Research statement

Ronen Eldan
Weizmann Institute of Science

Summary

My research focuses on the behavior of objects in a *high-dimensional* setting, in search of phenomena that arise as the dimension (or the number of degrees of freedom) of the system tends to infinity. It spans several domains of mathematics: probability, metric geometry, functional analysis, mathematical physics, combinatorics, learning theory and optimization. The study of phenomena in high dimensions, which originally stemmed from the local theory of Banach spaces, has led to a broad and rapidly growing theory. It turns out that whether probabilistic, geometric or combinatorial, the behavior a high dimensional object is often dictated by several common unifying principles and concepts, such as concentration of measure. These concepts play a role in several areas of mathematics, statistics and computer science, and the underlying theory has seen a boost in applications in recent years, in accordance with the explosion of interest in data science and machine learning, two fields of which this theory is a cornerstone.

The core of my research can be roughly divided into three general directions: The first direction puts focus on open problems in high dimensional probability, geometry and mathematical physics, and in particular, on problems reflecting *dimension-free* phenomena, meaning that the behavior of a high dimensional object is dictated by marginals of fixed dimension. The second direction aims to develop new methods towards understanding high dimensional phenomena via an emerging connection with the theory of stochastic calculus (or pathwise analysis), and the third direction has to do with the application of concepts from high dimensional theory to more applied directions, such as learning theory and optimization.

In the first part of this statement, I review some topics in high dimensional probability and geometry, describing some of the main open problems in which I have been interested. We begin with the role of convexity in concentration inequalities, focusing on three central conjectures regarding the distribution of mass on high dimensional convex bodies: the Kannan-Lovász-Simonovits (KLS) conjecture, the variance conjecture and Bourgain's hyperplane conjecture as well as emerging connections with quantitative central limit theorems, entropic jumps and stability bounds for the Brunn-Minkowski inequality. Second, we discuss dimension-free inequalities in Gaussian space and on the Boolean hypercube: isoperimetric and noise-stability inequalities and robustness thereof, transportation-entropy and concentration inequalities, regularization properties of the heat-kernel and L_1 versions of hypercontractivity. Finally, I will discuss my attempt to develop new methods for the analysis of Gibbs distributions with a mean-field behavior, related to the new theory of nonlinear large deviations, and towards questions regarding interacting particle systems and the analysis of large networks.

In a few recent works, by developing a novel approach of pathwise analysis, my coauthors and I managed to make progress in several open problems in the theory. This emerging method relies on the introduction of a stochastic process which allows one to associate quantities and properties related to the high-dimensional object of interest to corresponding notions in stochastic calculus, thus making the former tractable through the analysis of the latter. In the second part of the statement I describe this approach; My goal is to extend this method towards tackling some of the questions mentioned here.

In the final part of this statement, I describe the applied directions of my research, which aim to utilize the recent developments in the theory of high-dimensional probability and geometry to computational geometry, learning theory and mathematical optimization. Some examples of those applications are to the bandit convex optimization (a problem in reinforcement learning), sampling algorithms from high dimensional measures, estimation of volume and covariance based on random samples and barrier methods in convex optimization.

1 Introduction

My domain of research spans several fields of mathematics: Probability theory, Metric geometry, Functional analysis, Mathematical physics, Computational geometry, Combinatorics, Learning theory and Optimization. The questions in which I am interested are tied by a single common theme: being of *high-dimensional* nature. In other words, I am interested in phenomena that arise as the number of degrees of freedom goes to infinity.

To name a few examples, problems of high dimensional nature may have to do with the geometry of convex sets in $\mathbb{R}^n$ as $n \to \infty$, with functional inequalities in (multidimensional) Gaussian space, with the behavior of measures or functions on the Boolean hypercube, with interacting particle systems (where the number of particles goes to infinity), with the behavior of random graphs whose number of vertices tends to infinity, with reinforcement learning where the action space is "large", or with the complexity of convex optimization as a function of the dimension.

Problems of this type have inspired a rapidly growing theory. It turns out that, though originating from seemingly different domains of mathematics, many of those problems share common underlying themes and principles. A remarkable aspect of the this theory is the abundance of methods and tools that it has inspired, arising from numerous different fields; in many such problems the path to the solution passes through a topic with no a-priori apparent connection to the original problem. For example, the Brunn-Minkowski inequality is tightly related to the theory of transportation of measure; the geometry of sections of high dimensional convex bodies is often understood using tools from Harmonic analysis and concentration inequalities rely on tools from potential theory, partial differential equations and geometric measure theory.

In particular, the use of probabilistic language as well as ideas and results from probability theory in the study of high dimensional objects has paved the way to a long list of discoveries, perhaps the most pioneering example being the proof of Dvoretzky's theorem (the language of probability was first used in the proof due to V. Milman), showing the existence of Euclidean subspaces of a normed space by taking a random subspace. Since then, the application of concepts and theorems from probability theory has reached increasing levels of sophistication, and tools such as martingale concentration bounds, moment generating functions and the second-moment method have become essential ingredients in the high-dimensional geometer's cookbook.

One of my central long-term projects is to develop an emerging connection with the theory of stochastic calculus and pathwise analysis. In a series of works ([Eld13, Eld16, Eld15b, EL15, DEZ15, EL14, BEL15, ELL16, Eld17, EMZ18]), coauthors and I have established connections between quantities related to high dimensional problems and corresponding notions from stochastic calculus, towards obtaining insights on the behavior of high dimensional objects.

In this research statement, I will first provide background regarding the theory of high dimensional objects and mention some central problems in high dimensional probability and geometry that I have been working on as well as research directions which I plan to pursue in the future. In the second section, I will describe the emerging connection with the field of stochastic calculus. The third section will cover another direction of research, in which we attempt to leverage the insights gained from the theory of high dimensional objects towards more applied topics: optimization, learning theory, statistical analysis and algorithms.

2 Phenomena in high dimensions

The intuition derived from low-dimensional examples in various fields such as PDE's, topology and statistical mechanics would perhaps lead us to believe that understanding the behavior of high-dimensional objects is rather hopeless, since their behavior very quickly becomes complex and intractable as the dimension increases. One reason for this meta-phenomenon, sometimes referred to as "the curse of dimensionality", is the simple fact that number of possible configurations of a system tends to grow exponentially with respect to the dimension.

The striking fact about the theory of high-dimensional phenomena is that in many examples this intuition is completely reversed: sometimes the multiplicity of degrees of freedom can be a cause for order and simplicity. The central limit theorem is a good example: summing a large number of independent random variables, when normalized correctly, will result in a universal distribution that does not depend on the distributions of the individual variables. Two striking examples of the emergence of universality in high-dimensional settings are Dvoretzky's theorem and Klartag's central limit theorem for convex sets, both of which rely on a convexity assumption rather than on independence of coordinates; The former result asserts that typical sections of a high-dimensional convex body are approximately Euclidean and the latter states that typical marginals are normally distributed. Some of these phenomena are at the heart of theories concerning several more applied fields such as statistical learning theory and can be witnessed in many real-world applications.

A remarkable feature that illustrates the emergence of simplicity is the *dimension-free* behavior of several quantities related to a high-dimensional object. Ideologically, dimension free results suggest that the behavior of a high dimensional space is dictated by that of subspaces or marginals of a fixed (and usually small) dimension. A setting where this feature is well illustrated is the Gaussian space which we discuss next.

