a standard Brownian motion in $\mathbb{R}^n$ with $W_0 = 0$. Define $(\widetilde{W}_s)_{s\in[0,\infty)}$ by

$$\widetilde{W}_s = sW_{\frac{1}{s}}$$

and $\widetilde{W}_0 = 0$. Then $(\widetilde{W}_s)_{s\in[0,\infty)}$ is a standard Brownian motion as well. Consequently, from (4.4) we see that

$$Y_s := s\widetilde{\theta}_{1/s} = X + \widetilde{W}_s. \tag{4.13}$$

Recalling that the tilt process $(\theta_t)_{t\geqslant 0}$ coincides in law with $(\widetilde{\theta}_t)_{t\geqslant 0}$, we conclude from (4.13) that the tilt process coincides in law with the time inversion of a Brownian motion with a starting point drawn from the distribution μ.

Applying this time inversion, we treat the time-inverted tilt process $(Y_s)_{s\geqslant 0}$ as just a Brownian motion with a random starting point. Working with it requires nothing more than the explicit expression for the Euclidean heat kernel; for instance, the distribution of Y_s is given by the probability density function $\rho_s = P_s\rho$ with $\rho = p_0$. Given a function φ on $\mathbb{R}^n$ and $t > 0$, the random variable

$$\int_{\mathbb{R}^n} \varphi p_t$$

associated to the tilt process coincides in law with the distribution of $Q_s\varphi$ under the measure μ_s, for $s = 1/t$. Moreover,

$$Q_s\varphi(y) = \mathbb{E}\left[\varphi(X)|\theta_t = \theta\right] \qquad \text{for } s = 1/t > 0, \theta = ty \in \mathbb{R}^n. \tag{4.14}$$

It is this elementary, "functional analytic" perspective on the measures $p_{t,\theta}$ (or, in the new variables, $p_{s,y}$) that is taken in Section 2, which makes no explicit use of stochastic localization, pathwise analysis, martingales or stochastic calculus at all.

4.2. Föllmer drift as a "time-compressed" version of stochastic localization

Föllmer drift is a well-known stochastic process which couples between an absolutely continuous measure μ and Wiener measure on path space over a *finite* time interval, without loss of generality $[0, 1]$. It is the same process referred to as the "h-process" in Cattiaux and Guillin [12], because of its relation to Doob's h-transform.

In brief, the Föllmer drift associated to μ is a Brownian motion conditioned to have law μ at time $t = 1$. The measure $\mathcal{P}^\mu$ on $\mathcal{V}_1 = C_0[0, 1]$ defining the Föllmer drift of μ is defined as the measure having Radon–Nikodym derivative

$$\frac{\mathrm{d}\mathcal{P}^\mu}{\gamma_1}(W) = \frac{\mathrm{d}\mu}{\mathrm{d}\gamma}(W_1) \qquad (W = (W_t)_{0\leqslant t\leqslant 1} \in \mathcal{V}_1), \tag{4.15}$$

where γ_1 on the left side of (4.15) is the Wiener measure on $C_0[0, 1]$, while γ on the right side of (4.15) is the standard Gaussian measure in $\mathbb{R}^n$. The Föllmer drift $\mathcal{P}^\mu$ turns out to have a certain energy-minimizing property, and its energy is precisely twice the relative entropy $H(\mu|\gamma)$, properties which make it quite useful for proving functional inequalities; see, e.g., [21, 32].

We can interpret Föllmer drift in a manner completely analogous to the tilt process: let X be a random variable drawn from μ and let $(B_t)_{0\leqslant t\leqslant 1}$ denote an independent standard Brownian bridge on $[0,1]$. Then the law of the process

$$X_t = tX + B_t \qquad (0 \leqslant t \leqslant 1) \tag{4.16}$$

is precisely that of the Föllmer drift associated to μ. Moreover, just as above, one may consider an observer who sees X_t but not X and define posterior probability distributions for X given $(X_s)_{s\leqslant t}$. These posterior probability distributions are the random measures μ_t in [21, §3], in a slightly different normalization (in [21] the measure μ_t is the posterior probability distribution of $(X - X_t)/\sqrt{1-t}$ given X_t, rather than the posterior probability distribution of X itself given X_t).

In fact, there is an even closer relationship between Föllmer drift X_t and the tilt process θ_t of Eldan's stochastic localization, which manifests in two separate ways. First of all, for $t \leqslant 1$, we may write

$$\theta_t = tX + W_t = t(X + W_1) + (W_t - tW_1). \tag{4.17}$$

Note that W_1 is independent of $W_t - tW_1$ as these are jointly Gaussian, centered and $\mathbb{E}W_1(W_t - tW_1) = 0$. Recall that $B_t = W_t - tW_1$ is one way to construct a Brownian bridge. From (4.16) and (4.17) we see that the tilt process $(\theta_t)_{t\in[0,1]}$ has the law of the Föllmer drift for the measure associated to $X + W_1$, namely $\mu * \gamma$. In the same fashion, one sees that for any $T > 0$, the process $(\theta_t)_{t\in[0,T]}$ is identical in law to the Föllmer drift of $\mu * \gamma_T$, with the time interval rescaled to $[0,T]$.

