Habilitation à diriger des recherches

Discipline : Mathématiques

Université de Lille, CNRS, UMR 8524, Inria,

École Doctorale SPI Université Lille Nord de France

présentée et soutenue publiquement le 13 novembre 2020 par

Adrien Hardy

Attractive Repulsion

Après avis de :

Pierre-Olivier AMBLARD	*DR CNRS* (GIPSA-lab, Université Grenoble Alpes)
Mathieu LEWIN	*DR CNRS* (CEREMADE, Université Paris Dauphine)
Sylvia SERFATY	*Professor* (CIMS, New York University)

Devant la commission formée de :

Pierre-Olivier AMBLARD	*DR CNRS* (GIPSA-lab, Université Grenoble Alpes)
Carlós BELTRÁN	*Profesor* (DMEC, Universidad de Cantabria)
Djalil CHAFAÏ	*Professeur* (CEREMADE, Université Paris Dauphine)
David DEREUDRE	*Professeur* (LPP, Université de Lille)
Mathieu LEWIN	*DR CNRS* (CEREMADE, Université Paris Dauphine)
Mylène MAÏDA	*Professeure* (LPP, Université de Lille)
Sandrine PÉCHÉ	*Professeure* (LPSM, Université de Paris)
Sylvia SERFATY	*Professor* (CIMS, New York University)

Remerciements

Un grand merci à Pierre-Olivier Amblard, Mathieu Lewin et Sylvia Serfaty d'avoir accepté de rapporter cette HDR et pour leurs remarques avisées qui l'ont améliorée ; c'est un grand honneur pour moi. Merci aussi à Carlos Beltrán, David Dereudre, Djalil Chafaï et Sandrine Péché d'avoir accepté de faire partie du jury, en plein confinement, sans oublier Mylène Maïda, ma garante invincible.

Je dois avouer que j'ai de la chance d'être tombé sur ces si belles personnes que sont mes coauteurs, qui sans exception sont devenus des amis quand ils ne l'étaient pas déjà. Rémi, Carlos[♣]*, Djalil, David, Walid, Thierry, Mylène, Gaultier, Thomas, Jamal et Adam, ce mémoire vous est dédié : je vous remercie pour ces moments partagés de combativité et de célébration. Continuons à interagir même à longue portée !*

Un grand remerciement va aux collègues du laboratoire Paul Painlevé, CRIStAL, PhLAM, INRIA et aux amis lillois & Co, d'intersection largement non vide, pour ces bons moments de tout ordre. J'ai apprécié que vous fassiez comme si j'étais un probabiliste sans avouer au reste du monde que je ne savais pas, en arrivant à Lille, ce qu'était une espérance conditionnelle, un beau couplage ou un test statistique. Les débats animés en salle de convivialité m'ont nourri d'une énergie positive même quand j'ai pu afficher une opposition avec théâtralité. Et merci à ceux qui m'ont accompagné tard dans la nuit pour refaire le monde scientifique, entre autres ; ils se reconnaîtront mais ne se souviendront peut-être pas de tous les détails techniques.

La communauté des matrices aléatoires (au sens large) qui m'a bercé est assez singulière par sa bienveillance envers les plus jeunes et par sa convivialité. J'y ai pu me délecter de tant d'évènements scientifiques de qualité ! Je profite de cet endroit pour y remercier tous ses acteurs, les amis parisiens, toulousains et en particulier les organisateurs de conférences et du séminaire MEGA à l'IHP.

Finalement, et surtout, merci à ma famille et à mes vieux amis de toujours m'épauler dans cette vie où d'obscurs hiéroglyphes sont gribouillés sur des pages blanches, pianotés sur une machine avec enthousiasme, et malgré cette incroyable prétention de vouloir clarifier un peu cette expérience que l'on fait ensemble de réalités communes.

♡

[♣]Carlos: Básicamente digo que me siento afortunado de tener a mis coautores como son, a los que considero amigos, y a mí que les dedico estas memorias. Disfruté de la lucha y las celebraciones relacionadas con nuestro trabajo y gracias por eso. ¡Espero seguir interactuando, incluso a larga distancia!

Overview

To pass the habilitation à diriger de recherches (HDR) is the final initiation rite in the French academic system, where the candidate has to summarize her/his post-PhD research production in a pedagogical manner. I've heard this is a good opportunity to take some distance from the initial motivations, to draw the "big picture". From this perspective I realized that the choice for my research projects were always guided by the motivation to learn new mathematics, specially when they have a potential for concrete applications or for a better understanding of physical systems. These research projects also had in common one feature: they involve interacting particle systems of repulsive nature, that I all find quite elegant and attractive. This repulsive feature has for consequence rather counterintuitive properties for the unaware probabilist. Some of my papers are about understanding better these particle systems, some of them are about using them for applications. One can split such interacting systems into two classes, determinantal point processes (DPPs) and Coulomb gases, although they have a non-empty intersection and overflow on other interesting classes, such as random matrices or Gaussian analytic functions, which will show up here and there.

After a brief teaser, Chapter 1 presents the necessary background for the repulsive particle systems of interest, including a quick presentation of the works summarized in this thesis, as well as the general notation. Then, the manuscript is split into two parts: Chapters 2, 3, 4 deal with DPPs whereas the chapters 5, 6, 7, 8 focus on Coulomb gases.

In order to keep the size of this manuscript at its minimum, the presentation focus on the scientific contributions from my coauthors and I, and the reader may complain about the lack of general references, but we refer her/him to the published works associated to each chapter for further information.

Moreover, the joint work [P14] with Walid Hachem and Shlomo Shamai, which explains how the stationary measure of a well chosen matrix valued process is useful for computing the Shannon mutual information for a large class of multi-antenna channel models in wireless communication, is not reviewed in this thesis since it is too far from DPPs and Coulomb gases.

Publications list

Publications as a PhD student (2009–2013).

[P1] A. Hardy, A.B.J. Kuijlaars, *Weakly admissible vector equilibrium problems*, Journal of Approximation Theory, Volume 164, Issue 6, June 2012, Pages 854-868.

[P2] A. Hardy, *A note on large deviations for 2D Coulomb gas with weakly confining potential*, Electronic Communications in Probability, Volume 17 (2012), paper no. 19, 12 pp.

[P3] A. Hardy, A.B.J. Kuijlaars, *Large deviations for a non-centered Wishart matrix*, Random Matrices: Theory and Applications, Vol. 02, No. 01, 1250016 (2013).

[P4] A. Hardy, *Average characteristic polynomials of determinantal point processes*, Annales de l'Institut Henri Poincaré - Probabilités et Statistiques, Volume 51, Number 1 (2015), 283-303.

Publications as a post-doc (2013–2015).

[P5] W. Hachem, A. Hardy, J. Najim, *Large complex correlated Wishart matrices: Fluctuations and asymptotic independence at the edges*, Annals of Probability, Volume 21 (2016), paper no. 1, 36 pp.

[P6] W. Hachem, A. Hardy, J. Najim, *Large complex correlated Wishart matrices: The Pearcey kernel and expansion at the hard edge*, Electronic Journal of Probability, Volume 21 (2016), paper no. 1, 36 pp.

[P7] W. Hachem, A. Hardy, J. Najim, *A survey on the eigenvalues local behavior of large complex correlated Wishart matrices*, ESAIM: Proceedings and surveys, October 2015, Vol. 51, p. 150-174.

Publications as a Maître de conférences (2015–2020).

[P8] R. Bardenet, A. Hardy, *Monte Carlo with determinantal point processes*, Annals of Applied Probability, Volume 30, Number 1 (2020), 368-417.

[P9] D. Chafaï, A. Hardy, M. Maïda, *Concentration for Coulomb gases and Coulomb transport inequalities*, Journal of Functional Analysis, Volume 275, Issue 6, 15 September 2018, Pages 1447-1483.

[P10] A. Hardy, *Polynomial ensembles and recurrence coefficients*, Constructive Approximation, Vol 48, pages 137–162 (2018).

[P11] C. Beltrán, A. Hardy, *Energy of the Coulomb gas on the sphere at low temperature*, Archive for Rational Mechanics and Analysis, Vol 231, pages 2007–2017 (2019).

[P12] R. Bardenet, A. Hardy, *Time-frequency transforms of white noises and Gaussian analytic functions*, Applied and Computational Harmonic Analysis, in press (2019).

[P13] D. Dereudre, A. Hardy, T. Leblé, M. Maïda, *DLR equations and rigidity for the Sine-beta process*, Communications on Pure and Applied Mathematics, to appear (2018).

[P14] W. Hachem, A. Hardy, S. Shamai, *Mutual information of wireless channels and block-Jacobi ergodic operators*, IEEE Transactions on Information Theory, Volume 65, Issue 11, Nov. 2019.

[P15] A. Hardy, G. Lambert, *CLT for circular beta-ensembles at high temperature*, Journal of Functional Analysis, to appear (2019).

Other (proceedings & newsletters).

[P16] R. Bardenet, A. Hardy, *From random matrices to Monte Carlo integration via Gaussian quadrature*, IEEE Statistical Signal processing workshop, 2018, Freiburg, Germany (2018)

[P17] A. Hardy, M. Maïda, *Processus ponctuels déterminantaux*, Gazette des Mathématiciens - SMF (2018). An english version has also been published in the Newsletter of the EMS (2019).

[P18] R. Bardenet, P. Chanais, J. Flamant, A. Hardy, *A correspondence between zeros of time-frequency transforms and Gaussian analytic functions*, 13th International conference on sampling theory and applications, SampTA 2019, Aug 2019, Bordeaux, France.

Contents

Chapter 1

Introduction & background

This chapter aims at a short and pedagogical presentation on determinantal point processes and Coulomb gases, including an overview of the works that will be detailed in the next chapters.

As an appetizer we start with an inspirational particle system known as the CUE, to which all forthcoming models and results can be related to. Some of the notation and concepts we use here will be detailed in Section 1.2 below.

1.1 A central uplifting example: The CUE

Consider the unitary group $\mathcal{U}_N(\mathbb{C})$ of $N \times N$ complex unitary matrices and equip it with its normalized Haar measure ν, that is the unique probability measure on $\mathcal{U}_N(\mathbb{C})$ which is invariant under left and right multiplication. If we denote by $\mathrm{e}^{\mathrm{i}\mathbf{x}_1}, \dots, \mathrm{e}^{\mathrm{i}\mathbf{x}_N}$ the eigenvalues of a random unitary matrix with distribution ν, a result attributed to Weyl[♠] yields that the joint law of the phases $\mathbf{x}_1, \dots, \mathbf{x}_N$ living on the one-dimensional torus $\mathbb{T} := (-\pi, \pi] \simeq \mathbb{R}/2\pi\mathbb{Z}$ has a density proportional to

$$\prod_{1 \le i < j \le N} \left| \mathrm{e}^{\mathrm{i}x_i} - \mathrm{e}^{\mathrm{i}x_j} \right|^2. \tag{1.1}$$

The particle system $(\mathbf{x}_j)_{j=1}^N$ is known as the Circular Unitary Ensemble (of size N), abbreviated CUE or $\mathsf{CUE}(N)$ in the following.

Macroscopic behavior. At the global level, the $\mathsf{CUE}(N)$ is uniformly distributed on $\mathbb{T}$ on average, i.e. the empirical measure $\hat{\mu}_N := \frac{1}{N}\sum_{j=1}^N \delta_{\mathbf{x}_j}$ has mean $\mathbb{E}\hat{\mu}_N = \frac{\mathrm{d}x}{2\pi}$. More precisely, since $\mathcal{U}_N(\mathbb{C})$ is a Riemannian manifold with explicit Ricci curvature one can use Backry-Émery or Gromov-Milman approach to concentration inequalities to obtain sub-Gaussian concentration in the Wasserstein W_1 metric [Meckes

[♠] Diaconis and Forrester [2017] notice this result can be traced back to [Hurwitz, 1897], the oldest known fossil of embryonic random matrix theory.

and Meckes, 2013]: there exists $c > 0$ such that, for every $N \geq 1$ and $r > 0$,

$$\mathbb{P}_{\mathsf{CUE}}\Big(W_1(\hat{\mu}_N, \tfrac{\mathrm{d}x}{2\pi}) \geq r\Big) \leq \mathrm{e}^{-cN^2r^2}. \tag{1.2}$$

What may be surprising to the probabilist reader is the rate N^2 instead of N, which would be the correct rate for N independent and identically distributed (i.i.d) uniform random variables on $\mathbb{T}$. Thus, as $N \to \infty$, the $\mathsf{CUE}(N)$ converges globally to the uniform measure much faster than i.i.d uniform random variables would, underlying fluctuations of much smaller order. The reason for this phenomenon is the repulsion between the particles of the CUE, which is clearly seen from the form of the density (1.1). In fact, Johansson [1988] noticed that the strong Szegö theorem, which establishes second order asymptotics for large Toeplitz determinants, can be recast into a central limit theorem (CLT) for the CUE: given any $H^{1/2}$-smooth test function $f : \mathbb{T} \to \mathbb{R}$, we have

$$N \int f \left(\mathrm{d}\hat{\mu}_N - \tfrac{\mathrm{d}x}{2\pi}\right) \xrightarrow[N\to\infty]{\text{law}} \mathcal{N}(0, \sigma_f^2) \tag{1.3}$$

where σ_f is the $H^{1/2}$-Sobolev seminorm of f. In particular the fluctuations have standard deviation $1/N$, which should be compared to $1/\sqrt{N}$ in the setting of independent random variables.

Microscopic considerations. At the microscopic level, namely at a scaling where the typical distance between the particles is of order one, the $\mathsf{CUE}(N)$ converges as $N \to \infty$ towards a universal object called the Sine_2 process. More precisely, for any $x \in \mathbb{T}$ the point process $(N(\mathbf{x}_j - x))_{j=1}^N$ has a weak limit as $N \to \infty$ which one can describe precisely. The simplest way to do so is to notice that (1.1) is the square of a Vandermonde determinant and thus can be written as

$$\det\Big[K_N(x_i, x_j)\Big]_{j,k=1}^N \tag{1.4}$$

for a kernel K_N associated with a finite rank N projection operator. This structure remains unchanged after the change of variables $(x_j)_{j=1}^N \mapsto (N(x_j - x))_{j=1}^N$ but with a new kernel $\tilde{K}_N$ that converges locally uniformly as $N \to \infty$ to the limiting kernel

$$K_{\mathsf{Sine}}(x, y) = \frac{\sin(\pi(x-y))}{\pi(x-y)}. \tag{1.5}$$

$\diamond$ The particular structure (1.4) is an occurence of what is called a determinantal point process (DPP), a class of point processes that will be the main topic of the first part of this thesis; DPPs are introduced in Section 1.3.

$\diamond$ Using the statistical physics formalism, one could alternately write (1.1) as a canonical ensemble $\mathrm{e}^{-\beta H_N}$ at inverse temperature $\beta = 2$ for an energy functional H_N that is the sum over the pair potentials $\log|\mathrm{e}^{\mathrm{i}x_i} - \mathrm{e}^{\mathrm{i}x_j}|^{-1}$, which turns out to

be the Coulomb potential describing two-dimensional electrostatics. From this perspective, the CUE is a particular case of (the restriction of) a Coulomb gas, a class of particle systems that will be the main character of the second part of this thesis and introduced in Section 1.4.

1.2 General notation & conventions

- In what follows and without further precision, Λ refers to an arbitrary Polish space with a distance d, and a "measure" always means a Borel measure for the topology coming with this distance.

- We denote the Lipschitz constant of a Lipschitz function $f : \Lambda \to \mathbb{C}$ by

$$\|f\|_{\mathrm{Lip}} := \sup_{\substack{x,y\in\Lambda \\ x\neq y}} \frac{|f(x)-f(y)|}{d(x,y)}.$$

- For $f : \mathbb{T} \to \mathbb{C}$ in L^2, we denote by $\hat{f}_k := \int_{\mathbb{T}} f(x)\,\mathrm{e}^{-\mathrm{i}kx}\frac{\mathrm{d}x}{2\pi}$ its Fourier coefficients and introduce, for any $s > 0$, its H^s-Sobolev seminorm

$$\|f\|_{H^s} := \sqrt{\sum_{k\in\mathbb{Z}} |k|^{2s}|\hat{f}_k|^2}.$$

- $\mathcal{P}(\Lambda)$ stands for the set of probability measures on Λ, which is always equipped with its weak topology coming by duality with the bounded continuous functions on Λ. This topology is metrizable, for instance by the bounded Lipschitz (or Fortet-Mourier) distance

$$\mathrm{d}_{\mathrm{BL}}(\mu,\nu) := \sup_{\substack{\|f\|_{\mathrm{Lip}}\leq 1 \\ \|f\|_\infty \leq 1}} \int f(x)(\mu-\nu)(\mathrm{d}x). \tag{1.6}$$

  We also consider, for any $p \geq 1$, the stronger[♠] topologies associated with the Kantorovich/Wasserstein distance of order p

$$W_p(\mu,\nu) := \Big(\inf \iint d(x,y)^p \pi(\mathrm{d}x,\mathrm{d}y) \Big)^{1/p}, \tag{1.7}$$

  where the infimum ranges over the probability measures $\pi \in \mathcal{P}(\Lambda\times\Lambda)$ with respective marginal distributions μ and ν. If X, Y are random variables with respective laws μ_X, μ_Y, we also set for convenience $W_p(X,Y) := W_p(\mu_X,\mu_Y)$.

[♠] W_1 dominates d_{BL} since the Kantorovich-Rubinstein dual representation of W_1 states that

$$W_1(\mu,\nu) = \sup_{\|f\|_{\mathrm{Lip}} f\leq 1} \int f(x)(\mu-\nu)(\mathrm{d}x).$$

- The relative entropy of $\mu \in \mathcal{P}(\Lambda)$ with respect to $\nu \in \mathcal{P}(\Lambda)$ is defined by

$$H(\mu|\nu) := \int \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \log \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \, \mathrm{d}\nu$$

  when μ is absolutely continuous with respect to ν and set to $+\infty$ otherwise.

- $\mathsf{Conf}(\Lambda)$ stands for the set of (simple) point configurations on Λ, that is the locally finite subsets of Λ. One can identify a point configuration with its counting measure on Λ so that we can make $\mathsf{Conf}(\Lambda)$ a Polish space, when equipped with the weak topology coming by duality with the continuous functions having compact support. Finally, a point process on Λ is a probability measure on $\mathsf{Conf}(\Lambda)$.

