ANALYSIS OF HIGH-DIMENSIONAL DISTRIBUTIONS USING PATHWISE METHODS

RONEN ELDAN

ABSTRACT

The goal of this note is to present an emerging method in the analysis of high-dimensional distributions, which exhibits applications to several mathematical fields, such as functional analysis, convex and discrete geometry, combinatorics, and mathematical physics. The method is based on pathwise analysis: One constructs a stochastic process, driven by Brownian motion, associated with the high-dimensional distribution in hand. Quantities of interest related to the distribution, such as covariance, entropy, and spectral gap, are then expressed via corresponding properties of the stochastic process, such as quadratic variation, making the former tractable through the analysis the latter. We focus on one particular manifestation of this approach, the *Stochastic Localization* process. We review several results which can be obtained using Stochastic Localization and outline the main steps towards their proofs. By doing so, we try to demonstrate some of the ideas and advantages of the pathwise approach. We focus on two types of results relevant to high-dimensional distributions: The first one has to do with dimension-free concentration bounds, manifested by functional inequalities which have no explicit dependence on the dimension. Our main focus in this respect will be on the Kannan–Lovász–Simonovits conjecture, concerning the isoperimetry of high-dimensional log-concave measures. Additionally, we discuss concentration inequalities for Ising models and expansion bounds for complex-analytic sets. The second type of results concern the decomposition of a high-dimensional measure into mixtures of measures attaining a simple structure, with applications to mean-field approximations.

MATHEMATICS SUBJECT CLASSIFICATION 2020

Primary 60-02; Secondary 60H30, 60G17, 60G44, 52A23

KEYWORDS

Concentration of measure, Isoperimetric inequalities, High dimensional phenomena, Mean-field models, Gaussian space

ICM INTERNATIONAL CONGRESS OF MATHEMATICIANS 2022 JULY 6—14

Proc. Int. Cong. Math. 2022, Vol. 6, pp. 4246–4270
DOI 10.4171/ICM2022/61

1. INTRODUCTION

This note is concerned with probability measures on high-dimensional spaces. The intuition derived from low-dimensional examples in various fields such as topology and partial differential equations may suggest that an attempt to understand the behavior of high-dimensional objects is futile, since a system's behavior quickly becomes complex and intractable as the dimension increases.

Nevertheless, a recently emerging theory of "high-dimensional phenomena" reveals that some important classes of distributions turn out to be surprisingly well-behaved (some introductory books on this theory are [2, 32, 43]). We focus on one particular facet of this theory which concerns *dimension-free* phenomena: It is often the case that the behavior of objects of interest is dictated by their marginals onto a fixed number of directions. This is manifested, for example, in the fact that several important functional inequalities have no explicit dependence on the dimension.

An exemplary illustration of this phenomenon is given by the *Gaussian isoperimetric inequality*. Consider the space $\mathbb{R}^n$ equipped with the standard Gaussian measure whose density is

$$\frac{d\gamma_n}{dx} := (2\pi)^{-n/2} \exp\bigl(-|x|^2/2\bigr),$$

which we refer to as the Gaussian space. A subset $H \subset \mathbb{R}^n$ is called a *half-space* if it has the form $H = \{x : \langle x, v\rangle \le b\}$ for some $v \in \mathbb{R}^n$ and $b \in \mathbb{R}$. For $A \subset \mathbb{R}^n$ and $\varepsilon > 0$, we define the ε-extension of A by $A_\varepsilon := \{x \in \mathbb{R}^n : \exists y \in A, \|y - x\|_2 \le \varepsilon\}$. Moreover, define

$$\Phi(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx,$$

the normal cumulative distribution function. The Gaussian isoperimetric inequality reads,

Theorem 1.1 (Borell, Sudakov–Tsirelson [8, 42]). *If $A \subset \mathbb{R}^n$ is a measurable set and $H \subset \mathbb{R}^n$ is a half-space satisfying $\gamma_n(A) = \gamma_n(H)$ then for all $\varepsilon > 0$ we have*

$$\gamma_n(A_\varepsilon) \ge \gamma_n(H_\varepsilon) = \Phi\bigl(\Phi^{-1}\bigl(\gamma_n(A)\bigr) + \varepsilon\bigr).$$

This theorem highlights an important metaproperty of Gaussian space: The extremizers of functional and geometric inequalities are one-dimensional objects, in the sense that they only depend on one direction. This is, for example, the case with the logarithmic-Sobolev inequality, Ehrhad's inequality, and Talagrand's transportation–entropy inequality (see [31, 32] for details). A recent breakthrough by Milman and Neeman [39] shows that the k-set analog of the isoperimetric inequality is saturated by partitions which only depend on $k - 1$ directions.

Is it reasonable to look for larger classes of measures which are Gaussian-like in the sense that they obey similar principles? Product measures are one natural candidate: By considering the harmonics, it is clear that several inequalities, such as the Poincaré inequality, will be saturated by one-dimensional functions. The central limit theorem ensures us that product distributions are Gaussian-like in the sense that, under mild conditions, marginals onto "typical" directions are close to a Gaussian.

In recent years, the class of measures which satisfy a convexity property, called *log-concave* measures, arose as another promising candidate. One remarkable result which supports this is Klartag's central limit theorem for convex sets [27], which asserts that typical one-dimensional marginals of such measures have an approximately normal law. In this note, we discuss the aspects of *isoperimetry* and *concentration of measure* in this class, in search of a counterpart to Theorem 1.1.

Log-concave measures and the Kannan–Lovász–Simonovits conjecture. A measure ν on $\mathbb{R}^n$ is *log-concave* if its density with respect to the Lebesgue measure is of the form $d\nu = e^{-V}dx$ where $V : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex. This class captures, for example, the Gaussian measure, as well as the uniform measure on a convex set.

For a set $A \subset \mathbb{R}^n$, define the surface area measure of A with respect to ν by

$$\nu^+(\partial A) = \limsup_{\varepsilon \to 0+} \frac{1}{\varepsilon} \nu(A_\varepsilon \setminus A). \tag{1.1}$$

(recalling that $A_\varepsilon = \{x \in \mathbb{R}^n : \exists y \in A, \|y - x\|_2 \le \varepsilon\}$). In analogy with the Gaussian isoperimetric inequality, we would like to obtain a lower bound on $\nu^+(\partial A)$ in terms of $\nu(A)$. This gives rise to the definition

$$\psi_\nu := \inf_{A \subset \mathbb{R}^n} \frac{\nu^+(\partial A)}{\nu(A)(1 - \nu(A))},$$

known as the *Cheeger constant* of the measure ν. Up to universal constants, ψ_ν^2 is equivalent to the Neumann spectral gap of ν, see [38]. It plays a central role in the theory of *concentration of measure* phenomena; a lower bound on ψ_ν implies, for example, that a Lipschitz function is typically close to its mean (see [32] and Section 2 below).

Since the problem is not scale invariant, it is hopeless to find a lower bound on ψ_ν that holds uniformly over all log-concave measures. Indeed, by considering the push-forward through the map $x \to \lambda x$ and replacing A by λA, the Cheeger constant scales as $\frac{1}{\lambda}$. We therefore need to assume that the measure is normalized in some way. A natural way to do so is to require that $\mathrm{Cov}(\nu) = \mathrm{Id}$ where $\mathrm{Cov}(\nu)$ is the covariance matrix of ν, defined by $\mathrm{Cov}(\nu)_{i,j} := \mathbb{E}_{X \sim \nu}[X_i X_j]$. A centered measure ν satisfying $\mathrm{Cov}(\nu) = \mathrm{Id}$ is called *isotropic*. It turns out that, as a consequence of the Brunn–Minkowski inequality, this normalization essentially corresponds to the fact that half-spaces satisfy an isoperimetric inequality.

Fact 1.2 (see [30, SECTION 2]). *Let ν by a log-concave measure in $\mathbb{R}^n$. Consider the quantity*

$$\alpha_\nu := \inf_{\substack{H \subset \mathbb{R}^n \\ \text{half-space}}} \frac{\nu^+(\partial H)}{\nu(H)(1 - \nu(H))}$$

(the difference from ψ_ν being that the infimum is only taken over half-spaces). Then,

$$\frac{1}{3}\alpha_\nu \le \left\|\mathrm{Cov}(\nu)\right\|_{\mathrm{OP}}^{-1/2} \le 3\alpha_\nu.$$

We are now ready to state the Kannan–Lovász–Simonivits conjecture.

Conjecture 1.3 (KLS conjecture, [25]). *There exists a universal constant $c > 0$ such that any isotropic, log-concave measure ν on $\mathbb{R}^n$ satisfies $\psi_\nu \ge c$.*

In light of Fact 1.2, the KLS conjecture equivalently asserts that for any log-concave measure,

$$c\alpha_\nu \leq \psi_\nu \leq \alpha_\nu$$

for a universal constant $c > 0$. In words, up to a constant independent of the measure or the dimension, the isoperimetric minimizer of any (not necessarily isotropic) log-concave measure is a half-space, hence the analogy to Theorem 1.1.

This conjecture has a wide array of implications in high-dimensional convex geometry and computational geometry, see [1, 33] for extensive reviews. Here, we only mention what is perhaps the most important of implications, a conjecture due to Bougain, known as the *hyperplane conjecture* or the *slicing problem*.

Conjecture 1.4. *There is a universal constant $c > 0$ such that, for every n and every convex $K \subset \mathbb{R}^n$ of unit volume, there exists an affine hyperplane H such that*

$$\mathrm{Vol}_{n-1}(K \cap H) > c. \tag{1.2}$$

For a survey on the hyperplane conjecture and other related problems, see [30]. Denote by $\psi_n = \inf_\nu \psi_\nu$ where the infimum is over all isotropic log-concave measures ν on $\mathbb{R}^n$ (so that the KLS conjecture states that $\psi_n \geq c$ for a universal constant $c > 0$), and by L_n the largest (possibly dimension-dependent) constant which can replace the constant c on the right-hand side of (1.2). It was shown by Klartag and the author [20], that $\psi_n \lesssim L_n$. In particular, Bourgain's hyperplane conjecture is implied by the KLS conjecture (see also [5]).