Another way to construct a standard Brownian bridge from a standard Brownian motion is by "time compression": the process

$$\widetilde{B}_t = (1-t)W_{\frac{t}{1-t}} \qquad (0 \leqslant t \leqslant 1)$$

coincides in law with the standard Brownian bridge. By inverting this operation, we can construct a Brownian motion from a Brownian bridge: $W_t = (1+t)\widetilde{B}_{\frac{t}{1+t}}$. Hence, the Föllmer drift for μ and the tilt process of Eldan's stochastic localization satisfy the reciprocal relations

$$X_t \simeq (1-t)\theta_{\frac{t}{1-t}} \qquad \text{and} \qquad \theta_t \simeq (1+t)X_{\frac{t}{1+t}},$$

where $\simeq$ means "coincides in law". This follows from the defining formulas $\theta_t = tX + W_t$, $X_t = tX + B_t$ and the corresponding relations for B_t and W_t. Thus, Föllmer drift is simply a time-compressed version of the tilt process.

Note added in revision

El Alaoui and Montanari [1] have independently arrived at many of the results of Section 4, almost simultaneously with the appearance of our paper on arXiv. Furthermore, in a recent paper by Klartag and Lehec [25], the functional analytic framework suggested above, and its relation to Eldan's Stochastic localization, were instrumental in showing that the KLS conjecture holds true up to a factor that is polylogarithmic in the dimension.

Appendix A. Discreteness of the spectrum for rapidly decreasing log-concave densities

The goal of this appendix is to prove the following proposition:

Proposition A.1. — *Let μ be a log-concave probability measure on $\mathbb{R}^n$ with smooth, positive density $\rho = e^{-V}$ such that $\frac{V(x)}{|x|} \to \infty$ as $x \to \infty$. Then the spectrum of L_μ is discrete.*

For $A \subseteq \mathbb{R}^n$ we write $C_c^\infty(A)$ for the class of smooth, compactly supported functions in $\mathbb{R}^n$ that are supported in the set A. As explained in [2, §4.10], in order to prove Proposition A it suffices to show the following:

(∗) For any $a > 0$ there exists $r > 0$ such that for any $f \in C_c^\infty(\mathbb{R}^n \setminus B_r)$,

$$\int_{\mathbb{R}^n} |\nabla f|^2 \mathrm{d}\mu \geqslant a \cdot \int_{\mathbb{R}^n} f^2 \mathrm{d}\mu.$$

Here $B_r = \{x \in \mathbb{R}^n \,;\, |x| \leqslant r\}$.

Consider first the one-dimensional case in which μ is supported on a half-line. Thus $\mathrm{d}\mu = \rho(x)\mathrm{d}x = e^{-W(x)}\mathrm{d}x$ is a measure on $[0, \infty)$ with $W : [0, \infty) \to \mathbb{R}$ smooth and convex. We will apply the Muckenhoupt criterion ([34]; see also [2, §4.5.1]), which we state as the following lemma:

Lemma A.2. — *Let $\rho : [0, \infty) \to (0, \infty)$ be such that*

$$C := \sup_{r>0} \int_r^\infty \rho \int_0^r \frac{1}{\rho} < \infty.$$

Then for every smooth, compactly supported function $f : [0, \infty) \to \mathbb{R}$,

$$\int_0^\infty f^2 \, \rho \leqslant 4C \int_0^\infty (f')^2 \, \rho.$$

Let $x_0 > 0$ be such that $a := W'(x_0) > 0$. Then $W'(x) \geqslant a$ for all $x > x_0$ by convexity. Hence for any $r, x > x_0$,

$$\rho(r) \leqslant e^{-a(r-x)}\rho(x) \quad \text{for } x < r,$$
$$\rho(x) \leqslant e^{-a(x-r)}\rho(r) \quad \text{for } x > r.$$

Therefore, for any $r > x_0$,

$$\int_{x_0}^{r} \frac{1}{\rho(x)}\,\mathrm{d}x \leqslant \frac{1}{\rho(r)} \int_{x_0}^{r} e^{-a(r-x)}\,\mathrm{d}x \leqslant \frac{1}{\rho(r)} \cdot \frac{1}{a}, \tag{A.1}$$

$$\int_{r}^{\infty} \rho(x)\,\mathrm{d}x \leqslant \rho(r) \int_{r}^{\infty} e^{-a(x-r)}\,\mathrm{d}x \leqslant \rho(r) \cdot \frac{1}{a}. \tag{A.2}$$

Thus we obtain $\sup_{r>x_0} \int_r^\infty \rho \int_{x_0}^r \frac{1}{\rho} \leqslant \frac{1}{a^2}$, so Muckenhoupt's criterion yields that for any $f \in C_c^\infty([x_0, \infty))$,

$$\int_{x_0}^{\infty} (f')^2\,\mathrm{d}\mu \geqslant \frac{a^2}{4} \int_{x_0}^{\infty} f^2\,\mathrm{d}\mu, \tag{A.3}$$

whenever $a = W'(x_0) > 0$.