- The empirical measure $\hat{\mu}_N$ is defined by $\hat{\mu}_N := \frac{1}{N}\sum_{j=1}^N \delta_{\mathbf{x}_j}$ for an associated point process $(\mathbf{x}_j)_{j=1}^N$ of a.s. N points that should be clear from the context. Its expectation $\mathbb{E}\hat{\mu}_N$ is defined by duality, $\int f \, \mathrm{d}\mathbb{E}\hat{\mu}_N := \mathbb{E}\int f \mathrm{d}\hat{\mu}_N$. We say that $\hat{\mu}_N$ converges almost surely (a.s) to μ when, *in any joint probability space*, we have the weak convergence $\hat{\mu}_N \to \mu$ with probability one. Note that this strong form of almost sure convergence is sometimes called complete convergence.

- $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ refers to the law of a centered complex Gaussian variable with variance σ^2, namely of $Z := \sigma(X + \mathrm{i}Y)/\sqrt{2}$ with X, Y i.i.d standard $\mathcal{N}(0,1)$ variables; note that the $\sqrt{2}$ is here to have $\mathbb{V}\mathrm{ar}(Z) := \mathbb{E}|Z|^2 - |\mathbb{E}(Z)|^2 = \sigma^2$.

1.3 Determinantal point processes

We now provide some background material on determinantal point processes (DPPs) as a preparation for the three next chapters.

DPPs have been initially popularized in statistical and mathematical physics, in relation with random matrix models and Fermionic systems from quantum physics. They indeed became famous particle systems when it has been understood that there are several instances of DPPs arising as universal microscopic limits for many unrelated particle systems, drawing the contours of a new universality class.

Another singular feature of DPPs is that most analytic quantities admit a closed formula in terms of the correlation kernel, allowing tractable computations and asymptotics (at least, in principle, and with enough bravery). Since exact sampling algorithms are available, DPPs thus yield a handy tool to model and use repulsive phenomena in applied mathematics.

Projection kernels. Let us fix some Polish space Λ with a reference measure μ. For our purpose it is enough[◇] to consider DPPs on Λ associated with projection

[◇] All the DPPs that appear here are associated with projection kernels. More generally one could consider DPPs with contraction kernels, but they can always be represented as a statistical mixture of projection kernel DPPs.

kernel only, namely kernels $K : \Lambda \times \Lambda \to \mathbb{C}$ satisfying

$$\int K(x,u)K(u,y)\,\mu(\mathrm{d}u) = K(x,y), \qquad x,y \in \Lambda. \tag{1.8}$$

If we see K as an operator acting on $L^2(\mu)$, that is $Kf(x) := \int K(x,y)f(y)\,\mu(\mathrm{d}y)$ and assume reasonable conditions on $K(x,y)$ for K to map $L^2(\mu)$ onto itself, then the condition (1.8) indeed yields that K is a projection operator. If we denote by $N \in \mathbb{N} \cup \{\infty\}$ the rank of the projection K and assume that it is Hermitian, then one can look for kernels of the form

$$K(x,y) = \sum_{k=1}^{N} \varphi_k(x)\overline{\varphi_k(y)} \tag{1.9}$$

for some orthonormal family (φ_k) of $L^2(\mu)$.

DPPs with finite rank projection kernel. A DPP on Λ with finite rank projection kernel K is a point process of a.s. $N := \mathrm{rank}(K)$ points $\mathbf{x}_1, \dots, \mathbf{x}_N$ on Λ with joint probability distribution

$$\mathrm{d}\mathbb{P}(x_1, \dots, x_N) = \frac{1}{N!} \det\Big[K(x_i,x_j)\Big]_{i,j=1}^{N} \prod_{j=1}^{N} \mu(\mathrm{d}x_j). \tag{1.10}$$

We implicitly (and will always) assume the kernel is positive definite so that $\mathbb{P}$ is indeed a positive (probability) measure. Note that this is always the case for Hermitian kernels (1.9).

For example, the $\mathsf{CUE}(N)$ is a DPP on the one-dimensional torus $\mathbb{T}$ with reference measure $\mu = \frac{\mathrm{d}x}{2\pi}$ and rank N Hermitian projection kernel (1.9) with $\varphi_k(x) = \mathrm{e}^{\mathrm{i}kx}$.

Correlation functions and infinite rank projections. By expanding the determinant in (1.10), integrating out the $N-k$ last variables for some $0 \le k < N$ and using (1.8), we obtain the identity

$$\frac{1}{(N-k)!} \int_{\Lambda^{N-k}} \mathbb{P}(x_1, \dots, x_k, \mathrm{d}x_{k+1}, \dots, \mathrm{d}x_N) = \det\Big[K(x_i,x_j)\Big]_{i,j=1}^{k}. \tag{1.11}$$

This yields that, for any nice test function $\varphi : \Lambda^k \to \mathbb{R}$,

$$\mathbb{E}\left[\sum_{i_1 \neq \dots \neq i_k} \varphi(\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_k})\right] = \int_{\Lambda^k} \varphi(x)\,\rho_k(x)\,\mu^{\otimes k}(\mathrm{d}x) \tag{1.12}$$

where we introduced

$$\rho_k(x_1, \dots, x_k) := \det\Big[K(x_i,x_j)\Big]_{i,j=1}^{k}, \tag{1.13}$$

the so-called k-th correlation function of the point process $\mathbf{x}_1, \dots, \mathbf{x}_N$.

When one considers an infinite rank projection kernel K, for which the definition (1.10) does not make sense anymore, a DPP is then defined as the point process uniquely characterized by the correlation functions given in (1.13) for all $k \geq 1$. Such a DPP yields infinite point configurations a.s.

An example is the Sine_2 process, which is the DPP on $\mathbb{R}$ with Lebesgue reference measure and infinite rank symmetric projection kernel K_{Sine} introduced in (1.5).

The hyperbolic GAF. In general, discovering a determinantal structure for some point process of interest is like finding a precious stone, or at least the promise to realize every computations you ever dreamed of for this model. One example lies in the study of invariant Gaussian analytic functions (GAFs), that are random analytic functions f such that $(f(z_1), \dots, f(z_m))$ is complex Gaussian vector for every $m \geq 0$ and every $z_1, \dots, z_m \in \Lambda$ for some definition domain $\Lambda \subset \mathbb{C}$. The so-called planar, hyperbolic and spherical GAFs are defined as the three unique one-parameter families of GAFs (up to the multiplication by a non-vanishing analytic function) whose distribution of their roots is invariant under the isometry of the complex plane $\mathbb{C}$, the hyperbolic/Poincaré disk $\mathbb{D}$, and the Riemann sphere $\mathbb{S}$ respectively [Hough et al., 2009]. An explicit formula exists for the these invariant GAFs, and it has the form

$$\mathsf{GAF}(z) = \sum_{k=0}^{\infty} \sqrt{\frac{c_k}{k!}}\, \xi_k\, z^k, \qquad z \in \Lambda, \tag{1.14}$$

where (ξ_k) is a sequence of i.i.d $\mathcal{N}_{\mathbb{C}}(0,1)$ random variables; the definition domain Λ and the deterministic sequence (c_k) is specified in the following table.

Isometries	Domain Λ	c_k	Parameter
$\mathbb{C}$	$\mathbb{C}$	ℓ^k	$\ell > 0$
$\mathbb{D}$	$\lvert z\rvert < 1$	$\Gamma(k+\alpha+1)$	$\alpha > -1$
$\mathbb{S}$	$\mathbb{C}$	$N(N-1)\cdots(N-k+1)\mathbf{1}_{k\leq N}$	$N \in \mathbb{N}^*$

A beautiful result of Peres and Virág [2005] is that the point process on the unit disk given by the zeros of the hyperbolic GAF with parameter $\alpha = 0$ is determinantal with Lebesgue reference measure and infinite rank projection Hermitian kernel

$$K_{\mathbb{D}}(x,y) := \frac{1}{\pi(1-x\bar{y})^2}\,, \tag{1.15}$$

which is the reproducing kernel for the Bergman space $A^2(\mathbb{D})$.

In Chapter 2 we will see how invariant GAFs, and in particular the hyperbolic one, show up in signal processing in relation to the denoising problem. This opens a door for new denoising algorithms based on this determinantal structure.

Hyperuniform estimators. If we set for convenience

$$\mathcal{X}_f^N := \sum_{j=1}^N f(\mathbf{x}_j) \tag{1.16}$$

for a point process $\mathbf{x}_1, \dots, \mathbf{x}_N$, then (1.3) states that $\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = \mathcal{O}(1)$ for the CUE provided that f is smooth enough. This should be compared to the $\mathcal{O}(N)$ order variance in the setting of i.i.d random variables. From this perspective one can say that the CUE is a hyperuniform[♠] point process. Let us stress that hyperuniform particle systems have a huge potential for applications, since the variance of estimators is often synonymous of precision error, and it is thus natural to wonder if DPPs can provide a handy class of such hyperuniform point processes.

For a general DPP $\mathbf{x}_1, \dots, \mathbf{x}_N$ with projection kernel K_N and reference measure μ on Λ, equations (1.12)–(1.13) with $k = 1$ and $k = 2$ yield together

$$\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = \int |f(x)|^2 K_N(x,x)\,\mu(\mathrm{d}x) - \iint f(x)\overline{f(y)} K_N(x,y)K_N(y,x)\,\mu(\mathrm{d}x)\mu(\mathrm{d}y) \tag{1.17}$$

for every reasonable function $f : \Lambda \to \mathbb{C}$. If the particles are scaled so that their macroscopic limit is non-trivial, then the leftmost integral is of order N and thus we must design K_N such that the rightmost integral is at least of order N to achieve hyperuniformity.

In Chapter 3 we provide a class of hyperuniform DPPs on the hypercube $[-1,1]^d$ for every dimension $d \geq 1$, constructed by means of multivariate orthogonal polynomials. We illustrate how this can be of interest for numerical integration.

Sampling projection DPPs. A nice feature of DPPs with rank N projection Hermitian kernels is that an exact sampling algorithm is available, see [Hough et al., 2009]. A modest contribution made in [P10] is a quick way to derive this algorithm that doesn't require the kernel K to be Hermitian: consider the mean probability distribution

$$\eta_1(\mathrm{d}x_1) := \frac{1}{N} K(x_1,x_1)\mu(\mathrm{d}x_1) \tag{1.18}$$

and the conditional distributions defined, for $1 \leq k < N$ and $x_1, \dots, x_k \in \Lambda$, by

$$\eta_{k+1}(\mathrm{d}x_{k+1}|x_1,\dots,x_k) := \frac{1}{N-k} \frac{\det\left[K(x_i,x_j)\right]_{i,j=1}^{k+1}}{\det\left[K(x_i,x_j)\right]_{i,j=1}^{k}} \mu(\mathrm{d}x_{k+1}). \tag{1.19}$$

[♠]The notion of hyperuniformity is usually defined at the microscopic scale [Torquato, 2016]: a point process on infinite configurations on $\mathbb{R}^d$ is hyperuniform when the variance of the number of particles in a box $[-L,L]^d$ is of smaller order than the volume L^d of that box, which is the variance for a Poisson process, as $L \to \infty$. Up to a smoothing of the characteristic function of that box, the macroscopic analog for the definition of hyperuniformity is thus to say that, for a point process $\mathbf{x}_1, \dots, \mathbf{x}_N$ on $\mathbb{R}^d$ scaled such that $\mathbb{E}\hat{\mu}_N$ has a non-trivial large N limit, we have $\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = o(N)$ as $N \to \infty$ for f a (smooth) indicator function of a compact set.

Using the Schur determinant identity yields a formula for the density of $\eta_{k+1}(\mathrm{d}x|x_1,\dots,x_k)$ with respect to $\mu(\mathrm{d}x)$ that is cheaper to implement, given by

$$\frac{1}{N-k}\left(K(x,x)-\begin{bmatrix}K(x,x_1)\\ \vdots \\ K(x,x_k)\end{bmatrix}^T\left(\Big[K(x_i,x_j)\Big]_{i,j=1}^{k}\right)^{-1}\begin{bmatrix}K(x_1,x)\\ \vdots \\ K(x_k,x)\end{bmatrix}\right). \tag{1.20}$$

By integrating (1.20) against $\mu(\mathrm{d}x)$ and using (1.8), we see the η_k's are indeed probability measures. Since (1.18)–(1.19) clearly yield for $\mathbb{P}$ as in (1.10),

$$\mathrm{d}\mathbb{P}(x_1,\dots,x_N)=\eta_1(\mathrm{d}x_1)\prod_{k=2}^{N}\eta_k(\mathrm{d}x_k|x_1,\dots,x_{k-1}),$$

sampling $\mathbf{x}_1,\dots,\mathbf{x}_N$ from $\mathbb{P}$ thus amounts to sample $\mathbf{x}_1$ with distribution η_1, then $\mathbf{x}_2$ with distribution $\eta_2(\,\cdot\,|\mathbf{x}_1)$, then $\mathbf{x}_3$ with distribution $\eta_3(\,\cdot\,|\mathbf{x}_1,\mathbf{x}_2)$, etc. This chain rule is the same as the one of [Hough et al., 2009], see [P8, Section 2.4].

Note that, by positivity of the kernel, *if K is Hermitian* then (1.20) is bounded from above by $\frac{1}{N-k}K(x,x)$, making the use of rejection sampling possible to sample from the η_k's once one knows how to sample from the mean distribution η_1. Thus, together with the matrix inversion in (1.20), to sample from a rank N projection kernel DPP costs $\mathcal{O}(N^3)$ plus N rejection sampling steps. If K is not Hermitian, the rejection sampling step is more tricky and depends on the model.

Let us mention the very nice library DPPy developed for Python by Gautier et al. [2019], which includes several sampling algorithms for DPPs and matrix models.

1.4 Coulomb gases

Another interesting class of repulsive particle systems are the Coulomb and log-gases, which represent canonical models in statistical physics for particle systems in electrostatic interaction.

Coulomb law. Electrostatics is the physical phenomena that makes cat's fur spiky when you rub amber against it. Since the ancient greek word for amber is "élektron", at least the ancient Greeks understood it was more a property of amber stones than of cats. Charles-Augustin Coulomb was however able quantify in 1784, using an electric torsion balance, the repulsion force between two charged objects: its amplitude is decaying as their squared distance (Coulomb's law)[♡], exactly like the gravitational attraction. Using the apparent isotropy of electrostatic interaction, this yields that the repulsion force created by a unit charge at $0\in\mathbb{R}^3$ felt by another unit charge at $x\in\mathbb{R}^3$ equals to

$$\vec{F}=\frac{x}{|x|^3}$$

[♡]To be fair to the ancient Greeks, in his first memoir Coulomb made *two* experiments and then concluded, with 8% error, that "la force répulsive de deux petits globes électrisés de la même nature d'électricité est en raison inverse du carré de la distance du centre des deux globes."

up to a multiplicative constant. This force derives from the potential $g(x) = |x|^{-1}$ in the sense that $\vec{F} = \nabla g(x)$. In particular, the energy of a system of N unit charges at positions $x_1, \dots, x_N \in \mathbb{R}^3$ in presence of an external potential $V : \mathbb{R}^3 \to \mathbb{R}$ reads

$$H_N(x_1, \dots, x_N) = \sum_{i<j} g(x_i - x_j) + \sum_{j=1}^{N} V(x_j). \tag{1.21}$$

The potential V is here to model an external interaction with the system of charges.

Canonical ensembles and Coulomb gas on $\mathbb{R}^3$. Given an energy functional (or Hamiltonian) $H : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and a constant $C \in (\inf H, \sup H)$, the micro-canonical ensemble refers in statistical physics to the probability distribution $\mathbb{P}$ on $\mathbb{R}^n$ that maximizes its Shannon entropy $\mathbb{E}_{\mathbb{P}}[\log \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}x}]$ within the probability measures having a given mean energy $\mathbb{E}_{\mathbb{P}}[H] = C$; it thus represents the state of a physical system at a given energy level with no extra symmetry/information. The solution of this optimization problem is unique for reasonable energies H and its density reads $\frac{1}{Z_\beta} \mathrm{e}^{-\beta H}$ for some $\beta > 0$, where the so-called partition function $Z_\beta > 0$ is a normalisation constant. The parameter β is interpreted as the inverse temperature of the system and there is a one-to-one correspondance between the constant energy value C and β. Using the dual formulation for this variational problem, its solution also minimizes the functional $\mathbb{P} \mapsto \mathbb{E}_{\mathbb{P}}[H] + \beta \mathbb{E}_{\mathbb{P}}[\log \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}x}]$, whose minimizer is rather called canonical ensemble in statistical physics. In the following, we thus refer to $\frac{1}{Z_\beta} \mathrm{e}^{-\beta H}$ as the *canonical ensemble* associated with the energy H at inverse temperature β.

The Coulomb gas is defined as the canonical ensemble associated with the energy (1.21) of N unit charges interacting according to the Coulomb law in an external potential V.

Coulomb potential/gas on $\mathbb{R}^d$. Alternately, in the spirit of the local form of Gauss law, one can recover the potential $g(x) = |x|^{-1}$ by solving the Poisson equation

$$\Delta g = -4\pi \delta_0$$

in the space of Schwarz distributions on $\mathbb{R}^3$. By extension, the Coulomb potential g of the Euclidean space $\mathbb{R}^d$ is defined for any $d \geq 2$ by

$$g(x) := \begin{cases} \log \dfrac{1}{|x|} & \text{if } d = 2, \\ \dfrac{1}{|x|^{d-2}} & \text{if } d \geq 3, \end{cases} \tag{1.22}$$

since one can check that it satisfies the equation

$$\Delta g = -c_d \, \delta_0 \tag{1.23}$$

where c_d is a positive constant given by

$$c_d := \begin{cases} 2\pi & \text{if } d = 2, \\ (d-2)\mathsf{Vol}(\mathbb{S}^{d-1}) & \text{if } d \geq 3, \end{cases} \quad \text{with} \quad \mathsf{Vol}(\mathbb{S}^{d-1}) := \frac{2\pi^{d/2}}{\Gamma(d/2)}.$$

The Coulomb gas on $\mathbb{R}^d$ is then defined as the canonical ensemble associated with the energy (1.21) where g is now replaced by the Coulomb potential of $\mathbb{R}^d$ and $V : \mathbb{R}^d \to \mathbb{R}$ is an appropriate external potential. It should be noticed that the Coulomb gas is a repulsive particle system since the energy (1.21) becomes arbitrary high when two particles get closer, due to the singularity of g at zero, and configurations with high energy are unlikely by definition of the canonical ensemble.

We will study the large N limit of the empirical measure $\hat{\mu}_N$ for such Coulomb gases in Chapter 5. More precisely, we explain how to derive fast concentration inequalities in the spirit of (1.2) for $\hat{\mu}_N$ around the equilibrium measure of the system, thanks to a Coulomb analog of the Talagrand transportation inequality (T_1).