2.0.1 Dimension-free phenomena on Gaussian space: Isoperimetry, noise stability and regularization under the heat flow

By Gaussian space, we simply mean the Euclidean space $\mathbb{R}^n$ equipped with the standard Gaussian measure $d\gamma = \frac{1}{(2\pi)^{n/2}} e^{-\frac{|x|^2}{2}} dx$. Perhaps the most fundamental example of a dimension-free inequality is the Gaussian isoperimetric inequality due to Borell and Sudakov-Tsirelson, which reads:

Theorem 2.1 *(Gaussian Isoperimetric inequality) Let $A \subset \mathbb{R}^n$ and let H be a half-space satisfying $\gamma(H) = \gamma(A)$. Then for all $0 < \rho < 1$,*

$$\gamma_+(H) \leq \gamma_+(A).$$

Here, γ_+ can be thought of as the $(n-1)$-dimensional Hausdorff measure multiplied by the Gaussian density. In other words, the theorem states that the isoperimetric minimizers in Gaussian space are effectively *one-dimensional* objects.

An extension of the latter is Borell's noise stability inequality [Bor85]. The noise stability of a set $A \subset \mathbb{R}^n$ is defined by the formula $\mathcal{S}_\rho(A) = \mathbb{P}(X \in A, Y \in A)$ where X, Y are ρ-correlated standard Gaussian random vectors (meaning that X, Y each have the law γ and $\mathbb{E}[X_i Y_j] = \rho \delta_{i,j}$). Borell's inequality asserts the following.

Theorem 2.2 *(Borell, [Bor85]) Let $A \subset \mathbb{R}^n$ and let H be a half-space having the same Gaussian measure as that of A, then for all $0 < \rho < 1$,*

$$\mathcal{S}_\rho(H) \geq \mathcal{S}_\rho(A).$$

This inequality has applications in numerous fields such as approximation theory, rearrangement inequalities, discrete analysis, game theory and complexity.

A few years back, Mossel and Neeman [MN15] proved *robustness* results for the aforementioned inequalities, showing that when the corresponding deficit is small, the set A has to be an approximate half-space in the total-variation metric sense. In [Eld15b], we have managed to improve their result, obtaining the optimal possible exponent in the dependence between the deficit and the distance. Moreover, we have found a seemingly more natural metric (namely, the distance between the centroids) which gives a two-sided bound for the deficit, up to logarithmic factors.

More precisely, defining $\delta(A, H) = \int_H P_t[1_H] d\gamma - \int_A P_t[1_A] d\gamma$ we were able to obtain the following bound.

Theorem 2.3 *(E., '15) If $A \subset \mathbb{R}^n$ and H is a half-space such that $\gamma(A) = \gamma(H)$ then (up to a logarithmic factor)*

$$\delta(A, H) \asymp \varepsilon(A, H)$$

where $\varepsilon(A,H) = \left|\int_H x d\gamma(x)\right|^2 - \left|\int_A x d\gamma(x)\right|^2$.

A third central dimension-free inequality is the hypercontractivity of the heat operator. Defining $P_t[f](x) := \int f(e^{-t/2}x + \sqrt{1-e^{-t}}y)d\gamma(y)$, the Ornstein-Uhlenbeck heat operator, the hypercontractivity of this operator refers to the fact that $\|P_t f\|_q \leq \|f\|_p$ whenever $t \geq \frac{1}{2}\log\frac{q-1}{p-1}$. This inequality can be seen as a quantitative bound reflecting the *regularizing* effect of the heat operator. Until recently, no dimension-free quantitative bound was known for general measures (or, for that matter, for the case $f \in L_1$). In this context, Talagrand [Tal89] conjectured that the smoothing effect should take place in the sense that $P_t[f]$ should satisfy an improved weak-L_1 estimate. Together with J. Lee [EL15], we have been able to prove this conjecture.

Theorem 2.4 *(E.-Lee, '15) Let X be a standard Gaussian random vector. There exists a function $g(\alpha)$ such that $\lim_{\alpha\to\infty} g(\alpha) = 0$ and such that for all non-negative functions f satisfying $\mathbb{E}[f(X)] = 1$ one has the improved Markov inequality*

$$\mathbb{P}(T_\rho[f](X) > \alpha) < \frac{C(\rho)}{\alpha} g(\alpha). \tag{1}$$

Very often, the proofs of dimension-free bounds follow a *tensorization* argument, which usually tends to be a simple step in the proof. In the two examples of results given above, tensorization does not seem to work. Instead, in each of these problems we found a one-dimensional stochastic process that manages to extract the "significant" direction in which the phenomenon takes place and, in a way described with more detail below, helps revealing the one-dimensional nature of the problem.

In the future, we intend to explore several related directions, towards extending these results in the following ways: (i) An ambitious question in the context of noise stability is to find the maximizers in the case of more than two sets, known as the standard simplices conjecture of Mossel and Isaksson [IM12]. (ii) Another natural question is to find the extremizers for isoperimetry and noise stability in the class of *symmetric* sets. (iii) We also hope to extend our methods to a discrete setting: Regarding the L_1 version of hypercontractivity, it is conjectured by Talagrand ([Tal89]) that an analogous result should hold in the setting of the discrete cube (the latter would imply our result via an application of the central limit theorem). We therefore propose to look for analogous methods on the discrete setting towards proving this conjecture. This may also be helpful towards a related conjecture by Coutrade and Kumar [CK14], regarding a version of noise stability where the L_2 norm is replaced by mutual information.

2.1 Concentration of mass on high-dimensional convex bodies

Next, we discuss measure spaces with a convexity property: the uniform measure on a convex set or more generally, *log-concave* measures. In this context we mainly focus on three central open problems in high dimensional convex geometry: the **hyperplane conjecture** (also known as the slicing problem) of Bourgain, the conjecture of Kannan-Lovász-Simonovits (in short, the **KLS conjecture**) related to the isoperimetric inequality on convex bodies and the **variance conjecture** (also known as the thin-shell conjecture). These problems have been open for approximately 35, 25 and 15 years, respectively, and a large body of literature has evolved around them. We refer to the books [AGB15, BGVV14].

It turns out that the assumption of convexity is natural in a high dimensional setting in the sense that it provides us with the suitable kind of regularity, that when combined with the effects of concentration of measure, we are able to obtain remarkable results. For example, according to Klartag's central limit theorem for convex sets, these objects exhibit universality in the sense that typical one-dimensional marginals are approximately Gaussian.

Let us now discuss another aspect in which we expect a universal and dimension-free behavior for these objects, namely in the form of attaining *concentration* bounds. To that end, let us begin by recalling that an immediate consequence of the Gaussian isoperimetric inequality (Theorem 2.1) is the concentration bound for Lipschitz functions under the Gaussian measure, which can be formulated as follows:

Theorem 2.5 *Let X be a standard Gaussian random vector on $\mathbb{R}^n$ and let φ be 1-Lipschitz. Then,*

$$\mathrm{Var}[\varphi(X)] \leq 1.$$

This is perhaps the most basic type of concentration. It has several equivalent formulations (see below) and it plays a major role in analytical and algorithmic aspects. One of the biggest open problems in convex geometry is to determine whether the same type of concentration is attained by the uniform measure on a convex body, or more generally, by any log-concave measure.

Let K be a convex body in $\mathbb{R}^n$ (hence a compact convex set with nonempty interior) and let X_K be a random vector uniformly distributed on K. We say that K *isotropic* if $\mathbb{E}[X_K] = 0$ and $\mathrm{cov}(X_K) = \mathrm{Id}$, where $\mathrm{cov}(X)$ denotes the covariance matrix of X. Define

$$G_n = \sup_{K,\varphi} \sqrt{\mathrm{Var}(\varphi(X_K))}$$

where the supremum is taken over all isotropic convex bodies $K \subset \mathbb{R}^n$ and all 1-Lipschitz functions $\varphi : \mathbb{R}^n \to \mathbb{R}$. The Kannan-Lovász-Simonovits conjecture reads,

Conjecture 2.6 *(KLS, [KLS95]) There exists a universal constant $C > 0$ such $G_n < C$.*

In other words, the KLS conjecture asserts that the concentration of all Lipschitz functions is dictated by that of linear functions.

