Research Statement

Ronen Eldan

December 6, 2014

1 Introduction

My domain of research lies in the intersection between the areas of Functional Analysis, Probability and Convex Geometry. Specifically, I am mainly (but not only) interested in high dimensional problems of probabilistic or geometric nature.

The primary object of focus in my Ph.D. thesis was the uniform measure over high dimensional convex bodies. I have been interested in establishing connections between geometric and analytic quantities such as covariance, entropy and spectral gap associated with this measure as well as studying the algorithmic aspects related to it. Since then, several new neighboring areas captured my interest, such as noise stability in the Gaussian and discrete spaces, the geometry of random structures, Gaussian fields, discrepancy of hypergraphs, random graphs, convex optimization and metric embeddings.

One common goal I have been trying to pursue in many of my recent works is the attempt to integrate new ideas and results from probability theory to the above topics. The proofs of many classical theorems in high dimensional convex geometry highly rely on the *probabilistic method*. As an example, Dvoretzky's theorem (described below) which has been known for more than forty years, does not have a known deterministic proof in the sense that this theorem establishes the existence of an object, no deterministic construction of which is known. This theorem, together with many others in its field, apply a large variety of probabilistic concepts such as concentration of measure, the theory of martingales and Gaussian isoperimetry.

However, it appears that some of the fundamental tools from probability theory may have yet to be fully used. In the past three years, I have been trying to find applications of additional areas from probability theory such as the theories of *stochastic calculus* and *Poisson point-processes* for proving theorems which have no *a-priori* connection to these theories thus taking the probabilistic method to new frontiers. A few examples of such projects can be found in [E1, E2, E7, EL1, EL2, DEZ], and are listed in more detail below. These examples hopefully indicate that the pursuit of more such applications, in which I am interested to carry on, can be fruitful.

In the following, I will list some specific directions of research and examples of topics in which I'm interested. For each topic I will suggest a few ideas and possible directions for related further research.

2 High-dimensional convex bodies

High dimensional problems of a probabilistic or geometric nature appear in various branches of mathematics, mathematical physics and theoretical computer science, and have been extensively studied in the last few decades. A better understanding of such objects may lead to important applications, as demonstrated by numerous results that appeared in the last years. These results have been applied in a variety of subjects such as statistical mechanics, signal processing, computer vision, tomography and machine learning. The general subject of my research is the theory behind some of these high dimensional objects.

A well-known phenomenon in high dimensional systems is the exponential growth of information and complexity with respect to the dimension. This phenomenon is sometimes referred to as ”the curse of dimensionality”, since, in many cases, analyzing a high dimensional system can be a very complicated task. However, as recent research suggests, in many cases the contrary is actually true. There is a rapidly growing theory that demonstrates that a high dimension can in fact be a blessing rather than a curse. When viewed correctly, some high dimensional objects appear to have more order and simplicity than low-dimensional ones. Two examples which illustrate this phenomenon are Dvoretzky's theorem and the Central Limit Theorem for Convex Bodies. Both theorems show that, in some sense, typical projections of a high dimensional convex body have a ”normal” or ”common” behavior: the first one states that any high-dimensional convex body has nearly-Euclidean sections of a high dimension, while according to the second, any high-dimensional convex body has approximately Gaussian marginals. There is a number of principles, such as the Lévy-Milman concentration of measure, which seem to compensate for the diversity. Convexity is one of the ways to take advantage of those principles in order to prove easy-to-formulate, nontrivial theorems. The main focus of my research so far has been on the role of convexity in the high dimensional setting.

2.1 The Hyperplane Conjecture, the Variance Conjecture and the KLS Conjecture

As mentioned above, one of the main subjects of focus in my Ph.D. was the uniform measure over a high dimensional convex body, an object which has received a significant amount of attention in recent years. Numerous papers study properties related to this measure, such as its covariance, entropy and spectral gap. A major part of this research revolves around three central conjectures related to this measure which remain open still: The **hyperplane conjecture** (also known as the slicing problem), the conjecture of Kannan-Lovasz-Simonovitz (in short, the **KLS conjecture**) related to the isoperimetric inequality on convex bodies and the **variance conjecture** (also known as the thin-shell conjecture), which have been open for approximately 30, 20 and 10 years respectively.