Coulomb potential on a compact manifold. We could also extend the definition of the Coulomb potential/gas to any Riemannian manifold M, where the Laplace-Beltrami operator replaces the usual Laplacian Δ. Note however that we have lost in general the translation invariance of the ambiant space. Moreover, the analog of the Poisson equation $\Delta g = -\delta_x$ for any fixed $x \in M$ has no solution when the manifold is compact (since $1 \in \mathcal{C}_c^\infty(M)$). This problem can be however tackled by adding a uniformly charged neutralizing background, that is to look for a so-called Green function $g(x, y)$ which is symmetric and satisfies $\Delta g(x, \cdot) = \sigma - \delta_x$ in distributions for every $x \in M$, where σ is the uniform probability measure on M. For instance, if $M = \mathbb{S}$ is the unit sphere of $\mathbb{R}^3$ then the Green function is given by $g(x, y) \propto \log \|x - y\|^{-1}$, where $\|\cdot\|$ stands for the Euclidean norm of $\mathbb{R}^3$. For higher dimensional spheres and higher complex projective spaces there are closed formulas but they are more involved [Beltrán et al., 2019]. However, for general compact manifolds the Green function $g(x, y)$ behaves like its Euclidean analog as $y \to x$.

As in $\mathbb{R}^d$, one can define the Coulomb gas on a compact Riemannian manifold by taking the canonical ensemble associated with the Coulomb energy (1.21) after replacing "$g(x_i - x_j)$" in its expression by the Green function $g(x_i, x_j)$ and some potential $V : M \to \mathbb{R}$.

In Chapter 6 we quantify how much the Coulomb gas on the sphere $\mathbb{S}$ can provide well-spaced points depending on the temperature. More precisely, we show that in the low temperature regime $\beta \sim N$ the N particles of the Coulomb gas on the sphere have minimal Coulomb energy up to a $\log N$ error, which is the precision required in Smale's 7th problem, with exponentially high probability.

Restriction of Coulomb gases to lower dimensional spaces. Another canonical ensemble popularized by random matrix theory is the log-gas, which a two-

dimensional Coulomb gas restricted to a one dimensional subset. For instance, given $\beta > 0$, the Gaussian-β-Ensemble (GβE) refers to the N particles on $\mathbb{R}$ with distribution

$$\frac{1}{Z_{N,\beta}} \prod_{1 \le i < j \le N} |x_i - x_j|^\beta \prod_{j=1}^{N} \mathrm{e}^{-N\beta x_j^2/4} \mathrm{d}x_j \,. \tag{GβE}$$

When $\beta = 1$ (resp. $\beta = 2$, resp. $\beta = 4$), this is the joint eigenvalue distribution induced by the standard Gaussian measure on the real vector space of $N \times N$ symmetric (resp. complex Hermitian, resp. quaternionic Hermitian) matrices.

Another interesting example is the CUE (again), which turns out to match with the restriction of the Coulomb gas of the 2-torus $\mathbb{T} \times \mathbb{T}$ (seen as a Riemannian submanifold of $\mathbb{R}^3$) with no potential at inverse temperature $\beta = 2$ restricted to $\mathbb{T} \simeq \mathbb{T} \times \{0\}$ [Borodin and Serfaty, 2013]. Similarly, the canonical ensemble on $\mathbb{T}$ obtained by this restriction at arbitrary inverse temperature $\beta > 0$ is called the Circular-β-Ensemble (CβE) and is given by

$$\frac{1}{Z_{N,\beta}} \prod_{1 \le i < j \le N} |\mathrm{e}^{\mathrm{i}x_i} - \mathrm{e}^{\mathrm{i}x_j}|^\beta \prod_{j=1}^{N} \frac{\mathrm{d}x_j}{2\pi}. \tag{CβE}$$

It turns out the CLT (1.3) also holds true for any fixed $\beta > 0$, with the same small order fluctuations. However, this breaks down at $\beta = 0$ where we recover the setting of i.i.d uniform random variables on $\mathbb{T}$, where the classical CLT holds instead.

In Chapter 7 we describe a transition that arise in the high temperature regime $\beta \sim 1/N$, where we establish a CLT with a limiting variance that interpolates between the fixed temperature setting ($H^{1/2}$) and the i.i.d setting (L^2).

At the microscopic level, we have seen that the CUE, or C2E, converges to the Sine_2 point process which is completely characterized by being the DPP with kernel K_{Sine} defined in (1.5). For general $\beta > 0$ there is also a microscopic limit for the CβE known as the Sine_β process, but it does not seem to be a DPP when $\beta \neq 2$, making its description quite uneasy.

In Chapter 8 we describe the Sine_β process as an infinite Gibbs measure associated with the logarithmic interaction on $\mathbb{R}$ by means of the Debrushin-Lanford-Ruelle (DLR) formalism. This yields further information on Sine_β, for instance that it is number-rigid for any $\beta > 0$.

Part I – Determinantal point processes

Chapter 2

DPPs for signal processing

Based on the joint work [P12] with Rémi Bardenet (CNRS, Université de Lille).

What is the musical score of a white noise? A central problem in signal processing is denoising, that is to extract the true signal from an ambiant noise, and a classical tool to do so is time-frequency analysis, which we briefly review; for more information see the reference books [Cohen, 1995; Flandrin, 1998; Gröchenig, 2001].

To start with an example, say that a signal is a member f of the Hilbert space $L^2(\mathbb{R}, \mathbb{C})$, where $\mathbb{R}$ refers to the time variable. A standard way to encode this signal is to use its spectrogram, which associates to each time and frequency a positive number, that is a continuous generalization of a musical score. It is however impossible to localise exactly a signal both in time and frequency, due to the Heisenberg uncertainty principle. One possible strategy is to localise the signal in time by multiplying it with a window function and then take its Fourier transform, that is to consider

$$V_g f(t,\omega) := \int_{\mathbb{R}} \overline{f(x)} g(x-t) \mathrm{e}^{\mathrm{i}\omega x} \mathrm{d}x = \langle f, M_\omega T_t g \rangle_{L^2}$$

where g is localized around zero, and where we introduced the multiplication and translation operators $M_\omega h = \mathrm{e}^{\mathrm{i}\omega \cdot} h$ and $T_t h = h(\cdot - t)$ for $t, \omega \in \mathbb{R}$. The linear transformation V_g yields the short-time Fourier transform (STFT) $V_g f$ of the signal f, and the spectrogram is finally defined by $|V_g f(t,\omega)|^2$. An interesting window function g is the standard Gaussian function: it is the most localized function in time and frequency in the sense that it achieves equality in the uncertainly principle inequality. It turns out it is more convenient to see the STFT as a function of one complex variable $z := t + \mathrm{i}\omega$, that we still denote by $V_g f$.

Next, given a white noise ξ on $L^2(\mathbb{R}, \mathbb{C})$, representing the ambiant noise, and $\sigma > 0$, suppose one observes the signal $f + \sigma\xi$. A prototypical approach to denoise this signal is to compute its STFT, then restrict it to regions of the complex plane where the spectrogram is larger than a given threshold, and finally use an inversion formula to recover the signal. However, to quantify the error made with this recipe one should know what is the typical amplitude of the white noise's part of the spectrogram, or the law of its extremal values, but such theoretical results don't

seem within reach yet. Still, most methodologies have focused on detecting and processing the local maxima of the spectrogram.

Zeros of spectrograms. Flandrin [2015] has recently observed that, when it is applied to a Gaussian white noise, the STFT with Gaussian window is a random analytic function identifiable by its zeros thanks to a Weierstrass-Hadamard product formula. Flandrin [2015] then proposed to use the point process formed by these random zeros in filtering and reconstruction of signals in noise. In particular, Flandrin [2015] empirically observed that the zeros of this spectrogram spread out very evenly on the time-frequency plane, with regular Voronoi tessellations. Pushing further Flandrin's investigation, Bardenet, Flamant, and Chainais [2018] proved in the same Gaussian setting that the zeros of the white noise's spectrogram have the same distribution as the zeros of the planar Gaussian analytic function (GAF). Bardenet et al. [2018] then used known probabilistic results on the planar GAF to inform the design of signal reconstruction procedures in the spirit of [Flandrin, 2015]. In fact, they were hunting for DPP, since this would have yield more precise signal reconstruction procedures due to the exact formulas available for the correlation functions, and they were very close but the invariant GAF whose zeros form a DPP is not the planar but the hyperbolic one. A natural problem is thus to find other natural transformations in signal processing that maps a Gaussian white noise onto the hyperbolic GAF, which we addressed in [P12], and invariant GAFs in general.

Zeros of scalograms and the hyperbolic GAF. Another class of signals of importance are the so-called analytic signals, which are complex signals $f \in L^2(\mathbb{R}, \mathbb{C})$ without negative frequencies, namely such that $\mathrm{Supp}(\hat{f}) \subset \mathbb{R}_+$ with $\hat{f}$ the Fourier transform of f. We denote by $H_2(\mathbb{R})$ the space of analytic signals, which can also be obtained as the boundary of the Hardy space $H_2(\mathbb{C}_+)$ of the upper half-plane $\mathbb{C}_+ := \{z \in \mathbb{C} : \ \mathfrak{Re}(z) > 0\}$.

The analog of the time-frequency transform of interest here is a time-scale transform, or continuous wavelet transform, popularized by Daubechies and Paul [1988] and defined by

$$W_\beta f(t,s) := \langle f, T_t D_s \psi_\beta \rangle_{L^2}$$

for $f \in H_2(\mathbb{R})$, $t \in \mathbb{R}$ and $s > 0$. Here, T_t is again the translation operator but now the dilation operator $D_s f := s^{-1/2} f(\cdot/s)$ comes into play, and ψ_β is the so-called mother wavelet defined by $\psi_\beta(x) := (x + \mathrm{i})^{-\beta}$. Again, we see $W_\beta f$ as a function of one complex variable $z = t + \mathrm{i}s$ on the upper half-plane $\mathbb{C}_+$. The next result can be found in [P12].

Theorem 2.1. *The wavelet transform $W_\beta \xi$ of a Gaussian white noise ξ on $H_2(\mathbb{R})$ is, up to an explicit conformal mapping $\mathbb{C}_+ \to \mathbb{D}$, the hyperbolic GAF of parameter $\alpha = 2\beta$. In particular, the point process given by the zeros of $W_\beta \xi$ is invariant under the isometries of the Poincaré half-plane and, when $\beta = 0$, it is determinantal with kernel*

$$K_{\mathbb{C}_+}(x,y) := \frac{-1}{\pi(x - \overline{y})^2}.$$

We note that $K_{\mathbb{C}_+}(x,y)$ is the reproducing kernel associated with the Bergman space $A^2(\mathbb{C}_+)$. A similar observation has been simultaneously and independently made by Abreu, Haimi, Koliander, and Romero [2018].

In fact, we provide in [P12] a systematic approach to derive this type of results.

A general framework. Let us take some distance to the initial problem and look at the general structure. Let $\mathcal{H}$ be a complex separable Hilbert space, representing the signals, and consider a linear transformation $\mathscr{L} : \mathcal{H} \to \mathcal{A}(\Lambda)$ taking its values in the space of analytic functions on a domain $\Lambda \subset \mathbb{C}$. To identify transformations such that $\mathscr{L}\xi$ is an invariant GAF, the approach from [P12] is quite simple at the formal level and goes as follows. Since for any orthonormal basis (f_k) of $\mathcal{H}$ a Gaussian white noise ξ on $\mathcal{H}$ can be represented as

$$\xi \overset{\text{law}}{=} \sum_k \xi_k \, f_k \tag{2.1}$$

with i.i.d $\mathcal{N}_{\mathbb{C}}(0,1)$ random variables (ξ_k), we have by linearity of $\mathscr{L}$,

$$\mathscr{L}\xi(z) \overset{\text{law}}{=} \sum_k \xi_k \, \mathscr{L}f_k(z). \tag{2.2}$$

To obtain an invariant GAF thus amounts to look for transformations $\mathscr{L}$ and basis (f_k) such that $\mathscr{L}f_k(z) = \sqrt{\frac{c_k}{k!}} z^k$. By looking for $\mathscr{L}$'s of the form $\mathscr{L}f(z) = \langle K_z, f\rangle$ for every $z \in \Lambda$, this leads to take

$$K_z := \sum_k \sqrt{\frac{c_k}{k!}} z^k \, f_k \tag{2.3}$$

to succeed, and hope there exists closed known formulas for (2.3).

To put this argument on a firm ground, we first have to deal with the problem that (2.1) is not true when $\mathcal{H}$ is infinite dimensional, since the right hand side diverges in $\mathcal{H}$ a.s; note that we did not specify what we meant by "Gaussian white noise" on $L^2(\mathbb{R},\mathbb{C})$ or $H_2(\mathbb{R})$ in the previous discussion. We moreover need to show the right hand side (2.2) is a well defined random analytic function. To do so we use Gross construction of abstract Wiener spaces and rigorously define ξ as a random distribution on the space Θ obtained as the closure of $\mathcal{H}$ with the weaker norm $\|f\|_\Theta^2 := \sum_k (1+k^2)^{-1} |\langle f, f_k\rangle|^2$. We also provide concentration estimates when approximating this abstract white noise by the finite sum obtained by truncation of the right hand side of (2.1) in the space Θ [P12, Section 5].

Connexion to classical orthogonal polynomials. To recover the planar GAF in the setting of the STFT as in [Bardenet et al., 2018], let's take $\mathcal{H} = L^2(\mathbb{R},\mathbb{C})$ with basis (f_k) given by the Hermite functions, constructed from the orthogonal polynomials (OPs) with respect to the Gaussian weight $\mathrm{e}^{-x^2/2}$ on $\mathbb{R}$. Up to change of variables involving a scaling parameter ℓ and non-vanishing multiplicative terms, the STFT V_g in a Gaussian window is know to be match the Bargmann transform,

that we denote by $\mathscr{L}$ (see the table below). Now, the fact that $\mathscr{L}$ has a kernel (2.3) with f_k given by the Hermite functions and $c_k = \ell^N$, namely that $\mathscr{L}$ maps the white noise onto the planar GAF, turns out to be equivalent to an identity satisfied by the generating function of the Hermite polynomials.

Similarly, in the setting of analytic signals, one takes $\mathcal{H} = H_2(\mathbb{R})$ and for basis (f_k) the inverse Fourier transform of the Laguerre functions, constructed with another family of classical OPs associated with the Gamma weight $x^{2\beta}\mathrm{e}^{-x}$ on $\mathbb{R}_+$. The Daubechies-Paul wavelet transform W_β, up to cosmetic modification again, can be represented as the Bergman transform $\mathscr{L}$ of the upper half-plane, and digging out the hyperbolic GAF reduces to a known identity for the generating function of Laguerre polynomials.

The reader familiarized with classical OPs and the Askey scheme should now want to try every generating function identities for OPs she/he knows to see what is going on, and this is what we partially did in [P12]. For instance, one recovers the spherical GAF by taking Krawtchouk polynomials, and it comes with a new discrete transform on the finite dimensional space $\mathcal{H} = \mathbb{C}^N$. We also recover again the planar and hyperbolic GAF from discrete transforms on the Hilbert space $\mathcal{H} = \ell^2(\mathbb{N}, \mathbb{C})$ thanks to Charlier and Meixner OPs; we summarize these findings in the next table. Let us stress that in practice a signal is eventually discretized to be treated by computer, and thus transformations that are discrete from the start may be of interest for the users.

$\mathcal{H}$	Transform $\mathscr{L}$		Basis (f_k)	GAF
$L^2(\mathbb{R}, \mathbb{C})$	$\dfrac{\mathrm{e}^{-z^2/2}}{\pi^{1/4}} \displaystyle\int_{\mathbb{R}} \overline{f(x)}\, \mathrm{e}^{\sqrt{2}xz - x^2/2} \mathrm{d}x$	(Bargmann)	Hermite	$\mathbb{C}$
$\ell^2(\mathbb{N}, \mathbb{C})$	$\displaystyle\sum_{x\in\mathbb{N}} \overline{f(x)} \frac{z^x}{\sqrt{x!}}$	(New)	Charlier	$\mathbb{C}$
$H_2(\mathbb{R})$	$\dfrac{1}{(1-z)^{2\beta+1}} \displaystyle\int_{\mathbb{R}_+} \overline{\hat{f}(x)}\, x^\beta \mathrm{e}^{-\frac{x}{2}\frac{1+z}{1-z}} \mathrm{d}x$	(Bergman)	Laguerre	$\mathbb{D}$
$\ell^2(\mathbb{N}, \mathbb{C})$	$\displaystyle\sum_{x\in\mathbb{N}} \overline{f(x)} \sqrt{\frac{\Gamma(x+\alpha+1)}{x!}}\, z^x$	(New)	Meixner	$\mathbb{D}$
$\mathbb{C}^{N+1}$	$\displaystyle\sum_{x=0}^{N} \overline{f(x)} \sqrt{\binom{N}{x}}\, z^x$	(New)	Krawtchouk	$\mathbb{S}$

Recognizing elementary signals. The picky reader may have noticed that the form of the three (new) discrete transforms in the previous table did not require any sophisticated constructions with OPs to be introduced, since one could have taken the canonical basis $e_k(x) = \mathbf{1}_{k=x}$ to make the findings of the associated GAF obvious. However, contrary to the the canonical basis, the bases constructed with OPs behave like sinusoidal signals and thus are more likely to represent real life signals, and the fact that the new transforms map these OP bases onto monomials

(up to an explicit conformal mapping) is remarkable. In particular, this makes the recognition of elementary signals f_k's on the "spectrograms" $|\mathscr{L} f_k(z)|^2$ quite easy.

Perspectives. First, it would be interesting to see if one can design an efficient denoising algorithm for the Daubechies-Paul wavelet transform based on the determinantal structure and the invariance under the hyperbolic isometries, which is just an exotic kind of stationarity. For instance the (gap) probability to have no zero of $W_0\xi(z)$ in a given compact can be evaluated and thus serve as a benchmark to detect signal in noise. It would be interesting to do the same for the discrete transforms we introduced above, especially the one on $\mathbb{C}^N$ since it bypasses the errors made in the discretization process.

At the theoretical level, it would be instructive to see if using other generating function identities leads to worthy results. For instance, one can check that a second identity involving the Laguerre OPs yields a transformation $\mathscr{L}$ known as the Hankel transform, which is the Fourier transform for rotationally invariant functions on $\mathbb{R}^d$, but the GAF obtained is not one of the three invariant GAFs. Perhaps it satisfies an invariance with respect to transforms that are more exotic than the isometries for the three complex geometries with constant curvature? Similarly, using an identity for Chebyshev OPs one recovers the free Bargmann transform, which is the analog of the Bargmann transform in free probability [Biane, 1997], and it maps a Gaussian white noise onto the hyperbolic GAF, and so do more generally transformations related to Jacobi OPs. But there are other generating identities of Jacobi OPs (see e.g. [Ismail, 2005]) leading to other hypergeometric GAFs with unknown invariance properties, at least to the author.

