



Stochastic methods in convexity

Pierre Bizeul

► To cite this version:

Pierre Bizeul. Stochastic methods in convexity. Functional Analysis [math.FA]. Sorbonne Université, 2023. English. NNT : 2023SORUS731 . tel-04577476

HAL Id: tel-04577476

<https://theses.hal.science/tel-04577476>

Submitted on 16 May 2024

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

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



École doctorale de sciences mathématiques de Paris centre

THÈSE DE DOCTORAT

Discipline : Mathématiques

présentée par

Pierre BIZEUL

Stochastic methods in convexity

Soutenue le 14 septembre 2023 devant un jury composé de :

Thierry	BODINEAU	Examineur et Président du Jury
Sergey	BOBKOV	Rapporteur
Arnaud	GUILLIN	Rapporteur
Omer	FRIEDLAND	Examineur
Dario	CORDERO-ERAUQUIN	Directeur de thèse
Joseph	LEHEC	Directeur de thèse

Institut de mathématiques de Jussieu-
Paris Rive gauche. UMR 7586.
Boîte courrier 247
4 place Jussieu
75 252 Paris Cedex 05

Université de Paris.
École doctorale de sciences
mathématiques de Paris centre.
Boîte courrier 290
4 place Jussieu
75 252 Paris Cedex 05

Remerciements

N'ayant jamais été en avance dans les démarches administratives, le temps me manque pour écrire des remerciements à la hauteur des gens qui m'ont entouré pendant cette thèse.

Merci à Dario et Joseph de m'avoir pris comme élève et d'avoir rédigé un projet sous 48h. Après tout un lot de tergiversations, je suis heureux d'avoir atterri à cet endroit. Merci pour votre investissement constant au cours de ces années, votre bienveillance et votre patience pour mes inconstances. Je mesure la chance que j'ai eu d'avoir pu vous regarder travailler, et de travailler avec vous.

Je tiens à remercier vivement Sergey Bobkov et Arnaud Guillin de s'être intéressés à mon travail et d'avoir pris le temps d'être rapporteurs de cette thèse. Special thanks to Sergey for introducing me to a very nice problem on our first meeting. Merci également aux examinateurs, Thierry Bodineau, avec qui j'ai eu la chance de discuter et qui m'a réservé un accueil des plus chaleureux, et Omer Friedland, dont la bonne humeur et l'enthousiasme m'ont accompagné tout au long de cette thèse.

A la loterie des tuteurs, j'ai eu la joie d'avoir Gilles Godefroy. Merci Gilles pour les nombreuses discussions, je n'aurais pas pu imaginer figure plus rassurante et agréable dans ce rôle. J'en profite pour remercier les nombreux chercheurs qui ont donné de leur temps pour me conseiller à différents moments de mon cursus, Benoît Stroh, Antoine Chambert-Loir, et plus particulièrement Djalil Chafaï et Michel Ledoux. La vie scientifique convexe parisienne n'aurait pas été la même sans les organisateurs et participants au CTOP, merci pour ça. En bon taupin, je salue aussi mes anciens professeurs, Jean-Christophe Feauveau et Eric Desmeules pour m'avoir donné le goût de la science, et des scènes de vieux couple pour commencer la journée. Une pensée homologique pour Alexandru Oancea, et merci Jamal, mon unique co-auteur jusqu'à présent, ce fut un véritable bonheur de travailler avec toi. Enfin, merci Emmanuel pour le post-doctorat idéal, j'ai hâte qu'on travaille ensemble.

A mes amis, les mathématicien.ne.s d'abord, merci de m'offrir tout le spectre d'intoxication mathématique, de Jeanne et Sholom à Johan, Nicolas au milieu et Adam quelque part ailleurs, le 72ème projet de collaboration sera le bon, mais d'abord je dois résoudre le TSP avec Meyer. Merci Chloé R. pour Godard, Marylou pour Rohmer, Bérangère pour *Camping*. Merci Coline, Grégoire Ben'Houz et tous les autres pour toutes les fêtes, 3,4 ! Merci ma soeur de les organiser, et de figurer parmi mes amis. Merci Pauline si un jour

tu me lis. Merci surtout à la troupe extraordinaire du quotidien, pour certains déjà cités, Jean-Khui la pinte de trop, Modjo au petit trot, Nicolò la phrase de trop, Madame Elo, Jeannot, 2Chlo -à Cannes l'an prochain, Marylove, Lolo qui ne s'est jamais appelée comme ça, Marylou a bientôt et Bebert 303. Et de plus en plus j'espère, Jedge, Salomé Nica. Merci Laetitia, sans tréma, j'espère que tu sais pourquoi, pour hier aujourd'hui et demain.

A tous ceux que j'ai vu moins souvent ces derniers temps, j'espère toujours vous revoir plus et ce soir peut-être, Guigui à la prochaine urbex, Ségal et toute l'équipe de Villejuif, Kevin Simon Emmeline Basile et tous les autres, tous mes anciens et toujours futurs camarades, Alice, Mathilde, Dam, Ninon Mahmoud et Samuel, produit de luxe à la rareté inégalée, et le revenant MC Kozo, qui un jour reviendra. A Saint-Affrique en 1871 naissait Emile Borel, et un peu plus d'un siècle plus tard, un paquet de collègues, Adri, Bosco, Balmé, Couta, Max, Garlenc et Gachou, les meilleures fêtes se font en Aveyron - et à Veyreau. Et bien sûr mon plus vieil ami, de Saint-Affrique à Toulouse à Paris, à la Cour des Miracles pour quelques point elo, mecton Cazottes.

Merci enfin à ma famille, au grand complet pour le modèle de convivialité, mes oncles tantes, cousins et grands-parents. Et pour finir, merci à mes parents de m'avoir toujours soutenu et encouragé, merci pour votre bonne humeur et merci d'être venu.

Contents

Introduction	7
0.1 Notations and mathematical preliminaries	12
0.1.1 Log-concavity	13
1 Entropy and Information jump for log-concave vectors	15
1.1 Introduction	15
1.2 A lemma of Ball-Nguyen	18
1.3 The Ornstein-Uhlenbeck process and proof of the theorems	21
2 The stochastic localization process	25
2.1 The localization method and the KLS bound	25
2.2 From localization to stochastic localization	26
2.3 The process and its basic properties	28
2.4 A functional-analytic description of the process	30
2.5 Approaching the KLS Conjecture	32
2.6 The logarithmic obstruction	39
2.7 From stochastic localization to random localization	43
3 On measures strongly log-concave on a subspace	47
3.1 Introduction	47
3.2 Restricted stochastic localization	52
3.3 Control of the covariance matrix	54
3.4 Proof of the main theorem	60
4 Log-Sobolev inequalities for subgaussian log-concave probabilities	61
4.1 Introduction and results	61
4.2 Background and Proof of Theorem 4.4	65
4.2.1 Sub-gaussian random variables	65
4.2.2 Sub-exponential random variables	67
4.2.3 Log-concave vectors	68
4.2.4 A short proof of Theorem 4.4	70

4.3	Proof of Theorem 4.5	71
4.3.1	The one dimensional case	71
4.3.2	A localization argument	73
4.4	An approach via stochastic localization	75
4.5	Some subclasses of subgaussian log-concave probabilities	79
4.5.1	Rotationally invariant measures	79
4.5.2	Tilt-stable measures	80
4.5.3	Proof of Theorem 4.10	84
5	Some additional remarks about tilt-stability and stochastic localization	87
5.1	More on tilt-stable measures	87
5.2	Relationship with the multiscale Bakry-Émery criterion	90
6	Positive solutions for large random linear systems	97
6.1	Introduction	97
6.2	Positive solutions: proof of Theorem 6.1	100
6.3	Stability: proof of Corollary 6.2	107
6.4	Heuristics at critical scaling, non-homogeneous systems and non-gaussian entries	108
	Perspectives	113

Introduction

This thesis is concerned with geometric and probabilistic phenomena arising in high-dimensional spaces. Typically, the study of high-dimensional objects become exponentially difficult as the dimension grows. For instance, in combinatorics, there are 2^n different binary labelings of a set with n elements. In another direction, if one wants to discretize the continuous hypercube $[0; 1]^n$ with a precision $\epsilon > 0$, it requires $(\frac{1}{\epsilon})^n$ points. This is known as the "curse of dimensionality" [15]

However, under additional assumptions, the curse can sometimes turn into a blessing [32]. A typical example is the well known Central Limit Theorem in probability. If $X_1 \dots X_n$ is a sequence of independent and identically distributed random variables, then their average $\frac{1}{n} \sum_{i=1}^n X_i$ converges to a normal distribution. From a geometric point of view, if we take X_1 to be uniform on the interval, then Central Limit Theorem is a statement about the behavior of the diagonal marginal of the distribution of mass of the continuous hypercube.

It turns out that the statistical context of iid random variables can be relaxed to a more geometric criterion of convexity. To be precise, if K is a convex body of the Euclidean space \mathbb{R}^n , and if μ is the uniform distribution on K , then for sufficiently large n , most of the marginals of μ are essentially normal. This is the content of the Central Limit Theorem for Convex bodies. The question has a long history, going back to Sudakov [80] and Diaconis and Freedman [31]. In [4] (see also [25]) it was reduced to estimating the variance of the Euclidean norm. To be a little bit more precise, we say that a probability μ is *isotropic* if it is centered, and satisfies

$$\int (x \cdot \theta)^2 d\mu = 1 \quad \text{for all } \theta \in S^{n-1}$$

Any non-degenerate probability can be made isotropic after a suitable affine transformation, so that this hypothesis is only a normalization. Anttila, Ball and Perissinaki [4] showed that the Central Limit Theorem reduced to the estimation of the so called *thin-shell* parameter :

$$\sigma_n^2 = \sup_{\mu} \sigma_{\mu}^2 = \sup_{\mu} \text{Var}|X|_2 \tag{0.0.1}$$

where X is distributed according to μ , and the supremum runs over all isotropic uniform probabilities on a convex body of \mathbb{R}^n . Since μ is isotropic, $\mathbb{E}|X|_2^2 = n$, so that a trivial bound is $\sigma_n = O(\sqrt{n})$. The first non-trivial bound was obtained by Klartag in [54]. The result actually holds not only in the class of uniform probabilities over convex bodies, but in the more general setting of log-concave probabilities, that is probabilities having convex support, and density of the form e^{-V} with V convex. It is not difficult to see that for an isotropic log-concave probability μ ,

$$\sigma_\mu^2 \simeq \frac{1}{n} \text{Var}_\mu |X|_2^2.$$

Thus, in the case where μ the standard Gaussian, or the uniform distribution on the hypercube, which are two instances of log-concave probabilities, by independence $\sigma_\mu \simeq 1$. The questions whether this holds for any isotropic log-concave probability is the content of the Variance, or thin-shell Conjecture.

Conjecture 0.1. *There exists a universal constant $C > 0$ such that*

$$\sigma_n \leq C.$$

An even stronger conjecture is the KLS Conjecture, formulated by Kannan Lovász and Simonovits [50]. We say that a probability μ satisfies a Poincaré inequality with constant $K > 0$ if for any locally Lipschitz function $f : \mathbb{R}^n \mapsto \mathbb{R}$, one has,

$$\text{Var}_\mu(f) \leq K^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\mu, \quad (0.0.2)$$

We denote by $c_P(\mu)$ the best constant such that (0.0.2) holds. Define

$$\Psi_n^2 = \sup_{\mu} c_P^2(\mu).$$

Notice that by plugging $f = |\cdot|_2^2$ in (0.0.2), one gets $\sigma_\mu \lesssim c_P(\mu)$, thus the following is a strengthening of Conjecture 0.1.

Conjecture 0.2. *There exists a universal constant $C > 0$ such that*

$$\Psi_n \leq C.$$

Plugging linear forms in (0.0.2) shows that $\Psi_n \geq 1$. For the standard Gaussian, it is well known that $c_P(\gamma) = 1$. In other words, the extremizers of the functional inequality (0.0.2) are simple one dimensional objects. The KLS conjecture asks whether this behavior generalizes, up to constant, to all log-concave probabilities. This is one instance of the (conjectural) simplicity of high-dimensional objects under a convexity assumption. The conjecture has a wide range of consequences both in computer science and convex geometry,

see [66] for a survey. One striking application is that it implies the famous slicing or hyperplane conjecture in convex geometry :

Conjecture 0.3. *There exists $C_1 > 0$ such that for any n , and any convex body K of \mathbb{R}^n of volume 1, there exists a hyperplane H such that*

$$\text{Vol}(K \cap H) \geq C_1.$$

Equivalently, there exists a constant C_2 such that for any isotropic log concave probability μ of \mathbb{R}^n with density f ,

$$f(0)^{1/n} \leq C_2$$

We define, for $n \geq 1$,

$$L_n = \sup_f f(0)^{1/n} \tag{0.0.3}$$

where the supremum runs over all densities f such that $d\mu = f(x)dx$ is isotropic and log-concave. The first connection between Ψ_n and L_n was announced by Ball, and proved in [11] where the authors show that $L_n \lesssim e^{c\Psi_n}$ for an absolute constant $c > 0$. The inequality was later improved drastically to $L_n \lesssim \Psi_n$ by Eldan and Klartag [36]. For a long time, the best estimate for the KLS constant was $\Psi_n \lesssim \sqrt{n}$, dating back to the original work of Kannan Lovász and Simonovits [50]. In 2013, Eldan introduced a new idea to tackle the conjecture, namely the stochastic localization process, which will be at the center of this manuscript. It allowed him to relate the KLS constant to the a priori weaker thin-shell constant, namely

$$\Psi_n \lesssim \log(n)\sigma_n,$$

obtaining the then best-bound $\Psi_n \lesssim n^{1/3} \log(n)$ by using the Guédon-Milman estimate for σ_n [47]. Later, Lee and Vampala obtained the bound $\Psi_n \leq n^{1/4}$ using a slightly simplified version of the process. A breakthrough was obtained by Chen, proving a subpolynomial estimate for Ψ_n [28], before a polylog estimate by Klartag and Lehec [62], and very recently the bound $\Psi_n \lesssim \sqrt{\log n}$ by Klartag [56].

Besides the Poincaré inequality (0.0.2), another functional inequality which shall be of interest to us is the logarithmic-Sobolev inequality. A Borel probability μ on \mathbb{R}^n is said to satisfy a logarithmic Sobolev inequality with constant $\rho > 0$ if for any locally Lipschitz function $f : \mathbb{R}^n \mapsto \mathbb{R}$, one has,

$$\text{Ent}_\mu(f^2) \leq 2\rho^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\mu, \tag{0.0.4}$$

where for a non-negative function g , $\text{Ent}_\mu(g) = \mathbb{E}_\mu(g \log g) - \mathbb{E}_\mu(g) \log \mathbb{E}_\mu(g)$. We denote by $\rho_{LS}(\mu)$ the optimal constant ρ such that (0.0.4) holds. The log-Sobolev inequality (0.0.4) is actually stronger than the Poincaré inequality (0.0.2). This should make sense intuitively, since the Poincaré inequality controls the L^2 norm of centered function by the L^2 norm of

its gradient, while the log-Sobolev controls the stronger $L^2 \log(L^2)$ norm by the very same energy term. Concretely, by looking at small perturbations of inequality (0.0.4), one easily concludes that

$$c_P(\mu) \leq \rho_{LS}(\mu).$$

Much like the Poincaré inequality, the log-Sobolev inequality has strong consequences on the behavior of the measure μ , in particular it implies stronger concentration and isometric properties. Yet again, the standard Gaussian satisfies a log-Sobolev inequality with constant $\rho_{LS}(\gamma) = 1$, and the extremizers are one dimensional. The classical Herbst's argument shows that the log-Sobolev inequality implies that the tails of the μ decay faster than a Gaussian, thus the exponential distribution, for instance, does not satisfy inequality (0.0.4). In the spirit of the KLS conjecture, it is tempting to ask whether in the log-concave setting, the obstruction to satisfying inequality (0.0.4) only comes from the behavior of linear forms.

Conjecture 0.4. *There exists a constant $C > 0$ such that for any $n \in \mathbb{N}$ and any centered log-concave probability μ on \mathbb{R}^n ,*

$$\rho_{LS}(\mu) \leq C \sup_{\theta \in \mathbb{S}^{n-1}} |\langle \cdot, \theta \rangle|_{\Psi_2(\mu)}$$

where for a function $g : \mathbb{R}^n \mapsto \mathbb{R}$, $|g|_{\Psi_2(\mu)} = \inf \{t > 0 / \mathbb{E}_\mu [\exp(g^2/t^2)] \leq 2\}$.

We introduce the quantity

$$G_n = \sup_{\mu} \rho_{LS}(\mu) \tag{0.0.5}$$

where the supremum runs over all centered log-concave probabilities μ on \mathbb{R}^n such that

$$\sup_{\theta \in \mathbb{S}^{n-1}} |\langle \cdot, \theta \rangle|_{\Psi_2(\mu)} \leq 1.$$

By scaling, Conjecture 0.4 is equivalent to the boundedness of G_n . Little is known about Conjecture 0.4. Bobkov showed [19] that $G_n \lesssim \sqrt{n}$.

In Chapter 1 we revisit the article of Ball and Nguyen [11], about the so called entropy jump for log-concave vector. For a random vector X in \mathbb{R}^d , assuming integrability, we define

$$\text{Ent}(X) = \mathbb{E}(-\log(X)).$$

It is well known that when the covariance is fixed, the Gaussian vector with said covariance maximizes the entropy. For normalization, suppose that the covariance of X is I_d , we know that the sequence

$$S_n = \frac{1}{\sqrt{n}}(X_1 + \cdots + X_n)$$

where the X_i are iid copies of X , converge to a standard Gaussian. Thus a natural

question is whether the entropy is monotone along this sequence, that is, whether $\text{Ent}(S_n)$ is increasing. This was an important conjecture that has been solved by Arstein-Avidan, Ball, Barthe and Naor in 2004 [5]. The case $n = 2$ is known as the Shannon-Stam inequality:

$$\text{Ent}\left(\frac{X_1 + X_2}{\sqrt{2}}\right) \geq \text{Ent}(X).$$

When X is log-concave and has covariance I_d , Ball and Nguyen lower-bounded its entropy jump by

$$\text{Ent}\left(\frac{X_1 + X_2}{\sqrt{2}}\right) \geq \frac{c_P^2(X)}{4(1 + c_P^2(X))} (\text{Ent}(G) - \text{Ent}(X))$$

where G is a standard Gaussian. Using the same approach, we extend this result to the case where X_1 and X_2 are not identically distributed, and prove a similar result for the Fisher information.

Chapter 2 is an introduction to the stochastic localization process, which we use in Chapter 3, 4 and 5. We start by a brief exposition of the classical localization in convex geometry which is at the root of stochastic localization. We then proceed to defining the process, analyzing its elementary properties and review its use in the context of the KLS conjecture with the numerous breakthroughs that took place since the original article of Eldan. Finally, the last two subsections are devoted to folklore knowledge about the process, that is, the logarithmic obstruction, which was pointed out by Klartag, and how the process can be used to obtain a "random localization".

Chapter 3 is devoted to the study of log-concave probabilities that are strongly log-concave on a subspace. Write $d\mu = e^{-V(x)} dx$. It is well-known that if the potential V is strongly convex, namely

$$\nabla^2 V \geq \rho Id,$$

then

$$c_P(\mu) \leq \frac{1}{\rho}.$$

We investigate the case where the strong convexity assumption is only available on a subspace, that is

$$P_E \nabla^2 V P_E \geq \rho$$

where P_E is the orthogonal projection onto a given subspace E of dimension $n - k$. Using the stochastic localization process, we provide an estimate for the Poincaré constant of μ as well as for its concentration function.

In Chapter 4 we investigate Conjecture 0.4. We leverage our estimate on the concentration function proved in Chapter 3 in the case $E = 0$, which combined with a net argument allow us to improve on Bobkov's estimate. We show that

$$G_n \lesssim n^{1/4}.$$

We also study a subclass of subgaussian log-concave probabilities which satisfy

$$\text{Cov} \left(\frac{1}{Z_h} e^{h \cdot x} \mu \right) \leq \beta^2 I_n \quad \text{for all } h \in \mathbb{R}^n.$$

Equivalently, we require that the Hessian of the log-Laplace transform \mathcal{L}_μ of μ is uniformly bounded, which is stronger than requiring a quadratic bound on \mathcal{L}_μ . From there we provide an upper-bound for the quantity

$$\sup_{h \in \mathbb{R}^n, t \geq 0} \left| \text{Cov} \left(\frac{1}{Z_{t,h}} e^{-t|x|^2 + h \cdot x} \right) \right|_{op},$$

which can be shown to control the log-Sobolev constant of μ , by stochastic localization.

Chapter 5 we provide some additionnal remarks about tilt-stable measures and draw a connection between the stochastic localization process and the multiscale Bakry-Emery criterion of Bodineau and Bauerschmidt [14].

Finally, Chapter 6 is independent of the previous and makes no reference to the stochastic localization process. We still study an high-dimensionnal object, in this case a large random linear system :

$$\frac{1}{\alpha_n \sqrt{n}} A_n X_n + \mathbf{1}_n = X_n, \tag{0.0.6}$$

where A_n is a random matrix of size $n \times n$ with independent Gaussian entries, $\mathbf{1}_n$ is the vector with 1 entries and α_n is a positive sequence. Such equations arise for instance from the study of foodwebs, where X_n represents the population of the species involved and $\frac{1}{\alpha_n \sqrt{n}} A_n$ models their interaction. We are interested in the feasibility of the system, that is the quantity

$$p_n = \mathbb{P}(X_n \geq 0).$$

We show that p_n exhibits a cut-off phenomena. That is, there is a sharp transition from 0 to 1 at an explicit critical scaling α_n^* . We also discuss extensions to the non-Gaussian case.

0.1 Notations and mathematical preliminaries

For two positive functions $f, g : X \mapsto \mathbb{R}^+$, depending on parameters $x \in X$ we write $f \lesssim g$ if there exists a constant $C > 0$ such that $f \leq Cg$. We write $f \simeq g$ if $f \lesssim g$ and $g \lesssim f$. Depending on the context, f and g might be functionals over log-concave probabilities, or positive sequences for instance. Unless specified otherwise, $|\cdot|$ denotes the Euclidean norm, while $|\cdot|_{op}$ denotes the operator norm. That is, for a matrix $A \in M_n(\mathbb{R})$,

$$|A|_{op} = \sup_{\theta \in S^{n-1}} |A\theta|.$$

For a random variable taking values in a space E , $X : \Omega \mapsto E$ and a probability μ on E , we write $X \sim \mu$ if μ is the law of X .

Throughout the manuscript, with the notable exception of Chapter 1, we say that a probability μ satisfies a Poincaré inequality with constant $K > 0$ if for any locally Lipschitz function $f : \mathbb{R}^n \mapsto \mathbb{R}$, one has,

$$\mathrm{Var}_\mu(f) \leq K^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\mu, \quad (0.1.1)$$

We denote by $c_P(\mu)$ the best constant such that (0.1.1) holds. If $X \sim \mu$ we again write $c_P(X) = c_P(\mu)$. Notice that with this convention the Poincaré constant is 1-homogeneous:

$$\forall \lambda \in \mathbb{R} \quad c_P(\lambda X) = |\lambda| c_P(X).$$

In Chapter 1 however, we exceptionally write

$$c_X = c_P^2(X),$$

where c_X is again referred to as the Poincaré constant of X and is 2-homogeneous.

Similarly, throughout the manuscript, a Borel probability μ on \mathbb{R}^n is said to satisfy a logarithmic Sobolev inequality with constant $\rho > 0$ if for any locally Lipschitz function $f : \mathbb{R}^n \mapsto \mathbb{R}$, one has,

$$\mathrm{Ent}_\mu(f^2) \leq 2\rho^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\mu, \quad (0.1.2)$$

where for a non-negative function g , $\mathrm{Ent}_\mu(g) = \mathbb{E}_\mu(g \log g) - \mathbb{E}_\mu(g) \log \mathbb{E}_\mu(g)$. We denote by $\rho_{LS}(\mu)$ the optimal constant ρ such that (0.1.2) holds.

0.1.1 Log-concavity

Here we recall a sparse number of well-known facts about log-concave probabilities that we will use frequently in the manuscript. A good reference is [2] and references therein.

We say that a Borel probability μ is log-concave, if for all compact sets A and B and $\lambda \in [0, 1]$,

$$\mu((1 - \lambda)A + \lambda B) \geq \mu(A)^{1-\lambda} \mu(B)^\lambda.$$

Equivalently, $\mu = e^{-V(x)} dx$ where $V : \mathbb{R}^n \mapsto \mathbb{R} \cup \{\infty\}$ is convex. We say that a random variable X is log concave if its law is.

The celebrated Prékopa-Leindler inequality ensures that marginalization preserves log-concavity

Lemma 0.5. *Every marginal of a log-concave measure is log-concave.*

Consequently, since the product of two log-concave measures is again log-concave, convolutions also preserve log-concavity.

Lemma 0.6. *If X and Y are log-concave, so is $X + Y$.*

A fundamental property of log-concave probabilities is that we can compare their moments. This is known as Borell's Lemma:

Lemma 0.7. *Let X be a log-concave real random variable. There exists a constant C , independent of X , such that for all $p \geq 1$,*

$$\mathbb{E}(|X|^p)^{1/p} \leq C\mathbb{E}|X|$$

Poincaré and Cheeger constant of log-concave measures

We say that a probability μ on \mathbb{R}^n satisfies a Cheeger inequality with constant h if

$$\mu^+(A) \geq \frac{1}{h} \min(\mu(A), 1 - \mu(A)) \quad (0.1.3)$$

for all measurable set A , where $\mu^+(A) = \liminf_{\epsilon \rightarrow 0} \frac{\mu(A_\epsilon) - \mu(A)}{\epsilon}$ and $A_\epsilon = \{x \in \mathbb{R}^n, d(x, A) \leq \epsilon\}$ is the euclidean ϵ -extension of A . We denote by h_μ the best such constant.

In general, Cheeger[27] and Maz'ya[71] proved that a Cheeger inequality implies a Poincaré inequality.

Lemma 0.8. *Let μ be a probability on \mathbb{R}^n , then*

$$c_P(\mu) \lesssim h_\mu.$$

Buser[26] and Ledoux[65] showed that this hierarchy could be reversed when μ is log-concave.

Lemma 0.9. *Let μ be a log-concave probability, then*

$$c_P(\mu) \simeq h_\mu$$

Finally, we conclude this introduction with a result of Emmanuel Milman [72], which shows that it is enough to consider set of half measure in 0.1.3.

Theorem 0.10. *Let μ be a log-concave probability, then*

$$\frac{1}{h_\mu} = \inf_{A, \mu(A)=1/2} \frac{\mu^+(A)}{\min(\mu(A), 1 - \mu(A))}$$

Chapter 1

Entropy and Information jump for log-concave vectors

This chapter is a reproduction of

Pierre Bizeul. Entropy and information jump for log-concave vectors. *Comptes Rendus. Mathématique*, 361(G2):487–493, 2023

1.1 Introduction

Let X be a random vector distributed according to a measure μ in \mathbb{R}^d , with density f with respect to the Lebesgue measure. We will denote this by $X \sim \mu = f d\lambda$. If $\int f |\log f| < +\infty$, we define its entropy by

$$\text{Ent}_L(\mu) = \text{Ent}_L(X) = - \int_{\mathbb{R}^d} f \log f,$$

where the subscript L stands for "Lebesgue".

It should be noted that this entropy can be either positive or negative, and that for any invertible matrix A , $\text{Ent}_L(AX) = \text{Ent}_L(X) + \log(|\det A|)$. Entropy is also translation invariant, and it is classical that, when the covariance matrix is fixed, the Gaussian distribution maximizes entropy. It will be useful to normalize vectors so that they are centered, and have covariance matrix identity. Such a vector, as well as its distribution, is called *isotropic*.

The classical Shannon-Stam inequality asserts that taking a convolution increases entropy: for two iid random vectors X_1 and X_2

$$\text{Ent}_L(X_1) \leq \text{Ent}_L\left(\frac{X_1 + X_2}{\sqrt{2}}\right).$$

Moreover, there is equality if and only if X has a Gaussian distribution. Now, one can

wonder if that equality case is stable, meaning if the *entropy jump* $\text{Ent}_L(\frac{X_1+X_2}{\sqrt{2}}) - \text{Ent}_L(X)$ is small, does it imply that X is almost Gaussian ? The general answer is no, as one can convince himself by considering a well chosen double-bump Gaussian [30].

However, when the distribution of X admits a spectral gap, excluding double-bumped type distributions, some positive answers exist. Recall that X is said to have a spectral gap, or equivalently satisfy a Poincaré inequality, if there exists a constant $c > 0$ such that for any smooth enough function f , the variance of $f(X)$ can be controlled in terms of the euclidean norm of $\nabla f(x)$ as follows:

$$\text{var} f(X) \leq c \mathbb{E} [|\nabla f(X)|^2].$$

The smallest such constant c will be denoted c_X and called the Poincaré constant of X . Under a spectral gap assumption, it was proven by Ball, Barthe and Naor in [10] that for a one dimensional isotropic random variable X ,

$$\text{Ent}_L\left(\frac{X_1 + X_2}{\sqrt{2}}\right) - \text{Ent}_L(X_1) \geq \frac{1}{2(1 + c_X)}(\text{Ent}_L(G) - \text{Ent}_L(X_1)),$$

where G is a standard Gaussian.

We can rewrite the right-hand side as a Kullback-Leibler divergence. Recall that, if X is isotropic,

$$\text{Ent}_L(G) - \text{Ent}_L(X) = D(X||G) = \int_{\mathbb{R}^d} f_\gamma \log(f_\gamma) d\gamma \geq 0,$$

where f_γ is the relative density of X with respect to the Gaussian measure γ , that is $X \sim f_\gamma d\gamma$ and $G \sim \gamma$. In the sequel we shall use the notation $D(X) = D(X||G)$. This is a strong measure of closeness to the Gaussian; for instance the Pinsker-Csiszar-Kullback inequality states that

$$\left(\int_{\mathbb{R}^d} |f - g| \right)^2 \leq \frac{1}{2} D(X||G),$$

where f and g are the density of X and G , respectively. In 2012, Ball and Nguyen generalized the result to arbitrary dimension, assuming log-concavity of X ([11]). They use a semigroup approach, differentiating twice the entropy along the Ornstein-Uhlenbeck semigroup.

The Fisher information of a random vector with smooth density f is

$$I_L(X) = \int \frac{|\nabla f|^2}{f} = 4 \int |\nabla(\sqrt{f})|^2,$$

whenever those integrals are well defined. As for the entropy, for a fixed covariance matrix, the Gaussian is extremal; in this case, it has the smallest information. Information is classically the derivative of the entropy along the semi-group. In the spirit of the Shannon-Stam Inequality, the Blachman-Stamm inequality asserts that taking a convolution decreases the

information:

$$I_L(X_1) \geq I_L\left(\frac{X_1 + X_2}{\sqrt{2}}\right).$$

As before, we can define a relative information, notably to the Gaussian measure $d\gamma$. If $X \sim f_\gamma d\gamma$ is a random vector with smooth density, we will denote its relative information to $d\gamma$ by

$$I(X||G) = I_\gamma(X) = \int \frac{|\nabla f_\gamma|^2}{f_\gamma} d\gamma.$$

When X is isotropic, integrating by parts yields:

$$I(X||G) = I_L(X) - d = I_L(X) - I_L(G).$$

Consequently, for a measure μ on \mathbb{R}^d , we write $D(\mu||\gamma) = D(X||G)$, $I_L(\mu) = I_L(X)$ and $I(\mu||\gamma) = I(X||G)$ where $X \sim \mu$ is a random vector distributed according to μ .

In this note, we use the same strategy as in [11], but improve their result in two directions. First we generalize it to non identically distributed pairs of random vectors. For two measures μ and ν define

$$\delta_{E,\lambda}(\mu, \nu) = \text{Ent}_L(X_\lambda) - (1 - \lambda)\text{Ent}_L(X_0) - \lambda\text{Ent}_L(X_1),$$

where X_0 and X_1 are independent random vectors distributed according to μ and ν respectively, and $X_\lambda = \sqrt{1 - \lambda}X_0 + \sqrt{\lambda}X_1$. The Shannon-Stam inequality asserts that $\delta_\lambda(\mu, \nu) \geq 0$ and this quantity is precisely the deficit in the Shannon-Stam inequality.

Theorem 1.1 (Quantitative Shannon-Stam). *Let μ, ν be two log-concave isotropic measures with Poincaré constant respectively c_0 and c_1 , and $\lambda \in [0, 1]$. Then,*

$$\delta_{E,\lambda}(\mu, \nu) \geq \frac{\lambda(1 - \lambda)}{4 \max(c_0, c_1)} (D(\mu||\gamma) + D(\nu||\gamma)).$$

This should be compared with a recent result of Eldan and Mikulincer ([38], Theorem 3). They get a more general result, allowing μ and ν to have different covariance matrices, but in the case where μ and ν have the same covariance matrix, they get a worst dependence on the Poincaré constant.

Secondly, we get a same kind of inequality for the information, yielding a stability result for the Blachman-Stam inequality. Define this time the information deficit of a pair of measures by

$$\delta_{I,\lambda}(\mu, \nu) = (1 - \lambda)I_L(X_0) + \lambda I_L(X_1) - I_L(X_\lambda)$$

where X_0 and X_1 are independent random vectors distributed according to μ and ν re-

spectively, and $X_\lambda = \sqrt{1-\lambda}X_0 + \sqrt{\lambda}X_1$

Theorem 1.2 (Quantitative Blachman-Stam). *Let μ, ν be two log-concave isotropic measures with Poincaré constant respectively c_0 and c_1 , and $\lambda \in [0, 1]$. Then,*

$$\delta_{I,\lambda}(\mu, \nu) \geq \frac{\lambda(1-\lambda)}{4 \max(c_0, c_1)} (I(\mu||d\gamma) + I(\nu||d\gamma)).$$

In the sequel, quantities computed with respect to the Lebesgue measure have a subscript “ L ” while quantities that are computed with respect to the Gaussian measure have none.

1.2 A lemma of Ball-Nguyen

Let X be a random vector with smooth density $f = e^{-\Psi}$ with respect to the **Lebesgue** measure. We define $\sigma_L(X)$ to be the random matrix $\sigma_L(X) := \nabla^2(\Psi)(X)$ and, we denote

$$K_L(X) = \mathbb{E} \left[\|\sigma_L(X)\|^2 \right],$$

where $\|\cdot\|$ denotes the Hilbert-Schmidt norm on matrices and the subscripts L yet again stands for "Lebesgue", which we take temporarily as the reference measure. Understanding this quantitiy will prove to be important later on, as it will appear in the second derivative of the entropy along the Ornstein-Uhlenbeck semigroup.

We recall a lemma of Ball-Nguyen [11], for which we provide a simple proof.

Lemma 1.3. (*Ball-Nguyen*) *Let X be a random vector in \mathbb{R}^d with smooth density, $E \subset \mathbb{R}^d$ be any subspace, p_E be the orthogonal projection onto E and $X_E = p_E(X)$. Then*

$$\sigma_L(X_E) \leq p_E \mathbb{E} [\sigma_L(X)|X_E] p_E^*, \quad a.s.,$$

for the partial order on symmetric matrices.