Let us briefly review some of the history around the KLS and hyperplane conjectures. In their original work, Kannan, Lovász, and Simonovits showed that $\psi_n \gtrsim n^{-1/2}$. The exponent $1/2$ was improved in several consecutive works, by Klartag [27] (relying on Bobkov [7]), by the author [16] (relying on Guédon–Milman [24]), and by Lee and Vempala [34], obtaining $\psi_n \gtrsim n^{-1/4}$. Regarding the hyperplane conjecture, Bourgain [12] showed that $L_n \gtrsim n^{-1/4} \log(n)^{-1}$. Until recently, the only improvement of this bound was $L_n \gtrsim n^{-1/4}$, due to Klartag [26].

A recent breakthrough by Chen makes a very significant improvement upon these bounds, nearly proving both conjectures.

Theorem 1.5 (Chen, [15]). *One has $\psi_n = n^{-o(1)}$. As a corollary, $L_n = n^{-o(1)}$.*

The pathwise approach and the Stochastic Localization scheme. Chen's proof is based on the so-called *Stochastic Localization* scheme, introduced in [16] and described in detail below. This scheme is one example of a more general metatechnique which we call pathwise analysis. In recent years, this metatechnique was proven useful in obtaining a variety of results that have to do with the analysis of high-dimensional distributions. The goal of this note is to highlight the main ideas behind it and review several applications thereof.

The use of ideas from diffusion and heat-flow to concentration inequalities dates back at least to the 1960s and to the seminal works of Nelson and Gross, which introduced the hypercontractivity property of heat semigroups and derived the log-Sobolev inequality for Gaussian space, respectively. In the following decades, heat flow (or semigroup) techniques

were realized to be a very powerful tool in proving concentration inequalities. These are, for example, the main ingredients in the celebrated Bakry–Emery theory [3]. These ideas rely on differentiation formulas for the heat semigroup which can alternatively be obtained via pathwise integration along the corresponding diffusion process.

The pathwise approach takes one more step and inspects the behavior of the process along a single path; it turns out that, when averaging over paths, quite a bit of information is lost, which can otherwise be revealed by using stochastic calculus. For example, bounds on the spectral gap and mixing times of diffusion processes can be obtained by coupling the paths of two diffusion processes. Some examples of works which manage to prove new bounds by direct analysis of the diffusion process are [10,11,13,36]. In this note we focus on a seemingly new type of pathwise proofs where, rather than considering the path of a diffusion process, one constructs an *evolution on the space of measures*, driven by Brownian motion, associated with a given distribution.

Structure of the paper. In what follows, in order to give an initial glimpse into pathwise techniques, in Section 2 we begin with a warm-up where we prove a concentration inequality for Lipschitz functions on Gaussian space using stochastic calculus. Then, in Section 3 we prove a generalization of the Gaussian isoperimetric inequality, due to Borell. In Section 4 we introduce the Stochastic Localization process and discuss the main ideas used in obtaining bounds for the KLS conjecture. Finally, in Section 5 we outline several other applications of Stochastic Localization towards (i) expansion bounds for complex-analytic sets, (ii) concentration inequalities for Ising models, and (iii) structure theorems which represent measures on the discrete hypercube as mixtures of product-like components.

2. A FIRST TASTE OF PATHWISE ANALYSIS: CONCENTRATION OF LIPSCHITZ FUNCTIONS IN GAUSSIAN SPACE

A useful property of Gaussian space, due to Maurey and Pisier, is the fact that Lipchitz functions have a sub-Gaussian tail:

Fact 2.1. *For any* 1*-Lipschitz function* $f : \mathbb{R}^n \to \mathbb{R}$*, we have*

$$\gamma_n\left(\left\{x : \left|f(x) - \int f d\gamma_n\right| > \alpha\right\}\right) \leq 4\big(1 - \Phi(\alpha)\big) \leq 2e^{-\alpha^2/2}, \quad \forall \alpha > 0.$$

This type of behavior is often referred to as *concentration of measure*. Put forth by V. Milman, such bounds have far-reaching applications and the behavior of this type is a cornerstone in the theory of high-dimensional phenomena (see, e.g., [41]). In this warm-up section, we provide a proof of this fact which will highlight some of the advantages of pathwise analysis. We assume that the reader has some familiarity with basic concepts in stochastic calculus.

Throughout the section, we fix a measurable function $f : \mathbb{R}^n \to \mathbb{R}$. Let $(B_t)_{t\geq 0}$ be a standard Brownian motion on $\mathbb{R}^n$; recall that B_1 has law γ_n. Consider the Doob martingale

$$M_t := \mathbb{E}\big[f(B_1) \mid B_t\big].$$

For a function $g : \mathbb{R}^n \to \mathbb{R}$, define $P_t[g](x) := \int_{\mathbb{R}^n} g(x + \sqrt{t}\, y)\gamma_n(dy)$. Since the law of B_1, conditioned on B_t, is $\mathcal{N}(B_t, (1-t)\mathrm{Id})$, we have

$$M_t = P_{1-t}[f](B_t). \tag{2.1}$$

Recall that, given a stochastic process $(X_s)_{s\geq 0}$, its *quadratic variation* is defined as

$$[X]_t := \lim_{\|\mathcal{P}=(t_0=0,t_1,\ldots,t_n=t)\|\to 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2,$$

where $\|\mathcal{P}\|$ denotes the mesh of the partition. Ito's isometry tells us that

$$\mathrm{Var}_{\gamma_n}[f] = \mathrm{Var}\big[f(B_1)\big] = \mathbb{E}[M]_1$$

(this is just the continuous version of the fact that for a discrete time martingale $X_0, X_1, \ldots,$ one has that $\mathrm{Var}[X_t] = \sum_{i=1}^{t} \mathbb{E}(X_i - X_{i-1})^2$).

In order to obtain a bound on $[M]_t$, using Itô's formula, we calculate

$$\begin{aligned} dM_t &\overset{(2.1)}{=} d\big(P_{1-t}[f](B_t)\big) \\ &= \big\langle \nabla P_{1-t}[f](B_t), dB_t \big\rangle + \frac{\partial}{\partial t} P_{1-t}[f](B_t)dt + \frac{1}{2}\Delta P_{1-t}[f](B_t)dt \\ &= \big\langle \nabla P_{1-t}[f](B_t), dB_t \big\rangle, \end{aligned}$$

where we used the identity $\frac{d}{ds}P_s[f] = \frac{1}{2}\Delta P_s[f]$. It follows that

$$\frac{d}{dt}[M]_t = \big\|\nabla P_{1-t}[f](B_t)\big\|_2^2 = \|V_t\|_2^2, \tag{2.2}$$

where $V_t := \nabla P_{1-t}[f](B_t)$. Since the operators ∇ and P_t commute, we also have

$$V_t = \mathbb{E}\big[\nabla f(B_1) \mid B_t\big],$$

which teaches us that V_t is a martingale. By the convexity of $\|\cdot\|_2^2$, we have that $\|V_t\|^2$ is a submartingale. We conclude that

$$\mathrm{Var}_{\gamma_n}[f] = \mathbb{E}[M]_1 = \int_0^1 \mathbb{E}\|V_s\|_2^2 ds \leq \mathbb{E}\|V_1\|_2^2 = \mathbb{E}_{\gamma_n}\|\nabla f\|_2^2.$$

This is precisely the Poincaré inequality. Alternatively, it can be easily proven using spectral methods. Moreover, instead of using Brownian motion, we could essentially repeat the argument directly using the semigroup P_t. Writing

$$\int f^2 d\gamma_n - \left(\int f d\gamma_n\right)^2 = P_1\big[f^2\big](0) - P_1[f](0)^2 = \int_0^1 \left(\frac{d}{dt} P_t\big[\big(P_{1-t}[f]\big)^2\big](0)\right) dt,$$

a simple calculation using integration by parts gives that

$$\frac{d}{dt} P_t\big[\big(P_{1-t}[f]\big)^2\big](0) = P_t\big[\big\|\nabla P_{1-t}[f]\big\|_2^2\big](0),$$

and an application of Jensen's inequality yields the Poincaré inequality.

Thus, in what we have seen so far, the "pathwise" aspect merely provides different viewpoint on a proof that can be carried out via elementary calculus. To see where it has a

real advantage, let us now assume that the function f is 1-Lipschitz. Under this assumption, with the help of Jensen's inequality, we learn that

$$\|V_t\|_2^2 = \big\|\nabla P_{1-t}[f](B_t)\big\|_2^2 \le P_{1-t}\big[\|\nabla f\|_2^2\big](B_t) \le 1, \quad \forall 0 \le t \le 1.$$

By equation (2.2), we have that $[M]_1 \le 1$ almost surely. Since M_1 has the same law as the push-forward of γ_n under f, Fact 2.1 follows as an immediate corollary of the following:

Proposition 2.2. *Let $(M_t)_{0\le t\le 1}$ be a martingale satisfying $[M]_1 \le 1$ almost surely. Then,*

$$\mathbb{P}\big(|M_1 - M_0| > \alpha\big) \le 4\big(1 - \Phi(\alpha)\big), \quad \forall \alpha > 0. \tag{2.3}$$

The key to the proof of this proposition is the Dambis/Dubins–Schwartz theorem which, roughly speaking, asserts that every continuous martingale can be represented as a time-changed Brownian motion. More formally, if M_t is a continuous martingale adapted to a filtration $\mathcal{F}_t$, then one can define a process $(W_t)_{t\ge 0}$ and a filtration $(\tilde{F}_t)_t$ over the same underlying probability space, such that:

(i) W_t is a Brownian motion with respect to the filtration $\tilde{F}_t$.

(ii) One has $M_t - M_0 = W_{[M]_t}$ and $\mathcal{F}_t = \tilde{\mathcal{F}}_{[M]_t}$ for all $t \ge 0$.

Next, we claim that $\tau := [M]_1$ is an $\tilde{\mathcal{F}}_t$-stopping time. Indeed, the claim that for all t, the event $\{\tau \le t\}$ is $\tilde{\mathcal{F}}_t$-measurable is equivalent to the claim that for all t the event $\{\tau \le [M]_t\}$ is $\mathcal{F}_t$-measurable, which is evident. Note that, by assumption, we have $\tau \le 1$ almost surely.