In order to formulate these conjectures, we introduce some notation. For a probability measure μ on $\mathbb{R}^n$ with finite second moments, we consider its barycenter $b(\mu) \in \mathbb{R}^n$ and covariance matrix $Cov(\mu)$ defined by

$$b(\mu) = \int_{\mathbb{R}^n} x d\mu(x), \qquad Cov(\mu) = \int_{\mathbb{R}^n} (x - b(\mu)) \otimes (x - b(\mu)) d\mu(x)$$

where for $x \in \mathbb{R}^n$ we write $x \otimes x$ for the $n \times n$ matrix $(x_i x_j)_{i,j=1,\dots,n}$. A log-concave probability measure μ on $\mathbb{R}^n$ is *isotropic* if its barycenter lies at the origin and its covariance matrix is the identity matrix. For an isotropic, log-concave probability measure μ on $\mathbb{R}^n$ we denote

$$L_\mu = L_f = f(0)^{1/n}$$

where f is the log-concave density of μ. Define

$$L_n = \sup_\mu L_\mu$$

where the supremum runs over all isotropic, log-concave probability measure μ on $\mathbb{R}^n$. The following question is known as the *hyperplane conjecture*:

Question 2.1 *Is it true that $L_n \leq C$, for a universal constant $C > 0$?*

In an equivalent form of a more geometric nature, the conjecture can be stated as follows:

Question 2.2 *Is there a universal constant $c > 0$ such that for any dimension n and any convex body $K \subset \mathbb{R}^n$ with $Vol_n(K) = 1$, there exist a hyperplane $H \subset \mathbb{R}^n$ for which $Vol_{n-1}(K \cap H) > c$?*

We continue with the KLS conjecture. Roughly speaking, This conjecture asserts that, up to a constant, the best way to cut a convex body into two parts is with a hyperplane. To be more precise, given a measure μ, Minkowski's boundary measure of a Borel set $A \subset \mathbb{R}^n$, is defined by,

$$\mu^+(A) = \liminf_{\varepsilon \to 0} \frac{\mu(A_\varepsilon) - \mu(A)}{\varepsilon}$$

where

$$A_\varepsilon := \{x \in \mathbb{R}^n; \ \exists y \in A, \ |x - y| \leq \varepsilon\}$$

is the ε-extension of A. Define,

$$G_n^{-1} := \inf_{\mu} \inf_{A \subset \mathbb{R}^n} \frac{\mu^+(A)}{\mu(A)} \tag{1}$$

where μ runs over all isotropic log-concave measures in $\mathbb{R}^n$ and $A \subset \mathbb{R}^n$ runs over all Borel sets with $\mu(A) \leq \frac{1}{2}$. The KLS conjecture is stated as follows:

Conjecture 2.3 *There exists a universal constant C such that $G_n < C$ for all $n \in \mathbb{N}$.*

It turns out that constant G_n has numerous algorithmic implications (see e.g., [V] and references therein). Last, we would like to formulate the *variance conjecture*. We say that a random vector X in $\mathbb{R}^n$ is isotropic and log-concave if it is distributed according to an isotropic, log-concave probability measure. Let $\sigma_n \geq 0$ satisfy

$$\sigma_n^2 = \sup_X \mathbb{E}(|X| - \sqrt{n})^2 \tag{2}$$

where the supremum runs over all isotropic, log-concave random vectors X in $\mathbb{R}^n$. Thus, the parameter σ_n measures the width of the "thin spherical shell" of radius $\sqrt{n}$ in which most of the mass of X is located. The following question is known as the *variance conjecture*:

Question 2.4 *Is it true that,*

$$\sigma_n \leq C \tag{3}$$

for a universal constant $C > 0$?

A considerable part of my research during my Ph.D. studies was devoted to establishing quantitative links between these three conjectures. In a joint work with B. Klartag, [EK2], we establish the bound,

$$L_n \leq C\sigma_n$$

for some universal constant $C > 0$. In particular, we show that a positive answer to the variance conjecture will imply a positive answer to the hyperplane conjecture. Furthermore, in [E1], I have shown that a thin-shell bound implies a respective (a-priori stronger) isoperimetric inequality. Namely,

$$G_n \leq C\sigma_n \log n.$$

This shows that, in particular, a positive answer to the variance conjecture will imply, up to a logarithmic factor, a positive answer to the KLS conjecture. The latter bound has a simple geometric formulation: the minimizers in the isoperimetric inequality for convex bodies are, up to a logarithmic factor, ellipsoids.

In the proof of the bound $L_n < C\sigma_n$, in addition to some techniques related to the log-laplace transform, we use a new construction of a certain Riemannian manifold related to a convex body. A possible direction for further research here could be to try to attain a deeper understanding of the geometry of this Riemannian manifold, which might lead to more progress around one of these conjectures. For example, it is interesting to investigate if one can attain any non-trivial bounds on the *curvature* of this manifold, which will, as demonstrated in the paper, can be used to derive thin-shell bounds. Moreover, it seems that a deeper understanding of the *Logarithmic Laplace transform* of a convex body may lead to new realizations about its mass distribution. As demonstrated by several proofs of the recent years, including the latter, the Logarithmic Laplace transform seems to be rather powerful tool in analyzing of the distribution of mass of convex bodies.

2.2 Stochastic Analysis-related methods in isoperimetric and concentration inequalities

For the proof of the connection between the variance conjecture and the KLS conjecture (in [E1]), a novel localization scheme is used. This localization scheme is based on a certain stochastic flow on the manifold of Gaussian functions. This construction is one example on how one can combine several tools from the theory of stochastic processes with existing ones from the theory of log-concave functions in order to establish volume-related bounds. Another result which applies stochastic processes-related tools to functional inequalities is the one of Lehec, [Leh], in which a new stochastic construction is the main tool in providing simple proofs to several classical functional inequalities. Recently, together with James Lee, we have managed to use some of the ideas in Lehec's construction in order to solve a conjecture of Talagrand (see below more more details).