Proof. Let $\Psi_E : E \rightarrow \mathbb{R}$ be such that $X_E \sim e^{-\Psi_E(x)} dx$. We have for $x \in E$

$$\Psi_E(x) = -\ln \int_{E^\perp} e^{-\Psi(x,y)} dy.$$

Then, setting $d\nu_x = \frac{e^{-\Psi(x,y)}}{\int_{E^\perp} e^{-\Psi(x,y)} dy} dy$, a straightforward computation shows that:

$$\forall x \in E, \quad \nabla^2 \Psi_E(x) = \int_{E^\perp} \nabla_{xx}^2 \Psi d\nu_x - \text{Cov}_{d\nu_x}(\nabla_x \Psi).$$

In particular,

$$\forall x \in E, \quad \nabla^2 \Psi_E(x) \leq \int_{E^\perp} \nabla_{xx}^2 \Psi d\nu_x,$$

which is the desired result □

If X_0 and X_1 are two independent random vectors in \mathbb{R}^d and $\lambda \in [0, 1]$, applying the previous lemma to the random vector (X_0, X_1) and the projection $p(x, y) = \sqrt{1 - \lambda}x + \sqrt{\lambda}y$ yields:

Lemma 1.4. *For any independent random vectors X_0, X_1 in \mathbb{R}^d , with smooth densities, and any $\lambda \in [0, 1]$*

$$\sigma_L(X_\lambda) \leq \mathbb{E}[(1 - \lambda)\sigma_L(X_0) + \lambda\sigma_L(X_1)|X_\lambda] \quad a.s.,$$

where $X_\lambda = \sqrt{1 - \lambda}X_0 + \sqrt{\lambda}X_1$.

If X_0 and X_1 are log-concave and independent, then so is X_λ , by Prékopa's theorem. Thus $\sigma_L(X_0), \sigma_L(X_1)$ and $\sigma(X_\lambda)$ are positive matrices, so as in Ball-Nguyen's article, the inequality above translates to an inequality on their norm.

Lemma 1.5. *For any log-concave independent random vectors X_0, X_1 in \mathbb{R}^d , with smooth densities, and any $\lambda \in [0, 1]$*

$$(1 - \lambda)K_L(X_0) + \lambda K_L(X_1) - K_L(X_\lambda) \geq \lambda(1 - \lambda)\mathbb{E}[\|\sigma_L(X_1) - \sigma_L(X_0)\|^2].$$

Remark : In particular, we have :

$$K_L(X_\lambda) \leq (1 - \lambda)K_L(X_0) + \lambda K_L(X_1),$$

which can be seen as a second-order Blachman-Stam inequality.

Proof. As explained, the matrices being positive, the inequality in Lemma 1.4 implies that:

$$\|\sigma_L(X_\lambda)\| \leq \|\mathbb{E}[(1 - \lambda)\sigma_L(X_0) + \lambda\sigma_L(X_1)|X_\lambda]\|.$$

Taking the expectation of the square and using Jensen's inequality then implies:

$$\begin{aligned} K_L(X_\lambda) &\leq \mathbb{E}[\|(1 - \lambda)\sigma_L(X_0) + \lambda\sigma_L(X_1)\|^2] \\ &= (1 - \lambda)^2 K_L(X_0) + \lambda^2 K_L(X_1) + 2\lambda(1 - \lambda)\mathbb{E}[\langle \sigma_L(X_0), \sigma_L(X_1) \rangle] \\ &= (1 - \lambda)K_L(X_0) + \lambda K_L(X_1) - \lambda(1 - \lambda)\mathbb{E}[\|\sigma_L(X_0) - \sigma_L(X_1)\|^2] \end{aligned}$$

since X_0 and X_1 are independent. □

Now we want to translate this result to the Gaussian setting. Assuming that X has density $f_\gamma = e^{-\varphi}$ with respect to the Gaussian measure, we similarly introduce:

$$\begin{aligned} \sigma(X) &:= \sigma_\gamma(X) := \nabla^2(\varphi)(X) \\ K(X) &:= K_\gamma(X) = \mathbb{E}[\|\sigma(X)\|^2] \end{aligned}$$

$$I(X) = I_\gamma(X) = \mathbb{E} \left[\|\nabla \varphi(X)\|^2 \right].$$

Note that the definition of I is consistent with the one given in the introduction. Then Lemma 1.5 becomes:

Lemma 1.6. *For any log-concave independent isotropic random vectors X_0, X_1 in \mathbb{R}^d , with smooth densities, and any $\lambda \in [0, 1]$, we have*

$$(1 - \lambda)M(X_0) + \lambda M(X_1) - M(X_\lambda) \geq \lambda(1 - \lambda)\mathbb{E} \left[\|\sigma(X_1) - \sigma(X_0)\|^2 \right], \quad (1.2.1)$$

where $M(X) = K(X) + 2I(X)$.

Proof. Let X be an isotropic log-concave random vector with density $e^{-\varphi}$ with respect to the Gaussian measure, and $e^{-\Psi}$ with respect to the Lebesgue measure. By definition we have:

$$\sigma(X) = \sigma_L(X) - Id.$$

Hence,

$$K(X) = K_L(X) - 2\mathbb{E} [\text{Tr}(\sigma_L(X))] + n.$$

Now, by integration by parts:

$$\begin{aligned} \mathbb{E} [\text{Tr}(\sigma_L(X))] &= \int_{\mathbb{R}^n} \text{div}(\nabla \Psi)(x) e^{-\Psi(x)} dx \\ &= \int_{\mathbb{R}^n} \nabla \Psi(x) \cdot \nabla \Psi(x) e^{-\Psi(x)} dx \\ &= \mathbb{E} \left[\|\nabla \Psi(X)\|^2 \right] = \mathbb{E} \left[\|\nabla \varphi(X) + X\|^2 \right] = I(X) + n. \end{aligned}$$

The lemma follows. □

The next lemma provides a lower bound for the right-hand side in the inequality (1.2.1).

Lemma 1.7. *For any log-concave independent isotropic random vectors X_0, X_1 in \mathbb{R}^d with smooth densities we have*

$$(1 - \lambda)M(X_0) + \lambda M(X_1) - M(X_\lambda) \geq \frac{\lambda(1 - \lambda)}{2 \max(c_0, c_1)} (I(X_0) + K(X_0) + I(X_1) + K(X_1)),$$

where c_0, c_1 are the Poincaré constants of X_0 and X_1 , respectively.

Proof. We denote by c_0 the Poincaré constant of X_0 . We condition on X_1 and apply the Poincaré inequality for X_0 to the function $\nabla \varphi_0(X_0) - \nabla^2 \varphi_1(X_1)X_0$ which is centered and

we use the fact that X_0 and X_1 are isotropic:

$$\begin{aligned} \mathbb{E} \left[\left\| \nabla^2 \varphi_0(X_0) - \nabla^2 \varphi_1(X_1) \right\|^2 \right] &\geq \frac{1}{c_0} \mathbb{E} \left[\left\| \nabla \varphi_0(X_0) - \nabla^2 \varphi_1(X_1) X_0 \right\|^2 \right] \\ &= \frac{1}{c_0} \mathbb{E} \left[\left\| \nabla \varphi_0(X_0) \right\|^2 + \left\| \nabla^2 \varphi_1(X_1) \right\|^2 \right] \\ &= \frac{1}{c_0} (I(X_0) + K(X_1)). \end{aligned}$$

By symmetry we get:

$$\mathbb{E} \left[\left\| \sigma_\gamma(X_1) - \sigma_\gamma(X_0) \right\|^2 \right] \geq \frac{1}{2 \max(c_0, c_1)} (I(X_0) + K(X_0) + I(X_1) + K(X_1)).$$

Plugging this into Lemma 1.6 concludes the proof. \square

1.3 The Ornstein-Uhlenbeck process and proof of the theorems

Let X be a random vector in \mathbb{R}^d with density f with respect to the **Gaussian** measure γ . Let L_γ be the diffusion operator defined by $L_\gamma f(x) = \Delta f(x) - \nabla f(x) \cdot x$. The differential equation associated to L_γ is the modified heat equation:

$$\frac{\partial f_t}{\partial t} = L_\gamma f_t ; \quad f_0 = f.$$

Its solution f_t is the relative density of the random vector $X_t = e^{-t}X + \sqrt{1 - e^{-2t}}G$, where G is a standard Gaussian independent of X , with respect to the Gaussian measure. From this description of X_t it is clear that X_t has a \mathcal{C}^∞ density, with integrability properties as good as f , and that the process commutes with convolutions, in the sense that, with the notations of the previous sections, for all independent X_0, X_1 random vectors, and any $\lambda \in [0, 1]$, we have

$$(X_\lambda)_t = (X_t)_\lambda \quad \text{in law.} \quad (1.3.1)$$

It is also useful to note that if X is such that $c_X \geq 1$, in particular if X is isotropic, then for all $t \geq 0$:

$$c_{X_t} \leq c_X.$$

Indeed, if X and Y are independent random vectors satisfying a Poincaré inequality and $\lambda \in [0, 1]$, then using the conditional variance formula, a few computations show that

$$c_{\sqrt{\lambda}X + \sqrt{1-\lambda}Y} \leq \lambda c_X + (1 - \lambda)c_Y,$$

which in our case yields $c_{X_t} \leq e^{-2t}c_X + (1 - e^{-2t}) \times 1 \leq c_X$.

We denote by $(P_t)_{t \geq 0}$ the semi-group defined by $P_t(f) = f_t$. The following computa-

tions are standard (see [9]): if X has finite entropy, then for $t > 0$

$$\frac{\partial}{\partial t} \text{Ent}(X_t) = I(X_t) \quad , \quad \frac{\partial I(X_t)}{\partial t} = -2I(X_t) - 2K(X_t) = -M(X_t) - K(X_t).$$

As a consequence, a linear inequality on the information can be integrated along the semi-group to get the same inequality for the entropy. We also get that $\frac{\partial I(X_t)}{\partial t} \leq -2I(X_t)$ which implies that $I(X_t) \leq e^{-2t}I(X_0)$. Moreover, the control role of L comes from the observation that

$$e^{2t} \frac{\partial(e^{-2t}I(X_t))}{\partial t} = -2M(X_t). \quad (1.3.2)$$

Now we are in position to prove Theorem 1.2, from which Theorem 1.1 will be an immediate corollary.

Let X_0 and X_1 be two isotropic random log-concave vectors in \mathbb{R}^d and $\lambda \in [0, 1]$, and $X_\lambda = \sqrt{1-\lambda}X_0 + \sqrt{\lambda}X_1$. Denote by $(X_0)_t$, $(X_1)_t$ and $(X_\lambda)_t$ their evolution along the Ornstein-Uhlenbeck semi-group. Further define:

$$I_0(t) = I((X_0)_t) \quad , \quad I_1(t) = I((X_1)_t) \quad , \quad I_\lambda(t) = I((X_\lambda)_t),$$

and similarly, define $K_0(t), K_1(t), K_\lambda(t), M_0(t), M_1(t), M_\lambda(t)$; using (1.3.2), the commutation property (1.3.1), the observation that the Poincaré constants only decrease along the semi-group and Lemma 1.7, we get the following:

Lemma 1.8. *With the previous notations, for all $t \geq 0$,*

$$\begin{aligned} -\frac{\partial}{\partial t}((1-\lambda)e^{-2t}I_0(t) + \lambda e^{-2t}I_1(t) - e^{-2t}I_\lambda(t)) &\geq \frac{\lambda(1-\lambda)e^{-2t}}{\max(c_0(t), c_1(t))}(I_0(t) + K_0(t) + I_1(t) + K_1(t)) \\ &\geq -\frac{\lambda(1-\lambda)}{2\max(c_0, c_1)}e^{-2t} \frac{\partial}{\partial t}(I_0(t) + I_1(t)). \end{aligned}$$

Proofs of Theorem 1.1 and Theorem 1.2.

Integrating the inequality of Lemma 1.8 from 0 to ∞ , we get:

$$\begin{aligned} (1-\lambda)I(X_0) + \lambda I(X_1) - I(X_\lambda) &\geq -\frac{\lambda(1-\lambda)}{2\max(c_0, c_1)} \int_0^{+\infty} e^{-2t} \frac{d}{dt} (I_0(t) + I_1(t)) dt \\ &= \frac{\lambda(1-\lambda)}{2\max(c_0, c_1)} \left(I(X_0) + I(X_1) - 2 \int_0^\infty e^{-2t} (I_0(t) + I_1(t)) dt \right) \\ &\geq \frac{\lambda(1-\lambda)}{4\max(C_0, C_1)} (I(X_0) + I(X_1)), \end{aligned}$$

where in the last inequality, we used the fact that $I_1(t) \leq e^{-2t}I(X_1)$ and $I_0(t) \leq e^{-2t}I(X_0)$. This proves Theorem 1.2.

Integrating Theorem 1.2 along the semi-group yields Theorem 1.1. \square

Chapter 2

The stochastic localization process

2.1 The localization method and the KLS bound

In their original paper, Kannan Lovász and Simonovits obtained the bound

$$c_P^2(\mu) \lesssim \text{Tr}(\text{Cov}\mu) \quad (2.1.1)$$

using a localization lemma of Lovász and Simonovits which we now describe. We state a slightly different version which is due to Bobkov [20].

Lemma 2.1. *Let $\alpha, \beta > 0$ and $(f_i)_{i=1,\dots,4}$ be non-negative continuous functions on \mathbb{R}^n such that for any segment $I \subset \mathbb{R}^n$ and any affine function ℓ on I ,*

$$\left(\int_I f_1 e^\ell \right)^\alpha \left(\int_I f_2 e^\ell \right)^\beta \leq \left(\int_I f_3 e^\ell \right)^\alpha \left(\int_I f_4 e^\ell \right)^\beta. \quad (2.1.2)$$

Then,

$$\left(\int_{\mathbb{R}^n} f_1 \right)^\alpha \left(\int_{\mathbb{R}^n} f_2 \right)^\beta \leq \left(\int_{\mathbb{R}^n} f_3 \right)^\alpha \left(\int_{\mathbb{R}^n} f_4 \right)^\beta \quad (2.1.3)$$

Now, we briefly explain how to deduce the KLS bound let μ be a log-concave probability with density p . Let A, B, C be a partition of \mathbb{R}^n and denote by $\epsilon = d(A, B)$. We apply Lemma 2.1 with $\alpha = \beta = 1$ and

$$f_1 = p1_A, \quad f_2 = p1_B, \quad f_3 = p1_C, \quad f_4 = \frac{gp}{\epsilon}$$

where g is a non-negative continuous function. We denote by μ_ℓ the probability on I having a density proportional to pe^ℓ . Inequality (2.1.3) rewrites

$$\mu(A)\mu(B) \leq \frac{\mu(C)}{\epsilon} \int g \, d\mu, \quad (2.1.4)$$

which holds if (2.1.2), that is the same inequality for μ_ℓ , holds. Now, we choose $B = (A_\epsilon)^c$ and let $\epsilon \rightarrow 0$. Notice that (2.1.4) takes the form of the isoperimetric inequality

$$\mu(A)\mu(A^c) \leq \mu^+(A) \int g \, d\mu,$$

which holds if the same inequality holds for μ_ℓ for any interval I . In summary, we reduced the isoperimetric inequality for μ to isoperimetric inequalities for one-dimensional log-concave probabilities μ_ℓ .

Corollary 2.2. *Let g be a non-negative continuous function on \mathbb{R}^n such that any one-dimensional log-concave probability ν on \mathbb{R}^n one has*

$$\frac{1}{h_\nu} \leq \int g \, d\nu.$$

Then, for any log-concave probability μ on \mathbb{R}^n ,

$$\frac{1}{h_\mu} \leq 2 \int g \, d\mu.$$

Note that the factor 2 comes from the inequality $2\mu(A)\mu(B) \geq \max(\mu(A), \mu(B))$. The choice of $g(x) = C|x|$ for a constant C ensuring that the one-dimensional inequality holds yields (2.1.1). Bobkov improved on this bound by choosing $g = C||x|^2 - \alpha^2|^{1/2}$ for an arbitrary α , to be optimized.

2.2 From localization to stochastic localization

Lemma 2.1 is the most used version of the localization lemma. It is easily deduced from the original localization lemma, which we now discuss.

Lemma 2.3. *Let f and g be two continuous function on \mathbb{R}^n such that*

$$\int f(x)dx > 0 \quad \text{and} \quad \int g(x)dx > 0$$

Then, there exists an interval I , and an affine function ℓ on I such that,

$$\int_I f(t)\ell(t)^{n-1}dt > 0 \quad \text{and} \quad \int_I g(t)\ell(t)^{n-1}dt > 0$$

Now, using a suitable choice of f and g it is not too difficult to deduce a variant of Lemma 2.1 involving weights $\ell(t)^{n-1}$ in the one-dimensional integrals. Passing to exponential weights require a bit more work, we refer to [50] for details. The proof of Lemma 2.3 relies on the so called bisection method. In a nutshell, one chooses a hyperplane H

such that, say,

$$\int_{H^+} f = \int_{H^-} f = \frac{1}{2} \int_{\mathbb{R}^n} f > 0.$$

where H^+ and H^- are the two complementary half spaces. One of the two, say H^+ , must satisfies

$$\int_{H^+} g > 0$$

One then repeats this procedure countably many times and check that at the end of the procedure, he is left with an integral over a segment (or a point). We refer to [69] for details. Note that this proof works verbatim when we take integral with respect to a general log-concave measure instead of the Lebesgue measure.

From the same method, it is possible to derive another point of view on localization. Starting from a measure μ and an integrable function f with say $\int f d\mu = 0$, by repeatedly bisecting the space and preserving the measure of f , we get in the limit

Lemma 2.4. *There exists a disintegration of μ :*

$$\mu = \int \mu_\omega dP(\omega)$$

where

- *Almost surely μ_ω is a "needle", that is a one dimensional log concave probability or a Dirac.*
- *$\int f d\mu_\omega = 0$ almost surely*

Let us briefly comment on the difference. Notice that when we work with two functions, the process is not symmetric in f and g , we take a bisecting hyperplane for f and then choose one of the two halfspaces where the integral of g is > 0 . Instead, here, we work with only one function, thus we keep all the information at each step, resulting in the decomposition Lemma 2.4. It is then easy to recover the result for two functions for μ . Starting with two functions with positive integrals f and g , one applies Lemma 2.4 to $f - \int f d\mu$ for instance, and get a disintegration of μ into needles μ_ω satisfying

$$\int f d\mu_\omega = \int f d\mu$$

almost surely. Now, since $\mu = \mathbb{E}\mu_\omega$, there exists at least one ω for which

$$\int g d\mu_\omega > 0.$$

In other words, instead of selecting the good part of the decomposition at each step, we may select it at the very end.

This is for example the point of view that is taken in Klartag's extension of the localization method to the Riemannian context [60]. We shall recover this variant of localization in section 2.7.

The drawback of this procedure is that one has little control over the resulting one-dimensional measures, since the algorithm is not really tractable. A typical idea to try and circumvent this problem is to randomize the procedure. This is the idea that inspired the development of the stochastic localization process. Now, a few leaps of faith are gonna be required. The operation of restricting an integral to a half-space $H = \{x, x \cdot \theta \geq 0\}$ is the multiplication of the integrand by the convex one-dimensional function

$$x \rightarrow 1_{x \cdot \theta \geq 0}.$$

Instead, we shall work the simplest convex one-dimensional function, namely the linear form $x \rightarrow x \cdot \theta$. Let us start by investigating the first step of the algorithm. Let p be a density on \mathbb{R}^n and consider

$$p_1(x) = p(x)(1 + (x - a) \cdot \theta) \quad (2.2.1)$$

where θ is a random variable to be specified, and $a \in \mathbb{R}^n$. We assume that p is compactly supported, so that p_1 remains positive if θ is small enough. The simple computation $\int_{\mathbb{R}^n} p_1 = 1 + \int_{\mathbb{R}^n} (x - a) \cdot \theta p(x) dx$ shows that if we want p_1 to be a density, we shall choose a to be the barycenter of p . In that case, if we further choose θ to be of zero expectation, we have decomposed our initial density p into a mixture of random densities p_1 which satisfy

$$\mathbb{E}p_1(x) = \mathbb{E}p(x) \quad \text{for all } x \in \mathbb{R}^n.$$

A natural choice is to take θ uniformly distributed on a small sphere, or again, $\theta \sim \mathcal{N}(0, \epsilon)$ for a small parameter $\epsilon > 0$. Letting $\epsilon \rightarrow 0$ leads to investigating the infinite system of stochastic differential equations

$$dp_t(x) = p_t(x)(x - a_t) \cdot dB_t \quad (2.2.2)$$

where $(B_t)_{t \geq 0}$ is a standard Brownian and a_t is the barycenter of p_t . This is the definition of the stochastic localization process, which we now make rigorous.

2.3 The process and its basic properties

Let μ be a compactly supported probability with density p . For $t \geq 0$ and $\theta \in \mathbb{R}^n$ let

$$Z_{t,\theta} = \int_{\mathbb{R}^n} p(x) e^{-\frac{t}{2}|x|^2 + \theta \cdot x}.$$

We define a probability density by

$$p_{t,\theta}(x) = \frac{1}{Z_{t,\theta}} p(x) e^{-\frac{t}{2}|x|^2 + \theta \cdot x},$$

and its barycenter and covariance matrix :

$$a(t, \theta) = \int_{\mathbb{R}^n} x p_{t,\theta}(x) dx$$

and

$$A(t, \theta) = \int_{\mathbb{R}^n} (x - a(t, \theta))^{\otimes 2} p_{t,\theta}(x) dx.$$

Theorem 2.5. *The stochastic differential equation*

$$d\theta_t = a(t, \theta_t) dt + dB_t \tag{2.3.1}$$

with B_t a standard Brownian motion admits a unique solution $(\theta_t)_{t \geq 0}$. Furthermore, denoting by $a_t = a(t, \theta_t)$, the probability density $p_t = p_{t,\theta_t}$ satisfies

$$dp_t(x) = p_t(x)(x - a_t) \cdot dB_t \tag{2.3.2}$$

In other words, the infinite system of SDEs (2.3.2) admits the solution

$$p_t(x) = \frac{1}{Z_t} p(x) e^{-\frac{t}{2}|x|^2 + \theta_t \cdot x} \tag{2.3.3}$$

where Z_t is a normalizing factor.

The fact that equation (2.3.1) admits a unique well defined solution follows from standard argument about SDE's. While the rigorous approach is to derive (2.3.2) from (2.3.1), going in the reverse way is more intuitive. Thus we start with equation (2.3.2), we assume existence and show that p_t is indeed a density taking the form (2.3.3). Denote by $m_t = \int p_t(x) dx$. Much like in the discrete setting (2.2.1), the term a_t in (2.3.2) ensures that $dm_t = 0$. Thus p_t integrates to 1 almost surely. Furthermore, (2.3.2) is close to being a logarithmic derivative, so we compute :

$$\begin{aligned} d \log p_t(x) &= \frac{dp_t(x)}{p_t(x)} - \frac{1}{2} \frac{d[p(x)]_t}{p_t(x)^2} \\ &= (x - a_t) \cdot dB_t - \frac{1}{2} |x - a_t|^2 \\ &= x \cdot (dB_t + a_t dt) - \frac{1}{2} |x|^2 + dz_t \\ &= x \cdot d\theta_t - \frac{1}{2} |x|^2 + dz_t. \end{aligned}$$

Where z_t regroups the terms that do not depend on x , making up for the normalizing

factor Z_t .

We denote by μ_t the density with density p_t . From now on we assume that μ is log-concave.

Proposition 2.6. • *The process $(\mu_t)_{t \geq 0}$ is a martingale, in the sense that for any compactly supported function φ , then*

$$\left(\int \varphi d\mu_t \right)_{t \geq 0}$$

is a martingale.

- *For any $t \geq 0$, μ_t satisfies the $CD(t, \infty)$ criterion, that is*

$$\nabla^2(-\log(p_t)) \geq tId \quad \text{uniformly}$$

- *$\mu_t \rightarrow \mu_\infty$ weakly, where $\mu_\infty = \delta_{a_\infty}$ is a Dirac measure and $a_\infty \sim \mu$*

Proof. The first point follow from (2.3.2), and the second from (2.3.3). The third point follows from the first two. \square

Now we compute the dynamics of the barycenter and covariance matrix, which shall prove useful.

Lemma 2.7. $da_t = A_t dB_t$

Proof. We simply compute

$$\begin{aligned} da_t &= d \int x p_t(x) dx \\ &= \left(\int x(x - a_t)^t p_t(x) dx \right) \cdot dB_t \\ &= A_t dB_t \end{aligned}$$

\square

A similar but tedious computations shows that

Lemma 2.8.

$$dA_t = \left(\int (x - a_t)^{\otimes 3} p_t(x) dx \right) dB_t - A_t^2 dt$$

2.4 A functional-analytic description of the process

Before explaining how stochastic localization may be used to tackle the KLS conjecture, we describe the functional-analytic point of view developed by Klartag and Putterman [59] which sheds some light on the process.

Let f be a measurable function and consider the martingale

$$M_t = \int f \mu_t = \frac{\int e^{-\frac{t}{2}|x|^2 + \theta_t \cdot x} f(x) p(x) dx}{\int e^{-\frac{t}{2}|x|^2 + \theta_t \cdot x} p(x) dx} \quad (2.4.1)$$

We are interested in expressing this martingale in terms of the heat semi-group. The heat semigroup is the collection of operators $(P_s)_{s \geq 0}$ which acts on square integrable function by

$$P_s g = g * \gamma_s$$

where γ_s is the normal density with covariance sId . By abuse of notations we shall also denote by γ_s the corresponding measure. It is classical that P_t is a self-adjoint operator from L^2 to L^2 . We need to make a convolution appear in (2.4.1), but that is not too difficult. Let g be a function, $t \geq 0$ and $\theta \in \mathbb{R}^n$. Finally let $Z_\gamma(s) = (2\pi s)^{n/2}$, we write :

$$\begin{aligned} \int e^{-\frac{t}{2}|x|^2 + \theta_t \cdot x} p(x) dx &= \int e^{-\frac{t}{2}|x - \theta/t|^2} e^{\frac{t}{2}|\frac{\theta}{t}|^2} p(x) dx \\ &= e^{\frac{1}{2t}|\theta|^2} Z_\gamma(1/t) P_{1/t} g \left(\frac{\theta}{t} \right) \end{aligned}$$

By substituting this into (2.4.1) we get that

$$M_{1/t} = \frac{P_t(fp)(t\theta_{1/t})}{P_t f(t\theta_{1/t})} = Q_t f(t\theta_{1/t}) \quad (2.4.2)$$

where we have defined an operator $Q_t g = \frac{P_t(gp)}{P_t g}$ which depends on μ . When the context is not clear, we may thus write Q^μ for the operator associated to μ or Q^ρ for the same operator associated to a function ρ . Let us describe what this operator is exactly. Let g and h be two functions such that the following integrals are well defined, and define $\mu_{\gamma_t} = P_t \mu$ to be the evolution of μ along the heat semi group, that is the density of μ_{γ_t} is simply $P_t p$ (we use this rather unusual subscript to avoid confusion with the stochastic localization process). Then by definition of Q_t ,

$$\begin{aligned} \int h P_t g d\mu &= \int P_t h g p \, dx \\ &= \int h P_t(gp) \, dx \\ &= \int h Q_t g P_t p \, dx \\ &= \int h Q_t g d\mu_{\gamma_t}. \end{aligned}$$

In other words, if we see P_t as an operator from $L^2(\mu_{\gamma_t})$ to $L^2(\mu)$, then

$$Q_t = P_t^*. \quad (2.4.3)$$

In particular, we may write

$$\int f d\mu = \int f P_t(1) d\mu = \int Q_t f d\mu_{\gamma_t}. \quad (2.4.4)$$

On the other hand, coming back to (2.4.2) and denoting by $Y_t = t\theta_{\frac{1}{t}}$, we have

$$\int f d\mu = \mathbb{E} M_{1/t} = \mathbb{E} Q_t f(Y_t) \quad (2.4.5)$$

This suggests that $Y_t \sim \mu_{\gamma_t}$. If that is the case, we can deduce an explicit expression for the tilt process :

$$\theta_t = tY_t \sim tX + tW_{1/t} \sim tX + W_t \quad (2.4.6)$$

where $(W_t)_{t \geq 0}$ is a standard Brownian. The last equality (in law) is the classical time reversal property of the Brownian motion, which is easily checked by computing the quadratic variation $[tW_{1/t}]_t = t$. Equation (2.4.6) is indeed true, as was shown in [59] using stochastic calculus.

Lemma 2.9. *The following equality holds in law*

$$(\theta_t)_{t \geq 0} \sim (tX + W_t)_{t \geq 0}$$

where $X \sim \mu$ and $(W_t)_{t \geq 0}$ is a standard Brownian.

To summarize, from this functional-analytic viewpoint, the stochastic localization method consists in writing, up to a time reversal $s = 1/t$,

$$\int f d\mu = \int Q_s f d\mu_{\gamma_s},$$

and analyzing the operator Q_s . This is essentially the method that is used by Bodineau and Bauerschmidt in [14]. More about this connection can be found in Chapter 5 .

2.5 Approaching the KLS Conjecture

In this subsection, we establish the classical lemma which relates the KLS constant Ψ_n^2 to the behavior of the covariance matrix A_t .

We start with a set S of half measure, $\mu(S) = \frac{1}{2}$. Using the martingale property, and the gaussian factor of μ_T , we have, at time $T > 0$:

$$\begin{aligned}\mu^+(S) &= \mathbb{E}\mu_T^+(S) \\ &\gtrsim \frac{1}{\sqrt{T}} \mathbb{E}(\mu_T(S)(1 - \mu_T(S))).\end{aligned}\tag{2.5.1}$$

Thus, we need to understand how the process $g_t = \mu_t(S)$ evolves over time. We know it starts from $\frac{1}{4}$ and ends at 0 since μ_∞ is a Dirac. We hope it stays say greater than $\frac{1}{8}$ for a sufficiently long time. Simple Itô calculus shows that, for all $T \geq 0$

$$d(g_t(1 - g_t)) = \text{martingale} - d[g]_t$$

In particular,

$$\mathbb{E}(g_T(1 - g_T)) = \frac{1}{4} - \mathbb{E}[g]_T.\tag{2.5.2}$$

Let us compute the quadratic variation.

Lemma 2.10.

$$d[g]_t \leq |A_t|_{op} dt$$

Proof.

$$dg_t = d \int_S p_t(x) dx = \left(\int_S (x - a_t) p_t(x) \right) \cdot dB_t.$$

Thus,

$$\begin{aligned}d[g]_t &= \left| \int_S (x - a_t) p_t(x) \right|^2 dt \\ &= \sup_{\theta \in S^{n-1}} \left(\int_S (x - a_t) \cdot \theta p_t(x) \right)^2 dt \\ &\leq \sup_{\theta \in S^{n-1}} \left(\int_{\mathbb{R}^n} ((x - a_t) \cdot \theta)^2 p_t(x) \right) dt \\ &\leq |A_t|_{op} dt\end{aligned}$$

□

Plugging Lemma 2.10 into (2.5.2) we get that

$$E(g_T(1 - g_T)) \geq \frac{1}{4} - \int_0^T \mathbb{E}|A_t|_{op} dt.\tag{2.5.3}$$

Plugging this into (2.5.1), we arrive at

Lemma 2.11. *Let μ be a log-concave probability, and μ_t its stochastic localization with*

covariance A_t . Let $T \geq 0$ such that

$$\int_0^T \mathbb{E}|A_t|_{op} dt \leq \frac{1}{8}.$$

Then,

$$\Psi_\mu^2 \lesssim \frac{1}{T}$$

As a first consequence, it is now straightforward to recover the KLS bound (2.1.1). From Lemma 2.8, we see that any linear functional of A_t has a decreasing expectation. In particular, since $|A_t|_{op} \leq \text{Tr}(A_t)$, we may choose $T = \frac{1}{8\text{Tr}(A)}$ in the previous Lemma 2.11 and conclude that

$$\Psi_\mu^2 \lesssim \text{Tr}(A).$$

Let us summarize what is known about the behavior of A_t . From now on, we assume that the initial measure μ is log-concave and isotropic. The general idea is to choose a smooth potential dominating the operator norm of A_t , typically

$$\Phi_t = \text{Tr}(A_t^p)^{1/p},$$

for some $p \geq 1$. As we have just seen, the case $p = 1$ is rather easy and gives the original KLS bound. For greater p , one typically computes the dynamics $d\Phi_t$ and resorts to Gronwall type arguments to show that for a suitable choice of t , the potential Φ_t has not changed too much.

Using $p = 2$, Lee and Vampala [67] were able to show that

$$\mathbb{E}\text{Tr}(A_t^2)^{1/2} \leq C\text{Tr}(A_0^2)^{1/2} = C\sqrt{n} \quad \text{for } t \leq T = \frac{1}{C\sqrt{n}}$$

Thus proving the then best bound on the KLS constant :

$$\Psi_n^2 \lesssim \sqrt{n}.$$

Let us introduce a new constant

$$\kappa_n = \sup_X |\mathbb{E}X_1(X \otimes X)|_{HS}^2. \quad (2.5.4)$$

where the norm is the Hilbert-Schmidt one, and the supremum runs over all isotropic log-concave vectors X . Then, although originally proved in a slightly different setting, using $p = \log(n)$, Eldan showed that [35] :

Lemma 2.12. *There exists a universal constant C such that for $t \leq T_0 = \frac{1}{C\kappa_n^2 \log(n)}$,*

$$\mathbb{E}|A_t|_{op} \leq C.$$

Plugging this into Lemma 2.11 brings

$$\Psi_n^2 \lesssim \kappa_n^2 \log(n) \quad (2.5.5)$$

Eldan then went on to show that

$$\kappa_n^2 \lesssim \sigma_n^2 \log n, \quad (2.5.6)$$

thus implying a reduction of the KLS conjecture to the Variance conjecture up to a logarithmic factor:

Theorem 2.13.

$$\psi_n^2 \lesssim \sigma_n^2 \log(n)^2.$$

Furthermore, it is easy to see that

$$\kappa_n^2 \leq 4\psi_n^2 \quad (2.5.7)$$

Indeed, write $B = \mathbb{E}X_1(X \otimes X) = \mathbb{E}X_1XX^T$. Then, using the Poincaré inequality for X

$$\begin{aligned} |B|_{HS}^2 &= \text{Tr}(BB^T) = \mathbb{E}X_1\langle BX, X \rangle \\ &\leq (\mathbb{E}X_1^2)^{1/2} \text{Var}(\langle BX, X \rangle)^{1/2} \\ &\leq 2c_P(X) \mathbb{E}(|BX|^2)^{1/2} \\ &\leq 2c_P(X) |B|_{HS} \end{aligned}$$

where in the last inequality we have used that X is isotropic. This provides the bound

Lemma 2.14. *There exists a universal constant C such that for $t \leq T_1 = \frac{1}{C\Psi_n^2 \log(n)}$,*

$$\mathbb{E}|A_t|_{op} \leq C.$$

Of course plugging this directly into Lemma 2.11 brings a rather uninteresting bound $\Psi_n^2 \lesssim \Psi_n^2 \log(n)$. However, Lemma 2.12 indicates that if the stochastic localization is able to estimate the KLS constant up to a logarithmic term. We shall see a bit later that this logarithmic obstruction is not an artefact of the proof and is somewhat inevitable. This lemma also opens the gate to a bootstrap type argument. Improving the bound on Ψ_n^2 improves the control on A_t which itself controls Ψ_n^2 . It was this idea which led to Chen's breakthrough [28]. The additional ingredient he brought was the following lemma, saying essentially that the covariance matrix cannot explode too fast :

Lemma 2.15. *For any $p \geq 3$, and any $t_2 > t_1 > 0$,*

$$\mathbb{E}(\text{Tr}A_{t_2}^p)^{1/p} \leq \left(\frac{t_2}{t_1}\right)^{2p} \mathbb{E}(\text{Tr}A_{t_1}^p)^{1/p}$$

Let $p \geq 3$, $T_1 = \frac{1}{C\Psi_n^2 \log(n)}$ and $T_2 > T_1$. We write $\Phi_t = \text{Tr}(A_t^p)^{1/p}$. Combining Lemma

2.14 and Lemma 2.15 we have

$$\begin{aligned}
\int_0^{T_2} \mathbb{E}|A_t|_{op} &\leq \int_0^{T_1} \mathbb{E}|A_t|_{op} dt + \int_{T_1}^{T_2} \mathbb{E}\Phi_t dt \\
&\leq CT_1 + \mathbb{E}(\Phi_{T_1}) \int_{T_1}^{T_2} \left(\frac{t}{t_1}\right)^{2p} dt \\
&\leq o(1) + Cn^{1/p} \frac{T_2^{2p+1}}{T_1^{2p}}
\end{aligned}$$

Recall that we want this integral to be less than $\frac{1}{8}$. We thus need $T_2^{2p+1} \leq c \frac{T_1^{2p}}{n^{1/p}}$ for an appropriate constant c . That is we choose

$$T_2 = c \frac{T_1^{2p/2p+1}}{n^{1/p(2p+1)}}$$

Lemma 2.11 then implies

$$\begin{aligned}
\Psi_n^2 &\leq \frac{1}{T_2} \\
&\lesssim n^{1/2p(p+1)} (\Psi_n^2 \log(n))^{2p/2p+1}
\end{aligned}$$

that is,

$$\Psi_n^2 \leq C^p n^{1/p} \log(n)^{2p}$$

for a universal constant $C > 0$ and any $p \geq 3$. Optimizing in p , we choose $p = \sqrt{\frac{\log n}{\log \log n}}$ yielding Chen's bound :

$$\Psi_n^2 \lesssim \exp(C\sqrt{\log n \log \log n}). \quad (2.5.8)$$

Notice that this implies in particular that $\Psi_n^2 = o(n^\alpha)$ for any $\alpha > 0$.