Other than being a very basic quantity, a positive answer would have numerous consequences in convex geometry and in particular in algorithmic aspects thereof. The magnitude of G_n is directly related to the rate of mixing of Markov chains used to sample from the uniform measure on the convex body K, which is by itself a central ingredient of many algorithmic and statistical tasks, such as covariance estimation. Other equivalent quantities include the spectral gap, isoperimetric profile and exponential concentration (see [Mil09]).

A related quantity is the *thin-shell* concentration of convex sets. Defining,

$$\sigma_n^2 := \sup_{\substack{X \sim \mathrm{U}(K) \\ K \subset \mathbb{R}^n \text{convex}}} \frac{\mathrm{Var}\,\|X\|_2}{\sup_{|\theta|=1} \mathrm{Var}[X \cdot \theta]},$$

It is clear that $\sigma_n \leq G_n$. The following special case of the KLS conjecture was suggested by Anttila-Ball-Perissinaki and Bobkov-Koldobsky [ABP03, BK03]:

Conjecture 2.7 *(Variance conjecture) One has $\sigma_n < C$ for a universal constant $C > 0$.*

We have reduced the KLS conjecture to the Variance conjecture up to a logarithmic term, proving:

Theorem 2.8 *([Eld13]) It holds that $G_n \leq C\sqrt{\log(n) \sum_{k=1}^{n} \frac{\sigma_k^2}{k}}$, for a universal constant $C > 0$.*

The main tool used for this reduction is our construction of the "stochastic localization scheme" described in Subsection 3.1 below. Recently, Lee and Vempala ([LV16]), relying on our method, have managed to obtain the state-of-the-art bound $G_n \leq Cn^{1/4}$ as well as an optimal log-Sobolev inequality for convex sets. A remarkable aspect of their proof is that it is almost completely self contained, and manages to bypass the need to use previous methods such as concentration of projections, centroid moment bodies or the Log-Laplace transform, giving a rather compact argument which recovers some of the central bounds regarding concentration of mass (such as thin-shell and Paouris' theorem) in a way that mostly involves the analysis of the aforementioned process.

Finally, we would like to address a fascinating and basic open problem in convex geometry which has attracted very significant efforts in the community over the past decades, the slicing problem. This problem is concerned with the quantity

$$L_n = \sup_K \left(\mathrm{Vol}(K)^{-1/n} \right)$$

where the supremum is taken over all isotropic convex bodies $K \subset \mathbb{R}^n$.

Conjecture 2.9 *(Hyperplane conjecture/Slicing problem) There exists a universal constant $C > 0$ such that $L_n \leq C$.*

The conjecture was first suggested by Bourgain, who came up with the bound $L_n \leq Cn^{1/4} \log n$. Twenty years later, Klartag improved the bound to $L_n < Cn^{1/4}$ [Kla06] which is best-known estimate up to date. Together with Klartag, we have also proved a reduction of the slicing problem to the variance conjecture:

Theorem 2.10 *([EK11]) One has $L_n \leq C\sigma_n$ for a universal constant $C > 0$.*

This followed a previous work of Ball and Nguyen ([BN12]) who found a reduction to the KLS conjecture. The reduction of Ball and Nguyen is based on an estimate on the *entropy jump* of log-concave random variables. In an ongoing work with my student Dan Mikulincer, we have found a stochastic approach to the proof of Ball and Nguyen which we believe may shed more light on these connections.

It is important to emphasize that the above relations only hold *globally*, in the sense that in order to deduce the concentration of Lipschitz functions on a specific convex body, a prior bound for the thin-shell concentration over *all* convex bodies needs to be known. A related result of Ball and Nguyen [BN12] also gives a *body-wise* link between the KLS conjecture and the slicing problem, with an exponential dependence between the respective quantities.

One specific direction which we suggest is to prove a *body-wise*, rather than a global, version of these connections:

Question 2.11 *Is it true that for every isotropic convex body $K \subset \mathbb{R}^n$ and every 1-Lipschitz function f, one has*

$$Var[f(X_K)] < \mathrm{Polylog}(n) Var[\|X_K\|]?$$

Question 2.12 *Is it true that for every isotropic convex body $K \subset \mathbb{R}^n$ one has*

$$Vol(K)^{-1/n} < \mathrm{Polylog}(n) \sqrt{Var[\|X_K\|]}?$$

In recent years, the questions in hand have been found to have close connections with several related aspects of the behavior of convex bodies: entropic jumps, rates of convergence in the entropic central limit theorem and stability of the Brunn-Minskowski inequality. Some of them are depicted in Figure 1. We discuss those connections next.

2.1.1 Entropy jumps and rates of convergence in the central limit theorem

The aforementioned reduction by Ball and Nguyen [BN12] builds on a related and independently interesting phenomenon referred to as *entropic jump*. For a random vector X define by $\mathrm{Ent}_\gamma(X)$ the relative entropy of X with

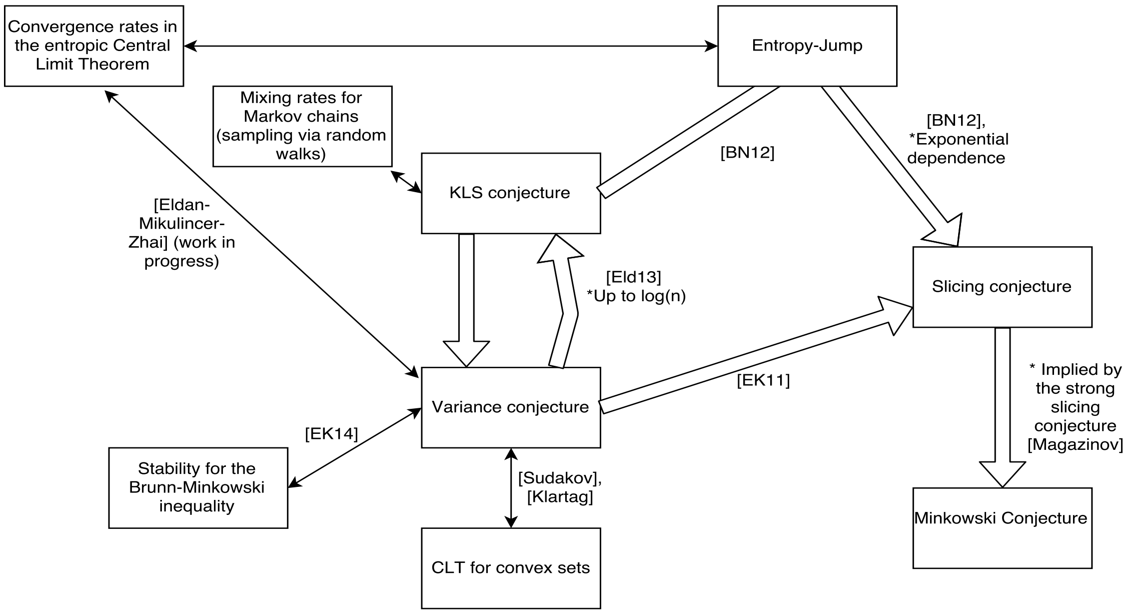

Figure 1: *The conjectures, with some of the relations described in Section 2.1.*

respect to the standard Gaussian measure. A classical result by Shannon and Stam states that if X is isotropic and has finite entropy and Y is an independent copy of X, then

$$\mathrm{Ent}_\gamma\left(\frac{X+Y}{\sqrt{2}}\right) \leq \mathrm{Ent}_\gamma(X).$$

In ([BBN03]) a quantitative version of this inequality in one dimension was established under the additional assumption that X admits a spectral gap. This was extended to higher dimensions in the work of Ball and Nguyen [BN12], when the underlying vector is log-concave. These inequalities have the form

$$\mathrm{Ent}_\gamma\left(\frac{X+Y}{\sqrt{2}}\right) \leq (1-c(X))\mathrm{Ent}_\gamma(X)$$

where $c(X)$ depends on the spectral gap of X. This constitutes the main step in the reduction of the slicing conjecture to the KLS conjecture. Moreover, an inequality of this type is closely related to the rate of convergence in the entropic central limit theorem, which concerns with the asymptotics of $\mathrm{Ent}_\gamma\left(\frac{X_1+...+X_N}{\sqrt{N}}\right)$ as $N \to \infty$, where $X_1, ..., X_N$ are independent copies of X.