In the last couple of years, together with a few co-authors, I have managed to take advantage of stochastic tools and ideas related to construction which appears in [E1] in order to obtain several other results in analysis ([DEZ, EL1, EL2, E7, E8]). Many older important proofs (see e.g., Bakry-Emery, [BE]) use diffusive processes in order to establish volumetric properties for log-concave measures. While these older proofs consider these diffusions from a PDE point of view, the results mentioned above suggest that assuming a stochastic point of view can be rather helpful by allowing us to make use of fundamental theorems of stochastic calculus such as Itô's formula and Girsanov's theorem as well as elementary probabilistic techniques such as conditioning and coupling.

Thus, I would suggest, as a possible direction for further research, to try to understand how to combine more methods from the world of stochastic processes to high-dimensional geometric and functional inequalities. Hopefully, there exists other stochastic constructions such as these which will be useful in combining these methods. I also think it is worthwhile investing more time in further studying the constructions described in [E1] and [Leh] and also in trying to find analogous definitions on more general settings such as Riemannian manifolds.

2.3 The central limit theorem for convex sets

The variance conjecture is highly related to a remarkable result, the *Central Limit Theorem for Convex Bodies*, due to Klartag, [K1], stating that, in some sense, typical low-dimensional marginals of a high dimensional log concave measure is approximately Gaussian.

A theorem which we have proven in [EK1] is a generalization of this result to multi-dimensional marginals with a pointwise metric. We have shown that for marginals of log concave measures with small enough (polynomial) dimension, with high probability, the density approximates the gaussian density up to a multiplicative constant.

While the central limit theorem for convex bodies states that **most** marginals of a convex body are approximately Gaussian with respect to the haar measure, in order to answer several related questions which are still open, it seems that one should also understand more about the **geometry** of this set of directions. One example of such question has to do with the behavior of scalar products of points taken uniformly from an isotropic convex body. Namely, Let X and $\tilde{X}$ be two independent copies of an isotropic log concave random vector. Denote $Y = \langle X, \tilde{X} \rangle$. Is it true that Y approximates, in some sense, a Gaussian random variable? The latter question is related to a second question of more algorithmic nature: given two convex bodies, how many points sampled from the uniform distribution are needed in order to distinguish between the two after applying a random rotation?

Many theorems in high dimensional convex geometry use concentration of measure related techniques (e.g Dvoretzy's theorem, the Central Limit Theorem, the existence of subgaussian directions) to prove the existence of large subsets of the Grasmannian, in which sections or projections of convex bodies have a "typical" behavior. However, these theorems usually give no information whatsoever about the geometry of those sets, and thus it may be very fruitful to find altenative techniques which provide a better understading of the geometry. As demonstrated by recent developments in CS, such techniques are also likely to be useful in improving algorithms, derandomization and compression schemes that take advantage of those high dimensional phenomena.

3 Gaussian and Boolean Noise stability

The *noise stability* of a set A in Gaussian space with noise parameter $0 < \rho < 1$ is defined as

$$\mathcal{S}_\rho(A) := \mathbb{P}(X \in A \text{ and } Y \in A)$$

where X, Y are standard Gaussian random vectors in $\mathbb{R}^n$ distributed according to a multivariate Gaussian distribution whose covariance structure is $\mathbb{E}[X_i Y_j] = \delta_{i,j}\rho$. An inequality of C.Borell states that if H is a half-space having the same Gaussian measure as that of A, then $\mathcal{S}_\rho(H) \geq \mathcal{S}_\rho(A)$. This inequality admits applications in numerous fields such as approximation theory, high-dimensional phenomena and rearrangement inequalities, and recently found also some surprising applications in discrete analysis and game theory.

Recently, Mossel and Neeman [MN] managed to prove a *robustness* result for this inequality, showing that if the deficit $\mathcal{S}_\rho(H) - \mathcal{S}_\rho(A)$ is rather small, then the set A has to be an approximate half-space in the total-variation metric sense. Very recently, in ([E7]), I have managed to improve their result obtaining the optimal possible exponent in the dependence between the deficit and the distance. Moreover, I have shown that if one considers a different metric to measure the distance between a set and its corresponding half-space (namely, the distance between the centroids) then one may actually improve the estimate to obtain a two-sided bound for the deficit, thus bounding it by essentially the same quantity from both below and above, up to logarithmic factors.

One possible direction in which this work may be continued is to try to characterize what the most noise-sensitive (e.g. least noise stable) sets are under a fixed surface area constraint. It seems that if one tries to maximize the noise sensitivity, the extremal answers will be genuinely high-dimensional sets, unlike the stability case. Another interesting direction is to try and refine the metric used in [E7] in order to obtain a precise two-sided estimate for the stability, removing the logarithmic factor. A possible definition of such metric is suggested as a conjecture in [E7].