After Chen's bound, a polylog bound was achieved by Klartag and Lehec [62]. Although still utilizing the stochastic localization process, Lemma 2.12 and Chen's Lemma 2.15 (extended to the case $p = 2$). Their proof is different and does not use exactly the same strategy. In particular they estimate the thin-shell parameter σ_n directly, by the looking at the dynamics of the barycenter a_t . Their proof yields

$$\sigma_n^2 \lesssim (\log n)^8$$

and thus via Eldan's Theorem 2.13,

$$\Psi_n^2 \lesssim (\log(n))^{10}.$$

The exponent can be improved as remarked later by Lee, Vampala and Jambulapati [49].

Finally, very recently, Klartag [56] obtained the bound.

$$\Psi_n^2 \lesssim \log n \quad (2.5.9)$$

To understand the proof we need to come back to the standard lemma, Lemma 2.11. Recall the strategy, we consider a set of half measure S and say that at time $T \geq 0$:

$$\begin{aligned} \mu^+(S) &= \mathbb{E} \mu_T^+(S) \\ &\geq \mathbb{E} \left(\sqrt{c_P(\mu_T)} \mu_T(S) (1 - \mu_t(S)) \right) \end{aligned} \quad (2.5.10)$$

Now, up until here we have established a criterion to ensure that the term $\mu_T(S)(1 - \mu_T(S))$ is essentially a constant for sufficiently small T and used the fact that $c_P(\mu_T) \leq \frac{1}{T}$. Let us use our best bound on $|A_t|_{op}$ which is Lemma 2.14. We set $T = \frac{1}{c_1 \Psi_n^2 \log(n)}$ for a suitable c_1 and we have for all $t \leq T$

$$\mathbb{E}|A_t|_{op} \leq C \quad (2.5.11)$$

for a universal constant C . Now, denoting again $g_t = \mu_t(S)$, we have

$$\begin{aligned} \mathbb{E}(g_T(1 - g_T)) &\geq \frac{1}{4} - \int_0^T \mathbb{E}|A_t|_{op} dt \\ &\geq \frac{1}{4} - \frac{c_2}{\Psi_n^2 \log(n)} \end{aligned} \quad (2.5.12)$$

where $c_2 = \frac{C}{c_1}$. As usual this ensures that we will be able to bound from below the term $g_T(1 - g_T)$. In particular, for a suitable choice of c_1 , we have

$$\mathbb{E}(g_T(1 - g_T)) \geq \frac{7}{32}$$

for instance. Since $g_T(1 - g_T) \leq \frac{1}{4}$ almost surely, a little computation shows that

$$\mathbb{P} \left(g_T(1 - g_T) \geq \frac{1}{8} \right) \geq \frac{3}{4} \quad (2.5.13)$$

Define the event

$$\mathcal{U} = \{|A_T|_{op} \leq 4C\}.$$

Recall that by (2.5.11), we have

$$\mathbb{P}(\mathcal{U}) \geq \frac{3}{4}. \quad (2.5.14)$$

In particular, we have

$$\mathbb{P} \left(\mathcal{U} \cap g_T(1 - g_T) \geq \frac{1}{8} \right) \geq \frac{1}{4}.$$

We denote by \mathcal{V} this event. We may thus write

$$\begin{aligned}\mu^+(S) &= \mathbb{E}\mu_T^+(S) \\ &\geq \mathbb{E}\left(\sqrt{c_P(\mu_T)} \mu_T(S)(1 - \mu_t(S))1_{\mathcal{V}}\right) \\ &\gtrsim \mathbb{E}\left(\sqrt{c_P(\mu_T)}1_{\mathcal{V}}\right)\end{aligned}\tag{2.5.15}$$

Looking at the last term, we have thus reduced the estimation of the Poincaré constant of an isotropic log-concave probability μ (or rather its Cheeger constant) to the estimation of the Poincaré constant of measures μ_T which satisfy

- $\text{Cov}(\mu_T) \leq 4C$ for a universal constant C .
- μ_T satisfies $CD(T, \infty)$, for $T = c_1 \Psi_n^2 \log(n)$.

At this point we did not do anything new, but just looked at the consequences of Lemma 2.14. In a nutshell, it tells us that up to a constant, when we have to estimate Ψ_n^2 , we might restrict our attention to the probabilities satisfying the two points above. In other words we have gained a curvature term for free. Now, from the Brascamp-Lieb inequality we know that

$$c_P(\mu_T) \leq \frac{1}{T}$$

while the KLS conjecture predicts that

$$c_P(\mu_T) \lesssim 1 \quad ?$$

A natural question is to try an intermediate bound which is the geometric mean : Do we have

$$c_P(\mu_T) \lesssim \frac{1}{\sqrt{T}}?$$

The answer is yes. This was the key new ingredient of Klartag in his latest paper [56], where using the L^2 method he shows that

Lemma 2.16. *Let $\nu = e^{-V(x)}dx$ be a log concave probability which satisfies $\nabla^2 V \geq \rho$. Then*

$$c_P(\nu) \leq \sqrt{\frac{|\text{Cov}(\nu)|_{op}}{\rho}}.$$

Plugging this into (2.5.15), we have

$$\begin{aligned}\mu^+(S) &\gtrsim \mathbb{E}\left(\sqrt{c_P(\mu_T)}1_{\mathcal{V}}\right) \\ &\gtrsim \frac{1}{T^{1/4}}\mathbb{P}(1_{\mathcal{V}}) \\ &\gtrsim \frac{1}{T^{1/4}} \gtrsim (\Psi_n^2 \log n)^{1/4}\end{aligned}$$

This being true for any isotropic log-concave probability μ , we arrive at

$$\Psi_n^2 \lesssim \Psi_n \sqrt{\log n}$$

and finally,

$$\Psi_n^2 \lesssim \log n. \quad (2.5.16)$$

2.6 The logarithmic obstruction

In this section, we describe the obstruction to proving the KLS conjecture via the standard scheme of proof that we just explained. We prove the following observation which is due to Klartag (unpublished).

Lemma 2.17. *There exists a universal constant c such that starting from the isotropic exponential product probability μ , one has*

$$\mathbb{E}|A_T|_{op} \geq c \log(n) \quad \text{for } T = \frac{1}{\log(n)}$$

The idea is that the stochastic localization process respects the product structure. In other words, if $\mu = \mu_1 \otimes \cdots \otimes \mu_n$ is a product measure, then μ_t is also a product measure, and $\mu_t = \mu_{1,t} \otimes \cdots \otimes \mu_{n,t}$, where $\mu_{i,t}$ is the one-dimensional stochastic localization of μ_i , for all $t \geq 0$. Indeed, let $1 \leq i \leq n$ from (2.3.3),

$$\mu_{i,t} \propto e^{-tx_i^2 + \theta_{t,i}x} \mu_i,$$

but from Lemma 2.9, $\theta_{t,i}$ has the same law $tX_i + W_t$ where W_t is a one dimensional Brownian motion. Then, the operator norm of the covariance matrix A_t is the maximum of n independent random variables, which in the case of the exponential measure result, as we shall see, in a logarithmic factor.

We take $\mu_1 = \cdots = \mu_n = \mu_{exp}$, the isotropic one sided exponential measure :

$$d\mu_{exp} = p_{exp}(x)dx = \frac{1}{e}e^{-x}1_{\{x \geq -1\}}dx. \quad (2.6.1)$$

For better readability, we drop the subscript and write $p = p_{exp}$.

$$\begin{aligned} p_t &= \frac{1}{e\tilde{Z}_t} e^{-\frac{t}{2}x^2 + \theta_t x} 1_{\{x \geq -1\}} \\ &= \frac{1}{Z_t} e^{-\frac{t}{2}\left(x - \frac{\theta_t - 1}{t}\right)^2} 1_{\{x \geq -1\}} \\ &= \frac{1}{Z(t, h_t)} e^{-\frac{t}{2}(x - h_t)^2} 1_{\{x \geq -1\}} \end{aligned}$$

where $h_t = \frac{\theta_t - 1}{t}$ and

$$Z(t, h) = \int_{-1}^{\infty} e^{-\frac{t}{2}(x-h)^2} dx = \int_{-1-h}^{\infty} e^{-\frac{t}{2}x^2} dx$$

We want to compute the variance of $\mu_{exp,t}$. We use the following standard fact about one-dimensional log concave densities

Lemma 2.18. *Let $\mu = f(x)dx$ be a one-dimensional log-concave probability, then*

$$\text{Var}_{\mu}(x) \simeq \frac{1}{(\max f)^2}$$

Note that this amounts to saying that the slicing conjecture holds in dimension 1.

From Lemma 2.18, we just need to compute the maximum of p_t . For $t \geq 0$ and $h \geq -1$ we write

$$p_{t,h}(x) = \frac{1}{Z(t, h)} e^{-\frac{t}{2}(x-h)^2} 1_{\{x \geq -1\}}$$

and $\mu_{t,h}$ the corresponding measure. Notice that if $h \geq -1$, the maximum of $p_{t,h}$ is precisely attained at h and has value

$$\sqrt{\frac{t}{2\pi}} \leq \frac{1}{Z(t, h)} \leq 2\sqrt{\frac{t}{2\pi}}$$

Thus, in that case

$$\text{Var}_{\mu_{t,h}}(x) \simeq \frac{1}{t} \tag{2.6.2}$$

Now, we determine the probability of the event $h_t \geq -1$ at time $t \geq 0$. Recall that $\theta_t \sim tX + B_t$ where X is distributed according to the one dimensional exponential measure and B_t is a standard Brownian.

$$\begin{aligned} \mathbb{P}(h_t \geq -1) &= \mathbb{P}(\theta_t \geq 1 - t) \\ &\geq \mathbb{P}(tX + B_t \geq 1) \\ &\geq \frac{1}{2} \mathbb{P}(X \geq \frac{1}{t}) \\ &\geq \frac{1}{2} e^{-1/t}. \end{aligned}$$

In particular, for $T = \frac{1}{\log(n)}$ we get that $\mathbb{P}(h_T \geq -1) \geq \frac{1}{2n}$ which by (2.6.2) implies that

$$\mathbb{P}\left(\text{Var}_{\mu_T}(x) \geq \frac{1}{cT}\right) \geq \frac{1}{2n}. \tag{2.6.3}$$

Now let $\mu^{\otimes n}$ be the n dimensional product measure, $\mu_t^{\otimes n}$ be the stochastic localization process starting from $\mu^{\otimes n}$ and A_t its covariance matrix. Recall that $\mu_t^{\otimes n} = \mu_1, t \otimes \cdots \otimes \mu_{n,t}$

where the $\mu_{i,t}$ are iid copies of μ . Thus by (2.6.3),

$$\begin{aligned}
\mathbb{E}|A_T|_{op} &= \mathbb{E} \max_i \text{Var}_{\mu_{T,i}}(x) \\
&\gtrsim T \mathbb{P} \left(\max_i \text{Var}_{\mu_{T,i}}(x) \geq \frac{1}{cT} \right) \\
&\gtrsim T \left(1 - \frac{1}{2n} \right)^n \\
&\gtrsim T \\
&\gtrsim \log(n).
\end{aligned}$$

This proves Lemma 2.17. The interested reader can check that the transition arises indeed around T , that is, there is constant $c > 0$ such that

$$\mathbb{E}|A_t|_{op} \lesssim 1 \quad \text{for } t \leq \frac{1}{c \log n}.$$

To conclude this subsection, we describe a similar phenomenon happening for the smallest eigenvalue $\lambda_{\min}(A_t)$. First, let us explain why we should care about it. Recall the entropic formulation of the slicing conjecture, for any isotropic log-concave vector X , we have that

$$L_X \leq e^{-c \frac{\text{Ent}(X)}{n}} \quad (2.6.4)$$

for a universal constant $c > 0$, see for instance [11]. Now, take X an isotropic log-concave vector, and write X_t for its stochastic localization. By concavity of $x \rightarrow -x \log(x)$

$$\begin{aligned}
\text{Ent}(X) &\geq \mathbb{E} \text{Ent}(X_t) \\
&= \mathbb{E} \text{Ent}(A_t^{1/2} A_t^{-1/2} X_t) \\
&= \mathbb{E} \text{Ent}(\tilde{X}_t) + \frac{1}{2} \mathbb{E} \log \det(A_t)
\end{aligned} \quad (2.6.5)$$

where $\tilde{X}_t = A_t^{-1/2}(X_t - a_t)$ is isotropic. Remark that

$$c_P^2(\tilde{X}_t) \leq |A_t^{-1}|_{op} c_P^2(X_t) \leq \frac{1}{t \lambda_{\min}(A_t)}. \quad (2.6.6)$$

Now we use the following Lemma, which is due to Ball and Nguyen [11].

Lemma 2.19. *Let Y be an isotropic log-concave vector, then*

$$\text{Ent}(Y) \gtrsim -n c_P^2(Y)$$

Plugging this into (2.6.5), we get

$$\text{Ent}(X) \gtrsim \mathbb{E} \frac{-n}{t\lambda_{\min}(A_t)} + \mathbb{E} \log \det A_t \quad (2.6.7)$$

$$\gtrsim -n \left(\mathbb{E} \frac{1}{t\lambda_{\min}(A_t)} - \mathbb{E} \log \lambda_{\min}(A_t) \right) \quad (2.6.8)$$

$$\gtrsim -n \mathbb{E} \left(\frac{1+t}{t\lambda_{\min}(A_t)} \right) \quad (2.6.9)$$

where in the last inequality we used the inequality $\log(u) \geq -\frac{1}{u}$. In particular, we get

Lemma 2.20. *If $T > 0$ is such that $\mathbb{E} \frac{1}{\lambda_{\min}(A_t)} \leq 2$, then*

$$L_X \lesssim e^{\frac{c}{T}}$$

Unfortunately, using the same arguments as for the greatest eigenvalue, one can show that for the exponential product measure, the smallest eigenvalue behaves as

$$\mathbb{E} \frac{1}{\lambda_{\min}} \gtrsim 1 + t \log(n). \quad (2.6.10)$$

Thus, it is not possible to get a good estimate for the slicing constant via Lemma 2.20 in general. In that case the situation is much worse because of the exponential dependence.

Remark : It is possible to improve the decomposition for entropy (2.6.5). Indeed we can compute

$$\begin{aligned} d\text{Ent}(X_t) &= -d \int f_t(x) \log(f_t)(x) dx \\ &= \text{martingale} + \frac{1}{2} \int \frac{d[f]_t}{f_t} dx \\ &= \text{martingale} + \frac{1}{2} \int |x - a_t|^2 f_t(x) dx dt \\ &= \text{martingale} + \frac{1}{2} \text{Tr}(A_t) dt \end{aligned}$$

Thus we can get an exact decomposition for the entropy as :

$$\begin{aligned} \text{Ent}(X) &= \mathbb{E} \text{Ent}(X_t) + \frac{1}{2} \int_0^t \mathbb{E} \text{Tr}(A_s) ds \\ &= \mathbb{E} \text{Ent}(\tilde{X}_t) + \frac{1}{2} \mathbb{E} \log \det(A_t) + \frac{1}{2} \int_0^t \mathbb{E} \text{Tr}(A_s) ds \end{aligned}$$

However, the last term that we previously neglected should not help too much, since it is less than $\frac{1}{2}tn$.

2.7 From stochastic localization to random localization

Here we discuss some possible variants of the stochastic localization process as defined by Theorem 2.5. Let $(\dot{C}_t)_{t \geq 0}$ be symmetric non-negative matrices and set $C_t = \int_0^t \dot{C}_s ds$. Starting with a density p we consider the evolution

$$dp_t(x) = p_t(x)(x - a_t) \cdot \dot{C}_t^{\frac{1}{2}} dB_t \quad (2.7.1)$$

In other words, instead of choosing a uniform random direction dB_t we insert a control matrix. So far we have used $\dot{C}_t^{1/2} = Id$. Much like previously, we get an explicit expression for the solution

Theorem 2.21. *The infinite system of SDEs (2.7.1) admits the solution*

$$p_t(x) = \frac{1}{Z_t} p(x) e^{-\frac{1}{2} x^T C_t x + \theta_t \cdot x} \quad (2.7.2)$$

where Z_t is a normalizing factor and $(\theta_t)_{t \geq 0}$ is the solution of the simple SDE

$$d\theta_t = \dot{C}_t a_t dt + \dot{C}_t^{1/2} dB_t.$$

Now let f be a measurable function and $M_t = \int f d\mu_t$. Using (2.7.1) we compute

$$\begin{aligned} dM_t &= d \int f(x) p_t(x) dx \\ &= \left(\dot{C}_t^{1/2} \int f(x)(x - a_t) p_t(x) dx \right) \cdot dB_t \end{aligned}$$

In particular, denoting $v_t = \int f(x)(x - a_t) p_t(x) dx$, if we choose $\dot{C}_t = P_{v_t^\perp} = P_t$ to be the orthogonal projector onto v_t^\perp ,

$$dM_t = 0$$

that is

$$\int f d\mu_t = \int f d\mu \quad a.s. \quad (2.7.3)$$

At time t we have then effectively decomposed the measure μ into a mixture of measures $\mu = \mathbb{E} \mu_t$ respecting the measure of a given function f . Let us denote by V_t the convex function such that $\mu_t = e^{-V_t(x)} dx$. By (2.7.2), we see that

$$\nabla^2 V_t \geq \int_0^t P_s ds \quad (2.7.4)$$

Now, for any $s \geq 0$, P_s is a projector of rank $n - 1$. In particular, for any projector Q of

rank 2

$$\dim(\Im Q \cap \Im P_s) \geq 2 - 1 = 1$$

thus,

$$\mathrm{Tr}(QP_sQ) \geq 1.$$

For a symmetric matrix M write $\lambda_1(M), \dots, \lambda_n(M)$ the decreasing sequence of eigenvalues of M . We have :

$$\begin{aligned} \lambda_{n-1}(\nabla^2 V_t) &\geq \frac{1}{2} \inf_Q \mathrm{Tr}(Q \nabla^2 V_t Q) \\ &\geq \inf_Q \int_0^t \mathrm{Tr}(QP_sQ) ds \\ &\geq \frac{t}{2} \end{aligned}$$

where the infimum runs over rank two projectors. This ensures all but one eigenvalues of $\nabla^2 V_t$ goes to infinity. As a consequence, in the limit, μ_∞ is either a Dirac or a one-dimensional log-concave measure. We have thus obtained a needle decomposition of μ ,

$$\mu = \mathbb{E} \mu_\infty$$

where μ_∞ are log-concave probabilities satisfying

$$\int f d\mu_\infty = \int f d\mu.$$

As an application, since μ_∞ is more log-concave than μ , if μ satisfies a curvature condition $\nabla^2(V) \geq \rho$ then so does μ_∞ . One can then prove inequalities like $c_P(\mu) \leq \frac{1}{\rho}$ by just proving it in dimension one.

In the general setting where no curvature condition is assumed, one can still say that

$$\begin{aligned} \mathrm{Var}_\mu(f) &= \mathbb{E}_\mu(f^2) - (\mathbb{E}_\mu(f))^2 \\ &= \mathbb{E} [\mathbb{E}_{\mu_\infty}(f^2)] - (\mathbb{E}_{\mu_\infty}(f))^2 \\ &= \mathbb{E} \mathrm{Var}_{\mu_\infty}(f) \\ &\lesssim \mathbb{E}_\mu(|A_\infty|_{op}) = \mathbb{E}_\mu(\mathrm{Tr}(A_\infty)) \end{aligned}$$

where in the last inequality we used the fact that the KLS conjecture holds in dimension one, and A_∞ has rank at most one. One might then hope to bound that quantity. For all we know, any needle decomposition (not necessarily the one given by the stochastic localization process) could satisfy $\mathbb{E} \mathrm{Tr}(A_\infty) \lesssim 1$, and thus prove the KLS conjecture.

The idea of using a projector matrix to get a desintegration of μ into low-dimensionnal measures has been used fruitfully notably by Klartag in [61] where the standard gaussian

measure is decomposed into lower-dimensionnal gaussian measures whose barycenter belongs to an analytic manifold, and by Eldan Koehler Zeitouni [37] who proved a localization type lemma for the Ising model. In the next section, we use a constant projector matrix to analyze measures which satisfies a curvature condition on a subspace of \mathbb{R}^n .

Chapter 3

On measures strongly log-concave on a subspace

This chapter is a reproduction of

Pierre Bizeul. On measures strongly log-concave on a subspace. *arXiv preprint arXiv:2203.09422*, 2022, to appear in *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*

3.1 Introduction

Let $V : \mathbb{R}^n \mapsto \mathbb{R}$ be a C^2 convex function, such that $d\mu(x) = e^{-V(x)}dx$ is a log-concave probability measure. It is well-known that if μ is t -strongly log-concave, that is V satisfies the Bakry-Émery condition :

$$\nabla^2 V \geq t \tag{3.1.1}$$

for some $t > 0$, where ∇^2 stands for the Hessian, it has good isoperimetric properties. In particular, its Poincaré constant is at most $\frac{1}{t}$. Recall that the measure μ is said to satisfy a Poincaré inequality with constant c if for all locally Lipschitz function f we have :

$$\text{Var}_\mu(f) \leq c^2 \mathbb{E}_\mu(|\nabla f|^2)$$

where here and in the sequel, $|\cdot|$ stands for the Euclidean norm. The best such constant is denoted by $c_P(\mu)$, the Poincaré constant of μ . The KLS conjecture [51] proposes that when μ is log-concave, its Poincaré constant is, up to a universal constant, less than the operator norm of its covariance matrix. Since Poincaré inequalities are homogeneous, we can state the conjecture only for normalized measures, without loss of generality. A measure μ is called isotropic if it is centered and its covariance matrix is the identity. Introduce

$$\Psi_n = \sup_{\mu} c_P(\mu),$$

where the supremum runs over all isotropic log-concave measures of \mathbb{R}^n . The KLS conjecture then reads :

$$\Psi_n \leq c$$

for some universal constant $c > 0$.

A related property of strongly log-concave probabilities is that they exhibit good concentration function. Recall that the concentration function of a measure μ is the function $\alpha_\mu : \mathbb{R}^+ \mapsto [0, 1/2]$ defined by :

$$\alpha_\mu(r) = \sup_{\{S, \mu(S)=1/2\}} \mu(S_r^c)$$

where $S_r = \{x \in \mathbb{R}^n / d(x, S) \leq r\}$ and $d(x, S)$ is the Euclidean distance between x and S . It follows from the Prékopa-Leindler inequality that if μ is t -strongly log-concave, then for all measurable sets S ,

$$\mu(S_r^c) \leq \frac{1}{\mu(S)} \exp\left(-\frac{tr^2}{4}\right). \quad (3.1.2)$$

In particular, it has a Gaussian-type concentration function :

$$\alpha_\mu(r) \leq 2 \exp\left(-\frac{tr^2}{4}\right), \quad (3.1.3)$$

see for instance [35] Proposition 2.6 and its proof. It was first observed by Gromov and Milman ([46] see also [64] Corollary 3.2 for a better constant) that a Poincaré inequality implies exponential concentration, that is :

$$\alpha_\mu(r) \leq \exp\left(-\frac{r}{3c_P(\mu)}\right). \quad (3.1.4)$$

The converse implication has been established in the log-concave case by E. Milman [73] where he shows that when μ is log-concave,

$$c_P(\mu) \lesssim \alpha_\mu^{-1}(1/4) \quad (3.1.5)$$

where for two expressions a, b depending on parameters, $a \lesssim b$ means there is a universal constant $c > 0$ such that $a \leq cb$. We also write $a \simeq b$ when $a \lesssim b$ and $b \lesssim a$.

In an attempt to tackle the KLS conjecture, Eldan [35] introduced a stochastic process, known as stochastic localization, which, roughly, decomposes, for all time $t \geq 0$, a log-concave measure μ into an average of measures $\mu_t(\omega)$ which are t -strongly log-concave. This strategy enabled Eldan to relate the KLS conjecture to the a priori weaker Variance conjecture, then Lee and Vempala [68] to obtain the then better bound on Ψ_n : $\Psi_n \lesssim n^{1/4}$. Recently, Chen obtained that $\Psi_n = o(n^\alpha)$ for every $\alpha > 0$, [28], and very recently, Klartag and Lehec obtained $\Psi_n = O(\log(n)^5)$ [57].

In this note, we propose a slight generalization of the criterion (3.1.1) allowing the potential to be flat in some directions. The observation is that stochastic localization behaves well when restrained to a subspace.

Our main result is the following :

Theorem 3.1. *Let $V : \mathbb{R}^n \mapsto \mathbb{R}$ be a C^2 convex potential such that $d\mu(x) = e^{-V(x)}dx$ is a probability measure. Suppose that there is $1 \leq k \leq n$, a subspace E of codimension k and $\eta > 0$ such that*

$$\nabla^2 V \geq \eta P_E$$

where P_E is the orthogonal projector onto E . Let K be the covariance matrix of μ . Define $Q = P_{E^\perp} K P_{E^\perp}$

$$(i) \quad c_P(\mu) \lesssim \max \left(\frac{1}{\sqrt{\eta}}, \|Q\|_{op}^{1/2} \Psi_k \sqrt{\max(\log(k), 1)} \right)$$

(ii) *There is a universal constant $c > 0$ such that for every A such that $\mu(A) = \frac{1}{2}$,*

$$\mu(A_r^c) \lesssim \exp \left(-c \min \left(\frac{r}{\|Q\|_{op}^{1/2}}, r^2 \min \left(\eta, \frac{1}{\Psi_k^2 \max(\log(k), 1) \|Q\|_{op}} \right) \right) \right)$$

In the particular case $E = \{0\}$, inequality (ii) implies a new bound for the concentration function of log-concave measures, which we state, without loss of generality, in the isotropic case.

Corollary 3.2. *For any isotropic log-concave measure μ and any $r > 0$, we have*

$$\alpha_\mu(r) \lesssim \exp \left(-c \min \left(r, \frac{r^2}{\Psi_n^2 \log(n)} \right) \right)$$

Remark : Note that (ii) implies (i). Indeed, choosing

$$r = c' \max \left(\frac{1}{\sqrt{\eta}}, \|Q\|_{op}^{1/2} \Psi_k \sqrt{\max(\log(k), 1)} \right),$$

for an appropriate choice of constant $c' > 0$, we get that $\mu(A_r^c) \leq \frac{1}{4}$. By (3.1.5), this implies (i). On the other hand it is easy to check that the exponential concentration obtained by combining (i) with (3.1.4) is weaker than (ii). *Remark :* The idea of evaluating concentration functions with stochastic localization already appears in the work of Lee and Vempala ([68], Theorem 16). To improve the Paouris deviation inequality for the Euclidean norm ([77]), they develop a more refined analysis of the process, using the so-called Stieltjes potential. They prove that for any L -Lipschitz function g , and any isotropic log-concave probability measure μ one has :

$$\forall t \geq 0 \quad \mathbb{P}(|g(X) - \bar{g}(X)| \geq Lt) \leq \exp \left(-\frac{ct^2}{t + \sqrt{n}} \right) \quad (3.1.6)$$

where $X \sim \mu$ and $\bar{g}(X)$ is the median or mean of $g(X)$. Notice that when g is the Euclidean norm, then by Borell's Lemma [22], $\mathbb{E}_\mu(|x|) \simeq \mathbb{E}_\mu(|x|^2)^{1/2} = \sqrt{n}$ since μ is isotropic. Plugging this into (3.1.6) yields

$$\forall t \geq 0 \quad \mathbb{P}(|X| \geq t\sqrt{n}) \leq \exp(-c \min(t, t^2)\sqrt{n}) \quad (3.1.7)$$

However, thanks to the new estimate of Chen, $\Psi_n = o(n^\alpha)$ for every $\alpha > 0$, we can obtain this result directly from Corollary 3.2. Indeed, for a general isotropic log-concave probability measure μ it asserts that for all measurable A such that $\mu(A) = 1/2$ and all $r > 0$,

$$\mu(A_r^c) \lesssim \exp\left(-c \min\left(r, \frac{r^2}{\Psi_n^2 \log(n)}\right)\right).$$

Let g be a L -Lipschitz function, and let $A = \{x \in \mathbb{R}^n, g(x) \leq \bar{g}(X)\}$, by definition of the median, $\mu(A) = 1/2$. Now set $G_r = \{x \in \mathbb{R}^n, g(x) \leq \bar{g}(X) + Lr\}$, then because g is L -Lipschitz, $A_r \subset G_r$, where A_r is the r -extension of A . We get that

$$\mu(G_r^c) \leq \mu(A_r^c) \lesssim \exp\left(-c \min\left(r, \frac{r^2}{\Psi_n^2 \log(n)}\right)\right),$$

For the Euclidean norm, which is 1-Lipschitz, this yields

$$\mathbb{P}(|X| \geq r\sqrt{n}) \lesssim \exp\left(-c \min\left(r\sqrt{n}, \frac{r^2 n}{\Psi_n^2 \log(n)}\right)\right) \lesssim \exp(-c \min(r, r^2)\sqrt{n})$$

where we used the fact that $\Psi_n^2 \log(n) = o(\sqrt{n})$ thanks to Chen's estimate. Notice that using the Lee-Vampala estimate $\Psi_n^2 = O(\sqrt{n})$ would lead to an extra logarithmic factor in the deviation estimate whose removal was the object of their work with the Stieltjes potential.

Lemma 3.3. *It is enough to prove Theorem 3.1 when $Q = I_k$.*

Proof. Let $d\mu(x) = e^{-V(x)}dx$ be a measure satisfying the hypothesis of Theorem 3.1 and let X be a random vector whose law is μ . Set $S = \begin{bmatrix} \|Q\|_{op}^{1/2} I_{n-k} & 0 \\ 0 & Q^{1/2} \end{bmatrix}$ where the matrix is expressed in a basis adapted to the splitting $\mathbb{R}^n = E \oplus E^\perp$. Define the random vector $\tilde{X} = S^{-1}X$, whose law is $d\tilde{\mu}(x) = e^{-\tilde{V}(x)}dx = |\det S| e^{-V(Sx)}dx$ and covariance matrix

$$\tilde{K} = S^{-1}KS^{-1}.$$

For a symmetric $n \times n$ matrix M , we denote by $\lambda_1(M) \geq \dots \geq \lambda_n(M)$ its ordered

eigenvalues. It is classical and easy to check that for every $r > 0$ one has

$$\alpha_\mu(r) \leq \alpha_{\tilde{\mu}}\left(\frac{r}{\lambda_1(S)}\right) = \alpha_{\tilde{\mu}}\left(\frac{r}{\|Q\|_{op}^{1/2}}\right) \quad (3.1.8)$$

However, with this choice of S , $\tilde{\mu}$ satisfies :

$$\lambda_{n-k}(P_E \nabla^2 \tilde{V} P_E) \geq \tilde{\eta} = \|Q\|_{op} \eta \quad \text{and} \quad \tilde{Q} = P_{E^\perp} \tilde{K} P_{E^\perp} = I_k.$$

We can then apply Theorem 3.1 to $\tilde{\mu}$ which, combined with (3.1.8), yields the result. \square

We conclude this introduction with a classical inequality, which essentially goes back to Freedman [40], that we will use for controlling deviation of martingales in the sequel.

Lemma 3.4. *Let M_t be a continuous local martingale starting from 0.*

$$\forall T > 0 \quad \mathbb{P}(M_T \geq a, [M]_T \leq b) \leq \exp\left(-\frac{a^2}{2b}\right)$$

Proof. For all $\lambda \in \mathbb{R}$, define the process $\mathcal{E}(\lambda M)$ by

$$\mathcal{E}(\lambda M)_t = \exp\left(\lambda M_t - \frac{\lambda^2}{2}[M]_t\right).$$

Elementary Itô calculus shows that $\mathcal{E}(\lambda M)$ is a local martingale. Moreover, it is positive, so by Fatou's lemma it is a supermartingale. In particular, for all $t \geq 0$, $\mathbb{E} \mathcal{E}(\lambda M)_t \leq \mathcal{E}(\lambda M)_0 = 1$, that is :

$$\forall t \geq 0, \quad \mathbb{E} \exp\left(\lambda M_t - \frac{\lambda^2}{2}[M]_t\right) \leq 1$$

Now, assume that $[M]_T \leq b$ almost surely. Then,

$$\begin{aligned} \mathbb{P}(M_T \geq a) &= \mathbb{P}(\mathcal{E}(\lambda M)_t \geq e^{\lambda a - \frac{\lambda^2}{2}[M]_T}) \\ &\leq \mathbb{P}(\mathcal{E}(\lambda M)_t \geq e^{\lambda a - \frac{\lambda^2}{2}b}) \\ &\leq \mathbb{E}(\mathcal{E}(\lambda M)_t) e^{\frac{\lambda^2}{2}b - \lambda a} \\ &\leq e^{\frac{\lambda^2}{2}b - \lambda a} \end{aligned}$$

Choosing the optimal $\lambda = \frac{a}{b}$ yields :

$$\mathbb{P}(M_T \geq a) \leq e^{-\frac{a^2}{2b}}.$$

The proof follows from applying this argument to the local martingale $M_t^\tau = M_{t \wedge \tau}$, where $\tau = \inf\{t \geq 0, [M]_t \geq b\}$ is a stopping time. Indeed, remark that $[M_{t \wedge \tau}]_t \leq b$ almost surely, and that

$$\mathbb{P}(M_T \geq a, [M]_T \leq b) \leq \mathbb{P}(M_T^\tau \geq a) \leq e^{-\frac{a^2}{2b}}.$$

□

3.2 Restricted stochastic localization

Let μ be a log-concave measure satisfying the hypothesis of Theorem 3.1 with $Q = I_k$. We denote by $P : \mathbb{R}^n \mapsto \mathbb{R}^n$ the orthogonal projection onto the k -dimensional subspace E^\perp . In the following we work in an orthonormal basis such that this subspace is spanned by the k first basis vectors. Let f be the density of μ , for all $x \in \mathbb{R}^n$, consider the following stochastic differential equations :

$$df_t(x) = (x - a_t)^T P dB_t f_t(x) \quad ; \quad f_0(x) = f(x) \quad (3.2.1)$$

where $a_t = \int_{\mathbb{R}^n} x f_t(x) dx$ is the barycenter of the measure μ_t , which we define here as having density f_t , and $(B_t)_{t \geq 0}$ is a standard Brownian motion on \mathbb{R}^n .