In ongoing works with my student, Dan Mikulincer and with Alex Zhai ([EMZ18]) we have developed a pathwise approach to understanding the above phenomena. First, by embedding a prescribed random vector as a martingale in the space of Brownian paths, we were able to find a new approach to proving quantitative versions of the central limit theorems, in particular we obtain new bounds for rates of convergence in entropy and in transportation distance. Second, our stochastic constructions seem to give a new and rather natural proof of the Ball-Nguyen reduction, which may shed new light on the connection between the entropy jump and the slicing problem, and also gives a new connection to thin-shell concentration.

In the future, we intend to pursue the following goals.

1. Find the optimal asymptotics for convergence in the entropic central limit theorem, in particular, we would like to generalize the Berry-Essen bounds for the entropic central limit theorem, obtained in ([BCG14]), to arbitrary dimension, and capture the correct dependence on the dimension. Currently, our methods apply under

the additional assumptions that the random vector in hand is either log-concave or bounded.

2. Capture the correct dependence on the dimension of convergence in entropy and transportation distance in the log-concave case. These dependencies seem to have a close connection to the asymptotics of the KLS constant G_n.

3. Find weaker conditions under which entropy jump is witnessed. Our approach to entropy jump gives connections with some new quantities associated with a convex body (related to the symmetry breaking of the stochastic localization process). We believe that further scrutiny of these quantities may lead to new results.

2.1.2 Stability of the Brunn-Minkowski inequality

In one of its forms, the Brunn-Minkowski inequality states that for two convex bodies K, T of unit volume, one has

$$Vol\left(\frac{K+T}{2}\right) \geq 1,$$

and an equality is attained if an only if T is a translation of K. A stability result for this inequality is a quantitative statement about the case that there is almost an equality in the above equation. In this case, it is reasonable to expect that K and T are approximately similar to each other with respect to a certain metric. Some examples of possible metrics are the Hausdorff distance, the Wasserstein distance and the volume of the symmetric difference between the bodies.

In [EK14], together with Klartag, we approached this topic from a high-dimensional point of view, trying to attain estimates that have a correct dependence on the dimension. As it turns out, the dependence in several metrics, including Wasserstein metric, are closely related to the constant σ_n. For two convex bodies K and T, the Wasserstein distance between K_1 and K_2 is defined as

$$W_2(K,T) = \inf_{\substack{(X_1,X_2) \\ X_i \sim \mathrm{U}(K_i)}} \sqrt{\mathbb{E}\|X_1 - X_2\|^2}$$

where the infimum is taken over all pairs of random vectors (X_1, X_2) such that X_i is uniform on K_i for $i = 1, 2$. In [Eld13], I showed that up to logarithmic terms, the variance conjecture is equivalent to the following.

Conjecture 2.13 *For all $\alpha > 0$ there exists a constant $C(\alpha)$ such that the following holds. Let K, T be two isotropic convex bodies, such that*

$$\mathrm{Vol}\left(\frac{K+T}{2}\right) \leq \alpha\sqrt{\mathrm{Vol}(K)\mathrm{Vol}(T)}.$$

Then,

$$W_2(K,T) \leq C(\alpha).$$

In our proof of this equivalence, we construct an explicit coupling between the uniform measures using martingales arising from the process described in Subsection 3.1. We hope to be able to combine these ideas with the existing methods in the literature, including optimal transportation (which seems to be the strongest tool for proving stability bounds known to date) towards conjecture 2.13.

2.2 Analysis of Boolean functions

A *Boolean function* is a function $f : \{-1,1\}^n \to \{-1,1\}$. The analysis of Boolean functions is a central pillar in theoretical computer science and graph theory. Concentration and structure theorems for these functions have a wide range of applications, spanning random percolations, algorithmic complexity, threshold phenomena for random graphs and constraint satisfaction problems.

Concentration inequalities usually relate the variance of f (say with respect to the uniform measure), to the way in which f oscillates locally. One important notion of local oscillations are the influences of f, defined as $I_i(f) = \mathbb{E}(\partial_i f(X))^2$ where X is uniformly distributed and ∂_i is the discrete derivative in the i-th direction.

Perhaps the simplest concentration inequality for Boolean functions is the Poincaré inequality which states that $\mathrm{Var}[f] \leq 2I(f)$, where $I(f) := \sum_i I_i(f)$. Two classical extensions of the Poincaré inequality are the the following: The Kahn-Kalai-Linial ([KKL88]) states that

$$\mathrm{Var}[f] \leq \frac{I(f)}{\min_i \log(1/I_i(f))},$$

improving on the Poincaré inequality by a logarithmic factor when all influences are small. Talagrand's sensitivity inequality [Tal94] states that

$$\mathrm{Var}[f] \leq \int |\nabla f(x)| d\mu(x),$$

thus being an L_1 version of the Poincaré inequality.

Each of these extensions exploits a different aspect of the fact that the function is Boolean, and their original proofs rely on rather different techniques. A 25-year-old conjecture of Talagrand [Tal96] suggests an inequality which strengthens both of these bounds, in the form

$$\mathrm{Var}[f] \leq \left(\int |\nabla f(x)| d\mu(x) \right) \left(\log \left(\frac{e}{\sum_i I_i(f)^2} \right) \right)^{-1/2}.$$

In a recent work with my student R. Gross [EG19], we have developed a new approach, based on pathwise analysis, towards proving concentration inequalities for Boolean functions. Using this approach, we settle Talagrand's conjecture, obtaining the above bound. This approach manages to bypass the use of hypercontractivity and produces several other new concentration results.

2.3 Localization and concentration of measures the discrete hypercube with applications to interacting particle systems and random graphs

Consider the Boolean hypercube $\{-1,1\}^n$, equipped with the uniform measure μ. In several recent works, we derive structure theorems which attempt to answer the following questions: Given a probability measure ν on $\{-1,1\}^n$,

Q1. Are there natural sufficient conditions on the measure ν such that it can be written as a convex combination

$$\nu = \sum_{i=1}^{N} \nu_i \tag{2}$$

where $N = o(2^n)$ and ν_i are probability measures which are in some sense close to a product measure?

Q2. What conditions on the measure ν ensure that it admits *concentration*, say in the form or either a spectral gap (with respect to a suitably chosen Dirichlet form), or in the form of concentration of Lipschitz functions?

A decomposition of the form (2) with the property that ν_i are localized is sometimes referred to in the physics literature as a *pure state decomposition*. We address the first question in two different works, considering two interpretations being close to product measures:

- A strong form of localization, considered in [Eld17] alludes to the existence of product measures ξ_i such that the W_1 transportation distance between ξ_i and ν_i is small.
- A weaker form of localization, considered in [Eld18] the pairwise correlations between coordinates of ν_i to be typically small.