We finally conclude the following: There exists a Brownian motion W_t adapted to a filtration $\tilde{\mathcal{F}}_t$ and an $\tilde{\mathcal{F}}_t$-stopping time τ such that $\tau \le 1$ almost surely and such that W_τ is equal in law to $M_1 - M_0$. At this point we can write

$$\mathbb{P}(M_1 - M_0 \ge \alpha) = \mathbb{P}(W_\tau \ge \alpha) \le \mathbb{P}\big(\exists t \in [0,1] \text{ such that } W_t \ge \alpha\big).$$

The proof of Proposition 2.2 is now concluded via the following "reflection principle."

Fact 2.3. *Let W_t be a standard Brownian motion. Then, for all $\alpha > 0$, we have*

$$\mathbb{P}\big(\exists t \in [0,1] \text{ such that } W_t \ge \alpha\big) = 2\mathbb{P}(W_1 \ge \alpha) = 2\big(1 - \Phi(\alpha)\big).$$

Proof. (sketch) Consider the stopping time $\tau = \inf\{t; W_t = \alpha\}$. Since, conditioned on $\tau \le 1$, we have that $W_1 - W_\tau$ has a symmetric law, we have $\mathbb{P}(W_1 \ge W_\tau | \tau \le 1) = \frac{1}{2}$. ∎

Fact 2.1 may be alternatively proven by combining the Gaussian isoperimetric inequality and the coarea formula, or by a direct coupling argument (see **[41, THEOREM 2.2]**). Nevertheless, the above proof highlights the advantage in considering the martingale M_t in a "path-by-path" manner, and the reason that this approach can reveal dimension-free phenomena: The process V_t extracts the "important" directions, in which the function f varies, and the law of $f(B_1)$ eventually only depends on the behavior of the *one-dimensional* process M_t. The reduction of the analysis of an n-dimensional function, or measure, to the behavior one-dimensional process will be a recurring motif later on.

3. THE GAUSSIAN ISOPERIMETRIC INEQUALITY AND NOISE-SENSITIVITY

As a next step towards demonstrating the pathwise technique, we provide a proof of the Gaussian isoperimetric inequality, Theorem 1.1. In fact, we prove a stronger statement, known as Borel's noise stability inequality [9].

Let B_t be a standard Brownian motion in $\mathbb{R}^n$, adapted to a filtration $\mathcal{F}_t$. Define $Z_t := \int_0^t e^{-s/2} dB_s$. Observe that $Z_\infty := \lim_{t\to\infty} Z_t$ has the law γ_n, since $\int_0^\infty e^{-t} dt = 1$. For all measurable $A \subset \mathbb{R}^n$ and $t > 0$, define

$$\mathrm{Sens}_t(A) := \mathbb{E}\big[\mathbb{P}(Z_\infty \in A \mid Z_t)\mathbb{P}(Z_\infty \notin A \mid Z_t)\big],$$

referred to as the t-noise sensitivity of A. From an analytic point of view, this quantity can be understood as the rate at which heat escapes the set A under the heat flow on Gaussian space, defined by the Ornstein–Uhlenbeck operator $\mathcal{L} = \Delta - x \cdot \nabla$.

A standard argument shows that noise-sensitivity is related to isoperimetry by

$$\gamma_n^+(\partial A) = \lim_{t\to 0} \frac{\mathrm{Sens}_t(A)}{\sqrt{t}}, \tag{3.1}$$

which holds, for example, under the assumption that A has finite perimeter.

Theorem 3.1 (Borell [9]). *If $A \subset \mathbb{R}^n$ is a measurable set and $H \subset \mathbb{R}^n$ is a half-space satisfying $\gamma_n(A) = \gamma_n(H)$, then for all $t \geq 0$,*

$$\mathrm{Sens}_t(A) \geq \mathrm{Sens}_t(H).$$

This theorem has far-reaching applications in statistics and theoretical computer science which we do not discuss here, but we refer the reader to [17,40] and references therein. Combined with equation (3.1), Theorem 1.1 follows as a corollary.

Towards proving Theorem 3.1, define for a set $A \subset \mathbb{R}^n$,

$$b(A) := \int_A x \gamma_n(dx),$$

the Gaussian first-moment of A. Moreover, we define for $s \in \mathbb{R}$,

$$q(s) = \int_{\Phi^{-1}(s)}^\infty t \gamma_1(dt).$$

Evidently, if $H \subset \mathbb{R}^n$ is a half-space then one has $\|b(H)\|_2 = q(\gamma_n(H))$. At the center of our proof lies the following simple fact.

Fact 3.2 (Level-1 inequality). *For any measurable $A \subset \mathbb{R}^n$,*

$$\big\|b(A)\big\|_2 \leq q\big(\gamma_n(A)\big), \tag{3.2}$$

with equality when A is a half-space.

This fact is referred to as the *level-1 inequality* since it characterizes the sets which maximize the L_2-energy on the first-order Hermite expansion. It constitutes the only inequality in the proof to come.

Proof. Set $\theta = \frac{b(A)}{\|b(A)\|_2}$. Let H be a half-space of the form $H = \{x; \langle x, \theta\rangle \geq \alpha\}$ with α chosen so that $\gamma_n(H) = \gamma_n(A)$. Note that, by definition,

$$q\big(\gamma_n(A)\big) = \left\| \int_H x\gamma_n(dx) \right\|_2 = \int_H \langle x, \theta\rangle \gamma_n(dx),$$

so we only need to show that

$$\int_H \langle x, \theta\rangle \gamma_n(dx) \geq \int_A \langle x, \theta\rangle \gamma_n(dx).$$

Since $\gamma_n(A) = \gamma_n(H)$, we may subtract α from both integrands, thus the above is equivalent to

$$\int_{\mathbb{R}^n} \big(\langle x, \theta\rangle - \alpha\big)(\mathbf{1}_{\langle x,\theta\rangle \geq \alpha} - \mathbf{1}_{x\in A})\gamma_n(dx) \geq 0,$$

which is evident. ∎

Define μ_t to be the law of Z_∞ conditioned on Z_t, which easily checked to be $\mathcal{N}(Z_t, e^{-t/2}\mathrm{Id})$, or in other words,

$$\mu_t(dx) = (2\pi)^{-n/2} e^{nt/2} \exp\left(-\frac{1}{2}e^t |x - Z_t|^2\right) dx.$$

Set $M_t := \mu_t(A) = \mathbb{P}(Z_\infty \in A \mid Z_t)$. Note that, by definition,

$$\mathrm{Sens}_t(A) = \mathbb{E}\big[M_t(1 - M_t)\big] = M_0(1 - M_0) - \mathrm{Var}[M_t]. \tag{3.3}$$

Consider a half-space H satisfying $\gamma_n(H) = \gamma_n(A)$ and, analogously, define $N_t = \mu_t(H)$. Equation (3.3) tells us that the statement of Theorem 3.1 is equivalent to the assertion that

$$\mathrm{Var}[M_t] \leq \mathrm{Var}[N_t], \quad \forall t > 0. \tag{3.4}$$

In order to compare the variances of the two processes, we first calculate the corresponding quadratic variations. Using Itô's formula, we write

$$\begin{aligned} dM_t &= d \int_A (2\pi)^{-n/2} e^{nt/2} \exp\left(-\frac{1}{2}e^t \|x - Z_t\|_2^2\right) dx \\ &= e^t \int_A \langle x - Z_t, dZ_t\rangle \mu_t(dx) \\ &= \big\langle b(A_t), dB_t \big\rangle, \end{aligned} \tag{3.5}$$

where $A_t := e^{t/2}(A - Z_t)$. Observing that $M_t = \mu_t(A) = \gamma_n(A_t)$, with the help of (3.2), we arrive at the inequality

$$\frac{d}{dt}[M]_t = \big\|b(A_t)\big\|_2^2 \leq q(M_t)^2. \tag{3.6}$$

Defining $H_t := e^{t/2}(H - Z_t)$, a similar calculation shows

$$\frac{d}{dt}[N]_t = \big\|b(H_t)\big\|_2^2 = q(N_t)^2. \tag{3.7}$$

On an intuitive level, equations (3.6) and (3.7) tell us that, in a certain sense, the martingale N_t is moving faster than M_t. Naively, we might hope that the above implies that $\mathbb{E}\frac{d}{dt}[N]_t \geq \mathbb{E}\frac{d}{dt}[M]_t$ for all t, which would conclude (3.4). This is not true, however.

Observe that $\mathbb{E}[N]_\infty = \mathrm{Var}[N_\infty] = \mathrm{Var}[M_\infty] = \mathbb{E}[M]_\infty$. The following lemma extracts the power of the pathwise approach. We can couple the two processes in a way that gives us the desired domination.

Lemma 3.3. *Let $v : \mathbb{R} \to [0, \infty)$ be a continuous function. Let $(M_t)_{t=0}^\infty$, $(N_t)_{t=0}^\infty$ be two continuous real-valued martingales such that $M_0 = N_0$, and such that*

$$\frac{d}{dt}[N]_t = v(N_t) \quad \text{and} \quad \frac{d}{dt}[M]_t \le v(M_t), \tag{3.8}$$

almost surely, for all $t \ge 0$. Then for all $t \ge 0$, one has

$$\mathrm{Var}[M_t] \le \mathrm{Var}[N_t].$$

Since equations (3.6) and (3.7) verify (3.8), an application of the above lemma yields (3.4), which concludes the proof of Theorem 3.1. It therefore only remains to prove this lemma.

Proof of Lemma 3.3 (sketch). Without loss of generality, assume $M_0 = N_0 = 0$. By the Dambis/Dubins–Schwartz theorem, there exist standard Brownian motions $B_t, \tilde{B}_t$ such that $N_t = B_{[N]_t}$ and $M_t = \tilde{B}_{[M]_t}$. By a standard disintegration theorem, the processes maybe defined on the same probability space in a way that $B_t = \tilde{B}_t$. In other words, there exist two martingales X_t, Y_t and a standard Brownian motion B_t, defined over the same probability space, such that X_t, Y_t have the same laws as N_t, M_t, respectively, and such that

$$X_t = B_{[X]_t} \text{ and } Y_t = B_{[Y]_t}, \quad \forall t \ge 0.$$

Let τ_X, τ_Y be the inverse functions of $[X]_t, [Y]_t$, respectively. Then the last display implies $X_{\tau_X(T)} = Y_{\tau_Y(T)} = B_T$, and by formula (3.8) we have

$$\frac{d}{dt}[X]_t|_{t=\tau_X(T)} = v(B_T) \ge \frac{d}{dt}[Y]_t|_{t=\tau_Y(T)}, \quad \forall T \ge 0,$$

which implies that $[X]_t \ge [Y]_t$ for all $t \ge 0$. By Itô's isometry, the lemma follows. ∎

4. STOCHASTIC LOCALIZATION AND THE KLS CONJECTURE

In this section we introduce the main technique discussed in this note, the Stochastic Localization process, and demonstrate how it can be used to produce lower bounds on the Cheeger constant of a log-concave measure.