This system of equation is the same as the usual stochastic localization, except for the addition of the matrix P which projects the random direction given by the Brownian onto the subspace where we need to bend the potential. The idea of adding a projector first appears in a paper of Klartag [55] for other purposes. The following facts and computations are very standard, and we refer the reader to [35] and [68] for a more detailed exposition. In particular, we need to assume that the support of μ is bounded to grant the existence and well-definedness of the process for all time $t \geq 0$ and then extend the result to arbitrary μ by approximation; we again refer to [35].

Proposition 3.5. • *Equation (3.2.1) defines a function-valued martingale f_t in the sense that for any continuous and compactly supported function ϕ :*

$$\int_{\mathbb{R}^n} \phi(x) f_t(x) dx \quad \text{is a martingale} \quad (3.2.2)$$

- f_t is a density and for all $x \in \mathbb{R}^n$,

$$f_t(x) = \frac{1}{Z_t} e^{-\frac{t}{2} x^T P x + c_t \cdot x} f(x) := e^{-V_t(x)} \quad (3.2.3)$$

where c_t is the solution of :

$$c_0 = 0, \quad dc_t = P dB_t + P a_t dt \quad (3.2.4)$$

in particular we see that $\nabla^2 V_t \geq \min(\eta, t)Id$

Proof. For the existence and well-definedness of the process, see the remark below. While it is possible to check that f_t as defined by (3.2.3) satisfy (3.2.1), we sketch a different proof to lighten the exposition. Let $m_t = \int_{\mathbb{R}^n} f_t(x)dx$ be the total mass at time t . Recall that $a_t = \frac{1}{m_t} \int_{\mathbb{R}^n} x f_t(x)dx$ is the barycenter of f_t . Then, by (3.2.1),

$$\begin{aligned} dm_t &= \left(P \int_{\mathbb{R}^n} (x - a_t) f_t(x) dx \right) \cdot dB_t \\ &= (P a_t (m_t - 1)) \cdot dB_t. \end{aligned}$$

It is easy to check that this simple stochastic differential equation admits a unique solution (see [76] §5.2). It is given by $m_t = 1$. To establish (3.2.3), we use (3.2.1) to compute:

$$\begin{aligned} d \log f_t(x) &= \frac{df_t(x)}{f_t(x)} - \frac{1}{2} \frac{d[f(x)]_t}{f_t(x)^2} \\ &= (P(x - a_t)) \cdot dB_t - \frac{1}{2} (x - a_t)^T P (x - a_t) dt \\ &= x \cdot (P dB_t + P a_t dt) - \frac{1}{2} x^T P x dt + dz_t \\ &= x \cdot dc_t - \frac{1}{2} x^T P x dt + dz_t \end{aligned}$$

where dz_t regroups the terms that do not depend on x . It encodes the normalizing factor Z_t . The expression (3.2.3) together with the proof that $m_t = 1$ ensures that f_t is a density. The martingale property (3.2.2) is straightforward since, for any ϕ compactly supported,

$$d \int_{\mathbb{R}^n} \phi(x) f_t(x) dx = \left(\int_{\mathbb{R}^n} \phi(x) P(x - a_t) f_t(x) dx \right) \cdot dB_t.$$

Finally, the lower-bound on the Hessian of V_t is a direct consequence of (3.2.3). □

Remark : Equation (3.2.1) defines an infinite system of stochastic differential equations. It is therefore a priori unclear whether a solution exists. However there is a simpler, although arguably less intuitive, way of defining the process. First notice that, given the initial data f , a_t is but a function of t and c_t defined as the barycenter of the density f_t (3.2.3). Hence, we can first define c_t by equation (3.2.4) and then f_t by equation (3.2.3), and only then compute $df_t(x)$.

The next two lemmas are standard and straightforward computations in stochastic localization which are obtained using Itô calculus. See [35] and ([68], Lemma 20). We denote by K_t the covariance matrix of μ_t . Since the computation for its infinitesimal change dK_t is a bit tedious, we omit it to lighten the exposition.

Lemma 3.6. $da_t = K_t P dB_t$

Proof. By Itô calculus and (3.2.1),

$$\begin{aligned} da_t &= d \int_{\mathbb{R}^n} x f_t(x) dx = \int_{\mathbb{R}^n} x(x - a_t)^T P f_t(x) dx dB_t \\ &= \int_{\mathbb{R}^n} x(x - a_t)^T f_t(x) dx P dB_t \\ &= \int_{\mathbb{R}^n} (x - a_t)(x - a_t)^T f_t(x) dx P dB_t = K_t P dB_t \end{aligned}$$

□

Lemma 3.7. $dK_t = \int_{\mathbb{R}^n} (x - a_t)(x - a_t)^T P(x - a_t)^T dB_t f_t(x) dx - K_t P K_t dt$

Now we want to have an estimate of the concentration function of μ . We first need to understand how the measure of a set evolves along the process.

Lemma 3.8. *Let $S \subset \mathbb{R}^n$ be a measurable set and define $s_t = \mu_t(S)$, then :*

$$d[s]_t \leq (\|PK_t P\|_{op}) dt$$

Proof.

$$ds_t = \int_S df_t(x) dx = \left\langle \int_S P(x - a_t) f_t(x) dx, dB_t \right\rangle$$

So the quadratic variation is

$$\begin{aligned} d[s]_t &= \max_{|\xi| \leq 1} \left(\int_S \xi^T P(x - a_t) f_t(x) dx \right)^2 dt \\ &\leq \max_{|\xi| \leq 1} \left(\int_S (\xi^T P(x - a_t))^2 f_t(x) dx \right) \left(\int_S f_t(x) dx \right) dt \\ &\leq \max_{|\xi| \leq 1} (\xi^T P K_t P \xi) dt \leq (\|PK_t P\|_{op}) dt \end{aligned}$$

□

To control the above quadratic variation, we need to control the norm of $Q_t = PK_t P$. This is the purpose of the next section.

3.3 Control of the covariance matrix

We will see that the matrix Q_t , seen as a $k \times k$ matrix, follows the same dynamics as the covariance matrix of the standard stochastic localization in \mathbb{R}^k . To be more precise, it is the covariance matrix of the marginal density, which follows a stochastic localization dynamics. Hence, to control its operator norm, we use the same strategy as Eldan.

Lemma 3.9. Define $g_t(y) = \int_{\mathbb{R}^{n-k}} f_t(y, x) dx$ the marginal density of the vector $Y_t = PX_t$, where X_t is the random vector with density f_t . The barycenter of Y_t is $b_t = Pa_t$. Then,

$$dg_t(y) = (y - b_t)^T dW_t g_t(y) \quad (3.3.1)$$

where W_t is a standard Brownian in \mathbb{R}^k . Moreover Q_t is the covariance matrix of Y_t and

$$dQ_t = \int_{\mathbb{R}^k} (y - b_t)(y - b_t)^T (y - b_t)^T dW_t g_t(y) dy - Q_t^2 dt \quad (3.3.2)$$

Proof. The lemma follows from straightforward computations. \square

Remark : Equation (3.3.1) is the definition of the stochastic localization process used by Lee and Vampala [68] and Klartag and Lehec [57]. It is also the process used by Chen [28] when the initial measure is isotropic. Eldan [35] has a slightly different definition, even if most of the ideas used to analyze one process transfer to the other. From now, the main purpose of this section is to show that the operator norm of Q_t is bounded by a constant up to time $T = \frac{c_0}{\Psi_k^2 \max(\log(k), 1)}$, with c_0 a universal constant, see Lemma 3.14 below. This result essentially goes back to Eldan [35], in a slightly different setting, and further appears in Lee-Vampala ([68], Lemma 58) and Chen ([28], Lemma 7). We provide a simplified exposition of the proof of Chen. Following Eldan, we use the potential $\Gamma_t = \text{Tr}(Q_t^p) = \sum_{i=1}^k \lambda_i^p$ for some $p \geq 1$ where $\lambda_1 \geq \dots \geq \lambda_k$ are the eigenvalues of Q_t . In the following, we denote by (e_1, \dots, e_k) a basis of eigenvectors of Q_t , where the dependence on t and ω is implicit.

Lemma 3.10.

$$d\Gamma_t = \sum_i p \lambda_i^{p-1} u_{ii} \cdot dW - \sum_i p \lambda_i^{p+1} dt + \sum_{i \neq j} p \lambda_i^{p-1} \frac{|u_{ij}|^2}{\lambda_i - \lambda_j} dt + \sum_i \frac{p(p-1)}{2} \lambda_i^{p-2} |u_{ii}|^2 dt \quad (3.3.3)$$

where for all i, j , $u_{ij} = \int_{\mathbb{R}^k} (y - b_t) \cdot e_i (y - b_t) \cdot e_j (y - b_t) g_t(y) dy$

Proof. The functional $\Phi : M \mapsto \text{Tr}(M^p)$ defined on symmetric matrices is C^∞ . On the dense open set U of matrices whose eigenvalues are pairwise distinct, the functionals $M \mapsto \lambda_i(M)$ are smooth by implicit value theorem. Let $Q \in U$, with eigenvalues $\lambda_1, \dots, \lambda_k$ and eigenvectors e_1, \dots, e_k and let $q_{i,j}$ be the entries of Q in the basis e . The following computations are standard, see ([34], Lemma 1.4.8)

$$\nabla \lambda_i(Q) = e_i e_i^T.$$

For the second derivative, the only non-zero terms are

$$\frac{\partial^2 \lambda_i}{\partial q_{i,j}^2} = \frac{2}{\lambda_i - \lambda_j}.$$

Combining this with (3.3.2) proves the result when Q_t belongs to U , it is easy to see that it extends to the general case. \square

Lemma 3.11.

$$d(\Gamma_t^{1/p}) = v_t \cdot dW_t + \delta_t dt \quad (3.3.4)$$

where

$$v_t = \left(\sum_i \lambda_i^p \right)^{\frac{1}{p}-1} \left(\sum_i \lambda_i^{p-1} u_{ii} \right) \quad (3.3.5)$$

and

$$\delta_t \leq (p-1) \left(\sum_i \lambda_i^p \right)^{\frac{1}{p}-1} \sum_{i,j} \lambda_i^{p-2} |u_{ij}|^2 \quad (3.3.6)$$

Proof. By Ito calculus, $d(\Gamma_t^{1/p}) = \frac{1}{p} \Gamma_t^{\frac{1}{p}-1} d\Gamma_t + \text{Itô term}$. But $x \rightarrow x^{1/p}$ is concave, so the Itô term is negative. Injecting equation (3.3.3) yields

$$v_t = \left(\sum_i \lambda_i^p \right)^{\frac{1}{p}-1} \left(\sum_i \lambda_i^{p-1} u_{ii} \right)$$

and

$$\delta_t \leq \frac{p-1}{2} \left(\sum_i \lambda_i^p \right)^{\frac{1}{p}-1} \sum_i \lambda_i^{p-2} |u_{ii}|^2 + \left(\sum_i \lambda_i^p \right)^{\frac{1}{p}-1} \sum_{i \neq j} \frac{\lambda_i^{p-1}}{\lambda_i - \lambda_j} |u_{ij}|^2.$$

Now, notice that $u_{ij} = u_{ji}$ so that

$$\begin{aligned} \sum_{i \neq j} \frac{\lambda_i^{p-1}}{\lambda_i - \lambda_j} |u_{ij}|^2 &= \frac{1}{2} \sum_{i \neq j} \frac{\lambda_i^{p-1} - \lambda_j^{p-1}}{\lambda_i - \lambda_j} |u_{ij}|^2 \\ &\leq \frac{1}{2} \sum_{i \neq j} (p-1) \max(\lambda_i, \lambda_j)^{p-2} |u_{ij}|^2 \\ &\leq \frac{p-1}{2} \sum_{i \neq j} (\lambda_i^{p-2} + \lambda_j^{p-2}) |u_{ij}|^2 \\ &\leq (p-1) \sum_{i \neq j} \lambda_i^{p-2} |u_{ij}|^2 \end{aligned}$$

which proves the lemma. \square

In the next two lemmas, we bound $|v_t|$ and δ_t in terms of $\Gamma_t^{\frac{1}{p}}$ in order to apply a Gronwall-type argument.

Lemma 3.12. *There is a universal constant $c > 0$ such that for all $t \geq 0$,*

$$|v_t| \leq c \left(\Gamma_t^{\frac{1}{p}} \right)^{3/2} \quad a.s. .$$

Proof. Let $\tilde{Y} = Y_t - b_t$ be distributed according to $g_t(y - b_t)dt$, where we drop the dependence in t for readability. Let $\tilde{Y}_1, \dots, \tilde{Y}_k$ its coordinates in the basis e_1, \dots, e_k . \tilde{Y} is a centered log-concave vector of \mathbb{R}^k of covariance Q_t and for all $1 \leq i \leq k$, $\mathbb{E}\tilde{Y}_i^2 = \lambda_i$. Note that for all $1 \leq i \leq k$, $u_{ii} = \mathbb{E}[\tilde{Y}_i^2 \tilde{Y}]$. Then, for all $\theta \in S^{k-1}$,

$$\begin{aligned} |u_{ii} \cdot \theta| &= |\mathbb{E}[\tilde{Y}_i^2 \tilde{Y} \cdot \theta]| \\ &\leq \mathbb{E}[\tilde{Y}_i^4]^{1/2} \mathbb{E}[(\tilde{Y} \cdot \theta)^2]^{1/2} \\ &\lesssim \mathbb{E}[\tilde{Y}_i^2] \|Q_t\|_{op}^{1/2} \\ &\leq \lambda_i \Gamma_t^{1/2p} \end{aligned}$$

where in the second inequality we used Borell's lemma ([22]). This proves the lemma. \square

Lemma 3.13. *For all $t \geq 0$,*

$$\delta_t \leq 4p\Gamma_t^{2/p}\Psi_k^2$$

Proof. With the same notations as in the previous lemma, for all $1 \leq i, j \leq k$, $u_{ij} = \mathbb{E}[\tilde{Y}_i \tilde{Y}_j \tilde{Y}]$. For all $1 \leq i \leq k$, we define the matrix $\Delta_i = \mathbb{E}[\tilde{Y}_i \tilde{Y} \tilde{Y}^T]$. Following Chen [28], we compute :

$$\begin{aligned} \sum_{i,j} \lambda_i^{p-2} |u_{ij}|^2 &= \sum_{i,j,k} \lambda_i^{p-2} \mathbb{E}(\tilde{Y}_i \tilde{Y}_j \tilde{Y}_k)^2 \\ &= \sum_i \lambda_i^{p-2} \text{Tr}(\Delta_i^2) \\ &= \sum_i \lambda_i^{p-2} \text{Tr}(\Delta_i \mathbb{E} \tilde{Y}_i \tilde{Y} \tilde{Y}^T) \\ &= \sum_i \lambda_i^{p-2} \mathbb{E}(\tilde{Y}_i \tilde{Y}^T \Delta_i \tilde{Y}) \\ &\leq \sum_i \lambda_i^{p-2} \mathbb{E}(\tilde{Y}_i^2)^{1/2} \text{Var}(\tilde{Y}^T \Delta_i \tilde{Y})^{1/2} \\ &\leq \sum_i \lambda_i^{p-2} \lambda_i^{1/2} c_P(\tilde{Y}) \left(4\mathbb{E}[|\Delta_i \tilde{Y}|^2]\right)^{1/2} \\ &= 2c_P(\tilde{Y}) \sum_i \lambda_i^{p-3/2} \text{Tr}(\Delta_i Q_t \Delta_i)^{1/2} \\ &\leq 2c_P(\tilde{Y}) \left(\sum_i \lambda_i^p\right)^{1/2} \left(\sum_i \lambda_i^{p-3} \text{Tr}(\Delta_i^2 Q_t)\right)^{1/2} \end{aligned}$$

Now,

$$\begin{aligned}
\sum_i \lambda_i^{p-3} \text{Tr}(\Delta_i^2 Q_t) &= \sum_i \lambda_i^{p-3} \sum_{j,k} \lambda_j (\Delta_i)_{j,k}^2 \\
&= \sum_{i,j,k} \lambda_i^{p-3} \lambda_j \mathbb{E} \left(\tilde{Y}_i \tilde{Y}_j \tilde{Y}_k \right)^2 \\
&\leq \sum_{i,j,k} \lambda_i^{p-2} \mathbb{E}(\tilde{Y}_i \tilde{Y}_j \tilde{Y}_k)^2
\end{aligned}$$

where in the last inequality, we used the convexity inequality : $\lambda_i^{p-3} \lambda_j \leq \frac{p-3}{p-2} \lambda_i^{p-2} + \frac{1}{p-2} \lambda_j^{p-2}$. Plugging this into the inequality above yields :

$$\sum_{i,j} \lambda_i^{p-2} |u_{ij}|^2 \leq 2c_p(\tilde{Y}) \left(\sum_i \lambda_i^p \right)^{1/2} \left(\sum_{i,j} \lambda_i^{p-2} |u_{ij}|^2 \right)^{1/2}$$

which implies

$$\sum_{i,j} \lambda_i^{p-2} |u_{ij}|^2 \leq 4c_p(\tilde{Y})^2 \left(\sum_i \lambda_i^p \right).$$

Plugging this into (3.3.6) remarking that $c_P(\tilde{Y}) = c_P(Y_t)$ yields :

$$\begin{aligned}
\delta_t &\leq 4p\Gamma_t^{1/p} c_P(Y_t)^2 \\
&\leq 4p\Gamma_t^{1/p} \|Q_t\|_{op}^2 \Psi_k^2 \\
&\leq 4p\Gamma_t^{2/p} \Psi_k^2.
\end{aligned}$$

□

We are now in position to control the growth of Γ_t by a Gronwall-type argument.

Lemma 3.14. *There are constants $c_0, c_1 > 0$ such that for any $t \leq T = \frac{c_0}{\Psi_k^2 \max(\log(k), 1)}$, we have :*

$$\mathbb{P} \left(\max_{s \in [0, t]} \|Q_s\|_{op} \geq 10 \right) \leq \exp \left(-\frac{c_1}{t} \right).$$

As a consequence, for any measurable set $S \subset \mathbb{R}^n$ of measure $\mu(S) = 1/2$, setting $s_t = \mu_t(S)$, we have :

$$\mathbb{P}([s]_t \geq 10t) \leq \exp \left(-\frac{c_1}{t} \right)$$

Proof. Set $p = \max(\log(k), 1)$, so that

$$\Gamma_0^{1/p} \leq e$$

as we will use repeatedly in the proof. Recall that

$$d(\Gamma_t^{1/p}) = v_t \cdot dW_t + \delta_t dt$$

and define the stopping time $\tau = \inf\{t \geq 0, \Gamma_t^{1/p} \geq 3\Gamma_0^{1/p}\}$. We denote by M_t the martingale term $M_t = \int_0^t v_s \cdot dW_s$. For all $t \geq 0$ we have :

$$\begin{aligned} \Gamma_{t \wedge \tau}^{1/p} &= \Gamma_0^{1/p} + M_{t \wedge \tau} + \int_0^{t \wedge \tau} \delta_s ds \\ &\leq \Gamma_0^{1/p} + \int_0^{t \wedge \tau} 4p \Gamma_s^{2/p} \Psi_k^2 ds + M_{t \wedge \tau} && (\text{By Lemma 3.13}) \\ &\leq \Gamma_0^{1/p} + 36p \Gamma_0^{2/p} \Psi_k^2 t + M_{t \wedge \tau} && \left(\Gamma_s^{2/p} \leq 9\Gamma_0^{2/p} \right) \\ &\leq \Gamma_0^{1/p} (1 + 36e \max(\log(k), 1) \Psi_k^2 t) + M_{t \wedge \tau} && \left(\Gamma_0^{2/p} \leq e\Gamma_0^{1/p} \right) \end{aligned}$$

We choose $c_0 \leq \frac{1}{36e}$, so that for all $t \leq T = \frac{c_0}{\Psi_k^2 \max(\log(k), 1)}$,

$$\Gamma_{t \wedge \tau}^{1/p} \leq 2\Gamma_0^{1/p} + M_{t \wedge \tau}.$$

Consequently, for all such t ,

$$\begin{aligned} \mathbb{P}(\tau \leq t) &= \mathbb{P}(\Gamma_{t \wedge \tau} = \Gamma_\tau) \\ &\leq \mathbb{P}\left(M_{t \wedge \tau} \geq \Gamma_0^{1/p}\right). \end{aligned}$$

Now, τ being a stopping time, $M_{t \wedge \tau}$ is a martingale, whose quadratic variation is

$$[M]_{t \wedge \tau} = \int_0^{t \wedge \tau} |v_s|^2 ds \leq c \int_0^{t \wedge \tau} 3 \left(\Gamma_0^{1/p}\right)^{3/2} ds \leq 3ce^{3/2} t = \tilde{c}_1 t$$

where in the first inequality we used Lemma 3.12. By Lemma 3.4 we get :

$$\begin{aligned} \mathbb{P}(\tau \leq t) &\leq \mathbb{P}\left(M_{t \wedge \tau} \geq \Gamma_0^{1/p}\right) = \mathbb{P}\left(M_{t \wedge \tau} \geq \Gamma_0^{1/p}, [M]_{t \wedge \tau} \leq \tilde{c}_1 t\right) \\ &\leq \exp\left(-\frac{\Gamma_0^{2/p}}{2\tilde{c}_1 t}\right) \\ &\leq \exp\left(-\frac{c_1}{t}\right). \end{aligned}$$

With $c_1 = \frac{1}{2\tilde{c}_1}$. Now notice that $3\Gamma_0^{1/p} \leq 3e \leq 10$ which proves the first statement. The second statement follows from Lemma 3.8

□

3.4 Proof of the main theorem

Take a subset S of measure $1/2$ and $r > 0$, for $t \leq T = \frac{c_0}{\Psi_k^2 \max(\log(k), 1)}$ we have :

$$\begin{aligned}
\mu(S_r^c) &= \mathbb{E} \mu_t(S_r^c) \\
&\leq \mathbb{E} \left[\mu_t(S_r^c) \mathbb{1}_{\mu_t(S) \geq \frac{1}{4}} \right] + \mathbb{P} \left(\mu_t(S) \leq \frac{1}{4} \right) \\
&\leq 4 \exp\left(-\frac{1}{4} \min(\eta, t) r^2\right) + \mathbb{P}(s_0 - s_t \geq \frac{1}{4}, [s]_t \leq 10t) + \mathbb{P}([s]_t \geq 10t) \quad (\text{By (3.1.2)}) \\
&\leq 4 \exp\left(-\frac{1}{4} \min(\eta, t) r^2\right) + \exp\left(-\frac{1}{320t}\right) + \exp\left(-\frac{c_1}{t}\right) \quad (\text{By Lemmas 3.4 and 3.14}) \\
&\leq 4 \left(\exp\left(-\frac{1}{4} \min(\eta, t) r^2\right) + \exp\left(-\frac{c_4}{t}\right) \right) \quad \left(\text{with } c_4 = \min(c_1, \frac{1}{320}) \right)
\end{aligned}$$

Define $\beta = \min(\eta, T)$ and choose $t(r) = \min(\eta, T, \frac{1}{r}) = \min(\beta, \frac{1}{r})$ we get that :

- If $r \geq \frac{1}{\beta}$,

$$\mu(S_r^c) \leq 8 \exp(-c_5 r) \quad (3.4.1)$$

where $c_5 = \min(1/4, c_4)$

- If $r \leq \frac{1}{\beta}$,

$$\mu(S_r^c) \leq 4 \left(\exp\left(-\frac{1}{4} \min(\eta, T) r^2\right) + \exp\left(-\frac{c_4}{\min(\eta, T)}\right) \right) \leq 8 \exp(-c_5 \beta r^2) \quad (3.4.2)$$

Overall this implies that for all $r > 0$,

$$\mu(S_r^c) \lesssim \exp(-\min(c_0 r, c_1 \beta r^2))$$

which is the desired result.

Chapter 4

Log-Sobolev inequalities for subgaussian log-concave probabilities

This chapter is based on a preprint, to appear on arXiv.

4.1 Introduction and results

A Borel probability μ on \mathbb{R}^n is said to satisfy a logarithmic Sobolev inequality with constant $\rho > 0$ if for any locally Lipschitz function $f : \mathbb{R}^n \mapsto \mathbb{R}$, one has,

$$\text{Ent}_\mu(f^2) \leq 2\rho^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\mu, \quad (4.1.1)$$

where for a nonnegative function g , $\text{Ent}_\mu(g) = \mathbb{E}_\mu(g \log g) - \mathbb{E}_\mu(g) \log \mathbb{E}_\mu(g)$ and $|\cdot|$ denotes the euclidean norm. We denote by $\rho_{LS}(\mu)$ the optimal constant ρ such that (4.1.1) holds. It is well known that the log-Sobolev inequality (4.1.1) implies gaussian concentration. Indeed, the Herbst argument implies a quadratic bound on the logarithmic Laplace transform of Lipschitz functions.

$$\log \int e^{sf} d\mu \leq \frac{\rho^2 s^2 |f|_{\text{Lip}}^2}{2} + s \int_{\mathbb{R}^n} f d\mu, \quad (4.1.2)$$

where $|f|_{\text{Lip}}$ is the Lipschitz constant of f . Markov's inequality then implies gaussian concentration of f around its mean,

$$\mu(|f - \mathbb{E}_\mu(f)| \geq t) \leq 2e^{-\frac{t^2}{|f|_{\text{Lip}}^2 \rho^2}}. \quad (4.1.3)$$

Taking f to be a linear form, we see that μ must have sub-gaussian tails.

In a related direction, we say that μ satisfy a Poincaré inequality with constant $K > 0$

if for any locally Lipschitz function $f : \mathbb{R}^n \mapsto \mathbb{R}$, one has,

$$\text{Var}_\mu(f) \leq K^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\mu, \quad (4.1.4)$$

where $\text{Var}_\mu(f) = \mathbb{E}_\mu(f^2) - \mathbb{E}_\mu(f)^2$ is the variance of f . We denote by $C_P(\mu)$ the optimal constant K such that (4.1.4) holds. It is classical that the log-Sobolev inequality (4.1.1) is stronger than (4.1.4),

$$C_P(\mu) \leq \rho_{LS}(\mu).$$

Not all measure may satisfy a Poincaré or log-Sobolev inequality. Even under good integrability assumptions, if the support of μ is disconnected, one may build a non-constant function whose gradient vanishes μ almost everywhere, violating (4.1.4), hence also (4.1.1). A general class of measure which avoids double-bump type distribution is the class of log-concave measures, that is measures that write $d\mu = e^{-V(x)}dx$, for some convex $V : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$. For such measures, the well-known KLS conjecture proposes that it is enough, up to a universal constant, to test linear functions in (4.1.4).

Conjecture 4.1 (KLS). *There exists a constant $C > 0$ such that for any $n \in \mathbb{N}$ and any log-concave probability μ on \mathbb{R}^n ,*

$$C_P(\mu)^2 \leq C \sup_{\theta \in \mathbb{S}^{n-1}} \text{Var}_\mu(\langle \cdot, \theta \rangle)$$

The KLS conjecture has attracted a lot of attention since its original formulation in [50], culminating in a polylog estimate by Klartag and Lehec [62]. By analogy, it is natural to conjecture that the log-Sobolev constant of log-concave probabilities should be controlled by the Ψ_2 norm of the coordinates.

Conjecture 4.2. *There exists a constant $C > 0$ such that for any $n \in \mathbb{N}$ and any centered log-concave probability μ on \mathbb{R}^n ,*

$$\rho_{LS}(\mu) \leq C \sup_{\theta \in \mathbb{S}^{n-1}} |\langle \cdot, \theta \rangle|_{\Psi_2(\mu)}$$

where for a function $g : \mathbb{R}^n \mapsto \mathbb{R}$, $|g|_{\Psi_2(\mu)} = \inf \{t > 0 / \mathbb{E}_\mu[\exp(g^2/t^2)] \leq 2\}$.

Let us slightly reformulate.

Definition 4.3. Let μ be a probability, and b_μ its barycenter. We say that μ is α sub-gaussian if

$$\sup_{\theta \in \mathbb{S}^{n-1}} |\langle x - b_\mu, \theta \rangle|_{\Psi_2(\mu)} \leq \alpha$$

for some $\alpha > 0$.

We introduce the quantity

$$G_n = \sup_{\mu} \rho_{LS}(\mu) \quad (4.1.5)$$

where the supremum runs over all 1 sub-gaussian log-concave measures μ on \mathbb{R}^n . By scaling, Conjecture 4.2 is equivalent to the boundedness of G_n .

Bobkov proved [19] that if μ is a centered log-concave probability on \mathbb{R}^n , then it satisfies a log-Sobolev inequality with constant of order $|X|_{\Psi_2}$ where X is distributed according to μ . We always have

$$|X|_{\Psi_2} \lesssim \sqrt{n} \sup_{\theta \in \mathbb{S}^{n-1}} |\langle \cdot, \theta \rangle|_{\Psi_2(\mu)}$$

where for two quantities a and b depending on parameters, we write $a \lesssim b$ when there exists a universal constant $c > 0$ such that $a \leq cb$ (see Proposition 4.19). This inequality is tight in general, since $|X|_{\Psi_2}^2 \geq \log(2)\mathbb{E}|X|^2$ which is obvious from the definition. Hence, Bobkov's result can be reformulated as

$$G_n \lesssim \sqrt{n}.$$

Our first result is the following improvement:

Theorem 4.4. *There exists a constant $C > 0$ such that for all $n \geq 1$*

$$G_n \leq C n^{1/4}.$$

In another direction, a well-known result of Bobkov [20] asserts that for a log-concave vector X ,

$$c_P^2(X) \lesssim \text{Var}(|X|^2). \quad (4.1.6)$$

In the case of the Euclidean ball, the inequality is tight, up to constant. In the same spirit, we show that

Theorem 4.5. *Let X be a log-concave vector, then*

$$\rho_{LS}^2(X) \lesssim ||X|^2 - \mathbb{E}|X|^2|_{\psi_1}$$

Yet again, the bound is tight, up to constants, when X is uniform over the Euclidean ball. As in [20], we use a localization argument to reduce the problem to dimension 1.

Next, we look at two subclasses of sub-gaussian log-concave probabilities. The first one is rotationally invariant log-concave probabilities. For this class we explain that Conjecture 4.2 holds. Using a result of Bobkov [21], we show that:

Theorem 4.6. *Let μ be a rotationally invariant log-concave probability. There exists a universal constant $C > 0$ such that,*

$$\rho_{LS}(\mu) \leq C |\langle \cdot, e_1 \rangle|_{\Psi_2(\mu)}$$

where e_1 is the first element of the canonical basis.

The second one is the class of tilt-stable log-concave probabilities.

Definition 4.7. We say that a measure μ is β tilt-stable, for some $\beta > 0$, if for all $h \in \mathbb{R}^n$

$$|\text{Cov}(\tau_h \nu)|_{op} \leq \beta^2 I_n.$$

where $\tau_h \mu = \frac{1}{Z_h} \mu e^{h \cdot x}$, and $|\cdot|_{op}$ denotes the operator norm.

Tilt-stability is a stronger requirement than sub-gaussianity.

Lemma 4.8. Let $\beta > 0$ μ be a β -tilt stable measure, then μ is $C\beta$ subgaussian.

We postpone its proof to Section 4.2 (see Lemma 4.18). As a consequence, Conjecture 4.2 implies the weaker

Conjecture 4.9. Let μ be a β tilt-stable log-concave probability, then

$$\rho_{LS}(\mu) \lesssim \beta.$$

We introduce the quantities

$$\tilde{G}_n = \sup_{\mu} \rho_{LS}(\mu)$$

and

$$\tilde{K}_n = \sup_{\mu} | |X| - \mathbb{E}|X| |_{\Psi_2}$$

where both suprema run over all log-concave probabilities μ of \mathbb{R}^n that are 1 tilt-stables. By scaling, Conjecture 4.9 then reads

$$\tilde{G}_n \lesssim 1.$$

Furthermore, since the norm is 1-Lipschitz, Proposition 4.15 below shows that

$$\tilde{K}_n \lesssim \tilde{G}_n. \tag{4.1.7}$$

We show the following reverse inequality:

Theorem 4.10.

$$\tilde{G}_n \lesssim n^{1/6} \tilde{K}_n^{1/3}$$

Remark : Plugging (4.1.7) into Theorem 14 only yields $\tilde{G}_n \lesssim n^{1/4}$, which is a corollary of Theorem 4.4. Any improvement over the inequality $\tilde{K}_n \lesssim n^{1/4}$ provides an improvement for \tilde{G}_n .

Let us say a word about the proof strategy.

Definition 4.11. Let $\beta > 0$. We say that a measure μ is β strongly tilt-stable if

$$\sup_{t>0} \sup_{h \in \mathbb{R}^n} \left| \text{Cov} \left(\frac{1}{Z_{t,h}} \mu e^{-t|x|^2 + h \cdot x} \right) \right|_{op} \leq \beta^2$$

By definition, strong tilt-stability implies tilt-stability. In Section 4.4, we show that

Lemma 4.12. *If μ is β strongly tilt-stable and log-concave, then*

$$\rho_{LS}(\mu) \lesssim \beta$$

Finally, we show that log-concave tilt-stable measures are strongly log-concave, but with an extra factor $n^{1/6} \tilde{K}_n^{1/3}$.

Lemma 4.13. *Let μ be a log-concave measure on \mathbb{R}^n that is 1 tilt-stable, then μ is $n^{1/6} \tilde{K}_n^{1/3}$ strongly tilt-stable.*

Combining Lemmas 4.12 and 4.13 proves Theorem 4.10.

Organization of the Chapter

In Section 4.2, we recall some background facts and prove Theorem 4.4. In Section 4.4, we investigate an approach to Conjecture 4.2 via stochastic localization. The only result from this section that we shall use later on is Corollary 4.35, which was first established in [29]. Section 4.3 is devoted to the proof of Theorem 4.5. Finally, in Section 4.5, we establish Theorem 4.6 and 4.10.

4.2 Background and Proof of Theorem 4.4

We start by recalling useful facts about sub-gaussian and sub-exponential random variables, for which a good reference is [81] and log-concave vectors.