A central application of Borell's inequality is to the so-called *Majority is the stablest* theorem, due to Mossel, O'Donnel and Oleszkiewicz, which is roughly a discrete version of the same result. This theorem states that the most stable subsets of the discrete cube (where noise now corresponds to switching each coordinate with a fixed, small probability) are sets of the form $\{\sum \alpha_i x_i \geq L\}$, subject to a low influence condition.

In an ongoing work with Mossel, it seems that the ideas which I have applied in the Gaussian case could be modified to provide a new method for proving inequalities on the discrete cube. So far, we are able to directly prove results of the same spirit as the Majority is stablest theorem, exploring new conditions under which this theorem holds.

As a possible direction for future research I would suggest looking into more refined versions of the stability theorem for the discrete cube. Rather than the low influence condition which roughly states that the L_2-norm of each component of the gradient is rather small, one could hope that weaker conditions such as assuming that all Fourrier coefficients are small. One more related conjecture which intrigues me is the Entropy-Influence conjecture. It is not clear whether or not the methods developed in [E7] and more generally, whether assuming a stochastic point of view, may be beneficial in the investigation of this conjecture.

4 Logarithmic anti-concentration inequalities and regularization of L_1 functions under convolution operators

Let γ be the standard Gaussian measure in $\mathbb{R}^n$. For $f \in L_2(\gamma)$, consider the Ornstein-Uhlenbeck convolution operator

$$T_\rho[f](x) := \mathbb{E}[f(\sqrt{1-\rho}x + \sqrt{\rho}\Gamma)]$$

where Γ is a standard Gaussian random vector. It is by now classical that this operator attains *hypercontractive* properties, hence that for every $1 < p < q$, there exists $0 < \rho_0 < 1$ and a constant $C(p,q)$ such that for every $\rho_0 < \rho < 1$ one has

$$\|T_\rho f\|_{L_q} < C(p,q)\|f\|_{L_p}.$$

On the contrary, little can be said about regularization (or smoothing) of functions in L_1 under this operator, even though it seems that some quantitative estimate should indeed hold true. In [T], Talagrand asked the following question:

Question 4.1 *(Talagrand) Does there exist a function $g(\alpha)$ such that $\lim_{\alpha\to\infty} g(\alpha) = 0$ and such that for all non-negative function f satisfying $\mathbb{E}[f(\Gamma)] = 1$ one has the following improved Markov inequality*

$$\mathbb{P}(T_\rho[f](\Gamma) > \alpha) < \frac{C(\rho)}{\alpha} g(\alpha)?$$

Talagrand, in fact, formulated a stronger version of this question; Instead of considering the Gaussian space, one can analogously consider the discrete cube $\{-1,1\}^n$ whose respective convolution operator corresponds to re-randomizing every coordinate, independently, with probability ρ and taking the expectation of the function evaluated at the new point. Exactly the same question can be asked regarding this discrete analogue of the convolution operator.

Except for a dimension-dependent bound which was proved in [BBBOW], little progress has been made towards answering this question until, in a very recent work with J. Lee, we answered this question positively. Our result follows as an immediate corollary of the following logarithmic anti-concentration bound that we establish:

Theorem 4.2 *Let $f(x) : \mathbb{R}^n \to \mathbb{R}$ be a non-negative function satisfying $\mathbb{E}[f(\Gamma)] = 1$ and*

$$\nabla^2 f(x) \succeq -\beta \mathrm{Id}, \ \ \forall x \in \mathbb{R}^n. \tag{4}$$

One has for all $\alpha > 2$,

$$\mathbb{E}\left[f(\Gamma)\mathbf{1}_{\{f(\Gamma)\in[\alpha,2\alpha]\}}\right] < \frac{C(\beta)}{(\log\alpha)^{1/4}}.$$

In order to see how Theorem 4.2 gives a positive answer to Question 4.1, note that $T_\rho[f]$ is a mixture of Gaussian densities satisfying condition (4) with constant $\beta = \frac{1}{1-\rho}$, and observe that this condition is preserved under summation, as a consequence of Artin's theorem, which states that a mixture of log-convex measures is log-convex.

Two main related questions are still left open. The first one is naturally to answer the question in the discrete setting. By applying the central limit theorem, it is not hard to check that the discrete version of this conjecture is stronger than the Gaussian one. Moreover, the discrete version is independently interesting since it has some applications to metric embeddings. The second question concerns with the exact asymptotics of the function $g(\alpha)$. It is natural to conjecture that, optimally, $g(\alpha) \sim \frac{1}{\sqrt{\log\alpha}}$ (which is the asymptotics attained when f is the normalized indicator of a half-space) whereas our result only yields $g(\alpha) \sim \frac{1}{(\log\alpha)^{1/4}}$.