The stronger form of localization relies on a notion of "complexity" of a measure put forth in [Eld17]. This notion is inspired by a related notion considered by Chatterjee and Dembo in their earlier work [CD16], which was motivated by questions regarding to large deviations subgraph counts on random graphs, and our work improves on the bounds given in this work. Both works [CD16] and our work [Eld17] produce a framework for investigating large-deviations in random graphs and mean-field approximations, but our work makes one more step by describing the structure of some large deviation events in additions to calculating the probabilities. An informal and qualitative statement of our result reads,

Theorem 2.14 *([Eld17]) Let ν be a Gibbs measure on $\{0,1\}^n$ with Hamiltonian f, which exhibits* low-complexity. *Then ν can be decomposed as a mixture,*

$$\nu = \int \nu_\theta dm(\theta),$$

where ν_θ are probability measures, so that the decomposition is entropy-efficient in the sense that

$$\mathrm{Ent}(\nu) \leq \int \mathrm{Ent}(\nu_\theta) dm(\theta) + o(n).$$

Moreover, there exists a set Θ with $m(\Theta) = 1 - o(1)$ such that for all $\theta \in \Theta$, there exists a product *measure ξ_θ such that $W_1(\nu_\theta, \xi_\theta) = o(n)$.*

It turns out that in order to obtain a weaker form of localization in the sense of the covariance matrix of the pure states being small, no assumption is needed on the measure ν. We prove the following theorem:

Theorem 2.15 *([Eld18]) For any measure ν on $\{0,1\}^n$ and any positive-definite matrix L, there exists a decomposition of the form*

$$\nu = \int \nu_\theta dm(\theta),$$

satisfying

$$\mathrm{Ent}(\nu) \leq \int \mathrm{Ent}(\nu_\theta) dm(\theta) + \log\det\left(\mathrm{Id} + \mathrm{Cov}(\nu) L\right) \tag{3}$$

and such that

$$\int \mathrm{Cov}(\nu_\theta) dm(\theta) \preceq L^{-1}.$$

More recently, we have managed to apply the same techniques towards obtaining sufficient conditions for the measure ν to satisfy concentration, addressing Q2 above. Our first result concerns with quadratic potentials, hence measures of the form $d\nu(x) \propto e^{-\langle x, Jx\rangle + \langle h, x\rangle} d\mu(x)$, with J being an arbitrary symmetric matrix and h an arbitrary vector. Together with Koehler and Zeitouni we show that the condition $\|J\|_{OP} \leq 1/4$ implies a Poincaré inequality with respect to the Dirichlet form given by the Glauber dynamics. A corollary is that the Glauber dynamics mixes in polynomial time over the Gibbs distribution of the Sherrington-Kirkpatrick spin glass model in high enough temperature. Our result reads,

Theorem 2.16 *([EKZ20]) If $\|J\|_O P \leq 1/4$, then for any test function φ, we have*

$$(1/4 - \|J\|_O P)\mathrm{Var}_\nu[\varphi] \leq \mathcal{E}_\nu(\varphi).$$

where $\mathcal{E}_\nu$ is the Dirichlet form corresponding to the Glauber dynamics.

A parallel work with my student Shamir tries to find more general convexity conditions under which the measure ν attains concentration, in search of a notion of log-concavity for functions on the discrete hypercube. A seemingly natural notion that we consider is in terms of the multi-linear extension of the density of ν into the solid hypercube. Recall that any function $f : \{-1, 1\}^n \to \mathbb{R}$ can be written uniquely in the form

$$f(x) = \sum_{A \subset [n]} \hat{f}(A) \prod_{i \in A} x_i,$$

also known as the Fourier-Walsh expansion of the function f. Since the above formula makes sense for any $x \in [-1, 1]^n$, we refer to this as the multi-linear extension of f into the solid cube $[-1, 1]^n$. We prove the following,

Theorem 2.17 *([ES20]) Let ν be a probability measure on $\{-1, 1\}^n$. Let f be the multi-linear extension of $\frac{d\nu}{d\mu}$. Suppose that*

$$\nabla^2 \log f(x) \leq \beta \mathrm{Id}, \quad \forall x \in [-1, 1]^n.$$

Then one has for any 1-Lipschitz function $\varphi : \{-1, 1\}^n \to \mathbb{R}$,

$$\mathrm{Var}_\nu[\varphi] \leq n^{2-c/\beta}.$$

One corollary of the above theorem obtains the first nontrivial concentration inequality for so-called *Rayleigh* distributions.

3 The Pathwise Method: from stochastic calculus to inequalities in high dimensions

Many of the aforementioned results are based on an emerging technique based on pathwise analysis, which has been at the center of my research. In this section we briefly describe this method.

The use of ideas from diffusion and heat-flow to concentration inequalities dates back at least to seminal works of Nelson and Gross, in the mid 70's, which introduced the hypercontractivity property of heat semigroups and derived the Log-Sobolev inequality for Gaussian space, respectively. In the following decades, semigroup and heat-flow 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 [BÉ85]. These ideas rely on differentiation formulas for the heat semigroup, which can in general, using the Feynmann-Kac equation be understood via pathwise integration along the corresponding diffusion.

Our method 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. This information can be revealed, with the help of the theory of stochastic calculus, by undertaking a pathwise approach.

The first applications of stochastic calculus to high dimensional inequalities known to us appeared in the early 2000's. To mention a few examples, a pathwise approach to heat semigroup proofs, aside from having a pedagogical value, has yielded several insights and results, see e.g., ([Cat04]). In works of C. Borell ([Bor00, Bor02]), ideas from stochastic control theory were used to prove concentration inequalities under a convex potential and also give a very simple and elegant proof of the Prékopa Leinder inequality. These methods, under a dual perspective, were used a decade later by Lehec ([Leh13]) to provide simple proofs of several entropy-related inequalities such as Talagrand's transportation-entropy inequality and the Shannon-Stam inequality. Recently, Van-Handel has extended the ideas of Borell, yielding a proof of the Ehrhard inequality ([VH17]).

Our general meta-technique can be described as follows. Given some probability space (Ω, μ) with $X \sim \mu$, we would like to sample the random variable X in a way that the random bits are generated infinitesimally, so that we can differentiate quantities related to X with the respect to the randomness (an example of such quantity is the conditional probability that X is in some set). This can be done by using a Brownian motion as the source of randomness, or in other words, by finding a *measure-preserving map* F from the space of paths of Brownian motion equipped with the Wiener measure to the space (Ω, μ). Given this mapping, we can use Ito calculus to differentiate with respect to the filtration associated with the Brownian motion.

Next, we describe two manifestations of this meta-technique (in other words, two constructions of a mapping F as described above), along with some of the applications of these constructions.

3.1 A stochastic localization scheme

One method for proving concentration inequalities is via *localization*. The idea is to iteratively reduce the size of the space while keeping some of the properties intact, until remaining set of focus has more tractable structure. This technique often allows one to reduce inequalities from arbitrary dimension to one-dimensional bounds. The first application of this method, known to us, was to derive isoperimetric inequalities on curved surfaces by Gromov and Milman ([GM87]). It has later been further developed by Lovász and Simonovits [LS93] in the setting of convex bodies.

These two iterate through *halving* the space via intersections with half-spaces. In ([Eld13]), we have introduced a new localization procedure which is conceptually different from existing localizations in several aspects: A. The localization is carried out on the space of functions rather than the space of sets, and can be thought of as a flow on measures. B. Randomness is introduced, and the process is driven by a Brownian motion. C. It is continuous in time.