4.2.1 Sub-gaussian random variables

Definition 4.14. Let X be a real random variable. Then the following properties are equivalent:

1. There exists $K_1 > 0$ such that

$$\mathbb{P}(|X| > t) \leq 2 \exp(-t^2/K_1^2) \quad \text{for all } t > 0.$$

2. There exists $K_2 > 0$ such that

$$\mathbb{E} \exp(X^2/K_2^2) \leq 2.$$

3. There exists $K_3 > 0$ such that

$$\mathbb{E} \exp(s(X - \mathbb{E}X)) \leq \exp(s^2 K_3^2) \quad \text{for all } s \in \mathbb{R}$$

Furthermore, the optimal constants in the three inequalities are all equivalent up to some universal constants. If X satisfies any of the above properties, we say that X is subgaussian, and we define :

$$|X|_{\Psi_2} = \inf \{t > 0 / \mathbb{E} [\exp(X^2/t^2)] \leq 2\}$$

Finally, for a measure μ on \mathbb{R}^n , and a function $f : \mathbb{R}^n \mapsto \mathbb{R}$, we define $|f|_{\Psi_2(\mu)} = |f(X)|_{\Psi_2}$ where X is distributed according to μ , $X \sim \mu$. In that terminology, the Herbst's argument (4.1.3) may be reformulated as :

Proposition 4.15. *Let μ be a probability satisfying a log-Sobolev inequality. Then for any centered Lipschitz function f ,*

$$|f|_{\Psi_2(\mu)} \lesssim \rho_{LS}(\mu) |f|_{\text{Lip}}$$

It is seen, by an application of Jensen's inequality, that the Ψ_2 norm of a variable controls its L^2 norm, as mentioned in the introduction.

Lemma 4.16. *Let X be a sub-gaussian random variable, then*

$$\text{Var}(X) \leq \mathbb{E}X^2 \leq \log(2)|X|_{\Psi_2}^2$$

As a consequence, using the triangle inequality, one can show that centering only improves the Ψ_2 behavior.

Lemma 4.17. *Let X be a real random variable, then*

$$|X - \mathbb{E}X|_{\Psi_2} \lesssim |X|_{\Psi_2}.$$

We are now in position to prove that tilt-stability implies subgaussianity.

Lemma 4.18. *Let X be a 1-tilt stable random vector, then for all $t \in \mathbb{R}$ and $\theta \in S^{n-1}$*

$$\log \mathbb{E} e^{t(X - \mathbb{E}X) \cdot \theta} \leq t^2/2$$

Proof. Without loss of generality, we assume that X is centered. Write μ for the law of X . Denote by \mathcal{L}_ν the log-Laplace transform of ν , that is for any $h \in \mathbb{R}^n$,

$$\mathcal{L}_\nu(h) = \log \int_{\mathbb{R}^n} e^{h \cdot x} d\nu.$$

It is classical, and easily seen by direct differentiation, that the derivatives of the log-Laplace involves the moments of the measure:

$$\nabla \mathcal{L}_\nu(h) = \text{bar}(\tau_h \nu) \quad (4.2.1)$$

$$\nabla^2 \mathcal{L}_\nu(h) = \text{Cov}(\tau_h \nu). \quad (4.2.2)$$

Since $\mathcal{L}_\nu(0) = 0$ and $\nabla \mathcal{L}_\nu(h) = \text{bar}(\nu) = 0$, integrating (4.2.2) finishes the proof. \square

Notice that this implies Lemma 4.8. Finally, we establish the following deviation bound for the norm of a vector with sub-gaussian marginals.

Proposition 4.19. *Let X be a random vector in \mathbb{R}^n , define $\sigma_{SG}(X) = \sup_{\theta \in \mathbb{S}^{n-1}} |X \cdot \theta|_{\Psi_2}$. Then, there exists a universal constant $c_0 > 0$ such that for all $t \geq 2c_0\sqrt{n}\sigma_{SG}(X)$*

$$\mathbb{P}(|X| \geq t) \leq \exp\left(-\frac{t^2}{2c_0\sigma_{SG}^2(X)}\right).$$

Proof. We use a simple net argument, and we work with sub-optimal constants to lighten the proof. Let \mathcal{N} be a $\frac{1}{2}$ -net of the sphere. That is a collection of points on the sphere such that any point on the sphere is at distance at most $\frac{1}{2}$ of \mathcal{N} . It is classical that we might choose \mathcal{N} such that

$$|\mathcal{N}| \leq e^{2n}$$

where in that context $|\cdot|$ stands for the cardinal of the set. Now, for any $x \in \mathbb{R}^n$, we have that

$$|x| \leq 2 \sup_{\theta \in \mathcal{N}} x \cdot \theta. \quad (4.2.3)$$

Now, we use a simple union bound to establish the property. Let $t \geq 0$,

$$\begin{aligned} \mathbb{P}(|X| \geq t) &\leq \mathbb{P}(\exists \theta \in \mathcal{N} / x \cdot \theta \geq t) \\ &\leq |\mathcal{N}| \exp(-t^2/c_0\sigma_{SG}^2(X)) \\ &\leq \exp(2n - t^2/c_0\sigma_{SG}^2(X)) \\ &\leq \exp\left(-\frac{t^2}{2c_0\sigma_{SG}^2(X)}\right) \quad \text{for } t \geq 2c_0\sqrt{n}\sigma_{SG}(X). \end{aligned}$$

where we chose $c_0 \geq 1$. \square

4.2.2 Sub-exponential random variables

Definition 4.20. Let X be a real random variable. We say that if X is sub-exponential if there exists $K > 0$ such that

$$\mathbb{E} \exp(|X|/K) \leq 2.$$

In that case, we denote by $|X|_{\psi_1}$ the lowest such K .

Just like the ψ_2 norm, the ψ_1 norm controls the L^2 norm.

Lemma 4.21. *Let X be a sub-exponential random variable, then*

$$\text{Var}(X) \leq \mathbb{E}X^2 \leq 2|X|_{\Psi_1}^2$$

Proof. The lemma follows from the real inequality $1 + x^2 \leq e^{2|x|}$. □

And we deduce the centering lemma

Lemma 4.22. *Let X be a real random variable, then*

$$|X - \mathbb{E}X|_{\Psi_1} \lesssim |X|_{\Psi_1}.$$

4.2.3 Log-concave vectors

For a probability μ , we introduce its concentration function α_μ defined for $r \geq 0$ by

$$\alpha_\mu(r) = \sup_{A, \mu(A)=\frac{1}{2}} \mu(A_r^c)$$

where A_r is the r -extension of A :

$$A_r = \{x, d(x, A) \leq r\}$$

where d denotes the Euclidean distance. It is classical that if μ satisfies a log-Sobolev inequality with constant ρ then

$$\alpha_\mu(r) \leq \exp(-r^2/\rho^2)$$

which is a reformulation of (4.1.3).

The following result, which reduces the study of log-Sobolev inequalities for log-concave vectors to the a priori weaker gaussian concentration has been established by E.Milman in a series of papers([72],[74],[75])

Theorem 4.23. *Let μ be a log-concave measure, and $K > 0$ such that*

$$\alpha_\mu(r) \leq \exp(-r^2/K^2)$$

then,

$$\rho_{LS}(\mu) \lesssim K.$$

The following is the celebrated Bakry-Emery criterion [8], which provides a quantitative bound on the log-Sobolev constant of strongly log-concave measures:

Theorem 4.24. *Let $\mu = e^{-V(x)}$ be a log concave probability and assume that $\nabla^2 V \geq tI_d$ for some $t > 0$, then*

$$\rho_{LS}^2(\mu) \leq \frac{1}{t}.$$

Finally, we shall need the following results about one-dimensional log-concave vectors

Lemma 4.25. *Let X be a log-concave real random variable with unit variance $\text{Var}(X) = 1$. Then there exists universal constants c_0, c_1, c_2 such that*

1.

$$\text{Var}(X^2)^{1/2} \geq c_0.$$

2.

$$\text{Var}((X - \mathbb{E}X)^2)^{1/2} \leq c_1.$$

3.

$$|X - \mathbb{E}X|_{\psi_1} \leq c_2.$$

Proof. The existence of c_1 is just the fact that the thin-shell conjecture holds true in dimension one. The existence of c_2 is a reformulation of Borell's lemma.

Finally, there are various ways of proving the existence of c_0 . We replicate the proof given in [20]. We use the following extension of Borell's inequality proved by Bourgain [24]: If Q is a polynomial of degree k and $p > 0$, there exists a universal constant $C(k, p)$ such that for all log concave random vector Z in any dimension,

$$|Q(Z)|_p = (\mathbb{E}|Q(Z)|^p)^{1/p} \leq C(k, p)|Q(Z)|_0 = C(k, p)e^{\mathbb{E}\log|Q(Z)|}.$$

In our case, we write :

$$\begin{aligned} \text{Var}(X^2)^{1/2} &= |X^2 - \mathbb{E}X^2|_2 = |(X + (\mathbb{E}X^2)^{1/2})(X - (\mathbb{E}X^2)^{1/2})|_2 \\ &\geq |(X + (\mathbb{E}X^2)^{1/2})(X - (\mathbb{E}X^2)^{1/2})|_0 \\ &= |(X + (\mathbb{E}X^2)^{1/2})|_0 |(X - (\mathbb{E}X^2)^{1/2})|_0 \\ &\geq \frac{1}{C(1, 2)^2} |(X + (\mathbb{E}X^2)^{1/2})|_2 |(X - (\mathbb{E}X^2)^{1/2})|_2 \\ &\geq \frac{1}{C(1, 2)^2} \text{Var}(X) \\ &= \frac{1}{C(1, 2)^2}. \end{aligned}$$

□

4.2.4 A short proof of Theorem 4.4

Here we give a very short proof of Theorem 4.4, the main ingredient is the following estimate on the concentration function of log-concave probabilities, proved in [16].

Theorem 4.26. *Let μ be a log-concave probability with covariance matrix A , then there exists a constant $c_1 > 0$, such that*

$$\alpha_\mu(r) \leq \exp \left(-c_1 \min \left(\frac{r}{|A|_{op}^{1/2}}, \frac{r^2}{|A|_{op} \Psi_n^2 \log(n)} \right) \right)$$

We are now in position to prove Theorem 4.4.

Proof. Let μ be a 1 sub-gaussian log-concave probability and let X be distributed according to μ . We assume, without loss of generality that μ is centered. Let A be the covariance of μ . By Lemma 4.16,

$$A = \text{Cov}(\mu) \leq I_n$$

By Theorem 4.23, it is enough to estimate α_μ . Let S be any set of half measure, $\mu(S) = 1/2$. By Markov's inequality,

$$\mathbb{P}(|X| \geq 2\sqrt{n}) \leq \mathbb{P}(|X| \geq 2\text{Tr}(A)^{1/2}) \leq \frac{1}{4},$$

so that S intersects the ball of radius $2\sqrt{n}$. Hence for any $r \geq 4\sqrt{n}$, $S_r^c \subset B(0, r/2)^c$. Using Proposition 4.19, we get that

$$\alpha_\mu(r) \leq \exp \left(-\frac{r^2}{c} \right) \quad \text{for } r \geq R_0 = c\sqrt{n}(X). \quad (4.2.4)$$

for some absolute constant $c > 0$. For the small values of r , remark that, for $r \leq R_0$, we have that $r \leq \frac{r^2}{R_0}$. Plugging this into Theorem 4.26 yields for all $r \leq R_0$,

$$\begin{aligned} \alpha_\mu(r) &\leq \exp \left(-\frac{c' r^2}{\max \left(\sqrt{n} |A|_{op}^{1/2}, |A|_{op} \Psi_n^2 \log(n) \right)} \right) \\ &\leq \exp \left(-\frac{c'' r^2}{\sqrt{n} |A|_{op}^{1/2}} \right) \end{aligned} \quad (4.2.5)$$

where we used the fact that $\Psi_n^2 \log(n) = O(\sqrt{n})$ which has been known since the breakthrough of Chen [28]. Combining (4.2.4) and (4.2.5) finally yields

$$\alpha_\mu(r) \leq \exp \left(\frac{-c''' r^2}{\max \left(1, \sqrt{n} |A|_{op}^{1/2} \right)} \right) \leq \exp \left(\frac{-c''' r^2}{\sqrt{n}} \right)$$

which concludes the proof. \square

4.3 Proof of Theorem 4.5

In this section, we prove Theorem 4.5. The proof consists in a reduction to dimension one via a localization argument, together with a solution for the one-dimensional case.

4.3.1 The one dimensional case

Our aim is to prove the following lemma.

Lemma 4.27. *Let X be a log-concave random variable on the real line, then*

$$\rho_{LS}^2(X) \lesssim |X^2 - \mathbb{E}X^2|_{\psi_1}.$$

The first step is to show that the right-hand-side is minimized, up to constants, when X is centered:

Lemma 4.28. *Let Y be a centered log-concave random variable on the real line, then*

$$|Y^2 - \mathbb{E}Y^2|_{\psi_1} \lesssim \inf_{a \in \mathbb{R}} |(Y + a)^2 - \mathbb{E}(Y + a)^2|_{\psi_1}$$

Proof. Let Y be as in the definition. By homogeneity we may assume that $\mathbb{E}Y^2 = 1$, that is, Y is isotropic. We temporarily adopt the notation

$$K = |Y^2 - \mathbb{E}Y^2|_{\psi_1}.$$

Let c_0, c_1, c_2 be the three constants from Lemma 4.25. Recall that the standard deviation is a lower bound for the ψ_1 norm (Lemma 4.21).

We distinguish two cases:

- If $K \leq 8c_1c_2$

Then for any $a \in \mathbb{R}$, the log-concave vector $Y + a$ has unit variance, thus

$$\text{Var}((Y + a)^2)^{1/2} \geq c_0.$$

Thus,

$$\begin{aligned} |(Y + a)^2 - \mathbb{E}(Y + a)^2|_{\psi_1} &\geq \frac{1}{2} \text{Var}((Y + a)^2)^{1/2} \\ &\geq \frac{c_0}{2} \\ &\geq \frac{c_0}{16c_1c_2} K \end{aligned}$$

- If $K \geq 8c_1c_2$

Let $a \in \mathbb{R}$, expanding the squares, we find that

$$|(Y + a)^2 - \mathbb{E}(Y + a)^2|_{\psi_1} = |Y^2 + 2aY - 1|_{\psi_1}.$$

We compute:

$$\begin{aligned} \text{Var}(Y^2 + 2aY - 1) &= \text{Var}(Y^2 + 2aY) \\ &\geq \left(\text{Var}(2aY)^{1/2} - \text{Var}(Y^2)^{1/2} \right)^2 \end{aligned}$$

Now, since Y is an isotropic log-concave random variable,

$$c_0 \leq \text{Var}(Y^2)^{1/2} \leq c_1.$$

Thus, we get that

$$\begin{aligned} |Y^2 + 2aY - 1|_{\psi_1} &\geq \frac{1}{2} \text{Var}(Y^2 + 2aY - 1)^{1/2} \\ &\geq \frac{1}{2} \left(\text{Var}(2aY)^{1/2} - \text{Var}(Y^2)^{1/2} \right) \\ &\geq \frac{1}{2} (2a - c_1). \end{aligned} \tag{4.3.1}$$

We again make a case disjunction.

– If $a \geq \frac{K}{4c_2} + \frac{c_1}{2}$, we get from (4.3.1) that

$$|Y^2 + 2aY - 1|_{\psi_1} \geq \frac{K}{4c_2}$$

– If $a \leq \frac{K}{4c_2} + \frac{c_1}{2}$ we simply use the triangle inequality :

$$\begin{aligned} |Y^2 + 2aY - 1|_{\psi_1} &\geq |Y^2 - 1| - |2aY|_{\psi_1} \\ &\geq K - 2a|Y|_{\psi_1} \\ &\geq K - 2ac_2 \\ &\geq K/2 - c_1c_2 \\ &\geq K/4 \end{aligned}$$

In the end, we get that

$$|(Y + a)^2 - \mathbb{E}(Y + a)^2|_{\psi_1} \geq \frac{K}{C}$$

with $C = \max(4, 4c_2, \frac{c_0}{16c_1c_2})$, which is the desired result.

□

Now we lower bound the quantity $|Y^2 - \mathbb{E}Y^2|_{\psi_1}$ when Y is one dimensional and centered.

Lemma 4.29. *Let Y be a centered log-concave random variable on the real line, then*

$$|Y^2|_{\psi_1} \lesssim |Y^2 - \mathbb{E}Y^2|_{\psi_1}.$$

Proof. By the triangle inequality,

$$\begin{aligned} |Y^2|_{\psi_1} &\lesssim |Y^2 - \mathbb{E}Y^2|_{\psi_1} + |\mathbb{E}Y^2|_{\psi_1} \\ &\lesssim |Y^2 - \mathbb{E}Y^2|_{\psi_1} + \mathbb{E}Y^2 \\ &\lesssim |Y^2 - \mathbb{E}Y^2|_{\psi_1} + \text{Var}(Y) \end{aligned}$$

Now, applying Lemma 4.25 one more time,

$$\begin{aligned} \text{Var}(Y) &\lesssim \text{Var}(Y^2)^{1/2} \\ &\lesssim |Y^2 - \mathbb{E}Y^2|_{\psi_1} \end{aligned}$$

which concludes the proof. \square

Now we are in position to prove Lemma 4.27. Let X be a log-concave real random variable, and let $Y = X - \mathbb{E}X$. Recall Bobkov's result :

$$\rho_{LS}(X) = \rho_{LS}(Y) \lesssim |Y^2|_{\psi_1}.$$

Combining Lemmas 4.29 and 4.28, we get that

$$\begin{aligned} \rho_{LS}(X) &\lesssim |Y^2|_{\psi_1} \\ &\lesssim |Y^2 - \mathbb{E}Y^2|_{\psi_1} \\ &\lesssim |X^2 - \mathbb{E}X^2|_{\psi_1}, \end{aligned}$$

which is Lemma 4.27.

4.3.2 A localization argument

We use the following geometric version of the localization lemma:

Lemma 4.30. *Let μ be a log-concave probability, and S be any measurable set. Then there exists a disintegration*

$$\mu = \mathbb{E}_{\mu_\infty(\omega)}$$

where almost surely

- μ_∞ is one dimensional and log-concave
- $\mu_\infty(S) = \mu(S)$

In the sequel we drop the dependence in ω .

Now, for a measure μ , we define the quantity

$$\frac{1}{k_\mu} = \inf_S \frac{\mu^+(S)}{\mu(S)(1 - \mu(S))\sqrt{\log \frac{1}{\mu(S)}}} = \inf_S \frac{\mu^+(S)}{\mathcal{E}_\mu(S)}$$

Ledoux ([65]) showed that for all log-concave measure μ ,

$$k_\mu \simeq \rho_{LS}(\mu).$$

Now, we fix a set S and a disintegration $\mu = \mathbb{E}\mu_\infty$ given by the localization lemma. We may write:

$$\begin{aligned} \mu^+(S) &= \mathbb{E}\mu_\infty^+(S) \\ &\geq \mathbb{E}\left(\frac{1}{k_{\mu_\infty}}\mathcal{E}_{\mu_\infty}(S)\right) \\ &= \mathbb{E}\left(\frac{1}{k_{\mu_\infty}}\right)\mathcal{E}_\mu(S) \end{aligned}$$

Thus we need to estimate $\frac{1}{k_{\mu_\infty}}$. Now, let $a \geq 0$ and denote by $K_a = \|X\|^2 - a|_{\Psi_1}$ where $X \sim \mu$. Write X_∞ for the vector having density μ_∞ . We have that

$$2 \geq \mathbb{E}e^{(|X|^2 - a)/K_a} = \mathbb{E}\left[\mathbb{E}e^{(|X_\infty|^2 - a)/K_a}\right]$$

Thus by Markov's inequality, with probability greater than $1/2$,

$$\mathbb{E}e^{(|X_\infty|^2 - a)/K_a} \leq 4.$$

We work on that event that we denote by \mathcal{U} . Now since X_∞ is one-dimensional, we may write $X_\infty = b_\infty + \xi_\infty\theta_\infty$ where ξ_∞ is log concave, $|\theta| = 1$ and $b_\infty \perp \theta_\infty$. Thus,

$$\mathbb{E}e^{(|X_\infty|^2 - a)/K_a} = \mathbb{E}e^{(|b_\infty|^2 + \xi_\infty^2 - a)/K_a} \leq 4$$

That is, we have shown that

$$|\xi_\infty^2 - (a - |b|^2)|_{\psi_1} \leq 2K_a$$

In particular, although that is not necessary,

$$|\xi_\infty^2 - \mathbb{E}\xi_\infty^2|_{\psi_1} \lesssim K_a \tag{4.3.2}$$

since the mean minimizes the ψ_1 norm up to a universal constant. Now, using Lemma 4.27, we may rewrite (4.3.2) as

$$\frac{1}{k_{\mu_\infty}} \gtrsim \frac{1}{K_a} \quad (4.3.3)$$

Finally putting all together we get

$$\begin{aligned} \mu^+(S) &\geq E\left(\frac{1}{k_{\mu_\infty}}\right) \mathcal{E}_\mu(S) \\ &\geq E\left(\frac{1}{k_{\mu_\infty}} 1_{\mathcal{U}}\right) \mathcal{E}_\mu(S) \\ &\gtrsim \frac{1}{K_a} \mathcal{E}_\mu(S) \end{aligned}$$

S being arbitrary, we get

$$k_\mu \lesssim K_a = \|X\|^2 - a\|_{\Psi_1}.$$

Taking $a = \mathbb{E}X^2$ concludes the proof.

4.4 An approach via stochastic localization

In this section we describe a general strategy to estimate the log-Sobolev constant of a log-concave probability using stochastic localization. We briefly recall the definition and basic properties of the process in the Lee-Vampla formulation [67], we refer to [35],[67],[28] and [62] for a more detailed exposition.

Let μ be log-concave probability on \mathbb{R}^n with density f . For $t \geq 0$ and $h \in \mathbb{R}^n$ we introduce the probability

$$\mu_{t,h} = \frac{1}{Z_{t,h}} \mu e^{-t|x|^2 + h \cdot x} \quad (4.4.1)$$

which density will be denoted by $f_{t,h}$ and where $Z_{t,h}$ is a normalizing constant. We further denote by

$$a_{t,h} = \int_{\mathbb{R}^n} x \, d\mu_{t,h}$$

and

$$A_{t,h} = \int_{\mathbb{R}^n} (x - a_t)(x - a_t)^T d\mu_{t,h}$$

the barycenter and covariance matrix of the measure $\mu_{t,h}$. Consider the stochastic differential equation:

$$h_0 = 0 \quad dh_t = a_{t,h_t} dt + dB_t \quad (4.4.2)$$

where $(B_t)_{t \geq 0}$ is a standard Brownian motion. The stochastic localization of μ is the measure-valued process $(\mu_{t,h_t})_{t \geq 0}$, which by a slight abuse of notations, we hereby denote by $(\mu_t)_{t \geq 0}$. Accordingly, we drop the dependence in h_t to denote by f_t, a_t and A_t the density, barycenter, and covariance matrix of the process. The following lemma is classical, and is

an alternate definition of the process.

Lemma 4.31. *For any $x \in \mathbb{R}^n$,*

$$df_t(x) = (x - a_t)f_t(x) \cdot dB_t$$

As an immediate consequence, we obtain

Lemma 4.32. *For any measurable function φ , the process $(\int_{\mathbb{R}^n} \varphi d\mu_t)_{t \geq 0}$ is a martingale.*

To avoid unnecessary constants, we introduce for a measure ν with barycenter b ,

$$\tilde{\sigma}(\nu) = \inf\{K > 0, \forall u \in \mathbb{R}^n, \mathbb{E}_{X \sim \nu} e^{(X-b) \cdot u} \leq e^{\frac{u^2 K^2}{2}}\}.$$

By Definition 4.14, we have that $\tilde{\sigma} \simeq \sigma_{SG}$. We define $\tilde{\sigma}_t = \tilde{\sigma}(\mu_t)$. The following couple of lemmas show that in order to bound the log-Sobolev constant of μ , it is enough to bound $\tilde{\sigma}_t$. Let g be a locally-Lipschitz function, and $M_t = \int g^2 d\mu_t$.

Lemma 4.33. *For all $T > 0$,*

$$\begin{aligned} \text{Ent}_\mu(g^2) &= \mathbb{E} \text{Ent}_{\mu_T}(g^2) + \text{Ent}(M_T) \\ &= \mathbb{E} \text{Ent}_{\mu_T}(g^2) + \mathbb{E} \int_0^T \frac{d[M]_t}{2M_t} \\ &\leq \frac{2}{T} \mathbb{E}_\mu(|\nabla g|^2) + \mathbb{E} \int_0^T \frac{d[M]_t}{2M_t} \end{aligned}$$

Proof. The first line follows from the martingale property (Lemma 4.32), while the second one is derived from straightfroward Itô calculus. Finally for the last inequality we used the fact that $\mu_t = \mu_{t,h_t}$ satisfies the Bakry-Emery condition (Theorem 4.24), which can be seen from the definition (4.4.1). \square

Lemma 4.34.

$$\frac{d[M]_t}{2M_t} \leq \tilde{\sigma}_t^2 \text{Ent}_{\mu_t}(g^2).$$

Proof. Let $t \geq 0$, we first compute

$$dM_t = d \int g^2 d\mu_t = \left(\int g^2 (x - a_t) \mu_t \right) \cdot dB_t.$$

Let $\lambda > 0$ a parameter to be determined later. From the previous computation,

$$\begin{aligned}
d[M]_t &= \left| \int g^2(x - a_t) \mu_t \right|^2 \\
&= \sup_{\theta \in \mathcal{S}^{n-1}} \left(\int g^2(x - a_t) \cdot \theta \mu_t \right)^2 \\
&= \lambda M_t^2 \sup_{\theta \in \mathcal{S}^{n-1}} \left(\int \frac{g^2}{M_t} \frac{(x - a_t) \cdot \theta}{\sqrt{\lambda}} \mu_t \right) \\
&\leq \lambda M_t^2 \sup_{\theta \in \mathcal{S}^{n-1}} \left(\text{Ent}_{\mu_t} \left(\frac{g^2}{M_t} \right) + \log \mathbb{E}_{\mu_t} e^{\frac{(x - a_t) \cdot \theta}{\sqrt{\lambda}}} \right)^2 \\
&\leq \lambda M_t^2 \left(\frac{1}{M_t} \text{Ent}_{\mu_t} (g^2) + \frac{\tilde{\sigma}_t^2}{2\lambda} \right)^2 \\
&\leq 2 \left(\lambda \text{Ent}_{\mu_t}^2 (g^2) + \frac{\tilde{\sigma}_t^4 M_t^2}{4\lambda} \right)
\end{aligned}$$

where in the last inequality, we used that $(a + b)^2 \leq 2(a^2 + b^2)$ for reals a, b . Finally, we get that for any $\lambda > 0$,

$$\frac{d[M]_t}{2M_t} \leq \frac{\lambda \text{Ent}_{\mu_t}^2 (g^2)}{M_t} + \frac{\tilde{\sigma}_t^4 M_t}{4\lambda}.$$

Choosing the optimal $\lambda = \frac{\tilde{\sigma}_t^2 M_t}{2 \text{Ent}_{\mu_t}^2 (g)}$ concludes the proof. \square

At this point, it is unclear whether a high-probability bound on $\tilde{\sigma}_t$ is enough to establish a log-Sobolev inequality for μ . We explain the difference with what happens in the context of the KLS conjecture, where one seeks to bound the variance of an arbitrary function φ . Denoting by $N_t = \int_{\mathbb{R}^n} \varphi d\mu_t$ the analogs of Lemmas 4.33 and 4.34 are the followings

$$\text{For all } T \geq 0 \quad \text{Var}_{\mu}(\varphi) \leq \frac{1}{T} \mathbb{E}_{\mu} (|\nabla \varphi|^2) + \mathbb{E} \int_0^{T_0} d[N]_t \quad (4.4.3)$$

and the control on $d[N]_t$:

$$d[N]_t \leq |A_t|_{op} \text{Var}_{\mu_t}(\varphi) \quad (4.4.4)$$

which is obtained by Cauchy-Schwarz (see for instance [29]). Now, suppose that μ is isotropic, for normalization sake, one typically proves a bound of the form :

$$\mathbb{P} \left(\sup_{0 \leq t \leq T_0} |A_t|_{op} \geq 2 \right) \leq \exp \left(-\frac{c}{T_0} \right) \quad (4.4.5)$$

for some $T_0 \geq 0$. Now, by a theorem of E. Milman, we might assume that φ is 1-Lipschitz. Furthermore, we might also assume that μ has a bounded support, of polynomial diameter D (actually one can assume that $D \lesssim \sqrt{n}$). In that case, one has the trivial almost sure upper-bound

$$\text{For all } t \geq 0, \quad \text{Var}_{\mu_t}(f) \leq D^2 \quad \text{a.s.} \quad (4.4.6)$$

Using this, one can bound the second term of the right hand side of (4.4.3) as:

$$\begin{aligned} \mathbb{E} \int_0^{T_0} d[N]_t &\leq 2 \int_0^{T_0} \mathbb{E} \text{Var}_{\mu_t} \varphi \, dt + D^2 \exp\left(-\frac{c}{T_0}\right) \\ &\leq 2T_0 \text{Var}_{\mu}(\varphi) + D^2 \exp\left(-\frac{c}{T_0}\right) \end{aligned}$$

Now, since D is polynomial, if $T_0 \leq \frac{c_1}{\log(n)}$, $D^2 \exp\left(-\frac{c}{T_0}\right) = o(1)$. Plugging this into (4.4.3) yields:

$$\begin{aligned} \text{Var}_{\mu}(\varphi) &\leq \frac{2}{T_0} \mathbb{E}_{\mu} (|\nabla \varphi|^2) + o(1) \\ &\leq \frac{2}{T_0} + o(1) \\ &\lesssim \frac{2}{T_0}. \end{aligned}$$

This finally implies that $C_P(\mu)^2 \lesssim \frac{1}{T_0}$.

In our case, although it is enough to prove gaussian concentration for 1-Lipschitz function by Theorem 4.23, no such reduction is available at the level of the log-Sobolev inequality, so that it is unclear whether a high-probability bound for $\tilde{\sigma}_t$ of the type (4.4.5) would be enough to conclude that $\rho_{LS}(\mu)^2 \lesssim \frac{1}{T_0}$. However, it is clear that an almost sure bound on $\tilde{\sigma}_t$ is enough. As a consequence, we retrieve the following corollary, which already appears in [29].

Corollary 4.35. *Let μ be a M strongly log-concave log-concave probability on \mathbb{R}^n*

$$\rho_{LS}(\mu) \leq 2M$$

Proof. Recall that by definition,

$$M^2 = \sup_{t>0, h \in \mathbb{R}^n} |\text{Cov}(\frac{1}{Z_{h,t}} \mu e^{-t|x|^2+h \cdot x})|_{op}.$$

Set $t > 0$ and $h_0 \in \mathbb{R}^n$. The fact that

$$\sup_{h \in \mathbb{R}^n} |\text{Cov}(\frac{1}{Z_{h,t}} \mu e^{-t|x|^2+h \cdot x})|_{op} \leq M^2$$

shows that the measure μ_{t,h_0} is M tilt-stable. From Lemma 4.18, this implies that $\tilde{\sigma}(\mu_{t,h_0}) \leq M$. By letting t and h_0 take arbitrary values, we see that

$$\tilde{\sigma}_t \leq M \quad \text{a.s.}$$

Plugging this into Lemma 4.34 then in Lemma 4.33 yields, for an arbitrary function g , and

$T > 0$:

$$\begin{aligned} \text{Ent}_\mu(g^2) &\leq \frac{2}{T} \mathbb{E}_\mu(|\nabla g|^2) + M^2 \int_0^T \mathbb{E} \text{Ent}_{\mu_t} dt \\ &\leq \frac{2}{T} \mathbb{E}_\mu(|\nabla g|^2) + M^2 T \text{Ent}_\mu(g^2) \end{aligned}$$

Choosing $T = \frac{1}{2M^2}$ yields the result. \square

4.5 Some subclasses of subgaussian log-concave probabilities

4.5.1 Rotationally invariant measures

We say that a measure μ is rotationally invariant if for any orthogonal transformation $R \in O(n)$ and any measurable set A , $\mu(RA) = \mu(A)$. When μ is absolutely continuous with respect to the Lebesgue measure then $d\mu = \lambda(|x|)dx$ for some positive integrable function λ . Now, it is easy to check that μ is log-concave if and only if λ is log-concave and nonincreasing.

In order for μ to satisfy a log-Sobolev inequality, its marginals must be sub-gaussian. This is not always the case, as one might consider a density proportional to $e^{-|x|}$ (remark that this is not the exponential product measure, since the norm is the ℓ_2 one. However the decay of the tails is still only exponential.)

By scaling, we assume that μ is *isotropic*, that is for any $\theta \in \mathcal{S}^{n-1}$, $\int_{\mathbb{R}^n} (x \cdot \theta)^2 d\mu = 1$. Equivalently,

$$\int_{\mathbb{R}^n} |x|^2 d\mu = n.$$

Bobkov [21] established the following estimate for the concentration function.

Proposition 4.36. *The concentration function of μ satisfy :*

$$\alpha_\mu(r) \leq e^{-cr^2} \quad \text{for } r \leq \sqrt{n}$$

for some universal constant $c > 0$.

As a consequence, we get that

$$\alpha_\mu(r) \leq e^{-\frac{c_1 r^2}{\sigma_{SG}(\mu)^2}} \quad \text{for } r \leq 2\sigma_{SG}(\mu)\sqrt{n}. \quad (4.5.1)$$

where we used the fact that $\sigma_{SG}(\mu) \gtrsim 1$ since μ is isotropic. Using Proposition 4.19, we conclude that

$$\alpha_\mu(r) \leq e^{-\frac{c_2 r^2}{\sigma_{SG}(\mu)^2}} \quad \text{for } r \geq 0$$

for some universal constant $c_2 > 0$. By Theorem 4.23, this implies Theorem 4.6.

4.5.2 Tilt-stable measures

Let ν be a centered probability on \mathbb{R}^n . For any $h \in \mathbb{R}^n$, we define the tilted measure $\tau_h \nu = \nu_{0,h} = \frac{1}{Z_h} \nu e^{h \cdot x}$. Recall that ν is said to be β tilt-stable if for any $h \in \mathbb{R}^n$

$$\text{Cov}(\tau_h \nu) \preceq \beta^2 I_n.$$

Or, equivalently,

$$\nabla^2 \mathcal{L}_\nu(h) \leq \beta^2 I_n$$

uniformly in h , where \mathcal{L}_ν is the log-Laplace transform of ν , that is for any $h \in \mathbb{R}^n$,

$$\mathcal{L}_\nu(h) = \log \int_{\mathbb{R}^n} e^{h \cdot x} d\nu.$$

Tilted measures and the log-Laplace transform are known to play a central role in convex geometry ([53], [58], [36]). In the context of the discrete hypercube, tilt-stable measures appear notably in the work of Eldan and Shamir [39], where they are shown to exhibit non-trivial concentration and Eldan and Chen [29].

Examples of tilt-stable measures include strongly log-concave measures, by the Brascamp-Lieb inequality which is a weaker form of Theorem 4.24, as well as product of tilt-stable measures. Indeed, if $\nu = \nu_1 \otimes \nu_2 \otimes \cdots \otimes \nu_k$ is a product measure, notice that for any $h = (h_1, \dots, h_k)$,

$$\tau_h \nu = \tau_{h_1} \nu_1 \otimes \cdots \otimes \tau_{h_k} \nu_k$$

so that if each component is β_k tilt-stable, ν is itself tilt-stable with constant $\beta = \max_{1 \leq i \leq k} \beta_i$. The uniform measure on the discrete or continuous hypercube is then easily seen to be tilt-stable for instance.

An interesting question is to give sufficient conditions for a log-concave probability μ to be tilt-stable. A natural question in that direction is whether all sub-gaussian log-concave probabilities are in fact tilt-stables.