So much for the one-dimensional case. Now let $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ be an n-dimensional log-concave measure, and consider the family of functions $f_R : S^{n-1} \to \mathbb{R}$ defined by $f_R(u) = \frac{V(Ru)-V(0)}{R}$. By convexity, f_R is monotone increasing in R, and by assumption f_R converges pointwise to infinity. Hence, applying Dini's theorem, we see that f_R converges uniformly to ∞. Denote $V_u(r) = V(ru)$ for $u \in S^{n-1}$ and $r \geqslant 0$. By convexity,

$$V_u'(r) \geqslant f_R(u).$$

Hence for every $a > 0$ there exists $\widetilde{R} > 0$ such that $V_u'(r) \geqslant a$ for all $r \geqslant \widetilde{R}$ and $u \in S^{n-1}$. Denoting $W_u(r) = V_u(r) - (n-1)\log r$, we see we see that W_u is convex in $(0, \infty)$ and that for any $a > 0$ there exists $R = R(a) > 0$ such that $W_u'(r) \geqslant a$ for all $r \geqslant R$ and $u \in S^{n-1}$. By integrating in polar coordinates and using (A.3) we conclude that for any $a > 0$ and $f \in C_c^\infty(\mathbb{R}^n \setminus B_R)$, denoting $f_u(r) = f(ru)$,

$$\begin{aligned}
\|\nabla f\|_{L^2(\mu)}^2 &\geqslant \int_{S^{n-1}} \int_0^\infty r^{n-1} (f_u'(r))^2 e^{-V_u(r)} \mathrm{d}r\mathrm{d}u \\
&\geqslant \int_{S^{n-1}} \left(\int_R^\infty (f_u'(r))^2 e^{-W_u(r)} \mathrm{d}r \right) \mathrm{d}u \\
&\geqslant \frac{a^2}{4} \int_{S^{n-1}} \left(\int_R^\infty f_u(r)^2 e^{-W_u(r)} \mathrm{d}r \right) \mathrm{d}u = \frac{a^2}{4} \|f\|_{L^2(\mu)}^2.
\end{aligned}$$

We have thus verified condition $(*)$ above, completing the proof of Proposition A.1.

Bibliography

[1] A. E. Alaoui & A. Montanari, "An Information-Theoretic View of Stochastic Localization", 2021.

[2] D. Bakry, I. Gentil & M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators*, Grundlehren der Mathematischen Wissenschaften, vol. 348, Springer, 2014.

[3] K. Ball, F. Barthe & A. Naor, "Entropy jumps in the presence of a spectral gap", *Duke Math. J.* **119** (2003), no. 1, p. 41-63.

[4] F. Barthe & B. Klartag, "Spectral gaps, symmetries and log-concave perturbations", *Bull. Hell. Math. Soc.* **64** (2020), p. 1-31.

[5] R. Bauerschmidt & T. Bodineau, "Log-Sobolev inequality for the continuum sine-Gordon model", 2019, `https://arxiv.org/abs/1907.12308`.

[6] S. G. Bobkov, "Isoperimetric and analytic inequalities for log-concave probability measures", *Ann. Probab.* **27** (1999), no. 4, p. 1903-1921.

[7] H. J. Brascamp & E. H. Lieb, "On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation", *J. Funct. Anal.* **22** (1976), no. 4, p. 366-389.

[8] S. Brazitikos, A. Giannopoulos, P. Valettas & B.-H. Vritsiou, *Geometry of isotropic convex bodies*, Mathematical Surveys and Monographs, vol. 196, American Mathematical Society, 2014.

[9] P. Buser, "A note on the isoperimetric constant", *Ann. Sci. Éc. Norm. Supér.* **15** (1982), no. 2, p. 213-230.

[10] L. A. Caffarelli, "Monotonicity properties of optimal transportation and the FKG and related inequalities", *Commun. Math. Phys.* **214** (2000), no. 3, p. 547-563.

[11] R. H. Cameron & W. T. Martin, "Transformations of Wiener Integrals under Translations", *Ann. Math.* **45** (1944), p. 386-396.

[12] P. Cattiaux & A. Guillin, "On the Poincaré constant of log-concave measures", in *Geometric aspects of functional analysis. Israel seminar (GAFA) 2017–2019. Volume 1*, Lecture Notes in Mathematics, vol. 2256, Springer, 2020, p. 171-217.