4.1. Construction of the process and basic properties

Let B_t be a standard Brownian motion in $\mathbb{R}^n$, adapted to a filtration $\mathcal{F}_t$. As in the previous section, define $Z_t := \int_0^t e^{-s/2} dB_s$ and let μ_t be defined as the law of Z_∞ conditioned on Z_t. The measure-valued process (μ_t) interpolates between the standard Gaussian measure, at time 0, and a Dirac measure at time ∞. A key formula in the previous section was (3.5), which can be restated as follows: Setting $p_t(x) := \frac{\mu_t(dx)}{dx}$, we have

$$\forall x \in \mathbb{R}^n, \quad dp_t(x) = e^{t/2} p_t(x) \left\langle x - \int x p_t(x) dx, dB_t \right\rangle. \tag{4.1}$$

Now, let ν be an arbitrary probability measure on $\mathbb{R}^n$. We would like to consider a similar evolution with ν taking the place of the Gaussian measure. Suppose that $(C_t)_{t\geq 0}$ is a stochastic process adapted to $\mathcal{F}_t$, such that for all t, C_t is an $n \times n$ positive semidefinite matrix. Inspired by (4.1), consider the system of stochastic differential equations

$$\forall x \in \mathbb{R}^n, \quad F_0(x) = 1, \quad dF_t(x) = F_t(x)\langle x - a_t, C_t dB_t\rangle, \tag{4.2}$$

where

$$a_t := \int x F_t(x)\nu(dx).$$

We can now define a measure-valued process, $(\nu_t)_{t\geq 0}$, by $\nu_t(dx) = F_t(x)\nu(dx)$. Note that $\nu_0 = \nu$. The choice $\nu = \gamma_n$ and $C_t = e^{t/2}\mathrm{Id}$ recovers the evolution defined by (4.1), so the process ν_t can be thought of as a generalization of μ_t.

We remark that the system (4.2) is an infinite system of stochastic differential equations, but as we will see below, it may instead be written as a finite system. Its existence and uniqueness is proven in [16]. Informally, we can think of equation (4.2) as

$$F_{t+dt}(x) = F_t(x)\big(1 + \big\langle x - a_t, \mathcal{N}\big(0, C_t^2 dt\big)\big\rangle\big),$$

so that process can be understood as a continuous version the following iterative procedure: Start with a density on $\mathbb{R}^n$, and at each iteration multiply this density by a linear function, which is equal to 1 at the center of mass of ν_t, and whose gradient is distributed according to an infinitesimal Gaussian.

Before we continue, let us point out several basic properties of this process. First, using Itô's formula, we calculate

$$d\log F_t(x) = \frac{dF_t(x)}{F_t(X)} - \frac{d[F(x)]_t}{2F_t(x)^2} \overset{(4.2)}{=} \langle x - a_t, C_t dB_t\rangle - \frac{1}{2}\big\|C_t(x - a_t)\big\|_2^2 dt.$$

Consequently, the measure ν_t attains the form

$$\nu_t(dx) = \exp\left(z_t + \langle v_t, x\rangle - \frac{1}{2}\langle G_t x, x\rangle\right)\nu(dx), \tag{4.3}$$

with $G_t := \int_0^t C_s^2 ds$, where $v_t \in \mathbb{R}^n$ is an Itô process adapted to $\mathcal{F}_t$ and z_t is a normalizing constant. In particular, if we choose $C_t = \mathrm{Id}$ for all t, we have

$$\nu_t(dx) = \exp\left(z_t + \langle v_t, x\rangle - \frac{t}{2}\|x\|_2^2\right)\nu(dx). \tag{4.4}$$

Next, we calculate

$$d\nu_t\big(\mathbb{R}^n\big) = \left\langle \int_{\mathbb{R}^n} C_t(x - a_t)F_t(x)\nu(dx), dB_t\right\rangle = 0.$$

Equation (4.3) shows that $F_t(x)$ is positive for all x and t, so we conclude that ν_t is almost surely a probability measure for all t and that a_t is its center of mass. Finally, it is evident from (4.2) that $F_t(x)$ is a martingale for every x, which immediately gives the following.

Fact 4.1. *For every measurable $W \subset \mathbb{R}^n$, the process $\nu_t(W)$ is a martingale.*

4.2. Isoperimetry for log-concave measures using Stochastic Localization

Fix a log-concave measure ν on $\mathbb{R}^n$ and a measurable set $A \subset \mathbb{R}^n$. We would like to use the process constructed above in order to produce a lower bound on $\nu^+(\partial A)$. Consider the process $(\nu_t)_t$ defined in (4.2) with the choice $C_t = \mathrm{Id}$.

Applying Fact 4.1 to the set $A_\varepsilon \setminus A$ gives

$$\nu^+(\partial A) = \mathbb{E}\nu_t^+(\partial A). \tag{4.5}$$

To continue the analogy with Section 3, we consider the martingale $M_t := \nu_t(A)$. Recall that the proof of the Gaussian isoperimetric inequality amounted to obtaining an upper bound on $\frac{d}{dt}[M]_t$, in terms of M_t, for all $t \geq 0$. In our case, we will only be able to establish such a bound for small enough values of t. This will be complemented by the fact that for large t, the measure ν_t satisfies an isoperimetric inequality, a consequence of the following:

Theorem 4.2 (Bakry–Ledoux [4]; see also [33, THEOREM 25]). *Let μ be a probability measure on $\mathbb{R}^n$ whose density is of the form*

$$d\mu(x) = e^{-V(x) - \frac{\alpha}{2}\|x\|_2^2} dx \tag{4.6}$$

where $V(x) : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex and $\alpha > 0$. Then for all $A \subset \mathbb{R}^n$, we have

$$\mu^+(\partial A) \geq \sqrt{\alpha}\mu(A)\big(1 - \mu(A)\big). \tag{4.7}$$

We sketch an alternative proof of this theorem in Section 5.1. In light of (4.4), we may apply the theorem to the measure ν_t with $\alpha = t$, which yields

$$\begin{aligned} \nu^+(\partial A) &\overset{(4.5)}{=} \mathbb{E}\nu_t^+(\partial A) \\ &\overset{(4.4)+(4.7)}{\geq} \mathbb{E}\big[\sqrt{t}\nu_t(A)\big(1 - \nu_t(A)\big)\big] \\ &= \sqrt{t}\big(M_0(1 - M_0) - \mathrm{Var}[M_t]\big). \end{aligned} \tag{4.8}$$

As in Section 3, our goal is once again to bound from above the quantity $\mathrm{Var}[M_t] = \mathbb{E}[M]_t$. To this end, we calculate

$$dM_t = d\int_A F_t(x)\nu(dx) \overset{(4.2)}{=} \int_A \langle x - a_t, dB_t\rangle \nu_t(dx), \tag{4.9}$$

implying that

$$\frac{d}{dt}[M]_t = \left\| \int_A (x - a_t)\nu_t(dx) \right\|_2^2. \tag{4.10}$$

The right-hand side can be bounded with the help of the following simple lemma.

Lemma 4.3. *For every probability measure μ on $\mathbb{R}^n$ and every measurable $A \subset \mathbb{R}^n$,*

$$\left\| \int_A \left(x - \int_{\mathbb{R}^n} x\mu(dx) \right) \mu(dx) \right\|_2^2 \leq \big\|\mathrm{Cov}(\mu)\big\|_{\mathrm{OP}}. \tag{4.11}$$

Proof. Define $\theta := \frac{\int_A x\mu(dx)}{\|\int_A x\mu(dx)\|_2}$ (if the denominator vanishes, there is nothing to show). Also, without loss of generality assume $\int_{\mathbb{R}^n} x\mu(dx) = 0$. Then we have

$$\left\|\int_A x\mu(dx)\right\|_2^2 = \left(\int_{\mathbb{R}^n} \langle x,\theta\rangle \mu(dx)\right)^2 \leq \int_{\mathbb{R}^n} \langle x,\theta\rangle^2 \mu(dx) = \left\langle \theta, \mathrm{Cov}(\mu)\theta\right\rangle \leq \left\|\mathrm{Cov}(\mu)\right\|_{\mathrm{OP}}. \qquad \blacksquare$$

In the process μ_t considered in Section 3 (which corresponds to ν_t, only with the measure ν replaced by the Gaussian measure), the matrix $\mathrm{Cov}(\mu_t)$ was deterministic. The crucial difference here is that we have to account for $\|\mathrm{Cov}(\nu_t)\|_{\mathrm{OP}}$.

Combining (4.10) and (4.11), we have

$$\mathrm{Var}[M_t] = \mathbb{E}[M]_t \leq \int_0^t \left\|\mathrm{Cov}(\nu_s)\right\|_{\mathrm{OP}} ds.$$

Together with (4.8), the state of events can be concluded by the following proposition.

Proposition 4.4. *Let ν by a log-concave measure on $\mathbb{R}^n$. Construct the process $(\nu_t)_{t=0}^{\infty}$ using equation (4.2). Suppose that for some $t, \alpha > 0$, one has*

$$\mathbb{E}\left[\int_0^t \left\|\mathrm{Cov}(\nu_s)\right\|_{\mathrm{OP}} ds\right] \leq \alpha. \tag{4.12}$$

Then, for every $A \subset \mathbb{R}^n$ such that $\nu(A)(1-\nu(A)) \geq 2\alpha$, we have the Cheeger-type inequality

$$\nu^+(\partial A) \geq \frac{1}{2}\sqrt{t}\,\nu(A)\big(1-\nu(A)\big).$$

The condition $\nu(A)(1-\nu(A)) \geq 2\alpha$ is not crucial; one can show that it may, in fact, be ignored (see **[38, THEOREM 1.8]**), so that a bound of the form (4.12) implies $\psi_\nu \geq \frac{1}{2}\sqrt{t}$. Our goal will therefore be to establish (4.12).