Finally, having in mind that Fourier theory can be understood as a particular case of representation theory, it would be interesting to see which groups are behind the discrete transforms we introduced above (especially the one on $\mathbb{C}^N$). Indeed, the two-parameter family $(M_\omega T_t g)_{t,\omega\in\mathbb{R}}$ appearing in the STFT in a Gaussian window is known to provide an irreducible representation of the (reduced) Heisenberg group acting on $L^2(\mathbb{R},\mathbb{C})$, and this structure is rich enough to recover most properties of Gaussian time-frequency analysis [Gröchenig, 2001, Chapter 9]. A similar setting arises in the Daubechies-Paul wavelet transform, but with the group of affine transformations of $\mathbb{R}$ instead of the Heisenberg group. The unveiled connexion with classical OPs makes believe that a similar rich group theoretic structure should be present for the three discrete transforms introduced above, which may provide further motivation to study them.

Chapter 3

DPPs for Monte Carlo algorithms

Based on the joint work [P8] with Rémi Bardenet (CNRS, Université de Lille).

Monte Carlo algorithms. Given a measure μ on a Polish space Λ and a class of test functions $f : \Lambda \to \mathbb{R}$, a quadrature algorithm outputs N nodes $\mathbf{x}_1, \dots, \mathbf{x}_N \in \Lambda$ and weights $w_1, \dots, w_N > 0$ such that the error $\mathcal{E}_N(f)$ in the approximation

$$\int_\Lambda f(x)\mu(\mathrm{d}x) \simeq \sum_{j=1}^N w_j f(\mathbf{x}_j) \tag{3.1}$$

is reasonably decaying as $N \to \infty$. The error typically depends on the regularity of the integrand f as well. When $\Lambda = [-1, 1]$ and $\mu(\mathrm{d}x) = \mathrm{d}x$ the simplest idea, at the heart of the Riemann integration, is to take for nodes a uniform grid with mesh $1/N$ and uniform weights $1/N$, since a Taylor expansion then yields $\mathcal{E}_N(f) = \mathcal{O}(\sigma_f\, N^{-1})$ provided that $\sigma_f := \|f'\|_{L^\infty} < \infty$. For the d-hypercube $\Lambda = [-1, 1]^d$, the same $1/N$-mesh discretization yields the same error order, but using N^d nodes. This means that $\mathcal{E}_N(f) = \mathcal{O}(\sigma_f\, N^{-1/d})$, which is not reasonable even for rather small dimensions; one refers to this phenomena as the "curse of dimensionality".

At this point one should stress that there is an important need in real world applications for being able to integrate functions on rather large dimensional spaces. An example, among others, is Bayesian statistics, which is used more and more often to draw conclusions in biology and medicine♣, and where an answer typically takes the form of an integral over the parameter space of the model under study.

One way to repel the curse of dimensionality is provided by the classical CLT, since taking i.i.d uniform random nodes $\mathbf{x}_j$'s on $[-1, 1]^d$ and uniform $1/N$ weights thus yields a typical error $\|\mathcal{E}_N(f))\|_{L^2} \sim \mathcal{O}(\sigma_f\, N^{-1/2})$ with $\sigma_f^2 := \mathbb{V}\mathrm{ar} f(\mathbf{x}_1)$ *for every dimension* $d \geq 1$, and moreover that $N^{1/2}\mathcal{E}_N(f) \to \mathcal{N}(0, \sigma_f^2)$ in law as $N \to \infty$, leading to asymptotic confidence intervals. Put otherwise, with the same number of points, i.i.d random nodes explore the d-hypercube more efficiently than a uniform grid as soon as $d \geq 3$. This observation is a vanilla version of Monte Carlo algorithms, which refer to quadrature algorithms with random nodes/weights and often

♣For instance, requesting "Bayesian statistic" on PubMed outputs more than 17000 articles.

designed by using appropriate Markov processes, like in the Metropolis-Hastings algorithm. However, an error of order $N^{-1/2}$ can be still not satisfying whenever N is relatively small, having in mind applications where evaluating a function at a node represents a costly physical or biological experiment (in time and/or money). Note that in dimension one, i.i.d nodes are not performing as well as a uniform grid, but nodes generated from the CUE do, according to (1.3). Thus DPPs seem to be an interesting class of random nodes to improve Monte Carlo algorithms beyond the $N^{-1/2}$ frontier. Indeed, aside from the CUE, it is reasonable to expect that repulsive particle systems tend to explore more efficiently the ambiant space than i.i.d nodes.

Hyperuniform DPPs. Recalling the introductory chapter and the notation (1.16), we are left to design DPPs with rank N projection kernel K_N and reference measure, say, μ such that $\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = o(N)$ for a reasonable class of test functions f. One way to do so is to assume for simplicity[♡] that K_N is Hermitian of the form (1.9), so that the variance formula (1.17) yields (with $\langle f, g\rangle := \int f\,\overline{g}\,\mathrm{d}\mu$)

$$\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = \sum_{k=1}^{N} \langle f\varphi_k, f\varphi_k\rangle - \sum_{k,\ell=1}^{N} \big|\langle f\varphi_k, \varphi_\ell\rangle\big|^2.$$

Assume further that $(\varphi_k)_{k\in\mathbb{I}}$ is an orthonormal basis of $L^2(\mu)$ for some discrete set of indices $\mathbb{I}$ that contains $\{1,\dots,N\}$. If the test function f is such that $f\varphi_k \in L^2(\mu)$ for all k, then we can plug the expansion $f\varphi_k = \sum_{\ell\in\mathbb{I}}\langle f\varphi_k, \varphi_\ell\rangle\,\varphi_\ell$ into the first sum of the previous equation so as to obtain

$$\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = \sum_{k=1}^{N} \sum_{\ell\notin\{1,\dots,N\}} \big|\langle f\varphi_k, \varphi_\ell\rangle\big|^2, \tag{3.2}$$

and a similar formula holds for $\mathbb{C}\mathrm{ov}(\mathcal{X}_f^N, \mathcal{X}_g^N)$ by polarization. Thus the variance/covariance asymptotics as $N\to\infty$ depends on the growth of the coefficients $\langle f\varphi_k, \varphi_\ell\rangle$ for indices k,ℓ in the window $\{1,\dots,N\}\times\mathbb{I}\setminus\{1,\dots,N\}$. Moreover, since

$$\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = \sum_{m,n\in\mathbb{I}} \hat{f}_m\overline{\hat{f}_n}\,\mathbb{C}\mathrm{ov}(\mathcal{X}_{\varphi_m}^N, \mathcal{X}_{\varphi_n}^N) \tag{3.3}$$

where $\hat{f}_m := \langle f, \varphi_m\rangle$, this shows that $\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N]$ is an explicit function of the generalized Fourier coefficients $\hat{f}_m$ and the product formula coefficients $\langle\varphi_m\varphi_k, \varphi_\ell\rangle$. Thus, when a product formula is known for the φ_k's, the asymptotic behavior of $\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N]$ can be a priori derived after some combinatorics and enough bravery.

For instance, for the CUE we have $\varphi_k(x) = \mathrm{e}^{\mathrm{i}kx}$, $\mathbb{I} = \mathbb{Z}$, and the product formula $\langle\varphi_m\varphi_k, \varphi_\ell\rangle = \mathbf{1}_{m+k=\ell}$, and this yields after a quick computation

$$\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = \sum_{n\in\mathbb{Z}} \min(|n|, N)|\hat{f}_n|^2\,.$$

[♡]The same line of argument works for non-Hermitian kernels of the form $\sum\varphi_k\otimes\psi_k$ with biorthogonal φ_k's and ψ_k's.

This leads to the order one variance as soon as f is $H^{1/2}$-Sobolev and, if f is less smooth (e.g. a step function), the variance may still be smaller than N (e.g. $\log N$).

This approach can be extended to the d-dimensional torus $\mathbb{T}^d$, since the Fourier basis $\varphi_{\boldsymbol{k}}(x) = \mathrm{e}^{\mathrm{i}\langle \boldsymbol{k}, x\rangle}$ indexed by $\mathbb{I} = \mathbb{Z}^d$ satisfies $\langle \varphi_{\boldsymbol{m}}\varphi_{\boldsymbol{k}}, \varphi_{\boldsymbol{\ell}}\rangle = \mathbf{1}_{\boldsymbol{m}+\boldsymbol{k}=\boldsymbol{\ell}}$. However to define a DPP of N points, we further need to specify an ordering on the multi-indices $\boldsymbol{k}$'s, namely to pick a bijection $\mathfrak{b} : \mathbb{N}^* \to \mathbb{N}^d$, and set $\varphi_k := \varphi_{\mathfrak{b}(k)}$. For instance one may consider the bijection associated with the graded lexicographic order defined by saying that for every $M \geq 1$ we have

$$\{\mathfrak{b}(1), \dots, \mathfrak{b}(M^d)\} = \mathcal{C}_M := \{(k_1, \dots, k_d) \in \mathbb{N}^d : \, 0 \leq k_1, \dots, k_d \leq M-1\}$$

and by filling the layer between $\mathcal{C}_M$ and $\mathcal{C}_{M+1}$ according to the lexicographic order. After more work one can show that $\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = \mathcal{O}(N^{1-1/d})$ provided $f \in \mathcal{C}^1(\mathbb{T}^d, \mathbb{R})$, leading to a hyperuniform DPP. The main contribution of [P8] is the following generalization to the products of general orthogonal polynomials on $[-1,1]^d$.

Theorem 3.1. *Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ be a product measure on $[-1,1]^d$, where each μ_j has all its moments and a Lebesgue decomposition $\mu_j(\mathrm{d}x) = \omega_j(x)\mathrm{d}x + \mu_j^s(\mathrm{d}x)$ such that $\omega_j(x) > 0$ for a.e. $x \in (-1,1)$. For every $N \geq 1$, let $\mathbf{x}_1, \dots, \mathbf{x}_N$ be the DPP with rank N projection symmetric kernel associated with the orthonormal polynomials (φ_k) for μ, ordered by the graded lexicographic ordering of their degrees. Then, for every $f \in \mathcal{C}^1([-1,1]^d, \mathbb{R})$, we have $\mathbb{V}\mathrm{ar}[\mathcal{X}_f^N] = \mathcal{O}(N^{1-1/d})$ and the CLT*

$$\sqrt{N^{1+1/d}} \int f \, \mathrm{d}(\hat{\mu}_N - \mathbb{E}\hat{\mu}_N) \xrightarrow[N\to\infty]{\text{law}} \mathcal{N}(0, \sigma_f^2)\,,$$

where the limiting variance σ_f^2 is explicit.

This theorem can be seen as a higher dimensional generalization of Johansson's CLT (1.3). In fact, the case $d = 1$ was already obtained by Breuer and Duits [2017] and we only investigated the setting $d \geq 2$ in [P8]. Coeurjolly, Mazoyer, and Amblard [2020] have further investigated the Fourier DPP on the d-torus $\mathbb{T}^d$.

DPPs for Monte Carlo estimators. Having in mind that exact sampling algorithms for DPPs are available, a last obstruction to turn the previous CLT into a Monte Carlo algorithm is that $\int f \, \mathrm{d}\hat{\mu}_N$ is not directly an estimator of $\int f \, \mathrm{d}\mu$, since $\mathbb{E}\int f \, \mathrm{d}\hat{\mu}_N = \int f(x) \, \frac{1}{N} K_N(x,x) \, \mathrm{d}\mu(x)$, and $\frac{1}{N} K_N(x,x)$ has no reason to converge to 1. A natural way to tackle this problem is to take instead of $\frac{1}{N}$ the weights

$$w_j := \frac{1}{K_N(\mathbf{x}_j, \mathbf{x}_j)} \tag{3.4}$$

so that the bias of the estimator $\mathbb{E}\mathcal{E}_N(f)$ is now zero. Theorem 3.1 does not yield the solution directly since the test function $f_N(x) := \frac{N}{K_N(x,x)} f(x)$ to which we want to apply it now depends on N and may not be smooth at the boundary of the hypercube in the limit $N \to \infty$. Under extra regularity and technical assumptions we are able to derive a CLT for this estimator as well.

Theorem 3.2. *Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ be a product measure on $[-1,1]^d$, where each measures μ_j is finite, has a $\mathcal{C}^1$ density that is positive on $(-1,1)$, and satisfies the extra technical assumption [P12, Assumption 1]. For every $N \geq 1$, take for nodes $\mathbf{x}_1, \ldots, \mathbf{x}_N$ the same DPP as in Theorem 3.1 and weights as in (3.4). Then, for every $f \in \mathcal{C}^1([-1,1]^d, \mathbb{R})$ that vanishes on a neighborhood of the boundary of $[-1,1]^d$, we have $\mathbb{E}\mathcal{E}_N(f) = 0$ and $\|\mathcal{E}_N(f)\|_{L^2} = \mathcal{O}(N^{-(1+1/d)/2})$. Moreover, a CLT holds*

$$\sqrt{N^{1+1/d}}\,\mathcal{E}_N(f) \xrightarrow[N\to\infty]{\text{law}} \mathcal{N}(0, \Omega_f^2)\,,$$

where the limiting variance Ω_f^2 is explicit.

The extra technical assumption holds true when the μ_j's are Beta distributions for example, that is with density $(1-x)^{\alpha_j}(1+x)^{\beta_j}$ on $[-1,1]$ for some $\alpha_j, \beta_j > -1$ [P12, Proposition 2.8]. We also have an importance sampling version of the previous theorem, that can be used when μ is an arbitrary finite measure on $[-1,1]^d$ with a $\mathcal{C}^1$ positive density [P12, Theorem 2.9].

As a final remark, it is worthwhile to note that in dimension one, if we take for nodes $\mathbf{x}_1, \ldots, \mathbf{x}_N$ the zeros of the degree N orthonormal polynomial φ_N with respect to μ and pick the weights w_j as in (3.4), then we recover the famous Gaussian quadrature. Thus, we can say that our Monte Carlo method is a random version of the Gaussian quadrature but which still makes sense in higher dimensions. Let us stress that the zeros of φ_N and the DPP associated with the projection onto the N first OPs for μ behave similarly for large N [P4].

Perspectives. When N is large, the implementation details of our Monte Carlo method remains to be improved to be truly competitive against well processed Markov Chain Monte Carlo methods. This ambition is at the crux of the ERC Starting Grant "Black Jack" that Rémi Bardenet obtained recently, including faster DPP simulations and using repulsion in algorithmic parallelization.

There are also natural extensions: can we do better by choosing other bases (φ_k) than OPs, like wavelets, or eigenfunctions of Laplacians on a manifold? How small can the variance be in terms of N and the dimension d of the ambiant space? It seems it depends more on the "dimension" of the set of indices $\mathbb{I}$ than the dimension of the ambiant space itself.

One could also wonder if concentration inequalities are available for DPPs, so as to provide non-asymptotic confidence intervals for our estimators. A nice answer is provided by [Breuer and Duits, 2014, Theorem 3.1], which yields subexponential concentration for all finite rank projection Hermitian DPP. A natural question is then to investigate what is the class of sub-Gaussian DPPs, in which the CUE belongs according to (1.2).

Chapter 4

DPPs for covariance matrices

Based on the joint works [P5, P6, P7] with Walid Hachem (CNRS, Université Gustave Eiffel) and Jamal Najim (CNRS, Université Gustave Eiffel).

The sample covariance matrix. Consider a centered random vector Z of $\mathbb{R}^n$ with (well defined) covariance matrix Σ. A classical problem arising in many applications is to find, given an independent sample $Z_1, \dots, Z_N$ from Z, a decent estimator for Σ. This question is at the heart of principal component analysis, for instance. Another setting of interest is wireless networks, where the random vector Z is now complex, i.e. taking its values in $\mathbb{C}^n \simeq \mathbb{R}^{2n}$. In both real and complex settings the law of large numbers and the CLT state that the sample covariance matrix

$$\hat{\Sigma}_N := \frac{1}{N} \sum_{j=1}^{N} Z_j {}^{\mathbf{t}}\overline{Z}_j$$

is a consistent estimator of Σ as the sample size N grows to infinity and n is fixed. However, a typical case in modern applications is that sometimes N is large but so is n (e.g. $n \sim 10^4$ for a human genome dataset). Or sometimes n is small but N is as well (e.g. $n \sim N \sim 10$ for MIMO wireless networks). To investigate these situations, a more refined analysis using Bernstein concentration inequality and a covering argument yields, at least when Z is sub-Gaussian,

$$\|\Sigma - \hat{\Sigma}_N\| \leq C \left(\sqrt{\frac{n}{N}} + \frac{n}{N} \right) \tag{4.1}$$

with probability at least $1 - 2\mathrm{e}^{-n}$, where $\|\cdot\|$ stands for the operator norm and $C > 0$ only depends on $\|\Sigma\|$ and the sub-Gaussian norm of Z [Vershynin, 2018]. Thus $\hat{\Sigma}_N$ may not be such a nice estimator of Σ when n is of the same order than N, a regime that we investigate in this chapter.

The assumptions on the matrix model. From now we assume that n grows with N and by "$N \to \infty$", we refer to the regime where, for some fixed $\gamma \in (0, +\infty)$,

$$n, N \to \infty\,, \qquad \frac{n}{N} \to \gamma.$$

This means that we are considering in fact a sequence of $n \times n$ covariance matrices Σ; we assume its eigenvalues $\lambda_1, \dots, \lambda_n$ remain in a compact subset of $(0, \infty)$ independent on N and have a weak limit $\nu \in \mathcal{P}(\mathbb{R}_+)$ in distribution, namely

$$\frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j} \xrightarrow[N \to \infty]{} \nu.$$

We moreover restrict the discussion to a centered *complex Gaussian* vector Z.

In the following we denote by $\mathbf{x}_1, \dots, \mathbf{x}_N$ the (non-negative) eigenvalues of the sample covariance matrix $\hat{\Sigma}_N$ and $\hat{\mu}_N$ the associated empirical measure.