[13] J. Cheeger, "A lower bound for the smallest eigenvalue of the Laplacian", in *Problems in analysis. A symposium in honor of Salomon Bochner*, Princeton Mathematical Series, vol. 31, Princeton University Press, 1970, p. 195-199.

[14] Y. Chen, "An Almost Constant Lower Bound of the Isoperimetric Coefficient in the KLS Conjecture", *Geom. Funct. Anal.* **31** (2021), no. 1, p. 34-61.

[15] T. A. Courtade, "Bounds on the Poincaré constant for convolution measures", *Ann. Inst. Henri Poincaré, Probab. Stat.* **56** (2020), no. 1, p. 566-579.

[16] T. A. Courtade & M. Fathi, "Stability of the Bakry-Émery theorem on $\mathbb{R}^n$", *J. Funct. Anal.* **279** (2020), no. 2, article no. 108523 (28 pages).

[17] Y. S. Davidovich, B. I. Korenblyum & B. I. Khatset, "A certain property of logarithmically concave functions", *Dokl. Akad. Nauk SSSR* **185** (1969), p. 1215-1218, English translation in *Sov. Math., Dokl.* **10** (1969), p. 477–480.

[18] G. De Philippis & A. Figalli, "Rigidity and stability of Caffarelli's log-concave perturbation theorem", *Nonlinear Anal., Theory Methods Appl.* **154** (2017), p. 59-70.

[19] R. Eldan, "Thin shell implies spectral gap via a stochastic localization scheme", *Geom. Funct. Anal.* **23** (2013), no. 2, p. 532-569.

[20] R. Eldan & R. Gross, "Concentration on the Boolean hypercube via pathwise stochastic analysis", in *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC 2020)*, ACM Press, 2020, p. 208-221.

[21] R. Eldan, J. Lehec & Y. Shenfeld, "Stability of the logarithmic Sobolev inequality via the Föllmer Process", *Ann. Inst. Henri Poincaré, Probab. Stat.* **56** (2020), no. 3, p. 2253-2269.

[22] P. Hartman, *Ordinary differential equations*, Birkhäuser, 1982.

[23] Y.-H. Kim & E. Milman, "A generalization of Caffarelli's contraction theorem via (reverse) heat flow", *Math. Ann.* **354** (2012), no. 3, p. 827-862.

[24] B. Klartag, "Eldan's stochastic localization and tubular neighborhoods of complex-analytic sets", *J. Geom. Anal.* **28** (2018), no. 3, p. 2008-2027.

[25] B. Klartag & J. Lehec, "Bourgain's slicing problem and KLS isoperimetry up to polylog", 2022, `https://arxiv.org/abs/2203.15551`.

[26] A. V. Kolesnikov, "On Diffusion Semigroups Preserving the Log-Concavity", *J. Funct. Anal.* **186** (2001), no. 1, p. 196-205.

[27] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, vol. 89, American Mathematical Society, 2001.

[28] ———, "Spectral gap, logarithmic Sobolev constant, and geometric bounds", in *Eigenvalues of Laplacians and other geometric operators*, Surveys in Differential Geometry, vol. 9, International Press, 2004, p. 219-240.

[29] J. M. Lee, *Introduction to smooth manifolds*, Graduate Texts in Mathematics, vol. 218, Springer, 2006.

[30] Y. T. Lee & S. S. Vempala, "Eldan's Stochastic Localization and the KLS Conjecture: Isoperimetry, Concentration and Mixing", in *58th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2017)*, IEEE Computer Society, 2017, p. 998-1007.

[31] ———, "The Kannan-Lovász-Simonovits Conjecture", in *Current developments in mathematics 2017*, International Press, 2019, p. 1-36.

[32] J. Lehec, "Representation formula for the entropy and functional inequalities", *Ann. Inst. Henri Poincaré, Probab. Stat.* **49** (2019), no. 3, p. 885-899.

[33] D. S. Lemons & A. Gythiel, "Paul Langevin's 1908 paper "On the Theory of Brownian Motion"", *Am. J. Phys.* **65** (1997), p. 1079-1081.

[34] B. Muckenhoupt, "Hardy's inequality with weights", *Stud. Math.* **44** (1972), p. 31-38.

[35] B. Øksendal, *Stochastic differential equations: an introduction with applications*, Springer, 2013.

[36] M. Reed & B. Simon, *Methods of Modern Mathematical Physics. Vol. 4: Analysis of Operators*, Academic Press Inc., 1978.

[37] D. W. Stroock, "An introduction to analysis on path space", in *Probability theory and applications*, IAS/Park City Mathematics Series, vol. 6, American Mathematical Society, 1999, p. 227-276.

[38] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, 2003.