Note that condition (4.12) does not involve the set A at all. In that sense, we have managed to reduce a statement with a quantifier "for every $A \subset \mathbb{R}^n$," to a bound which involves only the measure ν. We now need to produce an upper bound on $\|\mathrm{Cov}(\nu_s)\|_{\mathrm{OP}}$. The process $\mathrm{Cov}(\nu_s)$ becomes tractable thanks to a "moment-generating" property described in the next subsection.

Before we proceed, let us note the trade-off between two conflicting goals, namely controlling from above the variance of $\nu_t(A)$ (which corresponds to taking t small enough), and controlling from below the "uniform concavity" term in $\log \frac{d\nu_t}{d\nu}$ (which corresponds to taking t large enough). It is reasonable to expect that a clever choice of the matrix C_t could be fruitful: For a general choice of C_t, equation (4.9) becomes

$$d\nu_t(A) = \left\langle C_t b_t(A), dB_t\right\rangle, \quad \text{where } b_t(A) := \int_A (x - a_t)\nu_t(dx). \tag{4.13}$$

Equations (4.3) and (4.13) suggest that the choice of the driving matrix C_t allows a more intricate control of this trade-off. On the one hand, by taking C_t to be small in the direction of b_t, we gain more control of the variance of $\nu_t(A)$, but, on the other hand, we would like the matrix $\int_0^t C_s^2 ds$ to be large. In the context of the KLS conjecture, it is not known if this strategy can produce better bounds, however, in Section 5 we give several examples which crucially rely on a careful choice of C_t.

4.2.1. Stochastic Localization as a moment-generating process

Recall that a_t is the center of mass of ν_t. A calculation shows that

$$\begin{aligned} da_t &= d\int_{\mathbb{R}^n} x\nu_t(dx) \\ &\overset{(4.2)}{=} \int_{\mathbb{R}^n} x\langle x-a_t, C_t dB_t\rangle F_t(x)\nu(dx) \\ &= \left(\int_{\mathbb{R}^n} x\otimes(x-a_t)\nu_t(dx)\right)C_t dB_t = \operatorname{Cov}(\nu_t)C_t dB_t. \end{aligned} \tag{4.14}$$

In words, the time-differential of the first cumulant of ν_t is equal to its second cumulant multiplied by the generating increment $C_t dB_t$. A calculation of similar spirit gives

$$d\operatorname{Cov}(\nu_t) = \mathcal{M}^{(3)}[\nu_t]C_t dB_t - \operatorname{Cov}(\nu_t)C_t^2\operatorname{Cov}(\nu_t)dt, \tag{4.15}$$

where

$$\mathcal{M}^{(k)}[\nu_t] := \int (x-a_t)^{\otimes k}\nu_t(dx)$$

is the kth moment tensor of ν_t. In general, the time differential of $\mathcal{M}^{(k)}[\nu_t]$ will involve the term $\mathcal{M}^{(k+1)}[\nu_t]C_t dB_t$.

This property is reminiscent of the *logarithmic Laplace* transform, where derivatives with respect to the space parameter correspond to cumulants of a tilted measure. This fact has far-reaching applications in asymptotic geometric analysis, notably it has been used in several works of Klartag, in particular in his breakthrough on the slicing problem [26].

4.2.2. Obtaining a bound for the KLS conjecture

We now give an overview of the next steps needed to obtain a bound for the KLS conjecture. Going back to Proposition 4.4, such a bound is reduced to obtaining upper bounds on the growth of $\|\operatorname{Cov}(\nu_t)\|_{\mathrm{OP}}$. According to equation (4.15), the expression for the differential of $\operatorname{Cov}(\nu_t)$ involves the process $\mathcal{M}^{(3)}[\nu_t]$.

First let us consider a simple (but somewhat wasteful) way to obtain a bound for $\|\operatorname{Cov}(\nu_t)\|_{\mathrm{OP}}$, based on the fact that

$$\left\|\operatorname{Cov}(\nu_t)\right\|_{\mathrm{OP}} \le \operatorname{Tr}\left(\operatorname{Cov}(\nu_t)^2\right).$$

Equation (4.15) combined with Itô's formula gives

$$\frac{d}{dt}\mathbb{E}\operatorname{Tr}\left(\operatorname{Cov}(\nu_t)^2\right) \le \mathbb{E}\left\|\mathcal{M}^{(3)}[\nu_t]\right\|_{\mathrm{HS}}^2. \tag{4.16}$$

The quantity on the right-hand side involves third moments of the measure ν_t. On a conceptual level, at this point, the state of events is that we have the implications:

$$\text{Upper bound on } \mathcal{M}^{(3)}(\nu_t) \Rightarrow \text{Upper bound on } \left\|\operatorname{Cov}(\nu_t)\right\|_{\mathrm{OP}} \Rightarrow \text{Lower bound on } \psi_\nu. \tag{4.17}$$

One way to continue from here would be to look for bounds on $\mathcal{M}^{(3)}(\mu)$ in terms of $\operatorname{Cov}(\mu)$ which hold universally over all log-concave measures μ on $\mathbb{R}^n$. This would imply that the

rate of growth of $\mathrm{Cov}(\nu_t)$ is bounded by $\mathrm{Cov}(\nu_t)$ itself. Following this route, the work [16] used a priori estimates on the "thin-shell" constant, defined as

$$\sigma_n := \sup_{\mu} \mathrm{Var}_{X\sim\mu} \|X\|_2,$$

where the supremum runs over all isotropic, log-concave probability measures μ on $\mathbb{R}^n$. Upper bounds on σ_n imply the type of bounds on $\mathcal{M}^{(3)}[\nu_t]$ which, when plugged into the implications above, give a reduction, up to a logarithmic factor, from the KLS conjecture to a weaker conjecture called the *variance conjecture* stating that $\sigma_n = O(1)$ (see [1, 16]).

Later on, Lee and Vempala ([35, LEMMA 33]) realized that, when taking the driving matrix C_t to be the identity, one could instead use the bound

$$\mathbb{E}\left\|\mathcal{M}^{(3)}[\mu]\right\|_{\mathrm{HS}}^2 \lesssim \mathrm{Tr}\left(\mathrm{Cov}(\mu)^2\right)^{3/2} \tag{4.18}$$

which holds uniformly for all log-concave measures. Together with equation (4.16) and an application of Gronwall's inequality, this gives that $\mathbb{E}\,\mathrm{Tr}(\mathrm{Cov}(\nu_t)^2) = O(n)$ for all $t \le \frac{1}{\sqrt{n}}$. Plugging this into Proposition 4.4 yields the bound $\psi_\nu \gtrsim n^{-1/4}$.

Let us now briefly discuss the additional steps needed to produce Chen's bound, $\psi_\nu = n^{-o(1)}$. First of all, by rather direct arguments, one can reverse implication (4.17) in the sense that

$$\begin{array}{c}\text{Lower bound on } \psi_\mu \\ \text{for all } \mu \text{ isotropic, log-concave}\end{array} \quad \Rightarrow \quad \begin{array}{c}\text{Improved upper bounds on } \mathcal{M}^{(3)}[\mu] \\ \text{in terms of } \mathrm{Cov}(\mu).\end{array} \tag{4.19}$$

In other words, if we have a priori bounds for the KLS conjecture we can improve, in some sense, on (4.18). Lee and Vempala speculated that the implications (4.17) and (4.19) can be chained in a way that "bootstraps" an a priori bound on the KLS constant yield a better bound, however, they were not able to successfully implement this strategy.

Chen added another important ingredient to the mix: In light of equation (4.4), we know that for large enough t, the measure ν_t is not only log-concave, but is t-uniformly log-concave in the sense of (4.6). According to Theorem 4.2, it has concentration properties which do not a priori hold for log-concave measures. The main strategy is then to split the interval $[0, t]$, in Proposition 4.4, into two intervals: In the first interval, the "bootstrap" bound on ψ_n is used, whereas in the second, he manages to leverage on the uniform log-concavity of ν_t in order to find a version of the implication (4.19) which gives a yet stronger lower bound on ψ_n, thereby closing the implication circle of (4.17) and (4.19).

5. DECOMPOSITION OF MEASURES AND FURTHER APPLICATIONS OF STOCHASTIC LOCALIZATION

In this section we describe several additional applications of the Stochastic Localization process. Common to these applications is that they rely on the fact that this process gives rise to a decomposition scheme which expresses a given measure as a *mixture*. We fix a measure ν on $\mathbb{R}^n$ and use the same notation as in Section 4.1.

Recall that, according to Fact 4.1, for every fixed measurable set $A \subset \mathbb{R}^n$, the process $\nu_t(A)$ is a martingale, which implies in particular that $\nu = \mathbb{E}\nu_t$. More generally, the optional

stopping theorem implies that, for every $\mathcal{F}_t$-stopping time τ, one has $\nu = \mathbb{E}\nu_\tau$. Therefore, every such stopping time induces a *decomposition* of the measure ν, in the sense that it can be viewed as a mixture whose components are the measures ν_τ.

To summarize, every stopping time τ may be associated with a probability measure $m = m_\tau$ on an abstract index set $\mathcal{I}$, and every $\alpha \in \mathcal{I}$ may be associated with a probability measure ν_α on $\mathbb{R}^n$, so that

$$\nu(W) = \int_{\mathcal{I}} \nu_\alpha(W) m(d\alpha), \quad \forall W \subset \mathbb{R}^n \text{ measurable.} \tag{5.1}$$

Above, the random measure ν_α with $\alpha \sim m$ has the same distribution as ν_τ where ν_t is defined by equations (4.2). In the next sections, we review several applications of this decomposition.

5.1. Needle decompositions

At the heart of the argument found in the original paper of Kannan, Lovász, and Simonovits [25] lies a procedure that takes the uniform measure on a convex set in $\mathbb{R}^n$ and represents it as a decomposition into measures whose support is contained in a one-dimensional affine subspace (referred to as "needles"). This was done using an iterative scheme which repeatedly cuts the set via hyperplane bisections which preserve the relative volume of the set A. This type of scheme, referred to as "localization" generalizes an earlier lemma by Lovász and Simonovits [37], and is based on ideas going back to Gromov and Milman [23]. Klartag [28] gives a somewhat canonical construction which generalizes this concept to Riemannian manifolds (where needles are supported on geodesics).