Formally, for a measure μ on $\mathbb{R}^n$ our process is defined as the solution of the stochastic differential equation

$$F_0(x) = 1, \ dF_t(x) = \langle x - a(\mu_t), A(\mu_t)^{-1/2} dW_t \rangle F_t(x)$$

where W_t is a standard Brownian motion, $\frac{d\mu_t}{d\mu} = F_t$ and with $a(\nu), A(\nu)$ being the center of mass and covariance matrix of ν, respectively.

This process should be thought of as the continuous version of the following iterative scheme: at each time step, generate a random direction v uniformly in the sphere and define a new measure by multiplying the density of the current measure by a linear function whose slope is ϵv.

A useful property of this localization is that there is an easy correspondence between time derivatives (or Itô differentials) of quantities related to μ_t and its moments. These correspondences are in a sense analogous to formulas regarding time derivatives of heat semigroups, and also attain a certain similarity to the moment-generating properties of the Log-Laplace transform (these were used for example, by Klartag in his derivation of the best-known bound for the isotropic constant). The derivatives become tractable via the use of Itô's formula and using these relations, we were able to reduce concentration inequalities to the behavior of moments.

Another strong feature of this process is that after running it for a finite amount of time, the measure μ_t becomes "positively curved" in the sense that $-\nabla^2 \log \frac{d\mu_t}{dx}$ can be bounded from below by a multiple of the identity. Positively curved measures are known to admit concentration properties. Thus, this process is helpful in expressing a measure as a mixture of measures which admit concentration; for this reason we refer to this process as "localization". Moreover,

μ_t can be shown to converge, as $t \to \infty$, to a Dirac δ-measure, distributed according to μ. (Curiously, a decomposition into measures attaining good concentration bounds was also recently used by T. Austin [Aus17] to prove that any measure-preserving automorphism attains a weak-Pinsker property).

Roughly speaking, the stochastic localization process combines the advantages of the classical localization processes with moment-generating properties of the Log-Laplace transform and has tractable time differentiation formulas analogous to heat semigroups.

Our works [Eld13, EL14, Eld15b] all heavily rely on this process and, as do the recent work of Lee and Vempala [LV16], where slightly modified versions of the stochastic localization process are used to derive self contained state-of-the-art bounds for the KLS and thin-shell conjectures and of Klartag [Kla17] where isoperimetric inequalities for the Gaussian measure of high dimensional complex-analytic sets are derived.

3.2 Stochastic control theory and entropy minimization on path space

Our second main line of works builds on the ideas of Borell and Lehec [Bor02, Leh13], relying on a process from stochastic control theory which goes back at least to Föllmer and Boué-Dupuis ([Föl85, BD98]).

The main construction, in its most basic setting, is rather simple: Take a measure ν whose density with respect to the standard Gaussian measure γ is $\frac{d\nu}{d\gamma} =: f$. Now take a Brownian motion W_t with an underlying probability measure Q and define a new measure P by $\frac{dP}{dQ} = f(W_1)$. Clearly, $W_1 \sim \nu$ under P and moreover W_t is the process which minimizes the entropy with respect to the standard Brownian motion under that condition.

An application of Girsanov's formula that W_t can be alternatively defined in an adaptive manner, by the equation $dW_t = dB_t + v_t dt$ where $v_t = \nabla \log P_{1-t} f$ with P_s being the Gaussian convolution operator. The entropy-minimization property of the process is translated to several very useful properties of the associated drift and, as illustrated in the aforementioned works, the behavior of the measure ν can often be understood via scrutiny of this drift. A useful formula by Lehec states that the entropy of ν can be expressed as the "energy" $\frac{1}{2}\mathbb{E}\int_0^1 |v_t|^2 dt$, which turns out to be very useful in proving inequalities related to the entropy. For example, the Log-Sobolev in Gaussian space becomes an easy consequence of the fact that v_t is a martingale.

Natural analogues of this process exist beyond the scope of the Gaussian setting. In [Eld17], we have defined a related process on the discrete hypercube, and in [Leh15], Lehec constructed a counterpart on Riemannian manifolds. Moreover, in ([ELL16]) together with Lee and Lehec we show that related ideas can be fruitful in the context of local-to-global mixing bounds for Markov chains.

4 Algorithms, Learning Theory and Mathematical Optimization

In this section I discuss several applied directions of my research, mainly having to do with computational geometry and learning theory. The common goal shared by the examples below is to successfully apply ideas and theorems from the theory of high dimensional objects to these types of problems.

4.1 The Convex Bandit problem

The bandit convex optimization problem (see [BCB12] for background and a detailed description) is the following sequential game: a set $\mathcal{K}$ of actions is fixed throughout the game. At each time step $t = 1, \dots, T$, a player selects an action $x_t \in \mathcal{K}$, and simultaneously an adversary selects a convex loss function $\ell_t : \mathcal{K} \mapsto [0, 1]$. The player's feedback is its suffered loss, $\ell_t(x_t)$. We assume that the adversary is oblivious, that is the sequence of loss functions $\ell_1, \dots, \ell_T$ is chosen before the game starts. The player can select an action x_t based on the history $H_t = (x_s, \ell_s(x_s))_{s<t}$. The

player's perfomance at the end of the game is measured through the *regret*:

$$R_T = \sum_{t=1}^{T} \ell_t(x_t) - \min_{x\in\mathcal{K}} \sum_{t=1}^{T} \ell_t(x),$$

which compares their cumulative loss to the best cumulative loss which could have obtained in hindsight with a fixed action. The main line of research regarding this problem has been to find the correct asymptotics of the regret (under the optimal strategy) in terms of T. A central open problem since [Kle04, FKM05] has been to reduce the gap between the known $\sqrt{T}$-lower bound and the $T^{3/4}$-upper bound for the regret.

In dimension one (i.e., $\mathcal{K} = [0, 1]$) this gap was closed in [BDKP15]. Later on, together with S. Bubeck, we have settled the general (multi-dimensional) case [BE15]. Our proof is based on the construction of a multi-scale exploration process for convex functions which we believe to be independently interesting, and for which we hope to find other applications. This construction relies heavily on recent developments in the theory of high dimensional convex bodies. In particular, the notion of *isotropic position* which is closely related to Conjecture 2.9 plays a central role in the proof. However, our result is non-constructive in the sense that it only establishes the existence of a strategy with the optimal regret; it does not provide an explicit (efficient) algorithm.

More recently, together with S. Bubeck and Y.T. Lee, we managed to obtain the first polynomial time algorithm attaining $\sqrt{T}$-regret (and a polynomial dependence on the dimension) [BEL16]. However, the exponent of the dependence on the dimension ($n^{9.5}$) is far too high for practical purposes.

It seems that the dependence on the dimension in our results, though polynomial, is far from optimal. In the future, we hope to be able to integrate more knowledge from the theory of high dimensional convex bodies towards a solution of these problems.

4.2 Sampling from Log-concave measures via random walks

As mentioned above, the KLS constant G_n is related to the rate in which one can sample points from a convex body or log-concave measure, which is, in turn, an important ingredient in algorithms which perform computational tasks such as Principal Component Analysis (PCA), volume estimation and convex optimization. The most effective way to sample a point is via Monte Carlo Markov chains, or in other words via random walks. A classic example to this is the breakthrough work of Dyer-Frieze-Kannan [DFK91], who found the first polynomial time algorithm that estimates the volume of a high dimensional convex body given by a membership oracle.