In the following, given a tilt-stable log-concave probability μ we use a perturbation argument to show that it is strongly tilt-stable by estimating

$$\sup_{t>0, h \in \mathbb{R}^n} |\text{Cov}(\mu_{t,h})|_{op}.$$

The idea is that the tilt-stability of μ allows us to get rid of the tilts in the above supremum. This is justified by the next lemma.

Lemma 4.37. *Let μ be a sub-gaussian probability, $h \in \mathbb{R}^n$ and $t > 0$. Then there exists*

$h_0 \in \mathbb{R}^n$ such that

$$\mu_{t,h} = \frac{1}{Z} (\tau_{h_0} \mu) e^{-t|x - \text{bar}(\tau_{h_0} \mu)|^2}$$

where Z is a normalizing constant. In other words, all the measures $\mu_{t,h}$ may be obtained as **centered** gaussian perturbations of tilts of μ .

Proof. Developping the right hand-side shows that

$$\frac{1}{Z} (\tau_{h_0} \mu) e^{-t|x - \text{bar}(\tau_{h_0} \mu)|^2} = \mu_{t, h_0 + 2t \text{bar}(\tau_{h_0} \mu)}.$$

Hence, we need to show that for any $t > 0$, the function $F : h_0 \rightarrow h_0 + 2t \text{bar}(\tau_{h_0} \mu)$ is onto. By (4.2.2), its jacobian is $J_F(h) = I_n + 2t \text{Cov}(\tau_{h_0} \mu) \geq I_n$. This implies that F is open (sends open sets to open sets) and proper (pre-image of compacts are compacts), hence onto. \square

A perturbation result

According to the previous Lemma, we wish to upper-bound the covariance of measures of the type $\nu_t = \frac{1}{Z_t} \nu e^{-t|x|^2}$ in terms of ν . In our setting, ν will be a centered tilt-stable log-concave measure. In general we can say the following.

Lemma 4.38. *Let ν be a probability, then*

$$|\text{Cov}(\nu_t)|_{op} \leq \int |x|^2 d\nu_t \leq \int |x|^2 d\nu$$

In particular, if ν is centered,

$$|\text{Cov}(\nu_t)|_{op} \leq \text{Tr}(\text{Cov}(\nu)) \tag{4.5.2}$$

Proof. It suffices to remarks that

$$\frac{d}{dt} \int |x|^2 d\nu_t = - \int |x|^2 d\nu_t + \left(\int |x|^2 d\nu_t \right)^2 \leq 0$$

and $|\text{Cov}(\nu_t)|_{op} \leq \text{Tr}(\text{Cov}(\nu_t)) \leq \int |x|^2 d\nu_t$. \square

Our goal is to improve on (4.5.2) when ν is a centered sub-gaussian log-concave probability.

The following lemma is inspired by Barthe and Milman [12].

Lemma 4.39. *Let ν be a centered sub-gaussian probability on \mathbb{R}^n with sub-gaussian constant $\sigma_{SG}(\nu)$ and $t > 0$. Then the probability*

$$\nu_t = \frac{1}{Z_t} e^{-t|x|^2} \nu$$

is sub-gaussian with constant:

$$\sigma_{SG}^2(\nu_t) \lesssim \sigma_{SG}^2(\nu) (1 + \log K(t))$$

$$\text{where } K(t) = \frac{\int_{\mathbb{R}^n} e^{-2t|x|^2} d\nu}{\left(\int_{\mathbb{R}^n} e^{-2t|x|^2} d\nu\right)^2}$$

Proof. For any $\lambda > 0$, let E_λ be the event $E_\lambda = \{\frac{d\nu_t}{d\nu} \leq \lambda\}$. Let S be a measurable set, we simply write

$$\begin{aligned} \nu_t(S) &= \int_{S \cap E_\lambda} d\nu_t + \int_{S \cap E_\lambda^c} d\nu_t \\ &\leq \int_{S \cap E_\lambda} \frac{d\nu_t}{d\nu} d\nu + \mathbb{P}_{\nu_t}(E_\lambda^c) \\ &\leq \lambda \nu(S) + \mathbb{P}_{\nu_t}(E_\lambda^c). \end{aligned}$$

Next, by Markov's inequality,

$$\mathbb{P}_{\nu_t}(E_\lambda^c) = \mathbb{P}_{\nu_t}\left(\frac{d\nu_t}{d\nu} > \lambda\right) \leq \frac{\mathbb{E}_{\nu_t} \frac{d\nu_t}{d\nu}}{\lambda} = \frac{K(t)}{\lambda}.$$

Let $\theta \in \mathcal{S}^{n-1}$ and $r > 0$. Setting $S = S_{\theta,r} = \{x, |x \cdot \theta| \leq r\}$, we get :

$$\begin{aligned} \nu_t(\{x, |x \cdot \theta| \leq r\}) &\leq \lambda \nu(\{x, |x \cdot \theta| \leq r\}) + \frac{K(t)}{\lambda} \\ &\leq 2\lambda e^{-\frac{cr^2}{\sigma_{SG}^2(\nu)}} + \frac{e^{\log K(t)}}{\lambda} \end{aligned}$$

where we used the sub-gaussianity of ν , and $c > 0$ is a universal constant. Optimizing in $\lambda > 0$ yields :

$$\begin{aligned} \nu_t(\{x, |x \cdot \theta| \leq r\}) &\leq \max\left(1, 2\sqrt{2} \exp\left(-\frac{cr^2}{2\sigma_{SG}^2(\nu)} + \frac{\log K(t)}{2}\right)\right) \\ &\leq 2 \exp\left(-\frac{c_1 r^2}{\sigma_{SG}^2(\nu) (1 + \log K(t))}\right) \end{aligned}$$

where $c_1 > 0$ is a universal constant. This shows that

$$|\langle \cdot, \theta \rangle|_{\Psi_2(\nu_t)}^2 \lesssim \sigma_{SG}^2(\nu) (1 + \log K(t)).$$

The functional $\langle \cdot, \theta \rangle$ is a priori not centered for ν_t . Using Lemma 4.17 concludes the proof. \square

Now, we wish to estimate $K(t)$. We are mostly interested in small values of t , since

for large t , when ν is log-concave, which is the case of interest for us, we will simply use $\text{Cov}(\nu_t) \leq \frac{1}{t}$. When ν is a standard Gaussian, a quick computation shows that $K(t)$ behaves like $\log K(t) \lesssim nt^2 = \mathbb{E}(|G|^2)t^2$. We recover this estimate with an extra factor, the Ψ_2 norm of $|X| - \mathbb{E}|X|$ (see Lemma 4.41 below). We start with a preliminary lemma

Lemma 4.40. *Let X be a random vector with sub-gaussian norm, then there exists a universal constant $c_1 > 0$ such that for any $r > 0$,*

$$\mathbb{P}(|X|^2 \leq \mathbb{E}|X|^2 - r) \leq \exp\left(-\frac{c_1 r^2}{4\mathbb{E}|X|^2 M^2}\right)$$

Proof. To alleviate notations, we denote by $M = \||X| - \mathbb{E}|X|\|_{\Psi_2(\nu)}$. For any $0 < r < \mathbb{E}|X|^2$ we have,

$$\begin{aligned} \mathbb{P}(|X|^2 \leq \mathbb{E}|X|^2 - r) &= \mathbb{P}\left(\sqrt{|X|} \leq \sqrt{\mathbb{E}|X|^2 - r}\right) \\ &\leq \mathbb{P}\left(|X| \leq (\mathbb{E}|X|^2)^{1/2} - \frac{r}{2(\mathbb{E}|X|^2)^{1/2}}\right) \\ &\leq \mathbb{P}\left(|X| \leq \mathbb{E}|X| + cM - \frac{r}{2(\mathbb{E}|X|^2)^{1/2}}\right) \end{aligned}$$

where in the first inequality we used the concavity of the square-root function, and in the second one, we used that

$$\begin{aligned} \mathbb{E}|X|^2 &= (\mathbb{E}|X|)^2 + \text{Var}(|X|) \\ &\leq (\mathbb{E}|X|)^2 + c^2 M^2 \leq (\mathbb{E}|X| + cM)^2 \end{aligned}$$

for a universal constant $c > 0$. Using the gaussian concentration for $|X|$, we get that for any $r \geq 4cM (\mathbb{E}|X|^2)^{1/2}$,

$$\mathbb{P}(|X|^2 \leq \mathbb{E}|X|^2 - r) \leq \mathbb{P}\left(|X| \leq \mathbb{E}|X| - \frac{r}{4(\mathbb{E}|X|^2)^{1/2}}\right) \quad (4.5.3)$$

$$\leq \exp\left(-\frac{r^2}{4\mathbb{E}|X|^2 M^2}\right) \quad (4.5.4)$$

Combining this with the trivial bound $\mathbb{P}(|X|^2 \leq \mathbb{E}|X|^2 - r) \leq 1$ for small r , yields the result. \square

Lemma 4.41. *Under the same hypothesis as in Lemma 4.39, we have*

$$\log K(t) \lesssim \left(1 + t^2 \||X| - \mathbb{E}|X|\|_{\Psi_2(\nu)}^2 \mathbb{E}|X|^2\right)$$

Proof. By Jensen's inequality, $\int_{\mathbb{R}^n} e^{-2t|x|^2} d\nu \geq e^{-t\mathbb{E}|X|^2}$, so that

$$K(t) \leq e^{2t\mathbb{E}|X|^2} \int_{\mathbb{R}^n} e^{-2t|x|^2} d\nu. \quad (4.5.5)$$

Now, using Lemma 4.40, for any $t > 0$

$$\begin{aligned} \mathbb{E} \left[e^{-t|X|^2} \right] &= \int_0^\infty \mathbb{P}(|X|^2 \leq r) te^{-tr} dr \\ &\leq t \int_0^{\mathbb{E}|X|^2} \mathbb{P}(|X|^2 \leq r) e^{-tr} dr + e^{-t\mathbb{E}|X|^2} \\ &\leq t \int_0^{\mathbb{E}|X|^2} \mathbb{P}(|X|^2 \leq \mathbb{E}|X|^2 - r) e^{-t(\mathbb{E}|X|^2 - r)} dr + e^{-t\mathbb{E}|X|^2} \\ &\leq te^{-t\mathbb{E}|X|^2} \int_0^{\mathbb{E}|X|^2} \exp \left(-\frac{c_1 r^2}{4\mathbb{E}|X|^2 M^2} + tr \right) dr + e^{-t\mathbb{E}|X|^2} \\ &\leq te^{-t\mathbb{E}|X|^2} e^{\frac{t^2 \mathbb{E}|X|^2}{c_1 M^2}} \int_{\mathbb{R}} \exp \left(-\left(\frac{\sqrt{c_1} r}{2M\sqrt{\mathbb{E}|X|^2}} - \frac{tM\sqrt{\mathbb{E}|X|^2}}{\sqrt{c_1}} \right)^2 \right) dr + e^{-t\mathbb{E}|X|^2} \\ &\leq e^{-t\mathbb{E}|X|^2} \left(2t\sqrt{\frac{2M^2 \mathbb{E}|X|^2}{c_1}} e^{\frac{M^2 t^2 \mathbb{E}|X|^2}{c_1}} + 1 \right) \\ &\leq \tilde{c}_1 e^{-t\mathbb{E}|X|^2} \left(1 + e^{\frac{t^2 \mathbb{E}|X|^2 M^2}{c_1}} \right) \end{aligned}$$

where in the last inequality we used that $z \leq e^{z^2}$ valid for all $z \in \mathbb{R}$. Combining this with (4.5.5) yields

$$K(t) \lesssim 1 + e^{\frac{2t^2 \mathbb{E}|X|^2 M^2}{c_1}},$$

concluding the proof. \square

4.5.3 Proof of Theorem 4.10

Let μ be a 1 tilt-stable log-concave probability. Let $t > 0$ and $h \in \mathbb{R}^n$. By Lemma 4.37, there exists $h_0 \in \mathbb{R}^n$ such that

$$\mu_{t,h} = \frac{1}{Z} (\tau_{h_0} \mu) e^{-t|x - \text{bar}(\tau_{h_0} \mu)|^2}.$$

Denote by $A_{t,h} = \text{Cov}(\mu_{t,h})$. Remark that $\tau_{h_0} \mu$ is again a 1 tilt-stable log-concave probability. Then, combining Lemmas 4.39 and 4.41, applied to $\tau_{h_0} \mu$ after a centering, we get that

$$\begin{aligned} |A_{t,h}|_{op} &\lesssim \sigma_{SG}^2(\mu_{t,h}) \\ &\lesssim 1 + t^2 \tilde{K}_n^2 \text{Tr}(A_{t,h}) \end{aligned}$$

$$\lesssim 1 + t^2 \tilde{K}_n^2 n$$

On the other hand, since μ is log-concave, μ_t satisfies the Bakry-Emery criterion with constant t , so that

$$\begin{aligned} |A_{t,h}|_{op} &\lesssim \max \left(1 + t^2 \tilde{K}_n^2 n, \frac{1}{t} \right) \\ &\lesssim n^{1/3} \tilde{K}_n^{2/3}. \end{aligned}$$

Applying Corollary 4.35 finishes the proof.

Chapter 5

Some additional remarks about tilt-stability and stochastic localization

5.1 More on tilt-stable measures

Let μ be a 1-tilt stable probability, which we do not assume to be log-concave at this point. For a symmetric positive matrix J and $h \in \mathbb{R}^n$ we define the measure

$$\mu_{J,h} \propto e^{\frac{1}{2}\langle Jx,x \rangle + \langle h,x \rangle} \mu.$$

We show that if the operator norm of J is not too large, then $\mu_{J,h}$ is well-behaved. When μ is the uniform probability on the discrete hypercube $\{-1;1\}^n$, the measures $\mu_{J,h}$ are called Ising models, which is an important topic in statistical physics. The following result is due to Bodineau and Bauerschmidt [13]. Although they were mostly interested in Ising models, their proof extend verbatim to this general setting.

Theorem 5.1. *Let μ be a 1-tilt-stable probability, $0 \preceq J \prec I_n$ and $h \in \mathbb{R}^n$, then*

- (i) $\text{Cov}(\mu_{J,h}) \leq \frac{1}{1-\|J\|_{op}} I_n$
- (ii) $c_P^2(\mu_{J,h}) \leq \frac{1}{1-\|J\|_{op}} \sup_h c_P^2(\mu_{0,h})$
- (iii) $\rho_{LS}^2(\mu_{J,h}) \leq \frac{1}{1-\|J\|_{op}} \sup_h \rho_{LS}^2(\mu_{0,h})$

Before moving to the proof, let us comment the result. First, notice that all inequalities become equalities if μ is the standard Gaussian. Now the fact that (i) implies (ii) and (iii) was recovered by Eldan and Chen in [29]. Our purpose here is to show how their method, based on the stochastic localization process, also gives (i) for free. The idea is essentially to get rid of the quadratic part of $\mu_{J,h}$ using the stochastic localization process.

From now on, we fix a matrix J and a vector h and we write

$$\nu = \mu_{J,h}$$

We denote by $T = \|J\|_{op}$. Recall the general form of the stochastic localization process Theorem 2.21. We make the choice $\dot{C}_t = \frac{J}{\|J\|_{op}}$ and write h_t in place of θ_t for the tilt process. Thus we get a martingale on the space of measures $(\nu_t)_{t \geq 0}$, we write $F_t = \frac{d\nu_t}{d\nu}$, its dynamics is

$$dF_t(x) = \left\langle \frac{J^{1/2}(x - a_t)}{\|J\|_{op}^{1/2}}, dB_t \right\rangle F_t(x) \quad \text{for } x \in \text{supp}(\nu) \quad (5.1.1)$$

Furthermore, ν_t has the explicit expression:

$$\begin{aligned} d\nu_t &= \frac{1}{Z_t} e^{-\frac{1}{2}t \frac{J}{\|J\|_{op}} x \cdot x + h_t \cdot x} d\nu \\ &= \frac{1}{Z_t} e^{\frac{1}{2} \langle (J - \frac{t}{\|J\|_{op}} J) x, x \rangle + \langle h + h_t, x \rangle} d\mu \end{aligned}$$

In other words, we have

$$\nu_t = \mu_{J(1 - \frac{t}{T}), h + h_t}$$

In particular, at time $t = T$, $\nu_T = \mu_{0, h+T}$. In other words we have decomposed the measure $\nu = \mu_{J,h}$ into a mixture of tilts of μ .

Now, let f be a function, and write $M_t = \int f d\nu_t$. Since M_t is a martingale, we have for all time $t \geq 0$,

$$\begin{aligned} \text{Var}_\nu(f) &= \mathbb{E} \text{Var}_{\nu_t}(f) + \text{Var}(M_t) \\ &= \mathbb{E} \text{Var}_{\nu_t}(f) + \mathbb{E} \int_0^t d[M]_s \end{aligned} \quad (5.1.2)$$

Now from (5.1.1), we compute for all time $t \geq 0$

$$dM_t = d \int f F_t d\nu = \left(\dot{C}_t^{1/2} \int f(x)(x - a_t) d\nu_t \right) \cdot dB_t$$

And using the Cauchy-Schwarz inequality,

$$\begin{aligned} d[M_t] &= |\dot{C}_t^{1/2} \int f(x)(x - a_t) d\nu_t|^2 dt \\ &= \sup_{\theta \in S^{n-1}} \left| \int f(x) C_t^{1/2}(x - a_t) \cdot \theta d\nu_t \right|^2 dt \\ &\leq \sup_{\theta \in S^{n-1}} \int \left(C_t^{1/2}(x - a_t) \right)^2 d\nu_t \text{Var}_{\nu_t}(f) dt \\ &\leq |C_t^{1/2} A_t C_t^{1/2}|_{op} \text{Var}_{\nu_t}(f) dt \end{aligned}$$

$$\leq |A_t|_{op} \text{Var}_{\nu_t}(f) dt$$

where in the last inequality we used the fact that $|C_t|_{op} = |\frac{J}{|J|_{op}}|_{op} = 1$. Plugging this into (5.1.2), we get that at time T

$$\text{Var}_{\nu}(f) \leq \mathbb{E} \text{Var}_{\nu_T}(f) + \mathbb{E} \int_0^T |A_t|_{op} \text{Var}_{\nu_t}(f) dt \quad (5.1.3)$$

We introduce the function

$$\alpha(u) = \sup_{J, |J|_{op}=u, h \in \mathbb{R}^n} \|\text{Cov}(\mu_{J,h})\|_{op}$$

Then if $f(x) = \langle x, \theta \rangle$ for some $\theta \in S^{n-1}$, (5.1.3) implies

$$\begin{aligned} \text{Var}_{\mu_{J,h}}(\langle x, \theta \rangle) &\leq \alpha(0) + \int_0^T \alpha^2(T-t) dt \\ &\leq \alpha(0) + \int_0^T \alpha^2(t) dt \end{aligned}$$

Taking the supremum over J, h and θ such that $|J| = T$ on the left hand-side, we get the differential inequation

$$\alpha(T) \leq \alpha(0) + \int_0^T \alpha^2(t) dt$$

If we call $y(T) = \alpha(0) + \int_0^T \alpha^2(t) dt$ this rewrites

$$\begin{aligned} \sqrt{y'} &\leq y \\ \implies \frac{y'}{y^2} &\leq 1 \\ \implies d\left(\frac{1}{y}\right) &\geq -1 \\ \implies y(T) &\leq \frac{y(0)}{1 - Ty(0)}, \quad \forall T \leq \frac{1}{y(0)} \end{aligned}$$

But $y(0) = \alpha(0) = 1$, thus we get for all $0 \leq T \leq 1$,

$$\alpha(T) \leq y(T) \leq \frac{1}{1-T}$$

which is (i).

Now, proving (ii) is easy. For an arbitrary function f we get from (5.1.2) that

$$\begin{aligned} \frac{d}{dt} \mathbb{E} \text{Var}_{\nu_t}(f) &= -\mathbb{E} d[M]_t \\ &\geq -\mathbb{E} |A_t|_{op} \mathbb{E} \text{Var}_{\nu_t}(f) \\ &\geq -\alpha(T-t) \mathbb{E} \text{Var}_{\nu_t}(f) \end{aligned}$$

This integrates into

$$\begin{aligned} \text{Var}_{\nu}(f) &\leq e^{\int_0^T \alpha(T-t) dt} \mathbb{E} \text{Var}_{\nu_T}(f) \\ &\leq \frac{1}{1+T} \mathbb{E} \text{Var}_{\nu_T}(f) \\ &\leq \frac{1}{1+T} \sup_h c_P^2(\mu_{0,h}) \mathbb{E} [\mathbb{E}_{\nu_T} |\nabla f|^2] \\ &\leq \frac{1}{1+T} \sup_h c_P^2(\mu_{0,h}) \mathbb{E}_{\nu} |\nabla f|^2 \end{aligned}$$

where we used the fact that $\nu_T = \mu_{0,h+h_T}$ and the martingale property in the last line. The proof for the log-Sobolev constant is completely analogous.

5.2 Relationship with the multiscale Bakry-Émery criterion

In 2020, Bodineau and Bauerschmidt introduced a new criterion to derive log-Sobolev inequalities, which they called the multiscale Bakry-Emery criterion. We give an overview of their approach and explain the link with the stochastic localization. This connection has been described by Shenfeld [78], we give an alternate point of view.

The multiscale Bakry-Emery criterion in a nutshell

Let ν be a measure on, say \mathbb{R}^n , with density

$$d\nu = e^{-V_0(x) - Ax \cdot x}$$

where A is a positive symmetric matrix. The usual Bakry-Emery criterion asserts that if V_0 is convex, then $\rho_{LS}^2(\nu) \leq \frac{1}{\lambda_{\min}(A)}$. The proof consists in analyzing the time-homogeneous semi-group associated to ν . In contrast, Bodineau and Bauerschmidt uses a time dependent semi-group, which we now describe. For $t \geq 0$, we define

$$\dot{C}_t = Q_t^2 = e^{-tA}, \quad C_t = \int_0^t \dot{C}_s ds \quad (5.2.1)$$

Notice that $C_{\infty} = A^{-1}$. For a positive symmetric matrix B we write γ_B for the centered Gaussian with covariance B . Now we define

- the renormalised potential V_t by

$$e^{-V_t} = e^{-V_0} * \gamma_{C_t} \quad (5.2.2)$$

- An operator Q_t by

$$Q_t(f)(x) = e^{V_t(x)} (e^{-V_0} f * \gamma_{C_t})(x) \quad (5.2.3)$$

- A renormalised measure ν_t such that for any function f

$$\mathbb{E}_\nu(f) = \mathbb{E}_{\nu_t}(Q_t f) \quad (5.2.4)$$

We have the explicit expression for ν_t : $\nu_t = \frac{1}{Z} \gamma_{C_\infty - C_t} e^{-V_t}$. Notice that $\nu_0 = \nu$ and $\nu_\infty = \delta_0$

Now using the same approach as Bakry-Emery, given a function f , Bodineau and Bauerschmidt analyze the dynamics of $t \rightarrow \mathbb{E}_{\nu_t}(\Phi(Q_t f))$ where $\Phi(x) = x \log x$. They prove the following

Theorem 5.2. *Under mild assumptions on ν , if there are numbers $\dot{\mu}_t$ such that*

$$Q_t \nabla^2 V_t Q_t \geq \dot{\mu}_t Id$$

then

$$\rho_{LS}(\nu)^2 \leq \int_0^\infty e^{-t\lambda_{\min}(A) - 2\mu_t} dt$$

where $\mu_t = \int_0^t \dot{\mu}_s ds$.

Relationship with the stochastic localization process

To simplify the exposition, we consider the special case $A = Id$, so that $C_t = (1 - e^{-t})I$. We make the change of variable $r = 1 - e^{-t}$. By a slight abuse, we keep the same notations as previously. We thus work on $[0, 1]$ instead of \mathbb{R}_+ , and we have

-

$$e^{-V_r} = e^{-V_0} * \gamma_r = P_r(V_0). \quad (5.2.5)$$

-

$$Q_r(f)(x) = e^{V_r(x)} (e^{-V_0} f * \gamma_r)(x) = \frac{P_r(f e^{-V_0})}{P_r(e^{-V_0})}. \quad (5.2.6)$$

-

$$\mathbb{E}_\nu(f) = \mathbb{E}_{\nu_r}(Q_r f). \quad (5.2.7)$$

We have the explicit expression for ν_r : $\nu_r = \frac{1}{Z} \gamma_{1-r} e^{-V_r}$. Notice that $\nu_0 = \nu$ and $\nu_1 = \delta_0$

And Theorem 5.2 rewrites

Theorem 5.3. *Under mild assumptions on ν , if there are numbers $\dot{\mu}_r$ such that*

$$\nabla^2 V_r \geq \dot{\mu}_r,$$

then,

$$\rho_{LS}(\nu)^2 \leq \int_0^1 e^{-2\mu_r} dr$$

where $\mu_r = \int_0^r \dot{\mu}_s ds$.

Notice that if V_0 is convex, then so is V_r and we recover the fact that $\rho_{LS}(\nu)^2 \leq 1$. Now the operator Q_r that we defined by (5.2.6) is much like the operator Q_t^μ that we introduced for a measure μ in the functional-analytic description of the stochastic localization process in section 2.7. The only difference is that e^{-V_0} is not the density of ν , thus Q is not exactly Q^ν , however it is not too difficult to pass from one to the other. For a function f we write for $0 \leq r \leq 1$,

$$\begin{aligned} \int f d\nu &= \int f e^{-V_0} e^{-\frac{|x|^2}{2}} dx \\ &= \int f e^{-V_0} P_r \left(e^{-\frac{|x|^2}{2(1-r)}} \right) dx \\ &= \int Q_r^\rho(f) e^{-\frac{|x|^2}{2(1-r)}} P_r(e^{-V_0}) dx \end{aligned} \quad (5.2.8)$$

This is the equality (5.2.7) defining ν_r . However, we have not gained anything applying this trick. Indeed, for a function f , $x_0 \in \mathbb{R}^n$ and $s \geq 0$, expanding and refactoring the Gaussian factor shows that

$$\begin{aligned} P_s(f e^{-V_0 - \frac{|x|^2}{2}}) &= \frac{1}{Z} \int f(y) e^{-V_0(y)} e^{-\frac{y^2}{2} - \frac{|x_0 - y|^2}{2s}} \\ &= \frac{1}{(1+s)^n} e^{-\frac{|x_0|^2}{2(1+s)}} P_{\frac{s}{1+s}}(e^{-V_0}) \left(\frac{x_0}{1+s} \right) \end{aligned} \quad (5.2.9)$$

In particular, we get that for any $s \geq 0$ and any $x \geq 0$

$$Q_s^\nu f((1+s)x) = Q_{\frac{s}{1+s}}^\rho(x) \quad (5.2.10)$$

Thus, working with Q^ρ or Q^μ makes no difference. To summarize, in the functional analytic point of view on stochastic localization we write

$$\int f d\nu = \int Q_s^\mu(f) d(P_s \nu) \quad \text{for } s \geq 0,$$

while in the Bodineau Bauerschmidt framework, we write:

$$\int f d\nu = \int Q_r^\rho(f) e^{-\frac{|x|^2}{2(1-r)}} P_r(e^{-V_0}) dx \quad \text{for } 0 \leq r \leq 1.$$

The equality between the two right-hand sides is obtained by setting $r = \frac{s}{1+s}$ and using (5.2.10) and (5.2.9).

Finally, we are interested in reformulating Theorem 5.3 in terms of quantities relevant in the study of the localization process. The task is to compute

$$\nabla^2 V_r = \nabla^2(-\log P_r(e^{-V_0}))$$

A straightforward but slightly tedious computation shows that for all $y \in \mathbb{R}^n$

$$(\nabla^2 V_r)(y) = \frac{1}{t} Id - \frac{1}{t^2} \text{Cov} \left(\frac{1}{Z} \nu e^{-\frac{1-r}{2r}|x|^2 + \frac{1-r}{r}x \cdot y} \right) \quad (5.2.11)$$

where Z is a normalising constant. We introduce

$$\chi(t) = \sup_{h \in \mathbb{R}^n} |\text{Cov}(\frac{1}{Z} e^{-\frac{t}{2}|x|^2 + h \cdot x})|_{op}$$

Then, by (5.2.11), the best uniform bound on $\nabla^2 V_r$ is

$$\nabla^2 V_r \geq \frac{1}{r} - \frac{\chi(\frac{1-r}{r})}{r^2}$$

We set

$$\chi(t) = \frac{1}{1+t} + \lambda(t)$$

Notice that if V is convex, then $\lambda(t) \leq 0$ for all $t \geq 0$, thus λ measures in a sense the lack of convexity. The bound now reads

$$\nabla^2 V_r \geq \frac{-\lambda(\frac{1-r}{r})}{r^2} \quad (5.2.12)$$

For $t \geq 0$ we define

$$\Lambda(t) = \int_0^t \lambda(s) ds$$

and we assume that the limit $\Lambda(\infty)$ exists. Now we compute, for $r \geq 0$

$$\begin{aligned} \int_0^r \frac{-\lambda(\frac{1-s}{r})}{s^2} ds &= \int_0^r \frac{d}{ds} \left(\Lambda \left(\frac{1-s}{r} \right) \right) ds \\ &= \Lambda \left(\frac{1-r}{r} \right) - \Lambda(\infty) \end{aligned}$$

Plugging this into Theorem 5.3 yields after a change of variable

$$\rho_{LS}^2(\nu) \leq e^{2\Lambda(\infty)} \int_0^{+\infty} \frac{e^{-2\Lambda(u)}}{(1+u)^2} du \quad (5.2.13)$$

It can be checked that the bound gives an equality when ν is a Gaussian. Note that asking for $\Lambda(\infty)$ to exist is not very restrictive. For instance, if ν satisfies

$$-\nabla^2(\log(\frac{d\nu}{dx})) \geq -M$$

for a positive number M , then for $T \geq M$ we get

$$\chi(T) \leq \frac{1}{T-M} \quad (5.2.14)$$

which ensures that the limit exists.

Now we explain how to recover a variant of (5.2.13) using the stochastic localization process directly. We keep the same notations, and we do assume the existence of an $M > 0$ such that $-\nabla^2(\log(\frac{d\nu}{dx})) \geq -M$.

Theorem 5.4. *With the same notations as before*

$$\rho_{LS}^2(\nu) \leq \lim_{T \rightarrow \infty} \frac{e^{\int_0^T \chi(s) ds}}{T} = e^{\Lambda(\infty)}$$

.

Yet again, one can check that this is an equality for Gaussians. It is not clear how to compare this with (5.2.13). As a first step, notice that one implication of Theorem 5.4 is the bound

$$\chi(t) \geq \frac{1}{\chi(0) + t} \quad (5.2.15)$$

with equality when ν is a Gaussian. Otherwise we would get a contradiction by applying the result to a tilt of ν that almost realizes $\chi(0)$. As a consequence we get that for all $u \geq 0$

$$e^{\Lambda(u)} \geq \frac{\chi(0) + u}{\chi(0)(1+u)} \quad (5.2.16)$$

In particular, plugging this into (5.2.13) yields

$$\rho_{LS}^2(\nu) \leq \chi(0) e^{2\Lambda(\infty)}. \quad (5.2.17)$$

However, it can be seen that this new bound is weaker than Theorem 5.4, since by (5.2.16), $e^{\Lambda(\infty)} \geq \frac{1}{\chi(0)}$.

We move to the proof of Theorem 5.4. From now on ν_t refers to the stochastic localization of ν , and not the renormalised measure. Let f be a positive function, and denote

by $M_t = \int f d\nu_t$. Since M_t is a martingale, we have the following decomposition of the entropy:

$$\text{Ent}_\nu(f) = \mathbb{E}\text{Ent}_{\nu_t}(f) + \mathbb{E} \int_0^t \frac{d[M]_s}{2M_s}.$$

Furthermore, we proved in the previous chapter

$$\mathbb{E} \frac{d[M]_t}{2M_t} \leq \chi(T) \mathbb{E}\text{Ent}_{\nu_t}(f).$$

Thus, we get that

$$\frac{d}{dt} \mathbb{E}\text{Ent}_{\nu_t}(f) \geq -\chi(T) \mathbb{E}\text{Ent}_{\nu_t}(f),$$

which integrates to

$$\text{Ent}_\nu(f) \leq e^{\int_0^T \chi(s) ds} \mathbb{E}\text{Ent}_{\nu_T}(f) \quad (5.2.18)$$

Now, since we have assumed that $-\nabla^2(\log(\frac{d\nu}{dx})) \geq -M$, for all times $T > M$, the measure ν_T is almost surely $T - M$ strongly log-concave, thus for all $T > M$

$$\begin{aligned} \text{Ent}_\nu(f) &\leq e^{\int_0^T \chi(s) ds} \mathbb{E}\text{Ent}_{\nu_T}(f) \\ &\leq \frac{e^{\int_0^T \chi(s) ds}}{T - M} \mathbb{E}\mathbb{E}_{\nu_T} \left(\frac{|\nabla f|^2}{f} \right) \\ &\leq \frac{e^{\int_0^T \chi(s) ds}}{T - M} \mathbb{E}_\nu \left(\frac{|\nabla f|^2}{f} \right) \end{aligned}$$

Thus, letting T go to infinity, we get

$$\begin{aligned} \rho_{LS}^2(\nu) &\leq \lim_{T \rightarrow \infty} \frac{e^{\int_0^T \chi(s) ds}}{T - M} \\ &= \lim_{T \rightarrow \infty} \frac{1 + T}{T - M} e^{\int_0^T \chi(s) - \frac{1}{1+s} ds} \\ &= e^{\int_0^\infty \lambda(s) ds} = e^{\Lambda(\infty)} \end{aligned}$$

Chapter 6

Positive solutions for large random linear systems

This chapter is based on

Pierre Bizeul and Jamal Najim. Positive solutions for large random linear systems. *Proceedings of the American Mathematical Society*, 149(6):2333–2348, 2021

6.1 Introduction

Denote by A_n an $n \times n$ matrix with independent Gaussian $\mathcal{N}(0, 1)$ entries and by α_n a positive sequence. We are interested in the componentwise positivity of the $n \times 1$ vector \mathbf{x}_n , solution of the linear system

$$\mathbf{x}_n = \mathbf{1}_n + \frac{1}{\alpha_n \sqrt{n}} A_n \mathbf{x}_n, \quad (6.1.1)$$

where $\mathbf{1}_n$ is the $n \times 1$ vector with components 1.

It is well-known since Geman [42] that the limsup of the spectral radius of $\frac{A_n}{\sqrt{n}}$ is almost surely (a.s.) ≤ 1 , so that matrix $\left(I_n - \frac{A_n}{\alpha_n \sqrt{n}}\right)$ is eventually invertible as long as $\alpha_n \gg 1$. In this case, vector $\mathbf{x}_n = (x_k)_{k \in [n]}$, where we denote by $[n] = \{1, \dots, n\}$, is

$$\mathbf{x}_n = \left(I_n - \frac{A_n}{\alpha_n \sqrt{n}}\right)^{-1} \mathbf{1}_n \quad \text{with} \quad x_k = \mathbf{e}_k^* \left(I_n - \frac{A_n}{\alpha_n \sqrt{n}}\right)^{-1} \mathbf{1}_n,$$

where \mathbf{e}_k is the $n \times 1$ canonical vector and B^* is the transconjugate of B (or simply its transpose if B is real).

The positivity of the x_k 's is a key issue in the study of Large Lotka-Volterra (LV) systems, widely used in mathematical biology and ecology to model populations with interactions.