We will not discuss the aforementioned localization schemes in detail here. Instead, we describe an alternative way to obtain a "needle decomposition" (hence, a decomposition into measures with one-dimensional support) for a prescribed measure on $\mathbb{R}^n$, using the (generalized) stochastic localization equations. We first demonstrate this by outlining an alternative proof of the Gaussian isoperimetric inequality, as well as to Theorem 4.2.

Take $\nu = \gamma_n$ and consider the process generated by (4.2). The main idea is the following one: In view of equation (4.13), by choosing $C_t = \mathrm{Proj}_{b_t^\perp}$, we have that $\nu_t(A)$ remains constant along the process. By doing so, we obtain a decomposition of γ_n into measures which satisfy $\nu_\alpha(A) = \gamma_n(A)$. Next, we will argue that as $t \to \infty$, we obtain a decomposition into measures with one-dimensional support.

Denote $G_t = \int_0^t C_s^2 ds$. Since C_t is a projection matrix of codimension 1, we have $\mathrm{Tr}(G_t) = (n-1)t$ and $\|G_t\|_{\mathrm{OP}} \le t$. This implies that all but one of the eigenvalues of G_t are at least $t/2$. According to (4.3), the measure ν_t is a Gaussian measure whose covariance matrix $\mathrm{Cov}(\nu_t)$ converges, as $t \to \infty$, to a matrix M of rank at most 1. For all $A \subset \mathbb{R}^n$, define $\nu_\infty(A) = \lim_{t\to\infty} \nu_t(A)$ (the limit exists by the martingale convergence theorem). It is straightforward to show that ν_∞ is σ-additive and therefore a probability measure, and in fact, it is a Gaussian measure whose covariance is $\lim_{t\to\infty} \mathrm{Cov}(\nu_t)$. Moreover, since $\nu(A) = \mathbb{E}\nu_t(A)$ for all t and $A \subset \mathbb{R}^n$, by taking limits we have that $\nu(A) = \mathbb{E}\nu_\infty(A)$.

In light of equation (5.1) (taking $\tau = \infty$), we arrive at the following lemma.

Lemma 5.1. *For every measurable $A \subset \mathbb{R}^n$, there exist a probability measure m on an index set $\mathcal{I}$ and, for every $\alpha \in \mathcal{I}$, a probability measure ν_α on $\mathbb{R}^n$ such that*

$$\gamma_n(W) = \int_{\mathcal{I}} \nu_\alpha(W) m(d\alpha), \quad \forall W \subset \mathbb{R}^n \text{ measurable}. \tag{5.2}$$

Moreover, for every $\alpha \in \mathcal{I}$, the measure ν_α has a Gaussian law with covariance matrix C_α such that (i) $\mathrm{rank}(C_\alpha) = 1$, (ii) $\|C_\alpha\|_{\mathrm{OP}} \leq 1$, *and* (iii) $\nu_\alpha(A) = \gamma_n(A)$.

We now use this decomposition to show that the n-dimensional Gaussian measure "inherits" the isoperimetric properties of the one-dimensional Gaussian measure. Indeed, assuming the bound

$$\gamma_1^+(\partial W) \geq \mathrm{I}\big(\gamma_1(W)\big), \quad \forall W \subset \mathbb{R} \text{ measurable}, \tag{5.3}$$

for some function $\mathrm{I} : [0,1] \to [0,\infty)$, and given any measurable set $A \subset \mathbb{R}^n$, we can find a decomposition of γ_n as in (5.2) such that every ν_α is a one-dimensional Gaussian measure of variance at most 1 and $\nu_\alpha(A) = \gamma_n(A)$. We get that

$$\gamma_n(A_\varepsilon \setminus A) \overset{(5.2)}{=} \int_{\mathcal{I}} \nu_\alpha(A_\varepsilon \setminus A) m(d\alpha) \geq \int_{\mathcal{I}} \nu_\alpha\big(\big(A \cap \mathrm{Supp}(\nu_\alpha)\big)_\varepsilon \setminus A\big) m(d\alpha).$$

By taking limits,

$$\gamma_n^+(\partial A) \geq \int_{\mathcal{I}} \nu_\alpha^+(\partial A) m(d\alpha) \overset{(5.3)}{\geq} \int_{\mathcal{I}} \mathrm{I}\big(\nu_\alpha(A)\big) m(d\alpha) = \mathrm{I}\big(\gamma_n(A)\big).$$

We have therefore reduced the proof of the Gaussian isoperimetric inequality in dimension n to the same inequality in dimension 1.

If ν has density of the form $d\nu = \exp(-V(x) - \alpha|x|^2)dx$ with $\alpha > 0$ and $V : \mathbb{R}^n \to \mathbb{R}$ convex, then the same procedure gives rise to a decomposition into one-dimensional needles whose potential exhibits uniform convexity of a similar form. Thus an analogous argument gives a reduction of Theorem 4.2 to the one-dimensional case of the theorem (which has an elementary proof that we omit due to space considerations).

Next, we discuss a needle decomposition obtained by Stochastic Localization, in a different setting, where the role of convexity is replaced by complex-analyticity.

5.1.1. A waist inequality for complex-analytic functions

In [29], Klartag uses a decomposition of the Gaussian measure γ_n, via Stochastic Localization, to prove several expansion inequalities for complex-analytic sets. For example, he obtains the following bound.

Theorem 5.2 (Klartag [29]). *Let $f : \mathbb{C}^n \to \mathbb{C}^k$ be a holomorphic function such that $f(0) = 0$. Write $Z = f^{-1}(0)$. Then one has*

$$\gamma_n(Z_\varepsilon) \geq \gamma_k\big(\big\{x \in \mathbb{C}^k : \|x\|_2 \leq \varepsilon\big\}\big), \quad \forall \varepsilon > 0, \tag{5.4}$$

where Z_ε is the ε-extension of Z and γ_m is the complex standard Gaussian measure on $\mathbb{C}^m$.

The above may be thought of in context of Gromov's waist inequality [22], according to which, every continuous function $f : \mathbb{R}^n \to \mathbb{R}^k$ has a level set $Z = f^{-1}(a)$ which satisfies (5.4). The key to the proof is to find a decomposition of γ_n of the form

$$\gamma_n = \int_J \nu_\alpha m(d\alpha)$$

such that:

(i) The measures ν_α are Gaussian measures with covariance matrix of rank at most k and operator norm bounded by 1.

(ii) The center of mass of each ν_α lies on Z.

Such a decomposition effectively reduces the proof of the theorem to the trivial case $k = n$.

We give a high-level sketch of ideas used to obtain such a decomposition. Consider the Stochastic Localization process of equation (4.2) taking the background measure ν to be the Gaussian measure γ_n. Our goal is to find a control matrix C_t so that the two properties above hold. In order to obtain property (ii), the idea is to make sure that the $a_t \in Z$ for all $t \geq 0$ (the center of mass of ν_t). The evolution of a_t obeys the equation (as in (4.14))

$$da_t = \mathrm{Cov}(\nu_t) C_t dB_t,$$

where now B_t is a Brownian motion in $\mathbb{C}^n$ and C_t is an $n \times n$ Hermitian matrix. We want to make sure that $f(a_t)$ remains constant. The key observation is that, due to the fact that f is holomorphic, there will be no quadratic variation terms in the formula for $df(a_t)$, and we have that

$$df_i(a_t) = \nabla f_i(a_t)^T \,\mathrm{Cov}(\nu_t) C_t dB_t, \quad \forall 1 \leq i \leq k.$$

For each t, by dimension considerations, we can find a projection matrix C_t of rank $n - k$ such that $df(a_t) = 0$. With this choice of driving-matrix, all but k eigenvalues of the matrix $G_t = \int_0^t C_s^2 ds$ must converge to infinity as $t \to \infty$ and, in light of (4.3), we get that $\mathrm{Cov}(\nu_t)$ tends to a matrix of rank k, as required by property (i).

5.2. Measures on the discrete hypercube

Up to this point, we were focused on absolutely continuous measures on $\mathbb{R}^n$ (or $\mathbb{C}^n$). In this section, we discuss applications of Stochastic Localization to discrete measures, where there is no natural notion of convexity and heat-flow techniques typically do not apply.

5.2.1. Concentration for Ising models via decomposition into low-rank systems

An *Ising model* is a measure ν on the discrete hypercube $\{-1, 1\}^n$ whose potential is a quadratic function or, in other words, its density is of the form

$$\nu\big(\{x\}\big) = Z_\nu^{-1} \exp\big(\langle x, Jx\rangle + \langle h, x\rangle\big), \quad \forall x \in \{-1, 1\}^n \tag{5.5}$$

for some $n \times n$ symmetric matrix J (called an interaction matrix) and some $h \in \mathbb{R}^n$ (an "external field"), and where Z_ν is a normalization constant. An important question in statistical mechanics is to characterize the pairs (J, h) for which the model is in *high temperature*.

One interpretation of high temperature is that $\sqrt{\mathrm{Var}\langle X, Y\rangle} \ll n$ where X, Y are independent vectors with law ν.

It is a common belief that for most cases of interest, measures in the high-temperature regime will admit stronger forms of concentration. For example, it is expected that the so-called Glauber dynamics admits a polynomially-large spectral gap in the high-temperature regime, which implies the existence of a polynomial-time sampling algorithm for ν, see [21] for definitions and background.

In what follows, we outline a way to obtain a concentration inequality for high-temperature Ising models using Stochastic Localization. In order to keep things simple and avoid encumbering the reader with definitions, we will derive a weaker form of concentration than what the method allows. A function $\varphi : \mathbb{R}^n \to \mathbb{R}$ is 1-Hamming–Lipschitz (1-Lipschitz in short) if $|\varphi(x) - \varphi(y)| \leq \|x - y\|_1$ for all $x, y \in \{-1, 1\}^n$. We will show the following.

Theorem 5.3. *For every ν of the form (5.5) such that $\|J\|_{\mathrm{OP}} \leq 1/2$ and every 1-Lipschitz test function $\varphi : \{-1, 1\}^n \to \mathbb{R}$,*

$$\mathrm{Var}_\nu[\varphi] \leq \frac{n}{\frac{1}{2} - \|J\|_{\mathrm{OP}}}.$$

The above bound was first obtained as a corollary of a result by Bauerschmidt and Bodineau [6]. A modification of the argument below produces a stronger bound which also establishes polynomial mixing of the Glauber dynamics, see [21]. We now outline the proof.