Wishart random matrices. The simplest case to study is the uncorrelated setting, where $\Sigma = I_n$ for all $n \geq 1$, in which case Marcenko and Pastur [1967] obtained the a.s. convergence $\hat{\mu}_N \to (\gamma - 1)_+ \, \delta_0 + \rho(x)\mathrm{d}x$ for an explicit limiting density

$$\rho(x) = \frac{1}{2\pi x} \sqrt{(\mathfrak{a}_- - x)(x - \mathfrak{a}_+)} \, \mathbf{1}_{[\mathfrak{a}_-, \mathfrak{a}_+]}(x), \qquad \mathfrak{a}_\pm = (1 \pm \sqrt{\gamma})^2. \tag{4.2}$$

Note that here $\nu = \delta_1$ and hence the limiting spectrum of Σ and $\hat{\Sigma}_N$ are quite different. In particular, the right hand side of the bound (4.1) cannot be small with high probability in the regime $n \sim N$ and thus $\hat{\Sigma}_N$ is a poor estimator of the true covariance matrix Σ.

It is however instructive to study the random eigenvalues of $\hat{\Sigma}_N$ at the microscopic level, namely the fluctuations of the spectrum around its macroscopic limit. It turns out, due to the complex Gaussian assumption on Z, that the $\min(n, N)$ positive eigenvalues $\mathbf{x}_j$'s are random and form a DPP on $\mathbb{R}_+$ with symmetric projection kernel associated with the Laguerre functions. This structure and asymptotic formulas for the Laguerre polynomials yield together that the Sine_2 process, which we recall is the DPP with kernel (1.5) that already appeared in the microscopic limit of the CUE, shows up in the bulk after scaling with a zooming factor N. More precisely, the point process $(N(\mathbf{x}_j - x))$ converges for every $x \in (\mathfrak{a}_-, \mathfrak{a}_+)$ to Sine_2.

At the right edge of the spectrum another universal point process shows up: the Airy process, which is the DPP on $\mathbb{R}$ with kernel

$$\mathrm{K}_{\mathsf{Airy}}(x, y) = \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}(y)\mathrm{Ai}'(x)}{x - y} \tag{4.3}$$

where Ai is the Airy function. More precisely, the maximal eigenvalue $\mathbf{x}_{\max}$ of $\hat{\Sigma}_N$, once centered around $\mathfrak{a}_+$ and scaled at a zooming factor $N^{2/3}$, converge in law to those of the maximal particle of the Airy process, known as the Tracy-Widom distribution [Johansson, 2000].

At the left edge $\mathfrak{a}_-$ of the spectrum, the Tracy-Widom distribution still describes the fluctuations of the minimal random eigenvalue $\mathbf{x}_{\min}$ of $\hat{\Sigma}_N$ provided that $\gamma \neq 1$ [Borodin and Forrester, 2003]. Indeed, when $\gamma = 1$ we have $\mathfrak{a}_- = 0$, what is usually referred as the hard edge, and the density ρ blows up like $1/\sqrt{x}$ near that edge,

which differs from the previous square root vanishing. In this setting another DPP on $\mathbb{R}_+$ arises, associated to the Bessel kernel

$$\mathrm{K}^{(\alpha)}_{\mathsf{Bessel}}(x,y) = \frac{\sqrt{y}\, J_\alpha(\sqrt{x}) J'_\alpha(\sqrt{y}) - \sqrt{x}\, J'_\alpha(\sqrt{x}) J_\alpha(\sqrt{y})}{2(x-y)} \tag{4.4}$$

where J_α is the Bessel function of the first kind with parameter α. More precisely, if $n = N + \alpha$ for some fixed $\alpha \in \mathbb{Z}$, then $\mathbf{x}_{\min}$ converges in law after N^2 scaling to the minimal particle of the DPP with kernel $\mathrm{K}^{(\alpha)}_{\mathsf{Bessel}}$ [Forrester, 1993b].

General setting: Macroscopic limit. In the general setting where the limiting spectrum of Σ is an arbitrary measure ν, Marcenko and Pastur [1967] also obtained the a.s. convergence $\hat{\mu}_N \to (\gamma - 1)_+ \, \delta_0 + \rho(x)\mathrm{d}x$ for a density ρ that is not explicit anymore, but its Cauchy-Stieltjes transform nevertheless satisfies a fixed point equation. Its support $\mathrm{Supp}(\rho)$ is compact, although not necessarily connected anymore with possibly an infinite number of connected components. Silverstein and Choi [1995] observed that at each soft edge (i.e. positive edge) of the connected components of $\mathrm{Supp}(\rho)$ corresponds to a local extrema of the real extension of the inverse Cauchy-Stieltjes transform

$$g(z) = \frac{1}{z} + \gamma \int \frac{\nu(\mathrm{d}x)}{1 - zx},$$

and that the density ρ may have cusp points (that is points in the interior of $\mathrm{Supp}(\rho)$ where the density vanishes) associated with possible inflexion points of g. From now, we call *regular* a soft edge (resp. cusp point) such that the associated critical point $\mathfrak{c}$ of g satisfies $\limsup_N \min_j |\mathfrak{c} - \lambda_j^{-1}| > 0$. We show that if a soft edge (resp. cusp point) is regular, then the density $\rho(x)$ vanishes like a square root at that edge (resp. like $|x|^{1/3}$ at that cusp point), and we finally show that the leftmost edge is a hard edge if and only if $\gamma = 1$, in which case $\rho(x)$ blows up like $1/\sqrt{x}$ near zero [P6].

General setting: Microscopic behavior. A natural question is: how do the eigenvalues fluctuate near each edge of the limiting spectrum? First, we show that at each regular soft edge there is a well defined sequence of extremal eigenvalue associated to that edge. More precisely, for a regular right edge $\mathfrak{a}$ we provide the existence of a deterministic subsequence $\varphi(N)$ such that $\mathbf{x}_{\varphi(N)} \to \mathfrak{a}$ and $\liminf_N (\mathbf{x}_{\varphi(N)+1} - \mathfrak{a}) > 0$ a.s, and a similar result holds for regular left edges. We describe the fluctuations for these extremal eigenvalues [P5].

Theorem 4.1. *At each regular soft edge, the associated extremal eigenvalue converges in law, after explicit centering around that edge and $N^{2/3}$ scaling, towards the Tracy-Widom distribution. Moreover, the collection of rescaled extremal eigenvalues become asymptotically independent in the large N limit.*

The Tracy-Widom fluctuations for the maximal eigenvalue, provided the rightmost edge is regular and making the extra assumption that there is no outliers

jumping out from the limiting support, was previously obtained in [El Karoui, 2007; Onatski, 2008]. We also describe the fluctuations near a hard edge [P5] in this general setting and moreover provide a precise rate of convergence [P6], motivated by an open problem raised by Edelman, Guionnet, and Péché [2016].

Theorem 4.2. *If $n = N + \alpha$ for some fixed $\alpha \in \mathbb{Z}$, then $\mathbf{x}_{\min}$ converges in law after scaling N^2 towards the law of the minimal particle of the DPP with kernel $K_{\mathsf{Bessel}}^{(\alpha)}$. Moreover, we provide an explicit $\mathcal{O}(N^{-1})$ error term in the convergence of the repartition functions.*

By combining these theorems, we can describe the convergence and fluctuations for the condition number of $\hat{\Sigma}_N$ as $N \to \infty$ [P5, Proposition 3.2].

Finally, we investigate what is happening microscopically at a cusp point. This requires to introduce another universal point process known as the Pearcy process, which is a DPP on $\mathbb{R}$ with a non-Hermitian kernel involving the Pearcey-like integral special functions [P6].

Theorem 4.3. *At each regular cusp point, the point process $(\mathbf{x}_j)$ converges after centering and $N^{3/4}$ scaling towards the Pearcey process, depending on a speed parameter κ that comes with Σ.*

These theorems are based on the fact that the random eigenvalues $(\mathbf{x}_j)$ still form a DPP in this general setting, although associated to a non-Hermitian kernel, constructed with multiple Laguerre polynomials of the second kind for which asymptotic formulas are not available in general yet. However, this kernel has a complex contour representation that is adapted to a steepest descent asymptotic analysis [Baik, Ben Arous, and Péché, 2005]. Let us stress that in general this analysis requires the delicate construction of steepest descent contours, that is usually done explicitly after tedious computations. In [P5,P6] we instead developed a unified (abstract) method to provide the existence of appropriate contours, by means of the maximum principle for subharmonic functions.

Let us also stress that Knowles and Yin [2017] extends the validity of Theorem 4.1 beyond the Gaussian setting.

Perspectives. The regularity assumption seems to be protector of the universality phenomena, for instance that the Tracy-Widom distribution appears at every regular soft edge although the macroscopic limit can be quite different from, say, the uncorrelated setting. A small breach into this assumption leads to the phase transition unveiled by Baik, Ben Arous, and Péché [2005]. It seems that, in general, without (a slightly weaker form of) this assumption, the fluctuations will actually depend on ν, and hence lie outside of the random matrix universality class. Quite interestingly, the same phenomenon arise in the study of the additive deformation of a GUE random matrix [Capitaine and Péché, 2015] and random Gelfand-Tsetlin patterns [Duse and Metcalfe, 2017, 2020]. It would be interesting to see what are the possible limiting processes that arise outside of this universality class.

Part II – Coulomb gases

Chapter 5

Concentration for Coulomb gases

Based on the joint work [P9] with *Djalil Chafaï* (Université Paris Dauphine - PSL) and *Mylène Maïda* (Université de Lille).

A law of large numbers for Coulomb gases. The main character of this chapter is the Coulomb gas on $\mathbb{R}^d$, for $d \geq 2$, introduced in Section 1.4. More precisely, recall the Coulomb potential g was introduced in (1.22). Given an inverse temperature parameter $\beta > 0$ and a potential $V : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, we consider the joint probability distribution of N particles $\mathbf{x}_1, \dots, \mathbf{x}_N$ on $\mathbb{R}^d$ given by

$$\mathrm{d}\mathbb{P}^V_{N,\beta}(x_1, \dots, x_N) := \frac{1}{Z^V_{N,\beta}} \, \mathrm{e}^{-\beta H_N(x_1,\dots,x_N)} \prod_{j=1}^N \mathrm{d}x_j \tag{5.1}$$

where the energy is defined♠ by

$$H_N(x_1, \dots, x_N) := \sum_{i \neq j} g(x_i - x_j) + N \sum_{j=1}^N V(x_j). \tag{5.2}$$

To ensure that the model is well defined, that is to have $0 < Z^V_{N,\beta} < \infty$, we assume V is finite on a set of positive Lebesgue measure and satisfies the growth condition

$$\int_{\mathbb{R}^d} \mathrm{e}^{-\beta(V(x) - \mathbf{1}_{d=2} \log(1+|x|^2))} \mathrm{d}x < \infty. \tag{$\mathbf{H}_\beta$}$$

For reasonable potentials V, the system has a macroscopic limit: there exists $\mu_V \in \mathcal{P}(\mathbb{R}^d)$ independent on β such that, almost surely,

$$\hat{\mu}_N \underset{N\to\infty}{\longrightarrow} \mu_V. \tag{5.3}$$

The so-called equilibrium measure μ_V is characterized as the unique solution of a variational problem: μ_V is the unique minimizer of the functional

$$\mathcal{E}_V(\mu) := \iint g(x-y) \, \mu(\mathrm{d}x)\mu(\mathrm{d}y) + \int V(x) \, \mu(\mathrm{d}x)$$

♠Note that, in comparison to (1.21), we now added a prefactor N in front of the potential V, which turns out to be the appropriate scaling to have a non-trivial macroscopic limit, and for cosmetic reasons we also sum over "$i \neq j$" instead of "$i < j$", which induces a change of the "true" electrostatic energy by a factor 2. The last modification is different from the convention of [P9].

defined on $\mathcal{P}(\mathbb{R}^d)$. One can guess this result by a mean field argument: If

$$\mathcal{E}_V^{\neq}(\mu) := \iint_{x\neq y} g(x-y)\mu(\mathrm{d}x)\mu(\mathrm{d}y) + \int V(x)\mu(\mathrm{d}x), \tag{5.4}$$

then we have $H_N(\mathbf{x}_1,\dots,\mathbf{x}_N) = N^2\mathcal{E}_V^{\neq}(\hat{\mu}_N)$ and thus the interaction energy is of order N^2. Since the entropy term arising from the randomness is of order N (think about the i.i.d. setting), the particle system will try to minimize the quantity $\mathcal{E}_V^{\neq}(\hat{\mu}_N)$ and, if we bet that $\hat{\mu}_N$ has a weak limit with a density, then this limit should minimize $\mathcal{E}_V$. More rigorously, if V is admissible[?] and continuous, then for any $\beta > 0$ satisfying ($\mathbf{H}_\beta$) the weak convergence (5.3) can be obtained from the Γ-convergence of $\frac{1}{N^2}H_N$ towards $\mathcal{E}_V$ [Serfaty, 2015], or from the large deviation principle at speed βN^2 and strictly convex good rate function $\mathcal{E}_V - \mathcal{E}_V(\mu_V)$ satisfied by the sequence $(\hat{\mu}_N)$ [Chafaï, Gozlan, and Zitt, 2014].

Concentration inequalities. Now that the macroscopic limit is identified, one can ask how fast $\hat{\mu}_N$ converges to μ_V, and for instance if the repulsive character of the Coulomb gas yields a faster-than-independent concentration inequality like (1.2) for the CUE. A similar concentration inequality has been obtained by Maïda and Maurel-Segala [2014] for the log-gases on $\mathbb{R}$ in a general potential V, but the Coulomb gas setting wasn't explored yet at the global scale; let us stress that concentration inequalities for local test functions have already been obtained by Rougerie and Serfaty [2016]. We indeed obtain fast sub-Gaussian concentration inequalities in the bounded Lipschitz and Wasserstein W_1 metrics.

Theorem 5.1 (Concentration for Coulomb gases)**.** *Assume that V is $\mathcal{C}^2$ on $\mathbb{R}^d$ and that its Laplacian ΔV satisfies the following growth constraint*

$$\limsup_{|x|\to\infty}\left(\frac{1}{V(x)}\sup_{\substack{y\in\mathbb{R}^d\\ |y-x|<1}}\Delta V(y)\right) < 2(d+2). \tag{5.5}$$

If V is admissible, then there exist constants $A > 0$, $B \in \mathbb{R}$, and a function $C(\beta)$ such that, for any $\beta > 0$ satisfying ($\mathbf{H}_\beta$), any $N \geq 2$, and any $r > 0$, we have

$$\mathbb{P}_{N,\beta}^V\Big(\mathrm{d}_{\mathrm{BL}}(\hat{\mu}_N,\mu_V) \geq r\Big) \leq \mathrm{e}^{-A\beta N^2 r^2 + \mathbf{1}_{d=2}(\frac{\beta}{2}N\log N) + B\beta N^{2-2/d} + C(\beta)N}. \tag{5.6}$$

If there exists $\kappa > 0$ such that $\liminf_{|x|\to\infty}\dfrac{V(x)}{|x|^\kappa} > 0$, then we have

$$C(\beta) = \begin{cases} \mathcal{O}(\log\beta), & \text{as } \beta\to 0,\\ \mathcal{O}(\beta), & \text{as } \beta\to\infty. \end{cases} \tag{5.7}$$

[?] $V : \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ is *admissible* if it is lower semicontinuous, finite on a set of positive capacity, and satisfies the growth condition $\lim_{|x|\to\infty}\left(V(x) - 2\log|x|\,\mathbf{1}_{d=2}\right) = +\infty$.

If one further assumes that V grows at least quadratically,

$$\liminf_{|x|\to\infty} \frac{V(x)}{|x|^2} > 0, \tag{$\mathbf{H}_{W_1}$}$$

then (5.6) also holds true in the W_1 metric.

Concerning the hypotheses, the constraint that ΔV does not grow faster than V is technical. It allows potentials growing like $|x|^\kappa$ for any $\kappa > 0$ or $\exp(|x|)$, but not $\exp(|x|^2)$. As for the regularity condition, we assume for ease of exposition that V is $\mathcal{C}^2$ but much less is required [P9, Remark 1.10].

We also explain how to obtain W_p versions of Theorem 5.1 without much efforts using that $\max_{i=1}^N |\mathbf{x}_i|$ is exponentially tight [P9, Theorem 1.12] since the W_p metrics are equivalent to W_1 on a compact set. However, by proceeding this way we expect to lose the optimality in N for the concentration bounds.

Combined with the Borel–Cantelli lemma, Theorem 5.1 directly yields the a.s. convergence of $\hat{\mu}_N$ to μ_V in W_1 even when one allows β to depend on N, provided it does not go to zero too fast, thanks to (5.7).

Corollary 5.2 (W_1 convergence)**.** *Under the last set of assumptions of Theorem 5.1, there exists a constant $\beta_V > 0$ such that the following holds: Given any sequence of positive real numbers (β_N) satisfying*

$$\beta_N \geq \beta_V \frac{\log N}{N}$$

for every N sufficiently large, then under $\mathbb{P}^N_{V,\beta_N}$ we have

$$\lim_{N\to\infty} W_1(\hat{\mu}_N, \mu_V) = 0$$

with probability one, in any joint probability space.

Moreover, if one keeps $\beta > 0$ fixed and let $r \to 0$ with N, then Theorem 5.1 is precise enough to yield the convergence of $\hat{\mu}_N$ towards μ_V at the mesoscopic scale, that is after zooming on the particle system around any fixed $x \in \mathbb{R}^d$ at the scale N^{-s} for any $0 \leq s < 1/d$ [P9, Corollary 1.8].

Coulomb transport inequality. The idea of the proof for Theorem 5.1 is quite simple: A Jensen inequality argument yields

$$Z^V_{N,\beta} \geq \mathrm{e}^{-N^2\beta\,\mathcal{E}_V(\mu_V)+\mathcal{O}(N)}$$

and thus, recalling (5.4), we have for any distance dist on $\mathcal{P}(\mathbb{R}^d)$,

$$\mathbb{P}^V_{N,\beta}\Big(\mathsf{dist}(\hat{\mu}_N, \mu_V) \geq r\Big) \leq \int_{\mathsf{dist}(\hat{\mu}_N,\mu_V)\geq r} \mathrm{e}^{-\beta\left(\mathcal{E}^{\neq}_V(\hat{\mu}_N)-\mathcal{E}_V(\mu_V)\right)+\mathcal{O}(N)}\mathrm{d}x.$$

After a regularization of the empirical measure $\hat{\mu}_N$ one can replace $\mathcal{E}_V^{\neq}$ by $\mathcal{E}_V$ in the right hand side, up to new error terms. If we assume for simplicity that each particle is restricted to a compact subset $K \subset \mathbb{R}^d$, then it is enough show that $\mathsf{dist}(\mu, \mu_V)^2$ is dominated by the rate function $\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V)$ to obtain a sub-Gaussian concentration bound; the remaining error terms in the theorem come from this regularization procedure and because we have to deal with the lack of compactness. The functional domination is obtained in the next theorem, which may be of independent interest.