As mentioned above, the answer to Talagrand's question follows from the more general inequality in Theorem 4.2. This bound which appears to be a new type of functional inequality suggests a few directions of further research as well. One question to ask is whether it may be generalized to a broader class of geometries. For example, it is natural to investigate whether an analogous result holds in positively curved Riemannian manifolds or more generally in spaces attaining the Bakry-Emery Curvature-Dimension condition (see [BE] for a definition). Another question is whether it has additional applications to concentration inequalities which are not necessarily related to the convoltuion operator.

5 Gaussian fields

Recently, I have become interested in trying to understand the connections between several quantities related to Gaussian processes in a general setting. Given a *Gaussian process*, namely, a finite number of (one-dimensional) Gaussian variables with a certain covariance structure, there are several natural quantities related to this process such that the expectation and variance of its maximum, its large deviation behaviour, the stability of the maximizer with respect to small perturbations of the process, correlations between the support vectors, etc.

While the expected supremum of a Gaussian process is a rather well-understood quantity (studied in landmark works by Dudley, Fernique, Talagrand and numerous follow-up works) it seems that there is still much to be learnt about the other related quantities and the connections between those quantities. One remarkable work exploring such connections was carried out by Chatterjee ([C]). A better understanding of such connections will have numerous applications to related subjects in probability, such as spin systems.

In a very recent and ongoing work with J. Ding and A. Zhai, we have continued Chatterjee's line of research. For example, we show that if a Gaussian process has a rather large expected maximum, then its tail admits improved large deviations bounds (compared to the ones obtained by the isoperimetric inequality). We also show that a process admitting super-concentration (hence, a non-trivial bound for its standard deviation) must have a rather large expected supremum and, moreover, we derive exponential bounds for the number of multiple-peaks of such processes, thus improving Chatterjee's estimate in [C].

I am highly interested in further exploring such connections. For example, I think it could be extremely useful to try to further characterize processes admitting the super-concentration property. Moreover, as most of these quantities and connections have so far been studied from a probabilistic point of view, I would be very interested to try to find relevant related quantities of a more geometric nature, such as quantities related to the unit ball whose corresponding norm is the maximum of the Gaussian process. There are also several specific examples of Gaussian processes which attract my interest, such as the S-K spin-glass model and the first-passage percolation (described in [C]).

6 Algorithmic and Information theoretic bounds in Convex Geometry and Graph Theory

6.1 Complexity lower bounds using probabilistic constructions

In [E2] and [E5], I have managed to establish 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 [E2], 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 [E5], 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).

6.2 Dimensionality of Geometric Graphs

A random geometric Graph $G(n, p, d)$ is a random Graph on n vertices constructed as follows: First, a uniformly chosen random point on the d-dimensional sphere is assigned independently for each vertex, $v \in [n]$. Then, the edges are assigned so that two vertices are connected if and only if their corresponding scalar product exceed a threshold t, chosen such that the probability for any two vertices to be connected is p. By $G(n, p)$ we denote the Erdös-Rényi graph, namely the graph on n vertices such that each pair of vertices is connected with probability p, independently.

Random geometric graphs were first considered in the 90's. Recently, the analysis of these object is becoming more popular due to its connection to the theory of social networks: By considering points on the sphere as feature vectors which encapsulate, say, the interests of a person, the model in which two vertices are more likely to be connected when their feature vectors are close to each other makes sense in the context of a social network.

In a very recent work with Bubeck, Ding and Rácz [BDER], we find the precise asymptotic threshold for the graph $G(n, p, d)$ to be indistinguishable from $G(n, p)$, as the size of the graph goes to infinity, proving that for any $0 < p < 1$, the total variation between the two graphs goes to zero whenever $d \gg n^3$ and to one whenever $d \ll n^3$. Moreover, we establish that $d \sim n$ is the threshold for being able to recover the precise dimension and find an efficient algorithm for doing so.

This emerging theory offers many open research questions, which I would be interested in pursuing. Let me mention two of them: First, it would make more sense, at least in the context of social networks, to consider the "sparse regime", hence ask the same questions when p goes to zero with n. In particular, the case where $p \cdot n$ is equal to a constant seems natural. Second, it is interesting to investigate other related models where, for example, the underlying geometry is different or the rule for connecting vertices is modified in order to get a model that better reflects actual social networks.

6.3 Random attachment models

Another relatively new research direction in graph theory, which also seems to model several real-world networks is recursive constructions of graphs via uniform or preferential random attachment. Such graphs are constructed by starting with a deterministic initial graph, usually called a seed. Then, the construction is continued in a recursive manner by adding new vertices repeatedly and connecting them to the existing graph according to a certain rule. One example is the *uniform attachment* tree model in which each new vertex will be connected to the existing graph via a single edge whose other vertex is chosen uniformly at random and independently from the previous choices. The corresponding *preferential attachment* model is the one in which the probability to connect to an existing vertex is proportional to its degree raised to some power $\alpha > 0$. Two recent works dealing with such models are [BMR, CDKM].