Consider for instance a given foodweb and denote by $\mathbf{x}_n(t) = (x_k(t))_{k \in [n]}$ the vector

of abundances of the various species within the foodweb at time t . A standard way to connect the various abundances is via a LV system of equations

$$\frac{dx_k(t)}{dt} = x_k(t) \left(1 - x_k(t) + \frac{1}{\alpha_n \sqrt{n}} \sum_{\ell \in [n]} A_{k\ell} x_\ell(t) \right) \quad \text{for } k \in [n], \quad (6.1.2)$$

where the interactions $(A_{k\ell})$ can be modeled as random in the absence of any prior information. Here, the $A_{k\ell}$'s are assumed to be i.i.d. $\mathcal{N}(0, 1)$. At the equilibrium $\frac{d\mathbf{x}_n}{dt} = 0$, the abundance vector \mathbf{x}_n is a solution of (6.1.1) and a key issue is the existence of a *feasible* solution, that is a solution \mathbf{x}_n where all the x_k 's are positive. Dougoud et al. [33] based on Geman et al. [43] proved that a feasible solution is very unlikely to exist if $\alpha_n \equiv \alpha$ is a constant. In fact, the CLT proved in [43] asserts that for any fixed number M of components

$$(x_k - 1)_{k \in [M]} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Z \sim \mathcal{N}(0, \sigma_\alpha^2 I_M),$$

where $\xrightarrow{\mathcal{D}}$ (resp. $\xrightarrow{\mathcal{P}}$) stands for the convergence in distribution (resp. in probability) and where $\sigma_\alpha^2 = \mathcal{O}(1)$. As an important consequence, vectors \mathbf{x}_n with positive components will become extremely rare since

$$\mathbb{P}\{x_k > 0, k \in [M]\} \xrightarrow[n \rightarrow \infty]{} \left(\int_{-\frac{1}{\sigma_\alpha}}^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \right)^M \Rightarrow \mathbb{P}\{x_k > 0, k \in [n]\} \xrightarrow[n \rightarrow \infty]{} 0.$$

In this article, we consider a growing scaling factor $\alpha_n \rightarrow \infty$ and study the positivity of \mathbf{x}_n 's components in relation with α_n .

We find that there exists a critical threshold $\alpha_n^* = \sqrt{2 \log n}$ below which feasible solutions hardly exist and above which feasible solutions are more and more likely to exist. More precisely, we prove the following:

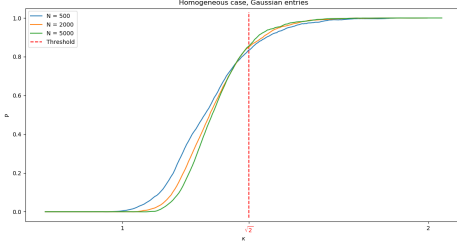
Theorem 6.1 (Feasibility). *Let $\alpha_n \xrightarrow[n \rightarrow \infty]{} \infty$ and denote by $\alpha_n^* = \sqrt{2 \log n}$. Let $\mathbf{x}_n = (x_k)_{k \in [n]}$ be the solution of (6.1.1).*

1. *If $\exists \varepsilon > 0$ such that $\alpha_n \leq (1 - \varepsilon)\alpha_n^*$ then $\mathbb{P}\{\min_{k \in [n]} x_k > 0\} \xrightarrow[n \rightarrow \infty]{} 0$,*
2. *If $\exists \varepsilon > 0$ such that $\alpha_n \geq (1 + \varepsilon)\alpha_n^*$ then $\mathbb{P}\{\min_{k \in [n]} x_k > 0\} \xrightarrow[n \rightarrow \infty]{} 1$.*

We illustrate the transition toward feasibility In Figure 6.1.

Remark : Proof of Theorem 6.1 is based on an analysis of the order of magnitude of the extreme values of the x_k 's, which heavily relies on sub-Gaussianness of Lipschitz functionals with Gaussian entries. This property remains true if the A_{ij} 's satisfies a logarithmic sobolev inequality - details are provided in Section 6.4. The case of discrete entries remains open although simulations (see Figure 6.1 (B)) indicate that a similar phase transition occurs.

Remark : Notice that $\frac{1}{\alpha_n^*}$ goes to zero extremely slowly. For modeling purposes, the threshold $\sigma_n^* := \frac{1}{\alpha_n^*}$ acts as an n -dependent upper bound of the standard deviation of the entries of $(\alpha_n^{-1}A_n)$, under which feasibility occurs.



(A) Gaussian entries.

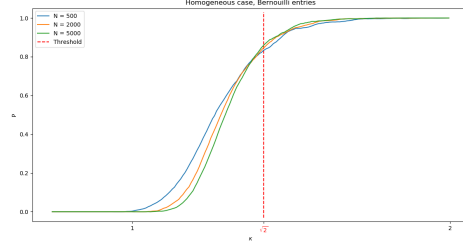
(B) Bernoulli ± 1 entries.

FIGURE 6.1: Transition toward feasibility. We consider different values of n , respectively 1000 (dashed line), 4000 (solid line). For each n and each κ on the x -axis, we simulate 10000 $n \times n$ matrices A_n and compute the solution \mathbf{x}_n of (6.1.1) at the scaling $\alpha_n(\kappa) = \kappa\sqrt{\log(n)}$. Each curve represents the proportion of feasible solutions \mathbf{x}_n obtained for 10000 simulations. The dotted vertical line corresponds to the critical scaling $\alpha_n^* = \sqrt{2\log(n)}$ for $\kappa = \sqrt{2}$.

To complement the picture, we provide the following heuristics at the critical scaling $\alpha_n^* = \sqrt{2\log n}$:

$$\mathbb{P}\left\{\min_{k \in [n]} x_k > 0\right\} \approx 1 - \sqrt{\frac{e}{4\pi \log n}} + \frac{e}{8\pi \log n} \quad \text{as } n \rightarrow \infty. \quad (6.1.3)$$

Aside from the question of feasibility arises the question of *stability*: for a large complex system, that is a system of coupled differential equations describing the time evolution of the abundances of the various species of a given foodweb, how likely a perturbation of the solution \mathbf{x} will return to the equilibrium? Gardner and Ashby [41] considered stability issues of complex systems connected at random. Based on the circular law for large matrices with i.i.d. entries, May [70] provided a complexity/stability criterion and motivated the systematic use of large random matrix theory in the study of foodwebs, see for instance Allesina et al. [1]. Recently, Stone [79] and Gibbs *et al.* [45] revisited the relation between feasibility and stability.

We complement the result of Theorem 6.1 by addressing the question of stability in the context of a LV system (6.1.2) and prove that under the second condition of the theorem feasibility and stability occur simultaneously.

Recall that the solution at equilibrium \mathbf{x}_n is stable if the Jacobian matrix \mathcal{J} of the Lotka-Volterra system evaluated at \mathbf{x}_n , that is

$$\mathcal{J}(\mathbf{x}_n) = \text{diag}(\mathbf{x}_n) \left(-I_n + \frac{A_n}{\alpha_n \sqrt{n}} \right) \quad (6.1.4)$$

has all its eigenvalues with negative real part.

Corollary 6.2 (Stability). *Let $\mathbf{x}_n = (x_k)_{k \in [n]}$ be the solution of (6.1.1). Assume that $\ell^+ := \limsup_{n \rightarrow \infty} \frac{\sqrt{2 \log n}}{\alpha_n} < 1$. Denote by \mathcal{S}_n the spectrum of $\mathcal{J}(\mathbf{x}_n)$. Then*

$$\max_{\lambda \in \mathcal{S}_n} \min_{k \in [n]} |\lambda + x_k| \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0 \quad \text{and} \quad \max_{\lambda \in \mathcal{S}_n} \operatorname{Re}(\lambda) \leq -(1 - \ell^+) + o_P(1). \quad (6.1.5)$$

Proof of Corollary 6.2 relies on standard perturbation results from linear algebra and on Theorem 6.1.

Organization of the paper

Theorem 6.1 is proved in Section 6.2, Corollary 6.2 in Section 6.3. In Section 6.4, elements supporting heuristics (6.1.3) are provided, together with extensions to non-homogeneous systems (where vector $\mathbf{1}_n$ in (6.1.1) is replaced by a generic deterministic vector \mathbf{r}_n) and non-Gaussian entries.

Acknowledgments

JN thanks Christian Mazza for introducing him to the study of large LV systems in theoretical ecology. The authors thank François Massol and Olivier Guédon for fruitful discussions, and also the referees, whose comments have improved the writing of the article.

6.2 Positive solutions: proof of Theorem 6.1

We will use the following notations for the various norms at stake: if \mathbf{v} is a vector then $\|\mathbf{v}\|$ stands for its euclidian norm; if A is a matrix then $\|A\|$ stands for its spectral norm and $\|A\|_F = \sqrt{\sum_{ij} |A_{ij}|^2}$ for its Frobenius norm. Let φ be a function from $\Sigma = \mathbb{R}$ or \mathbb{C} to \mathbb{C} then $\|\varphi\|_\infty = \sup_{x \in \Sigma} |\varphi(x)|$.

Some preparation and strategy of the proof

Denote by $Q_n = \left(I_n - \frac{A_n}{\alpha_n \sqrt{n}}\right)^{-1}$ the resolvent and by $s(B)$ the largest singular value of a given matrix B . Then it is well known that almost surely $s_n := s(n^{-1/2} A_n) \xrightarrow[n \rightarrow \infty]{} 2$ (see for instance [7, Chapter 5]) hence $s\left(\frac{1}{\alpha_n \sqrt{n}} A_n\right) \xrightarrow[n \rightarrow \infty]{} 0$. In particular, the solution

$$\mathbf{x}_n = (x_k)_{k \in [n]} = \left(I_n - \frac{A_n}{\alpha_n \sqrt{n}}\right)^{-1} \mathbf{1}_n = Q_n \mathbf{1}_n,$$

with I_n the $n \times n$ identity, is uniquely defined almost surely for all n large. In order to study the minimum of \mathbf{x}_n 's components, we partially unfold the above resolvent (in the

sequel, we will simply denote $A, \alpha, \mathbf{1}, Q$ instead of $A_n, \alpha_n, \mathbf{1}_n, Q_n$) and write:

$$\begin{aligned} x_k &= \mathbf{e}_k^* \mathbf{x} = \mathbf{e}_k^* Q \mathbf{1} = \sum_{\ell=0}^{\infty} \mathbf{e}_k^* \left(\frac{A}{\alpha \sqrt{n}} \right)^\ell \mathbf{1}, \\ &= 1 + \frac{1}{\alpha} \mathbf{e}_k^* \left(n^{-1/2} A \right) \mathbf{1} + \frac{1}{\alpha^2} \mathbf{e}_k^* \left(n^{-1/2} A \right)^2 Q \mathbf{1} = 1 + \frac{1}{\alpha} Z_k + \frac{1}{\alpha^2} R_k \end{aligned} \quad (6.2.1)$$

where

$$Z_k = \mathbf{e}_k^* \left(n^{-1/2} A \right) \mathbf{1} = \frac{1}{\sqrt{n}} \sum_{i=1}^n A_{ki} \quad \text{and} \quad R_k = \mathbf{e}_k^* \left(n^{-1/2} A \right)^2 Q \mathbf{1}. \quad (6.2.2)$$

Notice that the Z_k 's are i.i.d. $\mathcal{N}(0, 1)$.

Extreme values of Gaussian random variables

Consider the sequence (Z_k) of standard Gaussian i.i.d. random variables, recall that $\alpha_n^* = \sqrt{2 \log n}$ and let

$$M_n = \max_{k \in [n]} Z_k, \quad \check{M}_n = \min_{k \in [n]} Z_k \quad \text{and} \quad \beta_n^* = \alpha_n^* - \frac{1}{2\alpha_n^*} \log(4\pi \log n). \quad (6.2.3)$$

Denote by $G(x) = e^{-e^{-x}}$ the Gumbel cumulative distribution. Then the following results are standard, (i.e. [63, Theorem 1.5.3]): for all $x \in \mathbb{R}$

$$\mathbb{P} \{ \alpha_n^* (M_n - \beta_n^*) \leq x \} \xrightarrow{n \rightarrow \infty} G(x), \quad \mathbb{P} \{ \alpha_n^* (\check{M}_n + \beta_n^*) \geq -x \} \xrightarrow{n \rightarrow \infty} G(x). \quad (6.2.4)$$

Strategy of the proof

Eq. (6.2.1) immediately yields

$$\begin{cases} \min_{k \in [n]} x_k & \geq 1 + \frac{1}{\alpha} \check{M} + \frac{1}{\alpha^2} \min_{k \in [n]} R_k, \\ \min_{k \in [n]} x_k & \leq 1 + \frac{1}{\alpha} \check{M} + \frac{1}{\alpha^2} \max_{k \in [n]} R_k. \end{cases}$$

We rewrite the first equation as

$$\begin{aligned} \min_{k \in [n]} x_k &\geq 1 + \frac{\alpha_n^*}{\alpha_n} \left(\frac{\check{M} + \beta_n^*}{\alpha_n^*} - \frac{\beta_n^*}{\alpha_n^*} + \frac{\min_{k \in [n]} R_k}{\alpha_n^* \alpha_n} \right) \\ &= 1 + \frac{\alpha_n^*}{\alpha_n} \left(-1 + o_P(1) + \frac{\min_{k \in [n]} R_k}{\alpha_n^* \alpha_n} \right), \end{aligned} \quad (6.2.5)$$

where we have used the fact that $(\alpha_n^*)^{-1}(\check{M} + \beta_n^*) = o_P(1)$. Similarly,

$$\min_{k \in [n]} x_k \leq 1 + \frac{\alpha_n^*}{\alpha_n} \left(-1 + o_P(1) + \frac{\max_{k \in [n]} R_k}{\alpha_n^* \alpha_n} \right).$$

The theorem will then follow from the following lemma.

Lemma 6.3. *The following convergence holds*

$$\frac{\max_{k \in [n]} R_k}{\alpha_n \sqrt{2 \log n}} \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0 \quad \text{and} \quad \frac{\min_{k \in [n]} R_k}{\alpha_n \sqrt{2 \log n}} \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0 .$$

Proof of Lemma 6.3 requires a careful analysis of the order of magnitude of the extreme values of the remaining term $(R_k)_{k \in [n]}$. It is postponed to Section 6.2.

Lipschitz property and tightness of $R_k(A)$

Let $\eta \in (0, 1)$ and $\varphi : \mathbb{R}^+ \rightarrow [0, 1]$ be a smooth (infinitely differentiable) function with values

$$\varphi(x) = \begin{cases} 1 & \text{if } x \in [0, 2 + \eta] \\ 0 & \text{if } x \geq 3 \end{cases} ,$$

and strictly decreasing from 1 to zero as x goes from $2 + \eta$ to 3. Notice in particular that $\|\varphi'\|_\infty$ is finite. Recall that $s_n = s(n^{-1/2}A)$ is the largest singular value of the normalized matrix $n^{-1/2}A$ and denote by

$$\varphi_n := \varphi(s_n) = \varphi\left(s(n^{-1/2}A)\right) .$$

Notice that $\mathbb{P}\{\varphi_n < 1\} = \mathbb{P}\{s_n > 2 + \eta\} \xrightarrow[n \rightarrow \infty]{} 0$ (by the a.s. convergence of s_n to 2). Instead of working with R_k we introduce the truncated quantity:

$$\tilde{R}_k = \varphi_n R_k . \tag{6.2.6}$$

For a given $n \times n$ matrix A , we may consider its $2n \times 2n$ hermitized matrix $\mathcal{H}(A)$ defined as $\mathcal{H}(A) = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$. Notice that the singular values of A together with their negatives are the eigenvalues of $\mathcal{H}(A)$.

Lemma 6.4. *Let \tilde{R}_k be given by (6.2.6), then the function $A \mapsto \tilde{R}_k(A)$ is Lipschitz, i.e.*

$$\left| \tilde{R}_k(A) - \tilde{R}_k(B) \right| \leq K \|A - B\|_F , \tag{6.2.7}$$

where $\|A\|_F$ is the Frobenius norm and K is a constant independent from k and n .

Proof. Notice that $\varphi(s_n) = 0$ and $\varphi'(s_n) = 0$ for $s_n \geq 3$, which implies that one may consider the bound $s_n \leq 3$ in the following computations, for \tilde{R}_k or its derivatives would be zero otherwise. Recall the definition of the resolvent $Q = \left(I - \frac{A}{\alpha\sqrt{n}}\right)^{-1}$ then $Q^{-1}Q = I$ which yields $Q = I + \frac{A}{\alpha\sqrt{n}}Q$ from which we deduce that

$$\varphi_n \|Q\| \leq \varphi_n \left(1 - \frac{1}{\alpha} \|n^{-1/2}A\|\right)^{-1} \leq \frac{1}{1 - 3\alpha^{-1}} \leq 3 \tag{6.2.8}$$

for n large enough.

We first consider a matrix A such that $\mathcal{H}(A)$ has simple spectrum (i.e. with $2n$ distinct eigenvalues, each with multiplicity 1). We denote by $\partial_{ij} = \frac{\partial}{\partial A_{ij}}$ and prove that the vector $\nabla \tilde{R}_k(A) = \left(\partial_{ij} \tilde{R}_k(A), i, j \in [n] \right)$ satisfies

$$\|\nabla \tilde{R}_k(A)\| = \sqrt{\sum_{ij} \left| \partial_{ij} \tilde{R}_k(A) \right|^2} \leq K. \quad (6.2.9)$$

We may occasionally drop the dependence of \tilde{R}_k in A . We begin by computing

$$\begin{aligned} \partial_{ij} \tilde{R}_k &= \lim_{h \rightarrow 0} \frac{\tilde{R}_k(A + h \mathbf{e}_i \mathbf{e}_j^*) - \tilde{R}_k(A)}{h}, \\ &= (\partial_{ij} \varphi_n) R_k + \varphi_n \mathbf{e}_k^* \left(\partial_{ij} \left(n^{-\frac{1}{2}} A \right)^2 \right) Q \mathbf{1} + \varphi_n \mathbf{e}_k^* \left(n^{-\frac{1}{2}} A \right)^2 (\partial_{ij} Q) \mathbf{1} \\ &=: T_{1,ij} + T_{2,ij} + T_{3,ij}. \end{aligned}$$

Straightforward computations yield

$$\partial_{ij} \left(n^{-\frac{1}{2}} A \right)^2 = \frac{1}{n} (A \mathbf{e}_i \mathbf{e}_j^* + \mathbf{e}_i \mathbf{e}_j^* A) \quad \text{and} \quad \partial_{ij} Q = \frac{1}{\alpha \sqrt{n}} Q \mathbf{e}_i \mathbf{e}_j^* Q. \quad (6.2.10)$$

It remains to compute $\partial_{ij} \varphi_n = \varphi'(s_n) \partial_{ij} s_n$. Recall that $\mathcal{H}(A)$ has a simple spectrum and notice that $A \mapsto s_n(A)$ is differentiable. In fact, since s_n is simple, it is a simple root of the characteristic polynomial. In particular, it is not a root of its derivative and one can use the implicit function theorem to conclude its differentiability. Let \mathbf{u} and \mathbf{v} be respectively the left and right normalized singular vectors associated to $s(A)$. Then

$$\mathcal{H}(A) \mathbf{w} = s(A) \mathbf{w} \quad \text{with} \quad \mathbf{w} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \quad \text{and} \quad \|\mathbf{w}\|^2 = 2,$$

moreover \mathbf{w} is (up to scaling) the unique eigenvector of $s(A)$ since $s(A)$ is simple by assumption. We now apply [48, Theorem 6.3.12] to compute s_n 's derivative:

$$\partial_{ij} s(A) = \frac{1}{\|\mathbf{w}\|^2} (\mathbf{u}^* \mathbf{e}_i \mathbf{e}_j^* \mathbf{v} + \mathbf{v}^* \mathbf{e}_j \mathbf{e}_i^* \mathbf{u}) = \mathbf{u}^* \mathbf{e}_i \mathbf{e}_j^* \mathbf{v} \quad \text{hence} \quad \partial_{ij} s_n = \frac{1}{\sqrt{n}} \mathbf{u}^* \mathbf{e}_i \mathbf{e}_j^* \mathbf{v} \quad (6.2.11)$$

(recall that all the considered vectors are real). We first handle the term $T_{1,ij}$.

$$\begin{aligned} \sum_{ij} |T_{1,ij}|^2 &= \sum_{ij} \left| \mathbf{u}^* \mathbf{e}_i \mathbf{e}_j^* \varphi'(s_n) \mathbf{e}_k^* \left(n^{-1/2} A \right)^2 Q \frac{\mathbf{1}}{\sqrt{n}} \right|^2, \\ &\leq 3^6 \|\varphi'\|_\infty^2 \sum_i |\mathbf{u}^* \mathbf{e}_i|^2 \sum_j |\mathbf{e}_j^* \mathbf{v}|^2 \leq 3^6 \|\varphi'\|_\infty^2. \end{aligned}$$

We now handle the term $T_{2,ij}$.

$$\begin{aligned}
\sum_{ij} |T_{2,ij}|^2 &= \sum_{ij} \left| \varphi_n e_k^* \left(\frac{A}{\sqrt{n}} e_i e_j^* + e_i e_j^* \frac{A}{\sqrt{n}} \right) Q \frac{1}{\sqrt{n}} \right|^2 \\
&\leq 2\varphi_n^2 \sum_i \left| e_k^* \frac{A}{\sqrt{n}} e_i \right|^2 \sum_j \left| e_j^* Q \frac{1}{\sqrt{n}} \right|^2 + 2\varphi_n^2 \sum_i |e_k^* e_i|^2 \sum_j \left| e_j^* \frac{A}{\sqrt{n}} Q \frac{1}{\sqrt{n}} \right|^2, \\
&= 2\varphi_n^2 \left(e_k^* \frac{A}{\sqrt{n}} \frac{A^*}{\sqrt{n}} e_k \right) \left(\frac{1}{\sqrt{n}} Q^* Q \frac{1}{\sqrt{n}} \right) + 2\varphi_n^2 \left(\frac{1}{\sqrt{n}} Q^* \frac{A^* A}{n} Q \frac{1}{\sqrt{n}} \right) \leq 2^2 \times 3^4.
\end{aligned}$$

The term $T_{3,ij}$ can be handled similarly and one can prove $\sum_{ij} |T_{3,ij}|^2 \leq 3^8$. Gathering all these estimates, we finally obtain the desired bound:

$$\sqrt{\sum_{ij} |\partial_{ij} \tilde{R}_k|^2} \leq \sqrt{3 \sum_{ij} |T_{1,ij}|^2 + 3 \sum_{ij} |T_{2,ij}|^2 + 3 \sum_{ij} |T_{3,ij}|^2} \leq K,$$

where K neither depends on k nor on n .

Having proved a local estimate over $\|\nabla \tilde{R}_k(A)\|$ for each matrix A such that $\mathcal{H}(A)$ has simple spectrum, we now establish the Lipschitz estimate (6.2.7) for two such matrices A, B .

Let A, B such that $\mathcal{H}(A)$ and $\mathcal{H}(B)$ have simple spectrum and consider $A_t = (1 - t)A + tB$ for $t \in [0, 1]$. Notice first that the continuity of the eigenvalues implies that there exists $\delta > 0$ sufficiently small such that $\mathcal{H}(A_t)$ has a simple spectrum for $t \leq \delta$ and $t \geq 1 - \delta$. To go beyond $[0, \delta) \cup (1 - \delta, 1]$ and prove that $\mathcal{H}(A_t)$ has simple spectrum for the entire interval $[0, 1]$ except maybe for a finite number of points, we rely on the argument in Kato [52, Chapter 2.1] which states that apart from a finite number of t_ℓ 's: $t_0 = 0 < t_1 < \dots < t_L < t_{L+1} = 1$, the number of eigenvalues of $\mathcal{H}(A_t)$ remains constant for $t \in [0, 1]$ and $t \neq t_\ell, \ell \in [L]$. Since $\mathcal{H}(A_t)$ has simple spectrum for $t \in [0, \delta) \cup (1 - \delta, 1]$, it has simple spectrum for all $t \notin \{t_\ell, \ell \in [L]\}$.

We can now proceed:

$$\begin{aligned}
|\tilde{R}_k(A_{t_1}) - \tilde{R}_k(A)| &= \left| \lim_{\tau \nearrow t_1} \int_0^\tau \frac{d}{dt} \tilde{R}_k(A_t) dt \right| = \left| \lim_{\tau \nearrow t_1} \int_0^\tau \nabla \tilde{R}_k(A_t) \circ \frac{d}{dt} A_t dt \right|, \\
&\leq \lim_{\tau \nearrow t_1} \int_0^\tau \|\nabla \tilde{R}_k(A_t)\| \times \|B - A\|_F dt \leq K t_1 \|B - A\|_F.
\end{aligned}$$

By iterating this process, we obtain

$$\begin{aligned}
|\tilde{R}_k(B) - \tilde{R}_k(A)| &\leq \sum_{\ell=1}^{L+1} |\tilde{R}_k(A_{t_\ell}) - \tilde{R}_k(A_{t_{\ell-1}})| \\
&\leq \sum_{\ell=1}^{L+1} K(t_\ell - t_{\ell-1}) \|B - A\|_F = K \|B - A\|_F,
\end{aligned}$$

hence the Lipschitz property along the segment $[A, B]$.

The general property follows by density of such matrices in the set of $n \times n$ matrices and by continuity of $A \mapsto \tilde{R}_k(A)$. Let A, B be given and $A_\varepsilon \rightarrow A$ and $B_\varepsilon \rightarrow B$ be such that $\mathcal{H}(A_\varepsilon)$ and $\mathcal{H}(B_\varepsilon)$ have simple spectrum then:

$$\begin{aligned} \left| \tilde{R}_k(B) - \tilde{R}_k(A) \right| &\leq \left| \tilde{R}_k(B_\varepsilon) - \tilde{R}_k(B) \right| + K \|B_\varepsilon - A_\varepsilon\| + \left| \tilde{R}_k(A_\varepsilon) - \tilde{R}_k(A) \right| \\ &\xrightarrow{\varepsilon \rightarrow 0} K \|B - A\|_F. \end{aligned}$$

Proof of Lemma 6.4 is completed. \square

We now use concentration arguments to bound $\mathbb{E} \max_{k \in [n]} (\tilde{R}_k - \mathbb{E} \tilde{R}_k)$.

Proposition 6.5. *Let K be as in Lemma 6.4, then $\mathbb{E} \max_{k \in [n]} (\tilde{R}_k - \mathbb{E} \tilde{R}_k) \leq K \sqrt{2 \log n}$.*

Proof. By applying Tsirelson-Ibragimov-Sudakov inequality [23, Theorem 5.5] to $\tilde{R}_k(A)$, we obtain the following exponential estimate:

$$\mathbb{E} e^{\lambda(\tilde{R}_k(A) - \mathbb{E} \tilde{R}_k(A))} \leq e^{\frac{\lambda^2 K^2}{2}} \quad \forall \lambda \in \mathbb{R}.$$

We now estimate the expectation of the maximum (we drop the dependence in A).

$$\begin{aligned} \exp \left(\lambda \mathbb{E} \max_{k \in [n]} (\tilde{R}_k - \mathbb{E} \tilde{R}_k) \right) &\leq \mathbb{E} \exp \left(\lambda \max_{k \in [n]} (\tilde{R}_k - \mathbb{E} \tilde{R}_k) \right) \\ &\leq \sum_{k=1}^n \mathbb{E} e^{\lambda(\tilde{R}_k - \mathbb{E} \tilde{R}_k)} \leq n e^{\frac{\lambda^2 K^2}{2}}. \end{aligned}$$

Hence for $\lambda > 0$,

$$\mathbb{E} \max_{k \in [n]} (\tilde{R}_k - \mathbb{E} \tilde{R}_k) \leq \frac{\log n}{\lambda} + \frac{\lambda K^2}{2} =: \Phi(\lambda).$$

Optimizing in λ , we get $\lambda^* = \frac{\sqrt{2 \log n}}{K}$ and $\Phi(\lambda^*) = K \sqrt{2 \log n}$, the desired estimate. \square

Proposition 6.6. *We have $\mathbb{E} \tilde{R}_k(A_n) = \mathcal{O}(1)$ uniformly in $k \in [n]$.*

Proof. By exchangeability, we have $\mathbb{E} \tilde{R}_k = \frac{1}{n} \sum_{i=1}^n \mathbb{E} \tilde{R}_i$ and

$$\left| \frac{1}{n} \sum_{i=1}^n \mathbb{E} \tilde{R}_i \right| = \left| \frac{1}{n} \mathbb{E} \varphi_n \mathbf{1}^* \left(\frac{A}{\sqrt{n}} \right)^2 Q \mathbf{1} \right| \leq \left\| \frac{\mathbf{1}}{\sqrt{n}} \right\|^2 \mathbb{E} \left(\varphi_n \left\| \frac{A}{\sqrt{n}} \right\|^2 \|Q\| \right) = \mathcal{O}(1).$$

by (6.2.8). Proof of Proposition 6.6 is completed. \square

We are now in position to prove Lemma 6.3.

Proof of Lemma 6.3

Since the $\tilde{R}_i(A)$'s are exchangeable, $\mathbb{E}\tilde{R}_k(A) = \mathbb{E}\tilde{R}_1(A)$. Notice that $\max_{k \in [n]} \tilde{R}_k(A) - \tilde{R}_1(A)$ is nonnegative hence by Markov inequality,

$$\begin{aligned} \mathbb{P} \left\{ \frac{\max_{k \in [n]} \tilde{R}_k(A) - \tilde{R}_1(A)}{\alpha\sqrt{2\log n}} \geq \varepsilon \right\} &\leq \frac{\mathbb{E} \left(\max_{k \in [n]} \tilde{R}_k(A) - \tilde{R}_1(A) \right)}{\varepsilon\alpha\sqrt{2\log n}} \\ &= \frac{\mathbb{E} \left(\max_{k \in [n]} \left(\tilde{R}_k(A) - \mathbb{E}\tilde{R}_k(A) + \mathbb{E}\tilde{R}_1(A) \right) - \tilde{R}_1(A) \right)}{\varepsilon\alpha\sqrt{2\log n}}, \\ &= \frac{\mathbb{E} \left(\max_{k \in [n]} \left(\tilde{R}_k(A) - \mathbb{E}\tilde{R}_k(A) \right) \right)}{\varepsilon\alpha\sqrt{2\log n}} \leq \frac{K}{\varepsilon\alpha} \end{aligned}$$

by Proposition 6.5. This implies that

$$\frac{\max_{k \in [n]} \tilde{R}_k(A) - \tilde{R}_1(A)}{\alpha\sqrt{2\log n}} \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0. \quad (6.2.12)$$

We now prove that

$$\tilde{R}_1(A) / \left(\alpha\sqrt{2\log n} \right) \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0. \quad (6.2.13)$$

By Proposition 6.6, $\mathbb{E}\tilde{R}_1(A) = \mathcal{O}(1)$ hence $\mathbb{E}\tilde{R}_1(A)/(\alpha\sqrt{2\log(n)}) \rightarrow 0$. Applying Poincaré's inequality (cf. [23, Theorem 3.20] and its extension to Lipschitz functionals on p. 73) to the Lipschitz functional $A \mapsto \tilde{R}_1(A)$ (cf. Lemma 6.4), we can bound $\tilde{R}_1(A)$'s variance by K^2 and obtain

$$\mathbb{P} \left(\left| \frac{\tilde{R}_1(A) - \mathbb{E}\tilde{R}_1(A)}{\alpha\sqrt{2\log n}} \right| > \delta \right) \leq \frac{\text{var}(\tilde{R}_1(A))}{2\delta^2\alpha^2\log n} \leq \frac{K^2}{2\delta^2\alpha^2\log n} \xrightarrow[n \rightarrow \infty]{} 0.$$

This and Proposition 6.6 yield (6.2.13). Combining (6.2.12) and (6.2.13) finally yields:

$$\frac{\max_{k \in [n]} \tilde{R}_k(A)}{\alpha\sqrt{2\log n}} \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0.$$

In order to obtain the result for the untilded quantities, we write

$$\begin{aligned} \mathbb{P} \left\{ \left| \frac{\max_k R_k(A)}{\alpha\sqrt{2\log n}} \right| > \varepsilon \right\} &\leq \mathbb{P} \left\{ \max_k R_k(A) \neq \max_k \tilde{R}_k(A) \right\} + \mathbb{P} \left\{ \left| \frac{\max_k \tilde{R}_k(A)}{\alpha\sqrt{2\log n}} \right| > \frac{\varepsilon}{2} \right\}, \\ &= \mathbb{P}\{\varphi_n < 1\} + \mathbb{P} \left\{ \left| \frac{\max_k \tilde{R}_k(A)}{\alpha\sqrt{2\log n}} \right| > \varepsilon/2 \right\} \xrightarrow[n \rightarrow \infty]{} 0. \end{aligned}$$

One proves the second assertion similarly. This concludes the proof of Lemma 6.3.

6.3 Stability: proof of Corollary 6.2

In order to study the stability of large Lotka-Volterra systems, we are led to study the matrix

$$\mathcal{J}(\mathbf{x}_n) = \text{diag}(\mathbf{x}_n) \left(-I_n + \frac{A_n}{\alpha_n \sqrt{n}} \right).$$

We first establish the following estimates

$$\begin{cases} \min_{k \in [n]} x_k \geq 1 - \ell^+ - o_P(1), \\ \max_{k \in [n]} x_k \leq 1 + \ell^+ + o_P(1). \end{cases} \quad (6.3.1)$$

The first estimate immediately follows from (6.2.5) together with Lemma 6.3. From x_k 's decomposition (6.2.1) we have

$$\begin{aligned} \max_{k \in [n]} x_k &\leq 1 + \frac{M_n}{\alpha_n} + \frac{\max_{k \in [n]} R_k}{\alpha_n^2} = 1 + \frac{\alpha_n^*}{\alpha_n} \left(\frac{M_n - \beta_n^*}{\alpha_n^*} + \frac{\beta_n^*}{\alpha_n^*} + \frac{\max_{k \in [n]} R_k}{\alpha_n^* \alpha_n} \right) \\ &\leq 1 + \ell^+ + o_P(1), \end{aligned}$$

where the last inequality follows from Lemma 6.3 and the fact that $(\alpha_n^*)^{-1} (M_n - \beta_n^*) \xrightarrow{\mathcal{P}} 0$.

We now compare the spectra of matrices $\mathcal{D}(\mathbf{x}_n) = -\text{diag}(\mathbf{x}_n)$ and $\mathcal{J}(\mathbf{x}_n)$ by relying on Bauer and Fike's theorem [48, Theorem 6.3.2]: for every $\lambda \in \mathcal{S}_n$, there exists a component x_k of vector \mathbf{x}_n such that

$$\begin{aligned} |\lambda + x_k| &\leq \left\| \text{diag}(\mathbf{x}_n) \frac{A_n}{\alpha_n \sqrt{n}} \right\| \leq \frac{1}{\alpha_n} \|\text{diag}(\mathbf{x}_n)\| \left\| \frac{A_n}{\sqrt{n}} \right\| \\ &\stackrel{(a)}{\leq} \frac{1}{\alpha_n} (1 + \ell^+ + o_P(1)) (2 + o_P(1)) = o_P(1). \end{aligned}$$

where (a) follows from the second estimate in (6.3.1) and from the spectral norm estimate. Notice that the majorization above is uniform for $\lambda \in \mathcal{S}_n$. The first part of the corollary is proved. Finally,

$$\text{Re}(\lambda) + x_k \leq |\lambda + x_k| = o_P(1) \quad \Rightarrow \quad \text{Re}(\lambda) \leq -\min_{k \in [n]} x_k + o_P(1).$$

The estimate (6.1.5) finally follows from the first estimate in (6.3.1).