Theorem 5.3 (Coulomb transportation inequality for equilibrium measures)**.** *If V is admissible, then there exists $C_{\mathrm{BL}}^V > 0$ such that, for every $\mu \in \mathcal{P}(\mathbb{R}^d)$,*

$$\mathrm{d}_{\mathrm{BL}}(\mu, \mu_V)^2 \leq C_{\mathrm{BL}}^V \Big(\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V) \Big). \tag{5.8}$$

If V further satisfies the growth condition ($\mathbf{H}_{W_1}$) then there exists $C_{W_1}^V > 0$ such that, for every $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$W_1(\mu, \mu_V)^2 \leq C_{W_1}^V \Big(\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V) \Big). \tag{5.9}$$

Let us stress that a the growth assumption ($\mathbf{H}_{W_1}$) turns out to be a necessary condition for (5.9) to hold true for every $\mu \in \mathcal{P}(\mathbb{R}^d)$.

It is interesting to relate inequality (5.9) to Talagrand's transportation inequality (T_1) [Gozlan and Léonard, 2010; Villani, 2003]. Indeed, μ_V would satisfy inequality (T_1) if (5.9) holds for every $\mu \in \mathcal{P}(\mathbb{R}^d)$ after replacing the rate function $\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V)$ by the relative entropy $H(\mu|\mu_V)$, which is the rate function in the Sanov's large deviation principle for i.i.d. random variables with law μ_V. This functional inequality would imply (and in fact, is equivalent to) sub-Gaussian concentration for Lipschitz functions of i.i.d random variables with law μ_V [Marton, 1986; Bobkov and Götze, 1999]. From this perspective, (5.9) can be seen as an analog of Talagrand's transportation inequality (T_1) for the Coulomb interaction, hence the name of the previous theorem.

Let us also stress that inequalities of the same flavor for probability measures on the real line $\mathbb{R}$, linking the Wassertein W_1 or W_2 metrics with $\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V)$ when $d = 2$ have been previously obtained in the context of free probability by Biane and Voiculescu [2001]; Hiai, Petz, and Ueda [2004]; Ledoux and Popescu [2009]; Maïda and Maurel-Segala [2014]; Popescu [2013]. In this setting, these functionals are usually referred to as free (relative) entropies and are related to the large deviation principle due to Ben Arous and Guionnet [1997] for the one-dimensional log-gas associated to unitarily invariant ensembles in random matrix theory. Theorem 5.3 recovers the free transport inequality of [Maïda and Maurel-Segala, 2014; Popescu, 2013] as a particular case.

We have also a local version of the Coulomb transport inequality, which we use to prove it; the notation $\mathcal{E}(\mu)$ means $\mathcal{E}_0(\mu)$ (no potential) and we can legally extend its definition to signed measures under the assumptions we make.

Theorem 5.4 (Local Coulomb transportation inequality)**.** *For every compact subset $D \subset \mathbb{R}^d$, there exists a constant $C_D > 0$ such that, for every $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ supported in D with $\mathcal{E}(\mu) < \infty$ and $\mathcal{E}(\nu) < \infty$,*

$$W_1(\mu, \nu)^2 \leq C_D \, \mathcal{E}(\mu - \nu).$$

When $d = 2$ and $D \subset \mathbb{R}$, Theorem 5.4 yields [Popescu, 2013, Theorem 1]; we could say that Theorem 5.4 extends Popescu's local free transport inequality to higher dimensions.

The proof yields a non-optimal but explicit value for C_D given by 4^d times the volume of the smallest Euclidean ball that contains D, which can be bounded thanks to Jung's theorem by

$$C_D \leq \frac{1}{\Gamma(\frac{d}{2}+1)} \left(\mathsf{Diameter}(D) \sqrt{\frac{8\pi d}{d+1}} \right)^d .$$

More explicit constants. One may also look for a concentration inequality where the constants $A, B, C(\beta)$ are explicit in terms of V. We are able to derive such a statement when ΔV is bounded from above [P9, Theorem 1.9]. Let us only illustrate this result with an application to random matrix theory. Indeed, the next corollary follows since the law of the eigenvalues of a Ginibre random matrix is a two-dimensional Coulomb gas at inverse temperature $\beta = 1$ (in our convention) with potential $V(x) = \frac{1}{2}|x|^2$, for which the equilibrium measure is the uniform measure $\mu_\circ$ on the unit disc.

Corollary 5.5. *Let $\mathbf{M}_N$ be a $N \times N$ random matrix with i.i.d. standard complex gaussian $\mathcal{N}_{\mathbb{C}}(0,1)$ entries. If $\mathbb{P}_N$ stands for the joint law of the eigenvalues of $\frac{1}{\sqrt{N}}\mathbf{M}_N$ then, for any $N \geq 2$ and any $r > 0$, we have*

$$\mathbb{P}_N\Big(W_1(\hat{\mu}_N, \mu_\circ) \geq r \Big) \leq \mathrm{e}^{-\frac{1}{4C}N^2 r^2 + \frac{1}{2}N\log N + N[\frac{1}{C} + \frac{3}{2} - \log \pi]} \tag{5.10}$$

where $C := C_{W_1}^{|\cdot|^2}$ is the constant appearing in Theorem 5.3 for $d = 2$.

Let us stress that we do not know how to deduce this concentration inequality from the Gaussian nature of the entries of $\mathbf{M}_N$, like one typically does for Hermitian matrix models. The eigenvalues of a non-normal matrix are not Lipschitz functions of its entries, in contrast with the singular values for which the Courant–Fischer formulas and the Hoffman–Wielandt inequality hold [Bordenave and Chafaï, 2012].

One could also use the determinantal structure of the eigenvalues of the Ginibre random matrix to reach a concentration inequality, but the best general result available [Breuer and Duits, 2014] provides subexponential concentration bounds leading to a weaker inequality than Corollary 5.5. Again, it would be interesting to find sufficient conditions for DPPs to have sub-Gaussian concentration bounds.

Chapter 6

A Coulomb gas approach to Smale's 7th problem

Based on the joint work [P11] with *Carlos Beltrán* (Universidad de Cantabria).

Smale's 7th problem. In this chapter we consider the logarithmic energy of a configuration $x_1, \dots, x_N$ on the sphere $\mathbb{S} := \{x \in \mathbb{R}^3 : \|x\| = 1\}$ defined by

$$H_N(x_1, \dots, x_N) := \sum_{i \neq j} \log \frac{1}{\|x_i - x_j\|},$$

where $\|\cdot\|$ stands for the Euclidean norm of $\mathbb{R}^3$. For a given integer N, the problem of describing the configurations minimizing H_N, or equivalently maximizing the product $\prod_{i \neq j} \|x_i - x_j\|$, turns out to be extremely hard. This is however a natural problem: if you have the budget to construct, say, 17 orbiting communication satelittes, where would you place them so as to maximize your radio covering? As usual for real life problems, you may be satisfied with an approximate minimizing configuration. This is exactly what the Smale's 7th problem asks about, although his initial motivation was the problem of finding well-conditioned polynomials [Shub and Smale, 1993], that has been recently solved with a different approach by Beltrán, Etayo, Marzo, and Ortega-Cerdà [2020].

More precisely, the 7th problem from Smale [2000]'s list of mathematical problems for the next century asks to find for every $N \geq 2$ a configuration $x_1, \dots, x_N \in \mathbb{S}$ and a constant $c > 0$ independent on N such that

$$H_N(x_1, \dots, x_N) - \min_{\mathbb{S}^N} H_N \leq c \log N. \tag{6.1}$$

More precisely, quoting Smale: "*For a precise version one could ask for a real number algorithm in the sense of Blum, Cucker, Shub, and Smale [1996] which on input N produces as output distinct points $x_1, \dots, x_N$ on the 2-sphere satisfying (6.1) with halting time polynomial in N*".

Large N expansion. One difficulty in this problem is that the large N behavior of $\min_{\mathbb{S}^N} H_N$ is not even known up to precision $\log N$, so that it is hard to assess if any proposed configuration does satisfy Smale's requirement. Indeed, the actual knowledge is that, as $N \to \infty$,

$$\min_{\mathbb{S}^N} H_N = \left(\tfrac{1}{2} + \log \tfrac{1}{2}\right) N^2 - \tfrac{1}{2} N \log N + C_? \, N + o(N). \tag{6.2}$$

The exact value of the constant $C_?$ in (6.2) is still conjectural. In fact, using the renormalized energy approach introduced by Sandier and Serfaty [2012], Betermin and Sandier [2018] proved the existence of $C_?$ and the upper bound

$$C_? \leq C_* := 2\log 2 + \frac{1}{2}\log\frac{2}{3} + 3\log\frac{\sqrt{\pi}}{\Gamma(1/3)} = -0.0556053\ldots. \tag{6.3}$$

This upper bound is conjectured to be an equality by Brauchart, Hardin, and Saff [2012]. The tightest known lower bound is $C_? \geq -0.2232823526\ldots$ Dubickas [1996].

Thus, one can try deterministic algorithms that provide for every N a configuration of N points on $\mathbb{S}$, compute their logarithmic energy and compare it with the expansion (6.2) to see if they reach an error of order at most N. Hardin, Michaels, and Saff [2016] made many numerical experiments analyzing different constructions, but none of them seems to reach the upper bound for $C_?$ in (6.3), leading to error of order at least N and thus far away from the $\log N$ asked by Smale.

Another possible strategy to is to look for random configurations on $\mathbb{S}$ whose logarithmic energy could satisfy (6.1) on average, or with high probability. If one naively take i.i.d uniform configurations on $\mathbb{S}$, then a direct computation shows the mean energy equals the minimal one plus $\mathcal{O}(N \log N)$ correction. Armentano, Beltrán, and Shub [2011] suggested instead to try the repulsive zeros of the spherical GAF introduced in Chapter 2 and diminished the correction to an error $\mathcal{O}(N)$. Other attempts, including the DPP on the sphere known as the Spherical Ensemble [Alishahi and Zamani, 2015], or an elegant particle system called the Diamond ensemble [Etayo and Beltrán, 2020], reached the error $\mathcal{O}(N)$ as well. All these bounds (analytical and/or numerical) are still far from the upper bound in (6.3).

The Coulomb gas on the sphere. Another natural random configuration associated with this problem is the Coulomb gas on $\mathbb{S}$, which is the main character of this chapter. More precisely, let σ be the uniform measure on $\mathbb{S}$ normalized so that $\sigma(\mathbb{S}) = 1$. For any $N \geq 2$ and $\beta > 0$, consider the probability measure on $\mathbb{S}^N$

$$\mathbb{P}_{N,\beta}(\mathrm{d}x) := \frac{1}{Z_{N,\beta}} \, \mathrm{e}^{-\beta H_N(x_1,\ldots,x_N)} \prod_{j=1}^{N} \sigma(\mathrm{d}x_j)$$

with partition function

$$Z_{N,\beta} := \int_{\mathbb{S}^N} \mathrm{e}^{-\beta H_N(x_1,\ldots,x_N)} \prod_{j=1}^{N} \sigma(\mathrm{d}x_j). \tag{6.4}$$

As explained in Chapter 1.4, this is the Coulomb gas on $\mathbb{S}$, namely the canonical ensemble with (twice the) energy associated to the Green function $G(x,y) := \log \|x-y\|^{-1}$ that solve the Poisson equation

$$\Delta_{\mathbb{S}} \log \|x - \cdot\|^{-1} = 2\pi(\sigma - \delta_x) \tag{6.5}$$

for every $x \in \mathbb{S}$. Here $\Delta_{\mathbb{S}}$ stands for the Laplace-Beltrami operator coming with the usual Riemannian structure for $\mathbb{S} \subset \mathbb{R}^3$ inherited from the Euclidean structure of $\mathbb{R}^3$; the geodesic distance is thus given by $d_{\mathbb{S}}(x,y) := \arccos\langle x, y\rangle$.

Typical configurations of the Coulomb gas will try to minimize H_N because of its density distribution proportional to $\mathrm{e}^{-\beta H_N}$ and thus it is tempting to evaluate the energy H_N for such random configurations. Low β's (high temperature) are not interesting, since for instance $\beta = 0$ yields i.i.d. uniform random configurations. In fact, when $\beta = 1$ the Coulomb gas benefits from an integrable structure: up to stereographic projection, this is the spherical ensemble mentioned above and studied by Alishahi and Zamani [2015]. But the larger β is (low temperature) the more likely it is for $\mathbb{P}_{N,\beta}$ to generate a configuration close to a minimizer, although the determinantal structure is lost when $\beta \neq 1$ making exact computations out of reach.

The main contribution of this chapter is to show that the Coulomb gas on the sphere at temperature $\mathcal{O}(1/N)$ provides almost minimizing configurations in the sense of Smale's problem with exponentially high probability, as well as on average.

Theorem 6.1. *For any $N \geq 2$ and any $\beta \geq 1$, let $\mathbf{x}_1, \dots, \mathbf{x}_N$ be the random configuration on $\mathbb{S}$ with joint distribution $\mathbb{P}_{N,\beta}$. For any constant $c > 0$ we have*

$$H_N(\mathbf{x}_1, \dots, \mathbf{x}_N) - \min_{\mathbb{S}^N} H_N \leq c \log N \tag{6.6}$$

with probability at least $1 - \mathrm{e}^{-\kappa N}$, where

$$\kappa := c\,\frac{\beta}{N} \log N - \log \beta - 8 \log N.$$

Moreover, the mean energy satisfies

$$\mathbb{E}\Big[H_N(\mathbf{x}_1, \dots, \mathbf{x}_N)\Big] - \min_{\mathbb{S}^N} H_N \leq \frac{N}{\beta}\Big(\log \beta + 8 \log N\Big).$$

Note that given β and N, the constant c has to be chosen so that $\kappa > 0$ since otherwise the first result becomes empty. We reach the precision $\log N$ for any $N \geq 2$ when β is at least of order N. For example, by taking $\beta = N$ and $c = 10$ in Theorem 6.1, we obtain the following estimates.

Corollary 6.2. *For any $N \geq 2$, if the random configuration $\mathbf{x}_1, \dots, \mathbf{x}_N$ on $\mathbb{S}$ has for distribution the Coulomb gas $\mathbb{P}_{N,N}$ at inverse temperature $\beta = N$, then*

$$H_N(\mathbf{x}_1, \dots, \mathbf{x}_N) - \min_{\mathbb{S}^N} H_N \leq 10 \log N$$

with probability at least $1 - \mathrm{e}^{-N \log N}$. Moreover,

$$\mathbb{E}\Big[H_N(\mathbf{x}_1, \dots, \mathbf{x}_N)\Big] - \min_{\mathbb{S}^N} H_N \leq 9 \log N.$$

Idea of the proof. The strategy to obtain Theorem 6.1 is quite elementary: First, we notice that for any constant $C_{N,\beta} > 0$ satisfying

$$\log Z_{N,\beta} \geq -\beta \inf_{\mathbb{S}^N} H_N - C_{N,\beta} \tag{6.7}$$

we have, for every $\delta > 0$,

$$\mathbb{P}_{N,\beta}\Big(H_N(\mathbf{x}_1, \dots, \mathbf{x}_N) - \inf_{\mathbb{S}^N} H_N > \delta\Big) \leq \mathrm{e}^{-\beta\delta + C_{N,\beta}} \tag{6.8}$$

and moreover, using a convexity argument,

$$\mathbb{E}\Big[H_N(\mathbf{x}_1, \dots, \mathbf{x}_N)\Big] - \inf_{\mathbb{S}^N} H_N \leq \frac{C_{N,\beta}}{\beta}. \tag{6.9}$$

We stress that this approach remains valid for arbitrary energy H_N and is not tied to the Coulomb setting. The difficulty is thus to obtain such a constant $C_{N,\beta}$, which we achieve by a perturbative analysis for the Coulomb energy H_N near a minimizer. More precisely, we derive the following estimate, which improves the range of validity of [Beltrán, 2013, Theorem 1.8] with an alternative proof.

Proposition 6.3. *Let $N \geq 2$ and $(x_1^*, \dots, x_N^*) \in \mathbb{S}^N$ be any minimizer of H_N. If $(x_1, \dots, x_N) \in \mathbb{S}^N$ satisfies*

$$\max_{1 \leq i \leq N} d_{\mathbb{S}}(x_i, x_i^*) \leq \arcsin\left(\frac{s}{\sqrt{5}N^{3/2}}\right)$$

for some $0 \leq s \leq \sqrt{5N}/2$, then

$$H_N(x_1, \dots, x_N) \leq \min_{\mathbb{S}^N} H_N + s^2.$$

The proof of this estimate relies on the componentwise subharmonicity of the energy H_N and on a separation result of Dragnev [2002] for the minimizers of H_N.

Perspectives. First, it seems rather easy to generalize Theorem 6.1 with the same line of arguments to the Coulomb gas on $\mathbb{R}^d$ in a subharmonic potential, like $V(x) = |x|^2$, and to other (harmonic) compact Riemannian manifolds.

More interestingly, if one accepts stochastic algorithms as solutions for the precise version of Smale's 7th problem, Corollary 6.2 yields that it remains to show that one can sample a configuration from $\mathbb{P}_{N,N}$ in polynomial time to solve the problem, and there is extra margin for simulation error. More precisely, the problem amounts to sample random variables $\mathbf{y}_1, \dots, \mathbf{y}_N$ on $\mathbb{S}$ in polynomial time such that, if $\mathbf{x}_1, \dots, \mathbf{x}_N \sim \mathbb{P}_{N,N}$, then $H_N(\mathbf{y}_1, \dots, \mathbf{y}_N) \leq H_N(\mathbf{x}_1, \dots, \mathbf{x}_N) + C \log N$ with high probability for some universal constant $C > 0$. This makes the quantitative analysis for Monte Carlo simulations of the Coulomb gas, which started only recently [Chafaï and Ferré, 2019; Lu and Mattingly, 2019], an even more attractive topic for research.

Chapter 7

Fluctuations for the CβE

Based on the joint work [P15] with *Gaultier Lambert* (University of Zurich).