Without loss of generality, we may assume that J is positive semidefinite (we can always add a multiple of the identity without changing the distribution). Given an Ising model ν and a test function $\varphi : \{-1, 1\}^n \to \mathbb{R}$, consider the Stochastic Localization equations (4.2), with the matrix C_t to be defined later on. Define

$$b_t := \int_{\{-1,1\}^n} \varphi(x)(x - a_t)\nu_t(dx)$$

so that, by (4.13), we have

$$d \int \varphi(x)\nu_t(dx) = \langle C_t b_t, dB_t\rangle.$$

Set $J_t := J - \frac{1}{2}\int_0^t C_s^2 ds$. Equation (4.3) implies that ν_t is an Ising model with interaction matrix J_t. The idea now is to choose C_t to be the orthogonal projection on the intersection $\mathrm{Im}(J_t) \cap b_t^\perp$. By continuity, the matrix J_t is decreasing in the positive definite sense, but remains positive semidefinite. Since $C_t b_t = 0$, we have, using (4.13), that

$$\mathbb{E} \int \varphi d\nu_t = \int \varphi d\nu, \quad \forall t \geq 0.$$

By dimension considerations, C_t is nonzero as long as $\dim(\mathrm{Im}(J_t)) > 1$. By running the process until J_t is of rank at most 1 and using the decomposition (5.1), we arrive at the "needle decomposition" theorem formulated below. For every $u, h \in \mathbb{R}^n$, define

$$\nu_{u,h}\big(\{x\}\big) = Z_{u,h}^{-1} \exp\big(\langle x, u\rangle^2 + \langle h, x\rangle\big), \tag{5.6}$$

with $Z_{u,h}$ being a constant normalizing $\nu_{u,h}$ to be a probability measure.

Theorem 5.4. *Let ν be an Ising measure on $\{-1,1\}^n$ of the form (5.5) with J positive semidefinite, and let $\varphi : \{-1,1\}^n \to \mathbb{R}$. There exists a probability measure m on $\mathbb{R}^n \times \mathbb{R}^n$ such that ν admits the decomposition*

$$\nu = \int_{\mathbb{R}^n \times \mathbb{R}^n} \nu_{u,h} \, dm(u,h) \tag{5.7}$$

and such that m-almost surely the pair (u,h) satisfies $\int \varphi d\nu_{u,h} = \int \varphi d\nu$ and $\|u\|_2 \le \|J\|_{\mathrm{OP}}$.

This decomposition theorem allows us to show that an Ising measure inherits the concentration properties satisfied by rank-one Ising models whose interaction matrix has a corresponding norm. For models of rank-one, we rely on the following fact.

Fact 5.5 (see [21]). *For all $u, h \in \mathbb{R}^n$ such that $|u| < \frac{1}{2}$ and for all 1-Lipschitz $\varphi : \mathbb{R}^n \to \mathbb{R}$, we have*

$$\mathrm{Var}_{\nu_{u,h}}[\varphi] \le \frac{n}{1/2 - |u|}. \tag{5.8}$$

Now, given an Ising model ν with positive semidefinite interaction matrix J of norm at most $1/2$ and given a 1-Lipschitz test function φ, use Theorem 5.4 to find a measure m corresponding to ν, φ. We have, by the law of total variance,

$$\mathrm{Var}_\nu[\varphi] \overset{(5.7)}{=} \int_{\mathbb{R}^n \times \mathbb{R}^n} \mathrm{Var}_{\nu_{u,h}}[\varphi] dm(u,h) \overset{(5.8)}{\le} \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{n}{1/2 - |u|} dm(u,h) \le \frac{n}{1/2 - \|J\|_{\mathrm{OP}}}.$$

5.2.2. Entropy-efficient decomposition of discrete measures

In the previous subsections we saw how the Stochastic Localization process allows us to decompose a measure into well-behaved "needles." We now present a family of related applications which has proven useful in the context of interacting particle systems, random graphs, and large deviation theory.

We begin the discussion with a simple example referred to as the Curie–Weiss model: Fix $\beta > 0$ and consider the measure ν on $\{-1,1\}^n$, defined by

$$\nu(\{x\}) = Z_\beta^{-1} \exp\left(\frac{\beta}{n} \sum_{i \ne j} x_i x_j\right),$$

with Z_β a normalizing constant. Let $X \sim \nu$. It is well known that this measure exhibits the following phase transition: If $\beta < 1/2$, then $\mathrm{Cov}(X_1, X_2) \to 0$ as $n \to \infty$, whereas if $\beta > 1/2$ then $\mathrm{Cov}(X_1, X_2)$ is bounded away from 0 as $n \to \infty$ (and hence, also $\mathrm{Var}[\sum_i X_i] = \Omega(n^2)$). On the other hand, in the latter case, there exist two measures $\nu^\pm$ such that $\nu = \frac{1}{2}(\nu^+ + \nu^-)$ and such that $\nu^\pm$ are approximate product measures in the sense that $\|\mathrm{Cov}(\nu^\pm)\|_{\mathrm{OP}} = O(1)$ and, in fact, in a much stronger sense discussed later on.

This simple, yet somewhat prototypical example motivates the question of finding sufficient conditions on a measure ν on $\{-1,1\}^n$ under which it can be expressed as a decomposition $\nu = \sum_{i=1}^N \nu_i$ where the measures ν_i attain a simple form, and N is not too large. Here, we consider a more general form of decomposition where our goal is to express ν as

$$\nu = \int_{\mathcal{I}} \nu_\alpha m(d\alpha)$$

such that the ν_α's have a simple form. In this context, it is natural to replace the requirement that N is not too large by an upper bound on the *entropic-deficit* of the decomposition, defined as

$$\mathrm{Ent}[\nu] - \int_{\mathcal{I}} \mathrm{Ent}[\nu_\alpha] m(d\alpha),$$

where, for a measure μ on $\{-1,1\}^n$, we define $\mathrm{Ent}(\mu) := -\int_{\{-1,1\}^n} \log(\mu(\{x\}))\mu(dx)$.

Stochastic Localization is a useful tool in obtaining decompositions of this sort, via equation (5.1). The key is to analyze the evolution of the processes $\mathrm{Cov}(\nu_t)$ and $\mathrm{Ent}[\nu_t]$, which turn out to be quite tractable. As an initial idea of how it can be done, observe that choosing $C_t = \mathrm{Id}$ and taking expectations on both sides of equation (4.15), we have that

$$\frac{d}{dt}\mathbb{E}\,\mathrm{Cov}(\nu_t) = -\mathbb{E}\big[\mathrm{Cov}(\nu_t)^2\big].$$

One may interpret the last display as follows: The localization process "shrinks," in expectation, the large directions of the covariance matrix. Let us now outline an argument which builds on this intuition.

Fix a measure ν on $\{-1,1\}^n$ and consider the process ν_t obtained by running the process of equation (4.2) with the initial condition $\nu_0 = \nu$. For every t, take C_t to be the projection onto the span of the top eigenvector of $\mathrm{Cov}(\nu_t)$. Using (4.15), we have that

$$d\,\mathrm{Tr}\big(\mathrm{Cov}(\nu_t)\big) = -\big\|\mathrm{Cov}(\nu_t)\big\|_{\mathrm{OP}}^2 dt + \text{martingale term}.$$

A straightforward calculation using Itô's formula yields that

$$\begin{aligned} d\,\mathrm{Ent}(\nu_t) &= -\mathrm{Tr}\big(C_t\,\mathrm{Cov}(\nu_t)\big)dt + \text{martingale term} \\ &= -\big\|\mathrm{Cov}(\nu_t)\big\|_{\mathrm{OP}} dt + \text{martingale term}. \end{aligned}$$

By comparing that last two displays, we see that as long as $\|\mathrm{Cov}(\nu_t)\|_{\mathrm{OP}}$ is large, the trace of the covariance matrix of ν_t decays, in expectation, much faster than its entropy. Now fix $\lambda > 0$ and consider the stopping time

$$\tau := \min\big\{t; \big\|\mathrm{Cov}(\nu_t)\big\|_{\mathrm{OP}} \le \lambda\big\}.$$

By the above, we have that $\lambda\,\mathrm{Ent}(\nu_t) - \mathrm{Tr}(\mathrm{Cov}(\nu_t))$ is a submartingale up to the stopping time τ. Using the optional stopping theorem, we have that

$$\mathbb{E}\big[\mathrm{Ent}[\nu] - \mathrm{Ent}[\nu_\tau]\big] \le \frac{1}{\lambda}\mathbb{E}\big(\mathrm{Tr}\big(\mathrm{Cov}(\nu)\big) - \mathrm{Tr}\big(\mathrm{Cov}(\nu_\tau)\big)\big) \le \frac{\mathrm{Tr}(\mathrm{Cov}(\nu))}{\lambda} \le \frac{n}{\lambda}.$$

Using the decomposition (5.1), we arrive at the following theorem.

Theorem 5.6. *Let ν be a measure on $\{-1,1\}^n$. Then for every $\lambda \ge 1$, there exist a probability measure m on an index set $\mathcal{I}$ and a family of probability measures $\{\nu_\theta\}_{\theta\in\mathcal{I}}$ on $\{-1,1\}^n$ such that the measure ν admits the decomposition*

$$\nu(W) = \int_{\mathcal{I}} \nu_\theta(W) dm(\theta), \quad \forall W \subset \mathbb{R}^n \text{ measurable}, \tag{5.9}$$

such that

$$\big\|\mathrm{Cov}(\nu_\alpha)\big\|_{\mathrm{OP}} \le \lambda, \quad \forall \alpha \in \mathcal{I}$$

and

$$\mathrm{Ent}[\nu] - \int_J \mathrm{Ent}[\nu_\alpha] m(d\alpha) \le \frac{n}{\lambda}.$$

A related argument can also produce bounds on the Frobenius norm of $\mathrm{Cov}(\nu_\alpha)$. We refer the reader to [19] for other inequalities of this form, as well as application to mean-field approximation, which we do not discuss here.