One fundamental question to be asked about such models is whether the final distribution of the graph depends on the seed. In other words, one can ask if the total variation distance between two such graphs having any two different seeds does not go to zero as the number of vertices goes to infinity. In [CDKM], this question was answered in the case that $\alpha = 1$. In a very recent work with Bubeck, Mossel and Rácz, [BEMR], we have solved this question for the uniform attachment tree model, showing that the total variation distance indeed remains bounded away from zero. In another ongoing work with Bubeck, Rácz and T. Schramm, we have been able to make some progress in related questions about the non-tree uniform attachment model (in which each new vertex sends more than one edge to the existing graph).

However, many other interesting questions, which I would be happy to pursue, remain open still. One central direction is to better understand more aspects of the non-tree version of both uniform and preferential attachment models. Another is to better understand the dependence of the behavior of the preferential attachment model on the parameter α.

7 High-dimensional Random Walks and Brownian Motion

One more subject into which I have been looking is geometric properties of the convex hull of high dimensional random walks and Brownian motion ([E3], [E4]). I have been able to derive formulas quantifying some geometric properties of this convex hull such as the volume and surface area, the number of extremal points, the distribution of facets, etc. However, many questions about the behaviour of these objects are left unanswered. The investigation of those random convex bodies could be interesting for probabilists, but also, it may be beneficial from a high dimensional convex geometer's viewpoint, since there are relatively few concrete examples of high dimensional convex bodies, and the study of a new concrete convex body may help us attain a deeper understanding of the general theory, as well as provide counterexamples for conjectures relating to convex bodies. Moreover, as demonstrated by these works, it seems that when considering high dimensional random walks, tools taken from the world of high dimensional phenomena may be applied, and thus provide information about objects of probabilistic nature, like mixing times of random walks on the sphere. Some of these ideas, as well as some suggestions for further research are listed in [E3, section 6] and in [E4, section 6].

8 Geometry of random structures

I have also recently become interested in random geometric structures such as *diffusion-limited aggregation* and the *self-avoiding walk*. Some fundamental questions about these two models (such as the rate of growth of their diameter) seem to be notoriously hard and are open despite having occupied many probabilists in the last decades. Besides making an attempt at these fundamental questions, I am also interested in considering slight variations of these models which may be easier to analyze on one hand, and could hopefully contribute to our understanding of the original models on the other hand. Two related works from the past year are the following: in [E9] I proved that the diffusion-limited aggregation model, if defined on the hyperbolic space rather than in the Euclidean setting, has a positive density at time infinity (unlike what is conjectured in the Euclidean case), and in a work in progress with I. Benjamini, we show that a self avoiding walk on a cylindrical graph admits certain symmetry breaking phenomenon, and in some cases of cylinders has zero density.

9 Stability of the Brunn-Minkowski inequality

In one of its forms, the Brunn-Minkowski inequality states that for two convex bodies, K, T of volume 1, 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 deals with 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 in some sense, or in other words, close 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 [EK3], in a joint work with B.Klartag, we establish several such stability results. Unlike previous results, we approach the topic from a high-dimensional point of view, and try to attain estimates that have a correct dependence on the dimension. Our results demonstrate how, in some cases, the estimates may actually become better as the dimension grows. The techniques and ideas we use mostly from the theory of high-dimensional convex bodies, and many are related to concentration of mass results. Moreover, as we demonstrate in [EK3], there appears to be a strong link between certain stability bounds and bounds related to the distribution of mass on convex bodies.

This raises many natural questions regarding the actual dependence of existing stability results on the dimension. An example of a possible direction of research here is to determine the best dependence on the dimension in the result of [FMP]. Moreover, it seems that the search of refined bounds and new methods related to the Brunn-Minkowski inequality may give rise to new estimates regarding the distribution of mass of high-dimensional convex bodies.

10 Curvature conditions for graphs and Markov Chains

Another intriguing topic which recently attracted my attention is the attempt to find a parallel for the Bakry-Emery theory in the context of graphs or reversible Markov chains. Given Riemannian manifold with a measure defined over it, in the 1980s Bakry and Emery ([BE]) found a condition on the Ricci curvature of the Manifold and the Hessian of the potential of the measure, which is satisfied by many natural examples and which implies a variety of functional inequalities such as the log-Sobolev inequality, Talagrand's transportation-entropy inequality and concavity of the entropy functional. This condition is usually referred to as the Bakry-Emery Curvature-Dimension condition.

More recently, several attempts have been made to find a discrete analog for this condition, namely where the Riemannian manifold is replaced by a reversible Markov chain. However, it seems that a satisfactory condition that is universally agreed upon has yet to be found. For some examples of notable attempts, see [GRST] and references therein. One rather natural such condition, suggested by Y. Ollivier is Dobrushin's contraction property in transportation distance. This condition roughly says that given two states of the Markov chain, there exists a coupling between the random walks started from those states which contracts the distances by a multiplicative factor, in expectation. A natural conjecture suggested by Y. Peres is that such a condition may imply, for instance, a respective log-Sobolev inequality. Such results would have immediate consequences for many natural models which satisfy this curvature condition, like spin systems and Glauber dynamics on graph colorings.