Macroscopic limit and fluctuations of the CβE. Given an inverse temperature parameter $\beta > 0$, recall the CβE introduced in Section 1.4 is the system of N random particles on $\mathbb{T}$ obtained by restriction of Coulomb gas on $\mathbb{T} \times \mathbb{T}$, namely with distribution

$$\frac{1}{Z_{N,\beta}} \prod_{1 \le i < j \le N} |\mathrm{e}^{\mathrm{i}x_i} - \mathrm{e}^{\mathrm{i}x_j}|^\beta \prod_{j=1}^{N} \frac{\mathrm{d}x_j}{2\pi}.$$

The macroscopic behavior of the CβE presents the same features than the CUE presented in Section 1.1: $\hat{\mu}_N$ converges a.s. to the uniform mesure $\frac{\mathrm{d}x}{2\pi}$ on $\mathbb{T}$ and, for any smooth enough test function $f : \mathbb{T} \to \mathbb{R}$, Johansson [1988] proved♣ the CLT

$$N \int f \left(\mathrm{d}\hat{\mu}_N - \tfrac{\mathrm{d}x}{2\pi} \right) \xrightarrow[N\to\infty]{\text{law}} \mathcal{N}\Big(0, \frac{2}{\beta} \|f\|^2_{H^{1/2}}\Big), \tag{7.1}$$

where we recall the Sobolev semi-norm was defined by

$$\|f\|^2_{H^{1/2}} := 2 \sum_{k=1}^{\infty} k \, |\hat{f}_k|^2 \tag{7.2}$$

with $\hat{f}_k := \int_{\mathbb{T}} f(x) \, \mathrm{e}^{-\mathrm{i}kx} \, \frac{\mathrm{d}x}{2\pi}$ the Fourier coefficients of f.

The high temperature regime. First, notice that if we take $\beta = 0$, which corresponds to the infinite temperature setting, then the $\mathbf{x}_i$'s are independent random variables uniformly distributed on $\mathbb{T}$. Thus, the law of large numbers yields the a.s. convergence $\hat{\mu}_N \to \frac{\mathrm{d}x}{2\pi}$ as $N \to \infty$ and the classical CLT states that, for any L^2 function $f : \mathbb{T} \to \mathbb{R}$ such that $\hat{f}_0 = 0$,

$$\sqrt{N} \int f \left(\mathrm{d}\hat{\mu}_N - \tfrac{\mathrm{d}x}{2\pi} \right) \xrightarrow[N\to\infty]{\text{law}} \mathcal{N}\Big(0, \|f\|^2_{L^2}\Big). \tag{7.3}$$

♣ More precisely, the CLT in [Johansson, 1988] is stated for $\beta = 2$ but it is straightforward to check that the method still applies for any fixed $\beta > 0$ provided that the test function f is $\mathcal{C}^{1,\alpha}$ for some $\alpha > 0$, see also [Lambert, 2019, Theorem 1.2]. Note that although one may believe this CLT holds as soon as $\|f\|_{H^{1/2}} < \infty$, there are counterexamples for $\beta = 4$ [Lambert, 2019].

Note that, for such real valued test function f, the limiting variance can be written

$$\|f\|_{L^2}^2 := \int_{\mathbb{T}} |f(x)|^2 \frac{\mathrm{d}x}{2\pi} = 2\sum_{k=1}^{\infty} |\hat{f}_k|^2. \tag{7.4}$$

Note also the difference of normalization between (7.1) and (7.3).

The goal of this chapter is to present what is happening at the intersection of these two regimes. We start with a generalization of the concentration inequality (1.2) for arbitrary $\beta > 0$.

Theorem 7.1 (Concentration for $\mathsf{C}\beta\mathsf{E}$)**.** *For any $\beta > 0$, $N \geq 10$, and $r > 0$,*

$$\mathbb{P}_{\mathsf{C}\beta\mathsf{E}}\Big(\mathrm{W}_1(\hat{\mu}_N, \tfrac{\mathrm{d}x}{2\pi}) \geq r\Big) \leq \mathrm{e}^{-\frac{\beta}{2}(\frac{1}{8\pi}N^2r^2 - 5N\log N - C_0 N)}$$

where $C_0 := 2\log 2 + 17/2 + \pi^{-1} \simeq 10.2$.

This yields in particular the a.s. convergence $\hat{\mu}_N \to \frac{\mathrm{d}x}{2\pi}$ as $N \to \infty$ in the W_1 metric provided[♠] that $\beta \gg N^{-2}$. For the fluctuations, it turns out there is a critical temperature regime where the limiting variance interpolates between the Lebesgue L^2 and the Sobolev $H^{1/2}$ (semi-)norms, and this happens when $\beta \sim N^{-1}$.

Theorem 7.2 (CLT for $\mathsf{C}\beta\mathsf{E}$ at high temperature)**.** *Given $\theta > 0$, consider the $\mathsf{C}\beta\mathsf{E}$ at inverse temperature $\beta := \frac{2\theta}{N}$. Then, for every $f \in \mathcal{C}^5(\mathbb{T}, \mathbb{R})$ with $\hat{f}_0 = 0$, we have*

$$\sqrt{N}\int f\,(\mathrm{d}\hat{\mu}_N - \tfrac{\mathrm{d}x}{2\pi}) \xrightarrow[N\to\infty]{\text{law}} \mathcal{N}\Big(0, \sigma_\theta(f)^2\Big)$$

where the limiting variance is given by

$$\sigma_\theta(f)^2 = 2\sum_{k=1}^{\infty} \frac{1}{1+\theta/k}\,|\hat{f}_k|^2. \tag{7.5}$$

Recalling (7.2) and (7.4), it indeed follows that $\sigma_\theta(f) \to \|f\|_{L^2}$ as $\theta \to 0$ and that $\sqrt{\theta}\,\sigma_\theta(f) \to \|f\|_{H^{1/2}}$ when $\theta \to \infty$.

We note that a similar CLT for the $\mathsf{G}\beta\mathsf{E}$ was obtained [Trinh, 2017; Nakano and Trinh, 2018] thanks to the tridiagonal matrix representation for this specific model [Dumitriu and Edelman, 2002], although the limiting variance is not explicit there.

The two previous theorems are both particular cases of more general results we obtain for the log-gas on $\mathbb{T}$ in an arbitrary potential V that we introduce now.

[♠] This should be true even when $\beta \leq N^{-2}$, although one needs other techniques to prove it. Indeed, in this regime one can expect the particles are essentially independent (exponentially small correlations). See also the interesting change of behavior for the partition function of the $\mathsf{G}\beta\mathsf{E}$ when $\beta \sim N^{-2}$ pointed out in [Pakzad, 2018, Lemma 1.3].

Log-gas on $\mathbb{T}$ with an arbitrary potential V. For any $\theta \geq 0$ and any continuous potential $V : \mathbb{T} \to \mathbb{R}$, we now consider N random interacting particles on $\mathbb{T}$ with joint probability distribution

$$\mathrm{d}\mathbb{P}^V_{N,\theta}(x_1, \dots, x_N) := \frac{1}{Z^V_{N,\theta}} \prod_{i<j} |\mathrm{e}^{\mathrm{i}x_i} - \mathrm{e}^{\mathrm{i}x_j}|^{\frac{2\theta}{N}} \prod_{i=1}^{N} \mathrm{e}^{-V(x_i)} \frac{\mathrm{d}x_i}{2\pi} \tag{7.6}$$

where $Z^V_{N,\theta} > 0$ is a normalization constant. We set

$$\mu^0_V(\mathrm{d}x) := \mathrm{e}^{-V(x)} \frac{\mathrm{d}x}{2\pi} \tag{7.7}$$

and, by adding a constant to V if necessary, we assume that μ^V_0 is a probability measure on $\mathbb{T}$. In particular, (7.6) is the law of N i.i.d. random variables with distribution μ^0_V when $\theta = 0$.

The potential V has an effect on the macroscopic limit and its presence allows to see why something particular should happen in the regime $\beta \sim N^{-1}$. Indeed, the heuristic presented in Chapter 5 now concerns an interaction term $\prod_{i<j} |\mathrm{e}^{\mathrm{i}x_i} - \mathrm{e}^{\mathrm{i}x_j}|^{\frac{2\theta}{N}}$ with an exponential contribution proportional to N times energy while $(\mu^0_V)^{\otimes N}$ yields a contribution of N times entropy, and thus this is the only temperature regime where there is a fair competition between the energy and entropy. More precisely, if we introduce the logarithmic energy of $\mu \in \mathcal{P}(\mathbb{T})$ defined by

$$\mathcal{E}(\mu) := \iint \log \frac{1}{|\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{\mathrm{i}y}|} \mu(\mathrm{d}x)\mu(\mathrm{d}y) \tag{7.8}$$

and denote by $H(\mu|\mu^V_0)$ the relative entropy of μ with respect to μ^V_0, then the functional of interest here is

$$F^\theta_V(\mu) := \theta\, \mathcal{E}(\mu) + H(\mu|\mu^V_0). \tag{7.9}$$

The next result can be extracted from the literature.

Theorem 7.3. *Let $\theta \geq 0$ and assume $V : \mathbb{T} \to \mathbb{R}$ is continuous.*

(a) *The functional $F^\theta_V : \mathcal{P}(\mathbb{T}) \to [0, +\infty]$ has compact level sets $\{F^\theta_V \leq \alpha\}$, $\alpha \in \mathbb{R}$, and is strictly convex. In particular it has a unique minimizer μ^θ_V on $\mathcal{P}(\mathbb{T})$.*

(b) *The sequence $(\hat{\mu}_N)$ satisfies a large deviation principle at speed θN with rate function $\mu \mapsto F^\theta_V(\mu) - F^\theta_V(\mu^\theta_V)$, and in particular $\hat{\mu}_N \to \mu^\theta_V$ a.s. as $N \to \infty$.*

Indeed, when $\theta = 0$, this is Sanov's theorem for i.i.d random variables and elementary properties of the relative entropy [Dembo and Zeitouni, 2010]. Note that the unique minimizer of F^0_V is given by (7.7) and hence the notation is consistent. When $\theta > 0$, statement (a) is standard and (b) can be found in [Berman, 2018; García-Zelada, 2018].

When $V = 0$ we have $\mu^\theta_V = \frac{\mathrm{d}x}{2\pi}$ for every $\theta \geq 0$, and we show in general that μ^θ_V has a bounded density that is larger than a positive constant and is essentially as smooth as V is [P15, Proposition 2.1]. To complete the picture, we obtain the following concentration bounds.

Theorem 7.4. *Let $\theta > 0$ and assume $V : \mathbb{T} \to \mathbb{R}$ has a weak derivative V' in $L^2(\mathbb{T})$. Then, there exists $C = C(\mu_V^\theta) > 0$ such that, for every $N \geq 10$ and $r > 0$,*

$$\mathbb{P}_{N,\theta}^V\Big(\mathrm{W}_1(\hat{\mu}_N, \mu_V^\theta) \geq r\Big) \leq \mathrm{e}^{-\theta(\frac{1}{8\pi}Nr^2 - 5\log N - C)}.$$

Now we turn to our result for the fluctuations of the particle system around μ_V^θ, which needs some preparation. Let us also write μ_V^θ for the density of the equilibrium measure and introduce the operator $\mathscr{L}$ formally defined on $L^2(\mathbb{T})$ by

$$-\mathscr{L}\varphi = \varphi'' + 2\pi\theta\mathscr{U}(\mu_V^\theta\varphi') + (\log\mu_V^\theta)'\varphi'\,. \tag{7.10}$$

Here $\mathscr{U}$ stands for the Hilbert transform defined by

$$\mathscr{U}f(x) := -\text{p.v.}\int_{\mathbb{T}} \frac{f(t)}{\tan\left(\frac{x-t}{2}\right)}\frac{\mathrm{d}t}{2\pi} \tag{7.11}$$

with p.v. the Cauchy principal value. Note that when $\theta = 0$ the operator $\mathscr{L}\varphi = -\varphi'' + V'\varphi'$ is just a Sturm-Liouville operator. We show in general, for any $\theta > 0$ and any $V \in \mathcal{C}^{1,1}(\mathbb{T})$[?], that the operator $\mathscr{L}$ is positive self-adjoint defined (by Friedrichs extension) on the Hilbert space

$$\mathsf{H} := \Big\{f \in L^2(\mathbb{T}) :\ f' \in L^2(\mathbb{T}), \quad \int f\,\mathrm{d}\mu_V^\theta = 0\Big\}$$

equipped with the inner-product

$$\langle f, g\rangle_{\mathsf{H}} := \int f'\,g'\,\mathrm{d}\mu_V^\theta\,,$$

and moreover that the inverse $\mathscr{L}^{-1}$ is a densely defined trace-class operator there.

We are now in position the state our CLT, that essentially says that, in the sense of finite dimensional distributions, the measure $\sqrt{N}(\hat{\mu}_N - \mu_V^\theta)$ converges as $N \to \infty$ to a Gaussian process on H with covariance operator $\mathscr{L}^{-1}$; we also obtain an upper bound for its speed of convergence in the Wasserstein W_2 metric.

Theorem 7.5. *Let $\theta > 0$ and $V \in \mathcal{C}^{3,1}(\mathbb{T})$. Under the law $\mathbb{P}_{N,\theta}^V$ defined in (7.6), if $f \in \mathcal{C}^{2\kappa+1}(\mathbb{T})$ for some integer $\kappa \geq 2$, then we have*

$$\mathsf{Fluct}_N(f) := \sqrt{N}\int f\,\mathrm{d}(\hat{\mu}_N - \mu_V^\theta) \xrightarrow[N\to\infty]{\text{law}} \mathcal{N}\Big(0, \sigma_\theta^V(f)^2\Big)$$

where the limiting variance is given by

$$\sigma_\theta^V(f)^2 := \langle\Big(f - \int f\,\mathrm{d}\mu_V^\theta\Big), \mathscr{L}^{-1}\Big(f - \int f\,\mathrm{d}\mu_V^\theta\Big)\rangle_{\mathsf{H}}. \tag{7.12}$$

Moreover, there exists $C = C(\theta, V, f) > 0$ such that

$$W_2\Big(\mathsf{Fluct}_N(f),\, \mathcal{N}\Big(0, \sigma_\theta^V(f)^2\Big)\Big) \leq C\sqrt{\frac{\log N}{N^{\frac{\kappa-1}{\kappa+1}}}}.$$

[?] $\mathbb{C}^{m,1}$ means a function f having m derivatives such that $f^{(m)}$ is Lipschitz-continuous.

By taking $V = 0$, we recover Theorem 7.2 since the formula (7.5) for the limiting variance is obtained by observing that the operator $\mathscr{L}$ is easily diagonalized in the Fourier basis.

In this general setting we also show that $\sigma_\theta^V(f) \to \|f\|_{L^2(\mu_V^0)}$ as $\theta \to 0$ and we provide a sufficient condition so that $\sqrt{\theta}\,\sigma_\theta^V(f) \to \|f\|_{H^{1/2}}$ [P15, Section 8].

The proof of Theorem 7.5 relies on the concentration bounds in Theorem 7.4 and on a normal approximation technique introduced in [Lambert, Ledoux, and Webb, 2017], inspired from the Stein method [Ross, 2011]. To give an insight on why the operator $\mathscr{L}$ appears in this context, consider the "generator" $\mathbb{L} = \Delta + \frac{2\theta}{N}\nabla H_N \cdot \nabla$ acting on smooth functions $\mathbb{T}^N \to \mathbb{R}$, where H_N is the energy associated with the canonical ensemble (7.6). For smooth test function $f : \mathbb{T} \to \mathbb{R}$, we then show the approximate-commutation relation $\mathbb{L}\mathsf{Fluct}_N(f) = \mathsf{Fluct}_N(-\mathscr{L}f) + \mathsf{Error}(N)$ as $N \to \infty$, which in turn yields an asymptotic control for the limiting fluctuations by the Stein approach. Let us stress that Lambert et al. [2017] already used this approach to quantify the rate of convergence for the fluctuations of beta-Ensembles on $\mathbb{R}$ at fixed temperature, but there is a substantial technical difference in the high temperature regime due to the fact that the operator $\mathscr{L}$ has an extra Sturm-Liouville component. In particular, the spectral behavior of $\mathscr{L}$ is quite different, which is responsible for the different rate of convergence and limiting variance.

Let us also mention that a similar temperature regime has been studied in Guionnet and Bodineau [1999] for a two-component plasma model confined in a 2D box with different techniques, but with a similar limiting variance. An important progress has been recently made by Serfaty [2020] for the Coulomb gas on $\mathbb{R}^d$ in a temperature regime that is arbitrary close to this critical regime, including a CLT in dimension $d = 3$ that is new even at fixed temperature.

Perspectives. One reason to choose to work with the CβE model was to benefit from the compact support of the particle system and thus to avoid technical difficulties in identifying the limiting variance structure for the fluctuations. From this perspective, it would be interesting to describe the variance structure for higher dimensional Coulomb gas on a compact Riemannian manifold M at inverse temperature $\beta = \frac{2\theta}{N}$ and, say, without potential (one could even start with the 2-sphere as in the previous chapter). In this setting $\hat{\mu}_N$ converges a.s. to the uniform measure σ on M; concentration inequalities are already available from [García-Zelada, 2019]. Preliminary computations using similar techniques as we used for the CβE lead to a limiting variance defined, for any smooth test function f satisfying $\int f \mathrm{d}\sigma = 0$, by $\sigma_\theta^2(f) = \langle f, \mathscr{L}^{-1} f\rangle_{H^1}$, where H^1 stands for the usual Sobolev space and

$$\mathscr{L}f := -\Delta f + \theta f.$$

This would now yield an interpolation between the L^2 and H^1 spaces structure, in the sense that $\sigma_\theta(f) \to \|f\|_{L^2}$ when $\theta \to 0$ and $\sqrt{\theta}\,\sigma_\theta(f) \to \|f\|_{H^1}$ when $\theta \to \infty$. This supports the conjecture that the fluctuations of the Coulomb gas in any dimension at fixed temperature should be described by a Gaussian free field. However there are technical difficulties in the analysis that we did not overcome yet.

Chapter 8

DLR equations for Sine$_\beta$

Based on the joint work [P13] with *David Dereudre* (Université de Lille), *Thomas Leblé* (CNRS, Université de Paris) et *Mylène Maïda* (Université de Lille).

The Sine$_\beta$ process. We continue to investigate the CβE at a given (fixed) inverse temperature $\beta > 0$ but at the microscopic level: What is the limit in law of the point process $(N(\mathbf{x}_j - x))_{j=1}^N$ as $N \to \infty$ for some fixed reference point $x \in \mathbb{T}$, when $(\mathbf{x}_j)_{j=1}^N$ is distributed according to the CβE? As we have seen in the first chapter, when $\beta = 2$ the microscopic limit is the DPP associated with the sine kernel (1.5). More generally, when $\beta = 1, 2$ or 4, this limiting process is rather well understood due to the determinantal/Pfaffian structure available for their correlations functions [Deift and Gioev, 2009]; see also [Forrester, 1993a, Chapter 13] for other special cases with more involved formulas. For general $\beta > 0$, a limiting process exists and is called the Sine$_\beta$ process, although the mere existence of this limit is a non-trivial result which was obtained, together with a rather involved description of the limiting object, by Killip and Stoiciu [2009]. Similarly, Valkó and Virág [2009] obtained the existence and a sophisticated description of the microscopic limit of the GβE when the reference point x lies in the bulk[?], and it turns out to be Sine$_\beta$ as well[♣]. If one replaces the gaussian weight $\mathrm{e}^{-N\beta x_j^2/4}$ in (GβE) by a general weight $\mathrm{e}^{-N\beta V(x_j)}$, then there exists a large class of potentials V for which Sine$_\beta$ still arises in the bulk microscopic limit [Bourgade et al., 2014, 2012; Shcherbina, 2014; Bekerman et al., 2015]. One can also generalize the interaction term $|x_i - x_j|^\beta$ to $h(x_i - x_j)$ for some function h satisfying $h(x) \sim |x|^\beta$ as $x \to 0$ and smooth elsewhere without changing the microscopic limit [Venker, 2013]. From this perspective, the Sine$_\beta$ process is a universal object appearing as the bulk microscopic limit for a large class of interacting particle systems.