The measures ν_α given by the above theorem are close to product measures in a rather weak sense, and one may consider stronger notions of approximating a product measure. A particularly useful notion is defined in terms of the transportation distance to a product measure. For probability measures μ_1, μ_2 on $\{-1,1\}^n$, we define

$$\mathrm{W}(\mu_1, \mu_2) = \sup_{\|\varphi\|_{\mathrm{Lip}} \le 1} \left(\int_{\{-1,1\}^n} \varphi d\mu_1 - \int_{\{-1,1\}^n} \varphi d\mu_2 \right),$$

where $\|\cdot\|_{\mathrm{Lip}}$ denotes the Hamming–Lipschitz norm. This quantity is referred to as the (Wasserstein) transportation distance with respect to the Hamming metric. Given a probability measure μ on $\{-1,1\}^n$, let $\xi(\mu)$ be the unique product measure having the same center of mass of μ. Consider the quantity

$$\mathcal{P}(\mu) := \mathrm{W}\big(\mu, \xi(\mu)\big)$$

which quantifies how close μ is to a product measure. What conditions on a measure ν on $\{-1,1\}^n$ ensure that it admits a decomposition of the form (5.9) such that both the entropic deficit and $\mathcal{P}(\mu)$ are nontrivially small (say, both are $o(n)$)? The work [18] establishes this under a condition inspired by an earlier work of Chatterjee and Dembo [14] and referred to as *low complexity*. For a measure ν on $\{-1,1\}^n$, denote by f_ν its density with respect to the uniform measure. Define the *complexity* of ν by

$$\mathcal{D}(\nu) := \mathbb{E}_{\Gamma \sim \mathcal{N}(0,\mathrm{Id})} \sup_{x \in \{-1,1\}^n} \langle \nabla \log f_\nu, \Gamma \rangle$$

(which can be understood as the Gaussian-width of the gradient of its potential). The following decomposition theorem can be obtained via Stochastic Localization.

Theorem 5.7 ([18]). *For every measure ν on $\{-1,1\}^n$ and every $\varepsilon > 0$, there exists a decomposition of the form (5.9) such that its entropic deficit satisfies*

$$\mathrm{Ent}[\nu] - \int_J \mathrm{Ent}[\nu_\alpha] m(d\alpha) \le \varepsilon n$$

and such that

$$\int_J \mathcal{P}(\nu_\alpha) m(d\alpha) \lesssim \sqrt{\frac{n \mathcal{D}(\nu)}{\varepsilon}}.$$

Note that, as long as $\mathcal{D}(\nu) = o(n)$, we may obtain a decomposition with entropic deficit $o(n)$, such that $\mathcal{P}(\nu_\alpha) = o(n)$ for all but an $o(1)$ fraction of α's (with respect to m). This type of structure theorem has several applications, in particular to the emerging field of *nonlinear large deviations* pioneered by Chatterjee and Dembo in [14] and to mean-field approximations. We refer the reader to [18] for more details.

FUNDING

The author is supported by a European Research Council starting Grant (ERC StG), project number 803084.

REFERENCES

[1] D. Alonso-Gutiérrez and J. Bastero, *Approaching the Kannan–Lovász–Simonovits and variance conjectures*. Lecture Notes in Math. 2131, Springer, Cham, 2015.

[2] S. Artstein-Avidan, A. Giannopoulos, and V. D. Milman, *Asymptotic geometric analysis. Part I*. Math. Surveys Monogr. 202, American Mathematical Society, Providence, RI, 2015.

[3] D. Bakry and M. Emery, Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/1984*, pp. 177–206, Lecture Notes in Math. 1123, Springer, Berlin, 1985.

[4] D. Bakry and M. Ledoux, Lévy–Gromov isoperimetric inequality for an infinite dimensional diffusion generator. *Invent. Math.* **123** (1995), 259–281.

[5] K. Ball and V. H. Nguyen, Entropy jumps for isotropic log-concave random vectors and spectral gap. *Studia Math.* **213** (2012), no. 1, 81–96.

[6] R. Bauerschmidt and T. Bodineau, A very simple proof of the LSI for high temperature spin systems. *J. Funct. Anal.* **276** (2019), no. 8, 2582–2588.

[7] S. Bobkov, On isoperimetric constants for log-concave probability distributions. In *Geometric aspects of functional analysis. Israel Seminar 2004–2005*, pp. 81–88, Lecture Notes in Math. 1910, Springer, 2007.

[8] C. Borell, The Brunn–Minkowski inequality in Gauss space. *Invent. Math.* **30** (1975), no. 2, 207–216.

[9] C. Borell, Geometric bounds on the Ornstein–Uhlenbeck velocity process. *Z. Wahrsch. Verw. Gebiete* **70** (1985), no. 1, 1–13.

[10] C. Borell, Diffusion equations and geometric inequalities. *Potential Anal.* **12** (2000), no. 1, 49–71.

[11] C. Borell, Isoperimetry, log-concavity, and elasticity of option prices. In *New directions in mathematical finance*, edited by P. Wilmott and H. Rasmussen, pp. 73–91, Wiley, 2002.

[12] J. Bourgain, On the distribution of polynomials on high-dimensional convex sets. In *Geometric Aspects of Functional Analysis (1989–1990)*, pp. 127–137, Lecture Notes in Math. 1469, Springer, 1991.

[13] P. Cattiaux, A pathwise approach of some classical inequalities. *Potential Anal.* **20** (2004), no. 4, 361–394.

[14] S. Chatterjee and A. Dembo, Nonlinear large deviations. *Adv. Math.* **299** (2019), 396–450.

[15] Y. Chen, An almost constant lower bound of the isoperimetric coefficient in the KLS conjecture. *Geom. Funct. Anal.* **31** (2021), no. 1, 34–61.

[16] R. Eldan, Thin shell implies spectral gap up to polylog via a Stochastic Localization scheme. *Geom. Funct. Anal.* **23** (2012), no. 2, 532–569.

[17] R. Eldan, A two-sided estimate for the Gaussian noise stability deficit. *Invent. Math.* **201** (2015), no. 2, 561–624.

[18] R. Eldan, Gaussian-width gradient complexity, reverse log-Sobolev inequalities and nonlinear large deviations. *Geom. Funct. Anal.* **28** (2018), no. 6, 1548–1596.

[19] R. Eldan, Taming correlations through entropy-efficient measure decompositions with applications to mean-field approximation. *Probab. Theory Related Fields* **176** (2020), no. 3–4, 737–755.

[20] R. Eldan and B. Klartag, Approximately gaussian marginals and the hyperplane conjecture. In *Proc. of a workshop on "Concentration, Functional Inequalities and Isoperimetry"*, pp. 55–68, Contemp. Math. 545, Amer. Math. Soc., 2011.

[21] R. Eldan, F. Koehler, and O. Zeitouni, A spectral condition for spectral gap: fast mixing in high-temperature Ising models. 2020, arXiv:2007.08200.

[22] M. Gromov, Isoperimetry of waists and concentration of maps. *Geom. Funct. Anal.* **13** (2003), no. 1, 178–215.

[23] M. Gromov and V. D. Milman, Generalization of the spherical isoperimetric inequality to uniformly convex Banach spaces. *Compos. Math.* **62** (1987), 263–282.

[24] O. Guedon and E. Milman, Interpolating thin-shell and sharp large-deviation estimates for isotropic log-concave measures. *Geom. Funct. Anal.* **21** (2011), no. 5, 1043–1068.

[25] R. Kannan, L. Lovász, and M. Simonovits, Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.* **13** (1995), 541–559.

[26] B. Klartag, On convex perturbations with a bounded isotropic constant. *Geom. Funct. Anal.* **16** (2006), no. 6, 1274–1290.

[27] B. Klartag, A central limit theorem for convex sets. *Invent. Math.* **168** (2007), 91–131.

[28] B. Klartag, Needle decompositions in Riemannian geometry. *Mem. Amer. Math. Soc.* **249** (2017), no. 1180.

[29] B. Klartag, Eldan's Stochastic Localization and tubular neighborhoods of complex-analytic sets. *J. Geom. Anal.* **28** (2018), no. 3, 2008–2027.

[30] B. Klartag and V. D. Milman, The slicing problem by Bourgain. In *Analysis at Large, A Collection of Articles in Memory of Jean Bourgain*. Springer (to appear).

[31] M. Ledoux, Isoperimetry and Gaussian analysis. In *Lectures on probability theory and statistics*, pp. 165–294, Lecture Notes in Math., 1648. Springer, Berlin, Heidelberg, 1996.

[32] M. Ledoux, *The concentration of measure phenomenon*. Math. Surveys Monogr. 89, American Mathematical Society, Providence, RI, 2001.

[33] Y. T. Lee and S. Vempala, The Kannan–Lovász–Simonovits conjecture. In *Current developments in mathematics 2017*, pp. 1–36, Int. Press, Somerville, MA, 2017.

[34] Y. T. Lee and S. S. Vempala, Eldan's Stochastic Localization and the KLS hyperplane conjecture: an improved lower bound for expansion. In *2017 IEEE 58th annual symposium on foundations of computing (FOCS)*, pp. 998–1007, ACM, New York, 2017.

[35] Y. T. Lee and S. S. Vempala, Eldan's stochastic localization and the KLS conjecture: isoperimetry, concentration and mixing. 2018, arXiv:1612.01507.

[36] J. Lehec, Representation formula for the entropy and functional inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** (2013), no. 3, 885–899.

[37] L. Lovász and M. Simonovits, Random walks in a convex body and an improved volume algorithm. *Random Struct. Algebra* **4** (1993), 359–412.

[38] E. Milman, On the role of convexity in isoperimetry, spectral-gap and concentration. *Invent. Math.* **177** (2009), no. 1, 1–43.

[39] E. Milman and J. Neeman, The Gaussian multi-bubble conjecture. 2018, arXiv:1805.10961.

[40] E. Mossel and J. Neeman, Robust optimality of Gaussian noise stability. *J. Eur. Math. Soc. (JEMS)* **17** (2015), no. 2, 433–482.

[41] G. Pisier, Probabilistic methods in the geometry of Banach spaces. In *Probability and analysis (Varenna, 1985)*, pp. 167–241, Lecture Notes in Math. 1206, Springer, Berlin, 1986.

[42] V. N. Sudakov and B. S. Tsirel'son, Extremal properties of half-spaces for spherically invariant measures. *J. Sov. Math.* **9** (1978), no. 1, 9–18.

[43] R. Vershynin, *High-dimensional probability. An introduction with applications in data science. With a foreword by Sara van de Geer*. Camb. Ser. Stat. Probab. Math. 47, Cambridge University Press, Cambridge, 2018.

RONEN ELDAN

Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel, ronen.eldan@weizmann.ac.il