Perhaps the simplest Markov chain associated with a convex body is a random walk on a (fine) grid restricted to the body, which was considered in [DFK91]. Since then, several other types of random walks have been considered, and there has been an ongoing attempt to prove upper bounds on the rate of mixing of those chains. The best known mixing bounds known to date have been obtained by Lovász and Vempala [LV07], attained by the so-called hit-and-run chain. However, it remains an open question to find a chain with an optimal rate. The question of finding such a chain is related to the KLS constant G_n, which determines the rate of mixing of a reflecting Brownian motion in continuous time. However, even if this rate were known, two caveats would remain: 1. Discretization of time and space poses some new difficulties. 2. The KLS constant is related to the so-called "relaxation time" which assumes that one starts from a distribution of bounded density (sometimes known as a "warm start").

In a work with S. Bubeck and J. Lehec [BEL15] we propose a new and seemingly natural way to sample from a log-concave measure restricted to a convex body, using a chain which we call Reflected Langevin Monte Carlo. Our construction builds the work of Dalalyan [Dal14], and integrates additional ideas from the theory of stochastic calculus. While we prove that the mixing time is polynomial in the dimension, the dependence that we get is likely far from optimal. Experimental results, however, suggest that our method competes Markov chains with the best known mixing, such as the hit-and-run walk ([LV07]) and there is evidence that in some cases it is faster.

We propose to investigate some of the above sampling schemes in search of the optimal rate. It may be that the implication between the constant G_n defined in Subsection 2.1 and the rate of mixing can be reversed; perhaps introducing the correct chain its analysis can help prove new bounds on G_n.

4.3 Complexity lower bounds using probabilistic constructions

In [Eld11] and [Eld15a], I have established some information-theoretical lower bounds on the number of independent samples drawn from a high dimensional distribution, which is assumed to be unknown to us, needed in order to estimate some quantities related to that distribution, namely its *entropy* and its *covariance matrix*. In [Eld11], it is shown that in order to estimate the volume of a convex body, one needs a number of samples which is super-polynomial in the dimension, thus answering a question posed by László Lovasz. In [Eld15a], it is shown that in order to reconstruct a single entry in the inverse covariance matrix of a high-dimensional distribution, one needs a number of samples proportional to the dimension, thus answering a question raised by statisticians.

One of the main principles behind these lower bounds is the fact that in order to distinguish a "typical" high dimensional log concave measure from its spherical symmetrization, a very high number of random samples is needed. There are several interesting questions that arise around this principle: is it true that, given a polynomial number of samples, any randomly rotated high dimensional log-concave distribution (which is properly normalized) cannot be told apart from some spherically symmetric distribution? Under what extra assumptions can one reconstruct the covariance matrix of a log concave distribution with a small number of samples? An answer to the first question may provide us with a deeper understanding of the distribution of mass in high dimensional convex bodies. An answer to the second question may be applicable in many data analysis techniques (e.g. in principal component analysis, discriminant analysis, graphical models).

4.4 Gradient descent methods and self-concordant barriers

A *Self-concordant barrier function* is a central object in the theory of interior point methods, a class of algorithms that has revolutionized mathematical optimization. The main idea behind the definition of this function (which amounts to several relations satisfied by its derivatives) is that when such a function is added to a linear function, the outcome will attain a property that ensures lower bounds on the number of iterations it takes to converge to a minimal solution via Newton's method. See [Nes04] for an exact definition and for further details.

From a theoretical perspective, one of the most important results in this theory is the existence of a self-concordant barrier over any given convex domain in Euclidean space. The first such construction was suggested by Nesterov and Nemirovski (1994). Recently, together with S. Bubeck ([BE14]) we have introduced a new construction and have managed to attain the first improvement over their result. Using elementary techniques from convex geometry, we were able to construct an (arguably) much simpler barrier, which happens to be the first to attain optimal parameters.

The construction we suggest seems to open many new directions for further study. While the significance of this result is mainly of a theoretical nature, an algorithmic application to Simulated Annealing has already been found [AH15]. Naturally, we would like to find additional applications.

Our construction exhibits deep connections to several recent discoveries in the theory of the distribution of mass on high dimensional convex sets. Several results of Klartag, and in particular the *Central limit theorem for convex sets* [Kla07] suggest that the entropic barrier is expected to have certain universality properties, meaning that at typical points, the behavior of this barrier is in fact independent of the underlying convex body. In the future, we hope to take advantage of these connections in order to show prove bounds on the convergence rate of gradient descent which uses this barrier.

References

[ABP03] Milla Anttila, Keith Ball, and Irini Perissinaki. The central limit problem for convex bodies. *Trans. Amer. Math. Soc.*, 355(12):4723–4735, 2003.

[AGB15] David Alonso-Gutiérrez and Jesús Bastero. *Approaching the Kannan-Lovász-Simonovits and variance conjectures*, volume 2131 of *Lecture Notes in Mathematics*. Springer, Cham, 2015.

[AH15] Jacob Abernethy and Elad Hazan. Faster convex optimization: Simulated annealing with an efficient universal barrier, 2015. arXiv:1507.02528.

[Aus17] Tim Austin. Measure concentration and the weak pinsker property, 2017. arXiv:1705.00302.

[BBN03] Keith Ball, Franck Barthe, and Assaf Naor. Entropy jumps in the presence of a spectral gap. *Duke Math. J.*, 119(1):41–63, 2003.

[BCB12] S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. 5(1):1–122, 2012. Foundations and Trends in Machine Learning.

[BCG14] Sergey G. Bobkov, Gennadiy P. Chistyakov, and Friedrich Götze. Berry-Esseen bounds in the entropic central limit theorem. *Probab. Theory Related Fields*, 159(3-4):435–478, 2014.

[BD98] Michelle Boué and Paul Dupuis. A variational representation for certain functionals of Brownian motion. *Ann. Probab.*, 26(4):1641–1659, 1998.

[BDKP15] S. Bubeck, O. Dekel, T. Koren, and Y. Peres. Bandit convex optimization: $\sqrt{T}$ regret in one dimension. In *Proceedings of the 28th Annual Conference on Learning Theory (COLT)*, 2015.

[BÉ85] D. Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[BE14] Sébastien Bubeck and Ronen Eldan. The entropic barrier: a simple and optimal universal self-concordant barrier. *Preprint. (arXiv:1412.1587)*, 2014.

[BE15] Sébastien Bubeck and Ronen Eldan. Multi-scale exploration of convex functions and bandit convex optimization. *Proceedings of the Conference on Learning Theory (arXiv:1507.06580)*, 2015.

[BEL15] Sébastien Bubeck, Ronen Eldan, and Joseph Lehec. Sampling from a log-concave distribution with projected langevin monte carlo, 2015. Discrete and Computational Geometry, to appear (arXiv:1507.02564).

[BEL16] Sébastien Bubeck, Ronen Eldan, and Yin Tat Lee. Kernel-based methods for bandit convex optimization, 2016. arXiv:1607.03084.

[BGVV14] Silouanos Brazitikos, Apostolos Giannopoulos, Petros Valettas, and Beatrice-Helen Vritsiou. *Geometry of isotropic convex bodies*, volume 196 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2014.

[BK03] S. G. Bobkov and A. Koldobsky. On the central limit property of convex bodies. In *Geometric aspects of functional analysis*, volume 1807 of *Lecture Notes in Math.*, pages 44–52. Springer, Berlin, 2003.

[BN12] Keith Ball and Van Hoang Nguyen. Entropy jumps for isotropic log-concave random vectors and spectral gap. *Studia Math.*, 213(1):81–96, 2012.

[Bor85] Christer Borell. Geometric bounds on the Ornstein-Uhlenbeck velocity process. *Z. Wahrsch. Verw. Gebiete*, 70(1):1–13, 1985.