In an ongoing joint work with James Lee, we have so far managed to show that this condition implies a transportation-entropy inequality. Apart from trying to establish a log-Sobolev inequality under this condition, I am also intrigued by some other related questions that remain open, such as whether there is a natural way to interpolate measures on discrete spaces so that the entropy will be concave along the interpolation and whether there exists a natural curvature condition for a markov chain with an overlying measure.

11 Discrete Optimization via stochastic rounding schemes

Solving optimization problems over a discrete space is often exponentially hard with respect to the number of variables. One of the central ways to overcome this hardness is to consider a relaxation of this problem to a continuous setting, and then use a certain rounding technique in order to round the solution back to the original discrete space, obtaining an approximately-optimal solution.

One example where this idea has proven very effective is in finding the largest cut of the graph, or in general finding the cut-norm of a matrix. Given a matrix (a_{ij}) for $i \in I, j \in J$, one wants to find subsets

$I' \subset I, J' \subset J$ which maximize $|\sum_{i,j \in I' \times J'} a_{ij}|$. This problem is hard in general, but a method suggested by Alon-Naor [1] gives an algorithm which provides a constant approximation to this problem by two steps: the first step is to solve a Semi-definite programing problem to maximize $\sum_{i \in I, j \in J} a_{ij} u_i \cdot v_j$ where u_i, v_j are vectors in a high-dimensional sphere, and the second step is to "round" the vectors v_i, u_j to numbers $\{0, 1\}$ using a rounding scheme provided by Krivine in the 80's. This scheme ensures us that the bound obtained by the rounded solution will approximate the bound obtained by the continuous relaxation up to some universal constant.

This idea raises a very natural question, namely, given a continuous relaxation of a problem, what are the good rounding schemes and what are the corresponding ratios which they ensure? In particular, the best possible rounding scheme for the cut-norm problem is unknown, and is referred to as the Grothendieck constant. Finding this constant, which is considered a fundamental problem in functional analysis would have many algorithmic implications as well as a large impact on related questions in Functional analysis. It was conjectured by Krivine that a specific scheme called hyperplane-rounding is optimal (in the sense that it attains the Grothendieck constant). However, in a more recent paper, Braverman-Makarychev-Makarychev-Naor [2], show that this is not the case by demonstrating a more efficient rounding scheme.

The idea behind the improvement of [2] is to consider a family of rounding schemes associated with subsets of the plane, which contains Krivine's hyperplane rounding scheme as a point in this family, and prove that a specific perturbation of this scheme, inside the family, gives a better bound. Their simulations also found a fixed point referred to as the "Tiger pattern", a subset of $\mathbb{R}^2$ which is by itself intriguing. However, their method does not seem to provide the optimal bound.

While in [2], perturbations over a subsets of the plane are considered, my recent work [E7] implies that Krivine's method can also be considered as a point in a family of rounding schemes associated with stochastic processes. This suggests a new point of view on rounding schemes using random walks, and gives a new way to consider perturbations of the hyperplane scheme, which seem fairly tractable. In light of these new ideas, there are numerous related directions one may investigate. These could hopefully shed some light on the Grothendieck constant and provide new insights on rounding algorithms.

12 Barrier functions in convex optimization

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 certain bounds on the number of iterations it takes to converge to a minimal solution via Newton's method. See [N] 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). This has been the only known construction until recently, together with S. Bubeck ([BE]) 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. As mentioned above the significance of this result is mainly of a theoretical nature. It is not clear at this point if applying this construction to existing algorithms can improve their complexity, or whether one can find new algorithms which rely on this construction. Moreover, our universal barrier seems to have deep connections with the one of Nesterov-Nemirovski; it turns out that the two are the same on certain classes of convex bodies. However, this connection is at this point somewhat mysterious to us, and it may be fruitful to understand it better.

References

[1] N. Alon, A. Naor *Approximating the cut-norm via Grothendieck?s inequality* STOC '04.

[BE] D. Bakry and M. Emery, *Diffusions hypercontractives*. Séminaire de probabilités, XIX, 1983/84, vol. 1123 of Lecture Notes in Math., Springer, Berlin, 1985, pp. 177-206.

[BBBOW] K. Ball, F. Barthe, W. Bednorz, K. Oleszkiewicz and P. Wolff *L^1-smoothing for the Ornstein-Uhlenbeck semigroup.* Mathematika, Volume 59, Issue 01, January 2013, pp 160-168.

[2] M. Braverman, K. Makarychev, Y. Makarychev, A. Naor *The Grothendieck constant is strictly smaller than Krivine's bound* FOCS '11, arXiv:1103.6161.

[BDER] S. Bubeck, J. Ding, R. Eldan and M. Racz *Estimation in high-dimensional random geometric graphs.* Preprint, arXiv:1411.5713.