6.4 Heuristics at critical scaling, non-homogeneous systems and non-gaussian entries

A heuristics at the critical scaling

We provide here a heuristics to compute the probability that a solution \mathbf{x}_n is feasible at critical scaling $\alpha_n^* = \sqrt{2 \log n}$.

Heuristics 6.4.1. *The probability that a solution is feasible at the critical scaling α_n^* is asymptotically given by*

$$\mathbb{P}(x_k > 0, k \in [n]) \approx 1 - \sqrt{\frac{e}{4\pi \log n}} + \frac{e}{8\pi \log n} =: H_1(n). \quad (6.4.1)$$

In Figure 6.2, we compare the heuristics with results from simulations.

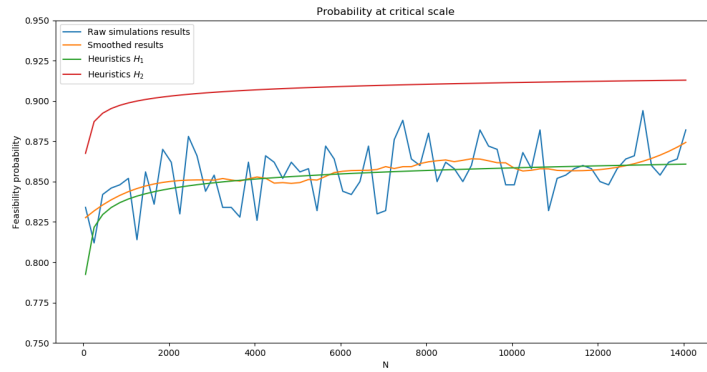


FIGURE 6.2: Probability at critical scaling. The solid curve corresponds to the proportion of feasible solutions at critical scaling α_n^* obtained for 10000 simulations (for n ranging from 50 to 3750 with a 200-increment) - notice the strong standard deviation. The dashed curve represents the heuristics H_1 defined in (6.4.1). The dotted curve represents the heuristics H_2 introduced in Remark 6.4. Notice the substantial discrepancy between H_1 and H_2 .

Arguments. Consider

$$x_k = 1 + e_k^* \frac{A_n}{\alpha_n^* \sqrt{n}} \mathbf{1}_n + \frac{R_k}{(\alpha_n^*)^2} = 1 + \frac{Z_k}{\alpha_n^*} + \frac{R_k}{(\alpha_n^*)^2} = 1 + \frac{1}{\alpha_n^*} \left(Z_k + \frac{R_k}{\alpha_n^*} \right).$$

Following Geman and Hwang [43, Lemma A.1], one could prove that Z_k and R_k are asymptotically independent centered Gaussian random variables, each with variance one. We thus approximate the quantity $Z_k + \frac{R_k}{\alpha_n^*}$ by a Gaussian random variable with distribution $\mathcal{N}\left(0, 1 + \frac{1}{(\alpha_n^*)^2}\right)$ and set $x_k \approx 1 + \left(\frac{1}{\alpha_n^*} \sqrt{1 + \frac{1}{(\alpha_n^*)^2}}\right) U_k$ where the U_k 's are i.i.d. $\mathcal{N}(0, 1)$.

6.4. HEURISTICS AT CRITICAL SCALING, NON-HOMOGENEOUS SYSTEMS AND NON-GAUSSIAN

Denote by $\check{M}_n^U = \min_{k \in [n]} U_k$ then

$$\mathbb{P}(x_k > 0, k \in [n]) \approx \mathbb{P}\left(1 + \left(\frac{1}{\alpha_n^*} \sqrt{1 + \frac{1}{(\alpha_n^*)^2}}\right) \check{M}_n^U > 0\right).$$

Recall that standard Gaussian extreme value convergence results yield

$$\mathbb{P}\{\alpha_n^*(-\check{M}_n^U - \beta_n^*) < x\} = \mathbb{P}\{\alpha_n^*(\check{M}_n^U + \beta_n^*) > -x\} \xrightarrow{n \rightarrow \infty} G(x) = e^{-e^{-x}}, \quad (6.4.2)$$

where β_n^* is defined in (6.2.3). Denote by $\Theta(\alpha) = \sqrt{1 + \alpha^{-2}}$ then

$$\mathbb{P}\left(1 + \Theta(\alpha_n^*) \frac{\check{M}_n^U}{\alpha_n^*} > 0\right) = \mathbb{P}\left(\alpha_n^*(\check{M}_n^U + \beta_n^*) > -\frac{(\alpha_n^*)^2}{\Theta(\alpha_n^*)} + \alpha_n^* \beta_n^*\right).$$

Notice that

$$-\frac{(\alpha_n^*)^2}{\Theta(\alpha_n^*)} + \alpha_n^* \beta_n^* = \frac{1}{2} - \frac{1}{2} \log(4\pi \log n) + \mathcal{O}\left(\frac{1}{(\alpha_n^*)^2}\right) = \frac{1}{2} + \log \frac{1}{\sqrt{2\pi} \alpha_n^*} + \mathcal{O}\left(\frac{1}{(\alpha_n^*)^2}\right).$$

Hence

$$\begin{aligned} \mathbb{P}\left(1 + \Theta(\alpha_n^*) \frac{\check{M}_n^U}{\alpha_n^*} > 0\right) &= \mathbb{P}\left(\alpha_n^*(\check{M}_n^U + \beta_n^*) > \frac{1}{2} + \log \frac{1}{\sqrt{2\pi} \alpha_n^*} + \mathcal{O}\left(\frac{1}{(\alpha_n^*)^2}\right)\right), \\ &\stackrel{(a)}{\approx} e^{-\exp\left(\frac{1}{2} + \log \frac{1}{\sqrt{2\pi} \alpha_n^*} + \mathcal{O}\left(\frac{1}{(\alpha_n^*)^2}\right)\right)} = e^{-\sqrt{\frac{e}{2\pi}} \frac{1}{\alpha_n^*} (1 + \mathcal{O}((\alpha_n^*)^{-2}))}, \\ &= 1 - \sqrt{\frac{e}{2\pi}} \frac{1}{\alpha_n^*} + \frac{1}{2} \frac{e}{2\pi} \frac{1}{(\alpha_n^*)^2} + \mathcal{O}\left(\frac{1}{(\alpha_n^*)^3}\right). \end{aligned} \quad (6.4.3)$$

We finally end up with the announced approximation

$$\mathbb{P}(x_k > 0, k \in [n]) \approx H_1(n) := 1 - \sqrt{\frac{e}{4\pi \log n}} + \frac{e}{8\pi \log n}.$$

Remark : A rougher approximation would have been to set $x_k \approx 1 + \frac{Z_k}{\alpha_n^*}$ with $Z_k \sim \mathcal{N}(0, 1)$ and to drop the next term $\frac{R_k}{(\alpha_n^*)^2}$ in the heuristics but this would have resulted in the following approximation

$$\mathbb{P}(x_k > 0, k \in [n]) \approx 1 - (4\pi \log(n))^{-1/2} + (8\pi \log(n))^{-1} =: H_2(n),$$

which is worse than $H_1(n)$, as illustrated in Figure 6.2.

□

Positivity for a non-homogeneous linear system

By homogeneous, we refer to a LV system where the intrinsic growth rate of species i is equal to 1. If not, the system is non-homogeneous (NH). The results developed so far

extend to a NH linear system where $\mathbf{1}_n$ is replaced by a vector \mathbf{r}_n with slight modifications. In particular, we identify a regime where feasibility and stability occur simultaneously. Denote by $\mathbf{r}_n = (r_k)$ a $n \times 1$ deterministic vector with positive components and consider the linear system

$$\mathbf{x}_n = \mathbf{r}_n + \frac{1}{\alpha_n \sqrt{n}} A_n \mathbf{x}_n. \quad (6.4.4)$$

Introduce the notations

$$r_{\min}(n) = \min_{k \in [n]} r_k, \quad r_{\max}(n) = \max_{k \in [n]} r_k \quad \text{and} \quad \sigma_{\mathbf{r}}(n) = \|\mathbf{r}/\sqrt{n}\| = \sqrt{n^{-1} \sum_{k \in [n]} r_k^2}.$$

Assume that there exist ρ_{\min}, ρ_{\max} independent from n such that eventually

$$0 < \rho_{\min} \leq r_{\min}(n) \leq \sigma_{\mathbf{r}}(n) \leq r_{\max}(n) \leq \rho_{\max} < \infty.$$

Then

Theorem 6.7 (Feasibility - NH case). *Let $\alpha_n \xrightarrow{n \rightarrow \infty} \infty$ and denote by $\alpha_n^* = \sqrt{2 \log n}$. Let $\mathbf{x}_n = (x_k)_{k \in [n]}$ be the solution of (6.4.4).*

1. *If $\exists \varepsilon > 0$ such that $\alpha_n \leq (1 - \varepsilon) \frac{\alpha_n^* \sigma_{\mathbf{r}}(n)}{r_{\max}(n)}$ then $\mathbb{P} \{ \min_{k \in [n]} x_k > 0 \} \xrightarrow{n \rightarrow \infty} 0$.*
2. *If $\exists \varepsilon > 0$ such that $\alpha_n \geq (1 + \varepsilon) \frac{\alpha_n^* \sigma_{\mathbf{r}}(n)}{r_{\min}(n)}$ then $\mathbb{P} \{ \min_{k \in [n]} x_k > 0 \} \xrightarrow{n \rightarrow \infty} 1$.*

Remark : Contrary to the homogeneous system where there is a sharp transition at $\alpha_n^* = \sqrt{2 \log(n)}$, the situation is not as clean-cut here and there is a buffer zone

$$\alpha_n \in \left[\frac{\sigma_{\mathbf{r}}(n)}{r_{\max}(n)} \sqrt{2 \log(n)}, \frac{\sigma_{\mathbf{r}}(n)}{r_{\min}(n)} \sqrt{2 \log(n)} \right]$$

in which the study of the feasibility is not clear.

This buffer zone is illustrated in Figure 6.3 where we simulate the transition toward feasibility for a non-homogeneous system (6.4.4) in the case where deterministic vector \mathbf{r}_n is equally distributed over $[1, 3]$, i.e.

$$\mathbf{r}_n(i) = 1 + \frac{2i}{n}, \quad \sigma_{\mathbf{r}}(n) = \sqrt{\frac{1}{n} \sum_{i \in [n]} r_n^2(i)} \xrightarrow{n \rightarrow \infty} \sqrt{\int_0^1 (1 + 2x)^2 dx} \quad (6.4.5)$$

We introduce the quantities

$$t_1 = \lim_N \frac{\sqrt{2} \sigma_{\mathbf{r}}(N)}{r_{\max}} \simeq 0.98 \quad \text{and} \quad t_2 = \lim_N \frac{\sqrt{2} \sigma_{\mathbf{r}}(N)}{r_{\min}} \simeq 2.94. \quad (6.4.6)$$

As one may notice, the transition region is wider than in the homogeneous case.

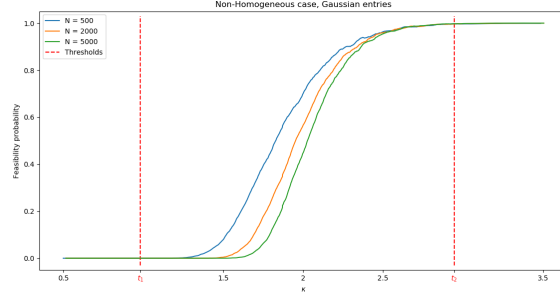


FIGURE 6.3: Transition toward feasibility for a NH system. The curves are obtained as for Figure 6.1 for \mathbf{r}_N defined in (6.4.5). The thresholds t_1 and t_2 are computed in (6.4.6)..

Elements of proof. We have

$$x_k = \mathbf{e}_k^* Q \mathbf{r}_n = r_k + \frac{1}{\alpha} \frac{\sum_{i=1}^n r_i A_{ki}}{\sqrt{n}} + \frac{1}{\alpha^2} \mathbf{e}_k^* \left(\frac{A}{\sqrt{n}} \right) Q \mathbf{r}_n = r_k + \frac{\sigma_{\mathbf{r}}(n)}{\alpha} U_k + \frac{1}{\alpha^2} R_k^{(r)}$$

where the U_k 's are i.i.d. $\mathcal{N}(0, 1)$. One can check by carefully reading the proof of Lemma 6.3 that the conclusions of the lemma apply to $R_k^{(r)}$. In particular, one may check that Proposition 6.6 holds uniformly in $k \in [n]$ in the non-homogeneous case. Denote by $\check{M} = \min_{k \in [n]} U_k$, then

$$\begin{aligned} \min_{k \in [n]} x_k &\leq r_{\max}(n) + \frac{\sigma_{\mathbf{r}}(n)}{\alpha} \check{M} + \frac{\max_{k \in [n]} R_k^{(r)}}{\alpha^2}, \\ &\leq r_{\max}(n) + \frac{\sigma_{\mathbf{r}}(n) \alpha^*}{\alpha} \left(\frac{\check{M} + \beta^*}{\alpha^*} - \frac{\beta^*}{\alpha^*} + \frac{\max_{k \in [n]} R_k^{(r)}}{\sigma_{\mathbf{r}}(n) \alpha^* \alpha} \right), \\ &= r_{\max}(n) + \frac{\sigma_{\mathbf{r}}(n) \alpha^*}{\alpha} (-1 + o_P(1)). \end{aligned}$$

The first statement of the theorem follows. The second statement follows similarly, noticing that $\min_{k \in [n]} x_k \geq r_{\min}(n) + \alpha^{-1} \sigma_{\mathbf{r}}(n) \check{M} + \alpha^{-2} \min_{k \in [n]} R_k^{(r)}$. Proof of Theorem 6.7 is completed. \square

A non homogeneous system (6.4.4) is associated to the following Lotka-Volterra system

$$\frac{dx_k(t)}{dt} = x_k(t) \left(r_k - x_k(t) + \frac{1}{\alpha_n \sqrt{n}} \sum_{\ell \in [n]} A_{k\ell} x_\ell(t) \right)$$

for $k \in [n]$ whose jacobian at equilibrium is still given by (6.1.4).

Theorem 6.8 (Stability - NH case). *Let $\mathbf{x}_n = (x_k)_{k \in [n]}$ be the solution of (6.4.4) and assume that*

$$\ell^+ := \limsup_{n \rightarrow \infty} \frac{\alpha_n^* \sigma_{\mathbf{r}}(n)}{\alpha_n r_{\min}(n)} < 1.$$

Denote by \mathcal{S}_n the spectrum of $\mathcal{J}(\mathbf{x}_n)$. Then for every $\lambda \in \mathcal{S}_n$,

$$\max_{\lambda \in \mathcal{S}_n} \min_{k \in [n]} |\lambda + x_k| \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0 \quad \text{and} \quad \max_{\lambda \in \mathcal{S}_n} \operatorname{Re} \lambda \leq -(1 - \ell^+) + o_P(1).$$

Beyond the Gaussian case

In this section, we extend the result to the class of random variables satisfying a logarithmic Sobolev inequality. The Gaussianity of the entries is used at three crucial steps, and in each case, it is enough that the entries satisfy a log-Sobolev inequality for some constant $\rho < \infty$:

- Gaussian entries immediately imply that the Z_k 's are independent standard Gaussian random variables, for which the study of the extrema is standard.

In the case where the entries are not Gaussian, the Z_k 's are no longer Gaussian but this issue can easily be circumvented since by the CLT the Z_k 's converge in distribution to a standard Gaussian. The extreme value study of such families of Z_k 's has been carried out in [3, Propositions 2 & 3]. In our case, it is easy to check that the $LSI(\rho)$ condition ensures that we can apply [3, Proposition 3].

- Gaussian concentration has been used to prove sub-Gaussianity of the Lipschitz functional $\tilde{R}_k(A)$. This remains valid under a log-Sobolev inequality assumption, the value of the constant being irrelevant to us. (The point is that, by tensorization, for any column i the log-Sobolev constant of the vector $(A_{i,j})_{1 \leq j \leq n}$ remains upper-bounded by ρ , which does not depend on n .)
- In the proof of Lemma 6.3 the Gaussian Poincaré inequality is used to prove that $\tilde{R}_1(A)/(\alpha\sqrt{2\log(n)})$ goes to zero in probability, but the log-Sobolev inequality implies the Poincaré inequality. Again the constant does not matter.

Hence we can extend Theorem 6.1 as :

Theorem 6.9. *Assume that the entries A_{ij} are i.i.d. centered, with finite variance equal to 1 and satisfy a logarithmic Sobolev inequality, then the conclusions of Theorem 6.1 hold.*

Perspectives

Let us briefly describe a sample of possible directions for future work.

Log-sobolev inequalities for log-concave measures

We strongly believe that the estimates proved in Chapter 4 can be improved.

A first goal is to elaborate a complete scheme of proof using stochastic localization, continuing the work initiated in Section 4.4, thus reducing the problem to controlling the evolution of the subgaussianity parameter $\sigma_{SG}(\mu)$ along the stochastic localization process.

In a related direction, consider a probability $\mu = e^{-V}$ satisfying $\nabla^2 V \geq \rho$ for some $\rho > 0$. In [56], the key ingredient proved by Klartag is an improved Lichnerowicz inequality:

$$c_P(\mu)^2 \leq \sqrt{\frac{|\text{Cov}(\mu)|_{op}}{\rho}}.$$

A natural question is to try and prove a similar inequality for the log-Sobolev constant:

$$\rho_{LS}(\mu)^2 \lesssim \sqrt{\frac{\sigma_{SG}^2(\mu)}{\rho}}.$$

Finally, much remains to be understood about tilt-stable log-concave measures. For a measure μ , write

$$d\mu_{t,h} = \frac{1}{Z_{t,h}} e^{-t|x|^2 + h \cdot x} d\mu$$

Recall that μ is 1-tilt stable if

$$\sup_{h \in \mathbb{R}^n} |\text{Cov}(\mu_{0,h})|_{op} \leq 1$$

In Section 4.5.2, we show that if μ is 1-tilt stable and log-concave it is C_n -strongly tilt stable, that is

$$\sup_{t>0, h \in \mathbb{R}^n} |\text{Cov}(\mu_{t,h})|_{op} \leq C_n,$$

for some constant C_n polynomial in the dimension. We recall that the strong tilt-stability constant controls the log-Sobolev constant. We strongly believe that C_n could be chosen to

be polylogarithmic in the dimension, thus providing a polylog estimate for the log-Sobolev constant of tilt stable log concave measures. This would require using a different approach than the perturbation argument of Section 4.5.2. Let us give a heuristic to support this belief. Fix $T = \frac{1}{c \log^2 n}$ for some constant $c > 0$. For any $t \leq T$, we know that

$$|\text{Cov}(\mu_{t,h_t})|_{op} \leq 2 \quad \text{with high probability.}$$

where $h_t = tX + B_t$. (This is a slightly stronger version of Lemma 2.14). However, in the spirit of Lemma 4.37, we may apply this argument not only to μ but to all its tilts $\mu_{0,h}$ for $h \in \mathbb{R}^n$. We could then hope to recover all the measures $\mu_{t,h}$ by this procedure, for $t \leq T$. To make it more clear, if instead of having a bound on the measures μ_{t,h_t} with high probability on h_t , we could derandomize the stochastic localization process and have a bound just for the 0 tilt, $\mu_{t,0}$, then applying Lemma 4.37 would immediately yield $C_n \lesssim \log^2 n$.

Finally, a very natural and important question, to our eyes, is to determine whether all sub-gaussian log-concave measures are in fact tilt-stables, up to a constant. A starting point is to investigate the one-dimensional case.

A KLS conjecture on the hypercube

Let μ be the uniform measure on the discrete hypercube $C_n = \{-1, 1\}^n$. For a positive symmetric matrix J and vector $h \in \mathbb{R}^n$, we introduce the measure

$$\mu_{J,h} = \frac{1}{Z_{J,h}} \mu e^{\frac{1}{2} Jx \cdot x + h \cdot x}$$

where $Z_{J,h}$ is a normalizing constant. As we have seen in Section 5.1,

Proposition 6.10.

$$\text{Cov}(\mu_{J,h}) \preceq \frac{1}{1 - |J|_{op}} I_n \quad \text{and} \quad C_P(\mu_{J,h}) \leq \frac{1}{1 - |J|_{op}}$$

where in that context C_P stands for the optimal constant in the discrete Poincaré inequality.

Using (a renormalized version of) the process introduced by Eldan and Chen[29], the measure $\mu_{J,h}$ can be decomposed into measures $\mu_t = \mu_{J_t, h_t}$ where $J_t = (1 - t)J$, for $0 \leq t \leq 1$. In particular, starting from a measure $\mu_{J,h}$ with $|J|_{op} \leq 1$, one gets that $c_P(\mu_t) \leq \frac{1}{t}$ almost surely. This is analogous to what happens for the stochastic localization of log-concave measures, and provides an approach to tackling the following KLS-type question:

Question 6.11. *Let $\mu_{J,h}$ be as above, with $|J|_{op} \leq 1$, is it true that*

$$C_P(\mu_{J,h}) \lesssim |\text{Cov}(\mu_{J,h})|_{op} \quad ?$$

Very similarly to the setting of the KLS Conjecture, one then needs to control the evolution of the covariance matrix along the process. The main difficulty is the absence of the traditional machinery for log-concave vectors, such as Borell's Lemma.

Shortest billiard trajectory of random polytopes

Finally, we describe a question which has a slightly different flavor and which is part of an ongoing discussion with Professor Ostrover.

Let K be a convex body of \mathbb{R}^n . We denote by $\xi(K)$ the length of the shortest periodic billiard trajectory in K , by $\text{inrad}(K)$ the inradius of K , and by $\text{width}(K)$ the minimal width of K . The following lower and upper-bound for $\xi(K)$ are due to Ghomi [44] and Ostrover and Artstein-Avidan[6], respectively.

$$4\text{inrad}(K) \leq \xi(K) \leq 2\text{width}(K). \quad (6.4.7)$$

Denote by $s(K) = \frac{\text{width}(K)}{\text{inrad}(K)}$. When K is symmetric, the two bounds coincide, so that $s(K) = 1$. However, in the case of the regular simplex Δ_n for instance, we have $s(\Delta_n) \simeq \sqrt{n}$. It is natural to wonder what happens in the case of random polytopes. More precisely Let ν be, for instance, the uniform measure on the sphere, define for $N \geq n$ the random polytope

$$P_N = \text{conv}(X_1, \dots, X_N)$$

where the X_1, \dots, X_N are iid and distributed according to ν . We propose to estimate $s(P_N)$. When $N = n$, P_N is close to being a regular simplex so that we expect $s(P_N)$ to be big. When N is (very) large however, P_N converges to the euclidean ball, so that $s(P_N)$ should be close to 1.

Bibliography

- [1] Stefano Allesina and Si Tang. The stability–complexity relationship at age 40: a random matrix perspective. *Population Ecology*, 57(1):63–75, 2015.
- [2] David Alonso-Gutiérrez and Jesús Bastero. *Approaching the Kannan-Lovász-Simonovits and variance conjectures*, volume 2131. Springer, 2015.
- [3] Clive W Anderson, Stuart G Coles, and Jürg Hüsler. Maxima of poisson-like variables and related triangular arrays. *The Annals of Applied Probability*, 7(4):953–971, 1997.
- [4] Milla Anttila, Keith Ball, and Irini Perissinaki. The central limit problem for convex bodies. *Transactions of the American Mathematical Society*, 355(12):4723–4735, 2003.
- [5] Shiri Artstein, Keith Ball, Franck Barthe, and Assaf Naor. Solution of shannon’s problem on the monotonicity of entropy. *Journal of the American Mathematical Society*, 17(4):975–982, 2004.
- [6] Shiri Artstein-Avidan and Yaron Ostrover. Bounds for minkowski billiard trajectories in convex bodies. *International Mathematics Research Notices*, 2014(1):165–193, 2014.
- [7] Zhidong Bai and Jack W Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.
- [8] Dominique Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84: Proceedings*, pages 177–206. Springer, 2006.
- [9] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grund. Math. Wiss.* Springer, 2014.
- [10] Keith Ball, Franck Barthe, and Assaf Naor. Entropy jumps in the presence of a spectral gap. *Duke Math. J.*, 119(1):41–63, 2003.
- [11] Keith Ball and Van Hoang Nguyen. Entropy jumps for isotropic log-concave random vectors and spectral gap. *Studia Math.*, 213(1):81–96, 2012.
- [12] Franck Barthe and Emanuel Milman. Transference principles for log-sobolev and spectral-gap with applications to conservative spin systems. *Communications in Mathematical Physics*, 323(2):575–625, 2013.

- [13] Roland Bauerschmidt and Thierry Bodineau. A very simple proof of the lsi for high temperature spin systems. *Journal of Functional Analysis*, 276(8):2582–2588, 2019.
- [14] Roland Bauerschmidt and Thierry Bodineau. Log-sobolev inequality for the continuum sine-gordon model. *Communications on Pure and Applied Mathematics*, 74(10):2064–2113, 2021.
- [15] Richard Bellman. The theory of dynamic programming. *Bulletin of the American Mathematical Society*, 60(6):503–515, 1954.
- [16] Pierre Bizeul. On measures strongly log-concave on a subspace. *arXiv preprint arXiv:2203.09422*, 2022.
- [17] Pierre Bizeul. Entropy and information jump for log-concave vectors. *Comptes Rendus. Mathématique*, 361(G2):487–493, 2023.
- [18] Pierre Bizeul and Jamal Najim. Positive solutions for large random linear systems. *Proceedings of the American Mathematical Society*, 149(6):2333–2348, 2021.
- [19] Sergey G Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *The Annals of Probability*, 27(4):1903–1921, 1999.
- [20] Sergey G Bobkov. On isoperimetric constants for log-concave probability distributions. In *Geometric Aspects of Functional Analysis: Israel Seminar 2004–2005*, pages 81–88. Springer, 2007.
- [21] SG Bobkov. Gaussian concentration for a class of spherically invariant measures. *Journal of Mathematical Sciences*, 167(3):326–339, 2010.
- [22] Christer Borell. Convex measures on locally convex spaces. *Arkiv för Matematik*, 12(1-2):239 – 252, 1974.
- [23] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.
- [24] Jean Bourgain. On the distribution of polynomials on high dimensional convex sets. In *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 1989–90*, pages 127–137. Springer, 2006.
- [25] Ulrich Brehm, Peter Hinow, Hendrik Vogt, and Jürgen Voigt. Moment inequalities and central limit properties of isotropic convex bodies. *Mathematische Zeitschrift*, 240(1):37–51, 2002.
- [26] Peter Buser. A note on the isoperimetric constant. In *Annales scientifiques de l’École normale supérieure*, volume 15, pages 213–230, 1982.

- [27] Jeff Cheeger. A lower bound for the smallest eigenvalue of the laplacian, problems in analysis (papers dedicated to salomon bochner, 1969), 1970.
- [28] Yuansi Chen. An almost constant lower bound of the isoperimetric coefficient in the kls conjecture. *Geometric and Functional Analysis*, 31:34–61, 2021.
- [29] Yuansi Chen and Ronen Eldan. Localization schemes: A framework for proving mixing bounds for markov chains. *arXiv preprint arXiv:2203.04163*, 2022.
- [30] Thomas A Courtade, Max Fathi, and Ashwin Pananjady. Quantitative stability of the entropy power inequality. *IEEE Transactions on Information Theory*, 64(8):5691–5703, 2018.
- [31] Persi Diaconis and David Freedman. Asymptotics of graphical projection pursuit. *The annals of statistics*, pages 793–815, 1984.
- [32] David L Donoho et al. High-dimensional data analysis: The curses and blessings of dimensionality. *AMS math challenges lecture*, 1(2000):32, 2000.
- [33] Michaël Dougoud, Laura Vinckenbosch, Rudolf P Rohr, Louis-Félix Bersier, and Christian Mazza. The feasibility of equilibria in large ecosystems: A primary but neglected concept in the complexity-stability debate. *PLoS computational biology*, 14(2):e1005988, 2018.
- [34] Ronen Eldan. *Distribution of Mass in Convex Bodies*. Tel Aviv University, 2012.
- [35] Ronen Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geometric and Functional Analysis*, 23(2):532–569, 2013.
- [36] Ronen Eldan and Bo’az Klartag. Approximately gaussian marginals and the hyperplane conjecture. *Concentration, functional inequalities and isoperimetry*, 545:55–68, 2011.
- [37] Ronen Eldan, Frederic Koehler, and Ofer Zeitouni. A spectral condition for spectral gap: fast mixing in high-temperature ising models. *Probability theory and related fields*, 182(3-4):1035–1051, 2022.
- [38] Ronen Eldan and Dan Mikulincer. Stability of the Shannon-Stam inequality via the Föllmer process. *Probab. Theory Related Fields*, 177(3-4):891–922, 2020.
- [39] Ronen Eldan and Omer Shamir. Log concavity and concentration of lipschitz functions on the boolean hypercube. *Journal of Functional Analysis*, 282(8):109392, 2022.
- [40] David A Freedman. On tail probabilities for martingales. *the Annals of Probability*, pages 100–118, 1975.

- [41] Mark R Gardner and W Ross Ashby. Connectance of large dynamic (cybernetic) systems: critical values for stability. *Nature*, 228(5273):784–784, 1970.
- [42] Stuart Geman. The spectral radius of large random matrices. *The Annals of Probability*, pages 1318–1328, 1986.
- [43] Stuart Geman and Chii-Ruey Hwang. A chaos hypothesis for some large systems of random equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 60(3):291–314, 1982.
- [44] Mohammad Ghomi. Shortest periodic billiard trajectories in convex bodies. *Geometric & Functional Analysis GAFA*, 14(2):295–302, 2004.
- [45] Theo Gibbs, Jacopo Grilli, Tim Rogers, and Stefano Allesina. Effect of population abundances on the stability of large random ecosystems. *Physical Review E*, 98(2):022410, 2018.
- [46] M. Gromov and V. D. Milman. A topological application of the isoperimetric inequality. *American Journal of Mathematics*, 105(4):843–854, 1983.
- [47] Olivier Guédon and Emanuel Milman. Interpolating thin-shell and sharp large-deviation estimates for isotropic log-concave measures. *Geometric and Functional Analysis*, 21(5):1043–1068, 2011.
- [48] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.
- [49] Arun Jambulapati, Yin Tat Lee, and Santosh S Vempala. A slightly improved bound for the kls constant. *arXiv preprint arXiv:2208.11644*, 2022.
- [50] Ravi Kannan, László Lovász, and Miklós Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete & Computational Geometry*, 13:541–559, 1995.
- [51] Ravi Kannan, László Lovász, and Miklós Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete & Computational Geometry*, 13(3):541–559, 1995.
- [52] Tosio Kato. *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media, 2013.
- [53] Bo’az Klartag. On convex perturbations with a bounded isotropic constant. *Geometric & Functional Analysis GAFA*, 16(6):1274–1290, 2006.
- [54] Bo’az Klartag. A central limit theorem for convex sets. *Inventiones mathematicae*, 168(1):91–131, 2007.

- [55] Bo'az Klartag. Eldan's stochastic localization and tubular neighborhoods of complex-analytic sets. *The Journal of Geometric Analysis*, 28(3):2008–2027, Jul 2018.
- [56] Bo'az Klartag. Logarithmic bounds for isoperimetry and slices of convex sets. *arXiv preprint arXiv:2303.14938*, 2023.
- [57] Bo'az Klartag and Joseph Lehec. Bourgain's slicing problem and kls isoperimetry up to polylog. *arXiv preprint arXiv:2203.15551*, 2022.
- [58] Bo'az Klartag and Emanuel Milman. Centroid bodies and the logarithmic laplace transform—a unified approach. *Journal of Functional Analysis*, 262(1):10–34, 2012.
- [59] Bo'az Klartag and Eli Putterman. Spectral monotonicity under gaussian convolution. *arXiv preprint arXiv:2107.09496*, 2021.
- [60] Bo'az Klartag. *Needle decompositions in Riemannian geometry*, volume 249. American Mathematical Society, 2017.
- [61] Bo'az Klartag. Eldan's stochastic localization and tubular neighborhoods of complex-analytic sets. *The Journal of Geometric Analysis*, 28:2008–2027, 2018.
- [62] Bo'az Klartag and Joseph Lehec. Bourgain's slicing problem and kls isoperimetry up to polylog. *Geometric and Functional Analysis*, 32(5):1134–1159, 2022.
- [63] Malcolm R Leadbetter, Georg Lindgren, and Holger Rootzén. *Extremes and related properties of random sequences and processes*. Springer Science & Business Media, 2012.
- [64] M. Ledoux. *The Concentration of Measure Phenomenon*. Mathematical surveys and monographs. American Mathematical Society, 2001.
- [65] Michel Ledoux. A simple analytic proof of an inequality by p. buser. *Proceedings of the American mathematical society*, 121(3):951–959, 1994.
- [66] Yin Tat Lee and Santosh S Vempala. The kannan-lov'asz-simonovits conjecture. *arXiv preprint arXiv:1807.03465*, 2018.
- [67] Yin Tat Lee and Santosh Srinivas Vempala. Eldan's stochastic localization and the kls hyperplane conjecture: an improved lower bound for expansion. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 998–1007. IEEE, 2017.
- [68] Yin Tat Lee and Santosh Srinivas Vempala. Eldan's stochastic localization and the KLS Hyperplane conjecture: An improved lower bound for expansion. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 998–1007, 2017.

- [69] László Lovász and Miklós Simonovits. Random walks in a convex body and an improved volume algorithm. *Random structures & algorithms*, 4(4):359–412, 1993.
- [70] Robert M May. Will a large complex system be stable? *Nature*, 238:413–414, 1972.
- [71] Vladimir Gilelevich Maz’ya. Classes of domains and imbedding theorems for function spaces. In *Doklady Akademii Nauk*, volume 133, pages 527–530. Russian Academy of Sciences, 1960.
- [72] Emanuel Milman. On the role of convexity in functional and isoperimetric inequalities. *Proceedings of the London Mathematical Society*, 99(1):32–66, 2009.
- [73] Emanuel Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Inventiones mathematicae*, 177(1):1–43, Jul 2009.
- [74] Emanuel Milman. Isoperimetric and concentration inequalities: equivalence under curvature lower bound. *Duke Mathematical Journal*, 154(2):207–239, 2010.
- [75] Emanuel Milman. Properties of isoperimetric, functional and transport-entropy inequalities via concentration. *Probability Theory and Related Fields*, 152(3):475–507, 2012.
- [76] Bernt Øksendal. Stochastic differential equations. In *Stochastic differential equations*, pages 65–84. Springer, 2003.
- [77] G. Paouris. Concentration of mass on convex bodies. *Geometric & Functional Analysis GAFA*, 16(5):1021–1049, Dec 2006.
- [78] Yair Shenfeld. Exact renormalization groups and transportation of measures. *arXiv preprint arXiv:2205.01642*, 2022.
- [79] Lewi Stone. The feasibility and stability of large complex biological networks: a random matrix approach. *Scientific reports*, 8(1):8246, 2018.
- [80] Vladimir Nikolaevich Sudakov. Typical distributions of linear functionals in finite-dimensional spaces of higher dimension. In *Doklady Akademii Nauk*, volume 243, pages 1402–1405. Russian Academy of Sciences, 1978.
- [81] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.