[Bor00] Christer Borell. Diffusion equations and geometric inequalities. *Potential Anal.*, 12(1):49–71, 2000.

[Bor02] Christer Borell. Isoperimetry, log-concavity, and elasticity of option prices. In *New directions in Mathematical Finance*, pages 73–91. Wiley, 2002. Edited by P. Wilmott and H. Rasmussen.

[Cat04] Patrick Cattiaux. A pathwise approach of some classical inequalities. *Potential Anal.*, 20(4):361–394, 2004.

[CD16] Sourav Chatterjee and Amir Dembo. Nonlinear large deviations. *Adv. Math.*, 299:396–450, 2016.

[CK14] Thomas A. Courtade and Gowtham R. Kumar. Which Boolean functions maximize mutual information on noisy inputs? *IEEE Trans. Inform. Theory*, 60(8):4515–4525, 2014.

[Dal14] A. Dalalyan. Theoretical guarantees for approximate sampling from smooth and log-concave densities. *Arxiv preprint arXiv:1412.7392*, 2014.

[DEZ15] Jian Ding, Ronen Eldan, and Alex Zhai. On multiple peaks and moderate deviations for the supremum of a Gaussian field. *Ann. Probab.*, 43(6):3468–3493, 2015.

[DFK91] M. Dyer, A. Frieze, and R. Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the ACM (JACM)*, 38(1):1–17, 1991.

[EG19] Ronen Eldan and Renan Gross. Concentration on the boolean hypercube via pathwise stochsatic analysis. *Proceedings of the 52nd ACM Symposium on Theory of Computing (STOC 2020), in press.*, 2019.

[EK11] Ronen Eldan and Bo'az Klartag. Approximately Gaussian marginals and the hyperplane conjecture. In *Concentration, functional inequalities and isoperimetry*, volume 545 of *Contemp. Math.*, pages 55–68. Amer. Math. Soc., Providence, RI, 2011.

[EK14] Ronen Eldan and Bo'az Klartag. Dimensionality and the stability of the Brunn-Minkowski inequality. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 13(4):975–1007, 2014.

[EKZ20] Ronen Eldan, Frederic Koehler, and Ofer Zeitouni. A spectral condition for spectral gap: Fast mixing in high-temperature ising models. *arXiv preprint arXiv:2007.08200*, 2020.

[EL14] Ronen Eldan and Joseph Lehec. Bounding the norm of a log-concave vector via thin-shell estimates. In *Geometric aspects of functional analysis*, volume 2116 of *Lecture Notes in Math.*, pages 107–122. Springer, Cham, 2014.

[EL15] R. Eldan and J. R. Lee. Regularization under diffusion and anti-concentration of temperature. *Duke Mathematical Journal, to appear (arXiv: 1410.3887)*, 2015.

[Eld11] Ronen Eldan. A polynomial number of random points does not determine the volume of a convex body. *Discrete Comput. Geom.*, 46(1):29–47, 2011.

[Eld13] Ronen Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geom. Funct. Anal.*, 23(2):532–569, 2013.

[Eld15a] Ronen Eldan. An efficiency upper bound for inverse covariance estimation. *Israel J. Math.*, 207(1):1–9, 2015.

[Eld15b] Ronen Eldan. A two-sided estimate for the Gaussian noise stability deficit. *Invent. Math.*, 201(2):561–624, 2015.

[Eld16] Ronen Eldan. Skorokhod embeddings via stochastic flows on the space of Gaussian measures. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(3):1259–1280, 2016.

[Eld17] Ronen Eldan. Gaussian-width gradient complexity, reverse log-sobolev inequalities and nonlinear large deviations, 2017. Geom. Funct. Anal., to appear.

[Eld18] Ronen Eldan. Taming correlations through entropy-efficient measure decompositions with applications to mean-field approximation, 2018. Probab. Theory and Rel. Fields, to appear. (arXiv:1811.11530).

[ELL16] Ronen Eldan, James R. Lee, and Joseph Lehec. Transport-entropy inequalities and curvature in discrete-space markov chains, 2016. Preprint. (arXiv:1604.06859).

[EMZ18] Ronen Eldan, Dan Mikulincer, and Alex Zhai. The clt in high dimensions: quantitative bounds via martingale embedding, 2018. arXiv:1806.09087.

[ES20] Ronen Eldan and Omer Shamir. Log concavity and concentration of lipschitz functions on the boolean hypercube. *arXiv preprint arXiv:2007.13108*, 2020.

[FKM05] A. Flaxman, A. Kalai, and B. McMahan. Online convex optimization in the bandit setting: Gradient descent without a gradient. In *In Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2005.

[Föl85] H. Föllmer. An entropy approach to the time reversal of diffusion processes. In *Stochastic differential systems (Marseille-Luminy, 1984)*, volume 69 of *Lecture Notes in Control and Inform. Sci.*, pages 156–163. Springer, Berlin, 1985.

[GM87] M. Gromov and V. D. Milman. Generalization of the spherical isoperimetric inequality to uniformly convex Banach spaces. *Compositio Math.*, 62(3):263–282, 1987.

[IM12] Marcus Isaksson and Elchanan Mossel. Maximally stable Gaussian partitions with discrete applications. *Israel J. Math.*, 189:347–396, 2012.

[KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions (extended abstract). pages 68–80, 1988.

[Kla06] B. Klartag. On convex perturbations with a bounded isotropic constant. *Geom. Funct. Anal.*, 16(6):1274–1290, 2006.

[Kla07] B. Klartag. A central limit theorem for convex sets. *Invent. Math.*, 168(1):91–131, 2007.

[Kla17] Bo'az Klartag. Eldan's stochastic localization and tubular neighborhoods of complex-analytic sets, 2017. arXiv:1702.02315.

[Kle04] R. Kleinberg. Nearly tight bounds for the continuum-armed bandit problem. *Advances in Neural Information Processing Systems (NIPS)*, 2004.

[KLS95] R. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.*, 13(3-4):541–559, 1995.

[Leh13] Joseph Lehec. Representation formula for the entropy and functional inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(3):885–899, 2013.

[Leh15] Joseph Lehec. Borell's formula for a riemannian manifold and applications, 2015. arXiv:1512.05992.

[LS93] L. Lovász and M. Simonovits. Random walks in a convex body and an improved volume algorithm. *Random Structures Algorithms*, 4(4):359–412, 1993.

[LV07] László Lovász and Santosh Vempala. The geometry of logconcave functions and sampling algorithms. *Random Structures Algorithms*, 30(3):307–358, 2007.

[LV16] Yin Tat Lee and Santosh S. Vempala. Eldan's stochastic localization and the kls hyperplane conjecture: An improved lower bound for expansion, 2016. arXiv:1612.01507.

[Mil09] Emanuel Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Invent. Math.*, 177(1):1–43, 2009.

[MN15] Elchanan Mossel and Joe Neeman. Robust optimality of Gaussian noise stability. *J. Eur. Math. Soc. (JEMS)*, 17(2):433–482, 2015.

[Nes04] Yuri Nesterov. *Introductory lectures on convex optimization: A basic course*. Kluwer Academic Publishers, 2004.

[Tal89] Michel Talagrand. A conjecture on convolution operators, and a non-Dunford-Pettis operator on L^1. *Israel J. Math.*, 68(1):82–88, 1989.

[Tal94] Michel Talagrand. On russo's approximate zero-one law. *Ann. Probab.*, 22(3):1576–1587, 07 1994.

[Tal96] Michel Talagrand. How much are increasing sets positively correlated? *Combinatorica*, 16(2):243–258, 1996.

[VH17] Ramon Van Handel. The borell–ehrhard game. *Probability Theory and Related Fields*, Feb 2017.