[BE] S. Bubeck, R. Eldan *The entropic barrier: a simple and optimal universal self-concordant barrier.* Preprint, arXiv:1412.1587.

[BEMR] S. Bubeck, R. Eldan, E. Mossel and M. Racz *From trees to seeds: on the inference of the seed from large trees in the uniform attachment model.* Preprint, arXiv:1410.3887.

[BMR] S. Bubeck, E. Mossel, and M. Z. Rácz. *On the infuence of the seed graph in the preferential attachment model.* arXiv preprint arXiv:1401.4849, 2014.

[CDKM] N. Curien, T. Duquesne, I. Kortchemski, and I. Manolescu. *Scaling limits and influence of the seed graph in preferential attachment trees.* arXiv preprint arXiv:1406.1758, 2014.

[BE2] I. Benjamini and R. Eldan *Self avoiding walks on cylinders.* Work in progress.

[C] S. Chatterjee, *Chaos, concentration, and multiple valleys* , arXiv:0810.4221 (2008).

[DEZ] J. Ding, R. Eldan and A. Zhai *On Multiple Peaks and Large Deviations of Gaussian Processes.* arXiv:1311.5592, to appear in Annals of Probability.

[E1] R. Eldan, *Thin shell implies spectral gap via a stochastic localization scheme*, Geometric and Functional Analysis. April 2013, Volume 23, Issue 2, pp 532-569.

[E2] R. Eldan, *A Polynomial Number of Random Points does not Determine the Volume of a Convex Body.*, Discrete and Computational Geometry Volume 46 Issue 1 (2009)

[E3] R. Eldan, *Extremal points of high dimensional random walks and mixing times of a Brownian motion on the sphere.* To appear in Annales de l'Institut Henri Poincare (B) Probability and Statistics (2012).

[E4] R. Eldan, *Volumetric Properties of the Convex Hull of an n-dimensional Brownian Motion*. arXiv:1211.2443, to appear in Electronic Journal of Probability (2012).

[E5] R. Eldan, *An efficiency upper bound for inverse covariance estimation.* To appear in Isr. Journal of Maths (2013).

[E6] R. Eldan, *Diffusion limited aggregation in the hyperbolic plane.* arXiv:1306.3129, to appear in Annals of Probability (2013).

[E7] R. Eldan, *A two-sided estimate for the Gaussian noise stability deficit.* arXiv:1307.2781 (2013).

[E8] R. Eldan, *Skorokhod Embeddings via Stochastic Flows on the Space of Measures.* arXiv:1303.3315 (2013).

[E9] R. Eldan, *Diffusion limited aggregation in the hyperbolic plane.* arXiv:1306.3129, to appear in Annals of Probability.

[EK1] R. Eldan, B. Klartag, *Pointwise Estimates for Marginals of Convex Bodies.*, Journal Functional Analysis 254, pp. 2275-2293 (2008).

[EK2] R. Eldan, B. Klartag, *Approximately gaussian marginals and the hyperplane conjecture.* Proc. of a workshop on "Concentration, Functional Inequalities and Isoperimetry", Contemporary Math., vol. 545, Amer. Math. Soc., 55–68 (2011).

[EK3] R. Eldan, B. Klartag, *Dimensionality and the stability of the Brunn-Minkowski inequality*, to appear in Annali della Scuola Normale Superiore di Pisa, Classe di Scienze (2012)

[EL1] R. Eldan and J. Lee, *Regularization under diffusion and anti-concentration of temperature*. Preprint (2014) (Available by request).

[EL2] R. Eldan and J. Lehec, *Bounding the norm of a log-concave vector via thin-shell estimates*. To appear in Inventiones Mathematicae.

[FMP] Figalli, A., Maggi, F., Pratelli, A., *A refined Brunn-Minkowski inequality for convex sets.* Ann. Inst. H. Poincaré Anal. Non Linéaire, vol. 26, no. 6, (2009), 2511-2519.

[GRST] N. Gozlan, C. Roberto, P. Samson and P. Tetali *Displacement convexity of entropy and related inequalities on graphs* Probability Theory and Related Fields, 2013.

[K1] Klartag, B., *A central limit theorem for convex sets.* Invent. Math., 168, (2007), 91–131.

[K2] Klartag, B., *A Berry-Esseen type inequality for convex bodies with an unconditional basis.* Probab. Theory Related Fields, vol. 145, no. 1-2, (2009), p.1-33.

[Leh] J. Lehec, *A representation formula for the entropy and functional inequalities.*, Annales de l'Institut Henri Poincare (B) Probability and Statistics (2010).

[MN] E. Mossel and J. Neeman. *Robust optimality of gaussian noise stability.* Arxiv : 1210.4126, 2012.

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

[T] M. Talagrand, *A conjecture on convolution operators, and a non-Dunford-Pettis operator on* L_1. Israel J. Math. 68 (1989), no. 1, 82?88.

[V] S. Vempala, *Recent progress and open problems in algorithmic convex geometry.* DOI: 10.4230/LIPIcs.FSTTCS.2010.42.