Describing Sine$_\beta$. The aformentioned descriptions of Sine$_\beta$ involve stochastic differential equations that yield the number of particles falling into a given interval; the

[?] The *bulk* commonly refers to the subset where the macroscopic limit has a positive density, which is here $(-2, 2)$ for the GβE.

[♣] See e.g. [Nakano, 2014] for the identification of these two limits.

starting point is the representation of the $\mathsf{C}\beta\mathsf{E}/\mathsf{G}\beta\mathsf{E}$ as the eigenvalues of a banded random matrix. Such representations turn out to be tractable enough to study fine properties of the point process, such as large gap probability estimates [Valkó and Virág, 2010], a CLT [Kritchevski et al., 2012] and large/maximum deviation estimates for the number of points in an interval [Holcomb and Valkó, 2015, 2017; Holcomb and Paquette, 2018], as well as the Poissonian behavior of Sine_β when $\beta \to 0$ [Allez and Dumaz, 2014].

More recently, the process has been alternately characterized by Valkó and Virág [2017] as the spectrum of a random infinite-dimensional operator, that allows a better understanding on the β-dependency of the process [Valkó and Virág, 2020].

The goal of this chapter is to present yet another description for Sine_β that may be more natural from a statistical physics perspective, by means of canonical Dobrushin-Lanford-Ruelle (DLR) equations. In short, we show that the Sine_β process is a natural infinite Gibbs measure at inverse temperature $\beta > 0$ associated with the logarithmic pair potential interaction, that we could informally write

$$\text{“ } \mathrm{d}\mathsf{Sine}_\beta(x_1, x_2, \ldots) = \frac{1}{Z_{\beta,\infty}} \prod_{1 \le i < j \le \infty} |x_i - x_j|^\beta \prod_{j=1}^{\infty} \mathrm{d}x_j \text{ . ”} \tag{8.1}$$

The Dobrushin-Lanford-Ruelle approach. Imagine that γ is a random point configuration on $\mathbb{R}$ with distribution the right hand side of (8.1); it does not make any sense but let's keep the reasoning at an informal level. Thus γ has a "density" proportional to $\mathrm{e}^{-\beta \mathsf{H}}$ with energy

$$\mathsf{H}(\gamma) := \sum_{\substack{x,y \in \gamma \\ x<y}} g(x-y)$$

associated with the logarithmic (or two-dimensional Coulomb) interaction

$$g(x) = \log |x|^{-1}.$$

Something that we could make sense of is the conditional law of the restriction $\gamma_\Lambda := \gamma \cap \Lambda$ of the random configuration to a fixed compact subset $\Lambda \subset \mathbb{R}$, knowing the exterior configuration $\gamma_{\Lambda^{\mathsf{c}}}$, where $\Lambda^{\mathsf{c}} := \mathbb{R} \setminus \Lambda$. More precisely, we are looking for a probability distribution $\rho_\Lambda(\,\cdot\,|\gamma_{\Lambda^{\mathsf{c}}})$ on the set $\mathsf{Conf}(\Lambda)$ of point configurations on Λ satisfying, for any bounded and measurable test function f,

$$\mathbb{E}_{\mathsf{Sine}_\beta}\Big[f(\gamma)\Big] = \mathbb{E}_{\mathsf{Sine}_\beta}\Big[\int_{\mathsf{Conf}(\Lambda)} f(\eta \cup \gamma_{\Lambda^{\mathsf{c}}})\, \rho_\Lambda(\mathrm{d}\eta|\gamma_{\Lambda^{\mathsf{c}}})\Big]. \tag{8.2}$$

To identify the conditional laws associated with (8.1), we consider the mutual energy of two point configurations γ, η defined by

$$\mathsf{H}(\gamma, \eta) := \sum_{\substack{x \in \gamma \\ y \in \eta}} g(x-y).$$

Since the energy formally splits as $\mathsf{H}(\gamma) = \mathsf{H}(\gamma_\Lambda)+\mathsf{H}(\gamma_\Lambda, \gamma_{\Lambda^c})+\mathsf{H}(\gamma_{\Lambda^c})$, the conditional density of γ_Λ knowing γ_{Λ^c} is thus proportional to the function $\eta \mapsto \mathrm{e}^{-\beta(\mathsf{H}(\eta)+\mathsf{H}(\eta,\gamma_{\Lambda^c}))}$ "defined" on $\mathsf{Conf}(\Lambda)$. Although for any $\eta \in \mathsf{Conf}(\Lambda)$ the finite sum $\mathsf{H}(\eta)$ makes sense it is not clear if $\mathsf{H}(\eta, \gamma_{\Lambda^c})$ does since the configuration γ_{Λ^c} is infinite. To deal with this problem we consider instead the move function $\mathsf{Move}_\Lambda(\eta, \gamma) := \mathsf{H}(\eta, \gamma_{\Lambda^c})-\mathsf{H}(\gamma_\Lambda, \gamma_{\Lambda^c})$, which represents the energetic change when moving the configuration γ_Λ into η inside Λ in presence of the exterior configuration γ_{Λ^c}. Now, one can hope that, due to possible cancellations or compensations, the infinite sum $\mathsf{Move}_\Lambda(\eta, \gamma)$ converges. This yields a candidate for the conditional density, defined up to a multiplicative constant by

$$\rho_\Lambda(\eta|\gamma_{\Lambda^c}) \propto \mathrm{e}^{-\beta(\mathsf{H}(\eta)+\mathsf{Move}_\Lambda(\eta,\gamma))}, \tag{8.3}$$

with respect to the reference measure given by the $\mathsf{N}_\Lambda(\gamma)$-fold product of the Lebesgue measure on Λ, where $\mathsf{N}_\Lambda(\gamma)$ stands for the number of particles of γ inside Λ.

In conclusion, the informal representation (8.1) for Sine_β can be rigorously recast by saying that its conditional densities satisfying (8.2) are given by (8.3), provided that the move functions are well defined, and this is indeed what we obtain.

Theorem 8.1 (DLR equations)**.** *For all $\beta > 0$ and for every compact $\Lambda \subset \mathbb{R}$:*

(a) *The move function $\mathsf{Move}_\Lambda(\eta, \gamma)$ exists[◇] for every $\eta \in \mathsf{Conf}(\Lambda)$ and for Sine_β-almost every $\gamma \in \mathsf{Conf}(\mathbb{R})$, .*

(b) *The DLR equations (8.2)–(8.3) hold true for the Sine_β process.*

This representation was already known when $\beta = 2$, where the determinantal structure is available [Bufetov, 2016].

Let us stress that the density (8.3) could have been taken with respect to another reference measure than the $\mathsf{N}_\Lambda(\gamma)$-fold Lebesgue measure. The fact that this is the appropriate reference measure here underlies a property called *tolerance* that we have implicitly obtained for the Sine_β process.

A technical part of the proof was to establish the existence of the move function, for which we needed delicate estimates for the variance of the number of points from Sine_β falling into a given compact set. Such estimates were available here thanks to the energetic approach developed by Leblé and Serfaty [2017].

Rigidity. Another interesting feature for Sine_β we obtain from the DLR equations is that it is number-rigid, in the sense that if $\gamma \sim \mathsf{Sine}_\beta$ then the knowledge of the exterior configuration γ_{Λ^c} determines a.s. the number $\mathsf{N}_\Lambda(\gamma)$ of particles of γ in Λ.

Theorem 8.2 (Number-rigidity)**.** *For all $\beta > 0$ and compact $\Lambda \subset \mathbb{R}$, there exists a measurable map $\mathsf{F} : \mathsf{Conf}(\Lambda^c) \to \mathbb{N}$ such that, if $\gamma \sim \mathsf{Sine}_\beta$, then $\mathsf{N}_\Lambda(\gamma) = \mathsf{F}(\gamma_{\Lambda^c})$.*

Theorem 8.2 has been obtained independently by Chhaibi and Najnudel [2018] without the use of the DLR formalism but following instead the initial approach of

[◇]More precisely, the limit of $\mathsf{Move}_\Lambda(\eta, \gamma \cap [-p, p]))$ exists when $p \to \infty$.

Ghosh and Peres [2017], which is to show that the variance of the smoothed number of points from Sine_β falling into a bounded interval can be made arbitrary small. To do so, they use variance estimates for the $\mathsf{C}\beta\mathsf{E}$ obtained by Jiang and Matsumoto [2015] based Jack's special functions identities that are tied to this specific model.

Our strategy relies instead on the following general result, that may be of independent interest [P13, Theorem 3.18]: Consider an interaction $g : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ in dimension $d \geq 1$ that has a very long range, in the sense that $g(x) \to -\infty$ as $|x| \to \infty$, and that is smooth away from the origin. Assume there is a class $\mathscr{C}$ of probability measures on $\mathsf{Conf}(\mathbb{R}^d)$ that is stable by disintegration and such that one can define P-a.s. the move functions associated with g for every $P \in \mathscr{C}$. Then any stationary distribution $P \in \mathscr{C}$ that satisfies the DLR equations is number-rigid.

Perspectives. A natural question, related to the existence of phase transitions in statistical physics, is to wonder if the Sine_β is the only stationary process satisfying the DLR equations. In general, uniqueness of the infinite Gibbs measure is expected in dimension one, although the general theory only applies to short range interactions [Friedli and Velenik, 2017, Section 6.5.5]. In the determinantal case $\beta = 2$, the uniqueness follows from [Kuijlaars and Miña-Díaz, 2019]. We expect this holds for general $\beta > 0$, and a possible strategy to prove this is to use Theorem 8.1 with $\Lambda = [-L, L]$ to represent the conditional law as a log-gas on $[-L, L]$ and to show this process converges to Sine_β as $L \to \infty$, perhaps by refining the transportation approach from [Shcherbina, 2014; Bekerman, Figalli, and Guionnet, 2015].

Next, it would interesting to see if our methods can be extended to non-stationary processes so as to tackle other one-dimensional universal processes like the Airy_β or the Bessel_β processes; the main problem is to find alternatives to the discrepancy estimates we used for Sine_β. When $\beta = 2$, both the Airy_2 process and the Bessel_2 process indeed satisfy well-defined DLR equations [Bufetov, 2016], and the uniqueness has been obtained for the Bessel_2 process [Molag and Stevens, 2019].

Another interesting problem would be to carry out this analysis to the two-dimensional Coulomb gas, since in this setting any solution of the DLR equations would be number-rigid thanks to our results. It is not clear however if there should be a unique solutions for the DLR equations or not, and the negative would entail a phase transition at a critical temperature(s?) that remains to be identified. There are indeed many numerical simulations providing evidence for this phase transition, see e.g. [Strandburg, 1988; Moore and Pérez-Garrido, 1999; Clark et al., 2009].

Bibliography

Abreu, L. D., Haimi, A., Koliander, G., and Romero, J. L. Filtering with wavelet zeros and Gaussian analytic functions. *Preprint arXiv:1807.03183*, 2018.

Alishahi, K. and Zamani, M. The spherical ensemble and uniform distribution of points on the sphere. *Electron. J. Probab.*, 20(23):1–27, 2015.

Allez, R. and Dumaz, L. From Sine kernel to Poisson statistics. *Electronic journal of probability*, 19, 2014.

Armentano, D., Beltrán, C., and Shub, M. Minimizing the discrete logarithmic energy on the sphere: The role of random polynomials. *Transactions of the American Math. Society*, 363(6):2955–2965, 2011.

Baik, J., Ben Arous, G., and Péché, S. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5):1643–1697, 2005. ISSN 0091-1798.

Bardenet, R., Flamant, J., and Chainais, P. On the zeros of the spectrogram of white noise. *Applied and Computational Harmonic Analysis*, 2018.

Bekerman, F., Figalli, A., and Guionnet, A. Transport maps for β-matrix models and universality. *Comm. Math. Phys.*, 338:589–619, 2015.

Beltrán, C. Harmonic properties of the logarithmic potential and the computability of elliptic Fekete points. *Const. Approx.*, 37:135–165, 2013.

Beltrán, C., Corral, N., and Criado del Rey, J. G. Discrete and continuous Green energy on compact manifolds. *J. Approx. Theory*, 237:160–185, 2019.

Beltrán, C., Etayo, U., Marzo, J., and Ortega-Cerdà, J. A sequence of polynomials with optimal condition number. *Preprint arXiv:1903.01356 (to appear in J. Amer. Math. Soc.)*, 2020.

Ben Arous, G. and Guionnet, A. Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields*, 108(4):517–542, 1997. ISSN 0178-8051.

Berman, R. J. On large deviations for Gibbs measures, mean energy and gamma-convergence. *Constructive Approximation*, 48(1):3–30, 2018.

Betermin, L. and Sandier, E. Renormalized energy and symptotic expansion of optimal logarithmic energy on the sphere. *Const. Approx.*, 47(1):39–74, 2018.

Biane, P. Segal-Bargmann transform, functional calculus on matrix spaces and the theory of semi-circular and circular systems. *J. Funct. Anal.*, 144:232–286, 1997.

Biane, P. and Voiculescu, D. V. A free probability analogue of the Wasserstein metric on the trace-state space. *Geom. Funct. Anal.*, 11(6):1125–1138, 2001.

Blum, L., Cucker, F., Shub, M., and Smale, S. Complexity and real computation: a manifesto. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 6(1):3–26, 1996.

Bobkov, S. G. and Götze, F. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.

Bordenave, C. and Chafaï, D. Around the circular law. *Probab. Surv.*, 9:1–89, 2012.

Borodin, A. and Forrester, P. J. Increasing subsequences and the hard-to-soft edge transition in matrix ensembles. *J. Phys. A*, 36(12):2963–2981, 2003. ISSN 0305-4470. Random matrix theory.

Borodin, A. and Serfaty, S. Renormalized energy concentration in random matrices. *Comm. Math. Phys.*, 320(1):199–244, 2013.

Bourgade, P., Erdös, L., and Yau, H.-T. Bulk universality of general β-ensembles with non-convex potential. *J. of Math. Phy., special issue in honor of E. Lieb's 80th birthday,*, 53, 2012.

Bourgade, P., Erdös, L., and Yau, H.-T. Universality of general β-ensembles. *Duke Mathematical Journal*, 163(6):1127–1190, 2014.

Brauchart, J. S., Hardin, D. P., and Saff, E. B. The next-order term for optimal Riesz and logarithmic energy asymptotics on the sphere. *Contemp. Math.*, 578 (2):31–61, 2012.

Breuer, J. and Duits, M. The Nevai condition and a local law of large numbers for orthogonal polynomial ensembles. *Advances in Mathematics*, 265:441–484, 2014.

Breuer, J. and Duits, M. Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients. *J. Amer. Math. Soc.*, 30(1):27–66., 2017.

Bufetov, A. I. Conditional measures of determinantal point processes. *arXiv preprint arXiv:1605.01400*, 2016.

Capitaine, M. and Péché, S. Fluctuations at the edge of the spectrum of the full rank deformed gue. *Probab. Theory Relat. Fields*, 165:117–161, 2015.

Chafaï, D. and Ferré, G. Simulating *c*oulomb gases and log-gases with hybrid *m*onte *c*arlo algorithms. *J. Stat. Phys.*, 174:692–714, 2019.

Chafaï, D., Gozlan, N., and Zitt, P.-A. First-order global asymptotics for confined particles with singular pair repulsion. *Ann. Appl. Probab.*, 24(6):2371–2413, 2014.

Chhaibi, R. and Najnudel, J. Rigidity of the Sine_β process. *arXiv preprint arXiv:1804.01216*, 2018.

Clark, B. K., Casula, M., and Ceperley, D. M. Hexatic and mesoscopic phases in a 2d quantum coulomb system. *Phys. Rev. Lett.*, 103(055701), 2009.

Coeurjolly, J.-F., Mazoyer, A., and Amblard, P.-O. Monte Carlo integration of non-differentiable functions on $[0,1]^i$, $i = 1, \dots, d$, using a single determinantal point pattern defined on $[0,1]^d$. *Preprint arXiv:2003.10323*, 2020.

Cohen, L. *Time-frequency analysis*, volume 778. Prentice Hall PTR Englewood Cliffs, NJ:, 1995.

Daubechies, I. and Paul, T. Time-frequency localization operators – a geometric phase space approach: Ii. the use of dilations. *Inverse problems*, 4:661–680, 1988.

Deift, P. and Gioev, D. *Random matrix theory: invariant ensembles and universality*, volume 18. Courant Lecture Notes in Mathematics. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, Rhode Island, 2009.

Dembo, A. and Zeitouni, O. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

Diaconis, P. and Forrester, P. J. Hurwitz and the origin of random matrix theory in mathematics. *Random Matrices Theory Appl.*, 6(1730001):(26 pages), 2017.

Dragnev, P. D. On the separation of logarithmic points on the sphere. *Approximation Theory, X (St. Louis, MO, 2001), Innov. Appl. Math.*, pages 137–144, 2002.

Dubickas, A. On the maximal product of distances between points on a sphere. *Liet. Mat. Rink.*, 36(3):303–312, 1996.

Dumitriu, I. and Edelman, A. Matrix models for beta ensembles. *J. Math. Phy.*, 43 (11):5830–5847, 2002.

Duse, E. and Metcalfe, A. Asymptotic geometry of discrete interlaced patterns: Part i. *International Journal of Mathematics*, 26(11):1550093, 2017.

Duse, E. and Metcalfe, A. Asymptotic geometry of discrete interlaced patterns: Part ii. *Annales de l'Institut Fourier*, 70(1):375–436, 2020.

Edelman, A., Guionnet, A., and Péché, S. Beyond universality in random matrix theory. *Ann. Appl. Probab.*, 26(3):1659–1697, 2016.

El Karoui, N. Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Ann. Probab.*, 35(2):663–714, 2007. ISSN 0091-1798.

Etayo, U. and Beltrán, C. The diamond ensemble: a constructive set of points with small logarithmic energy. *J. Complexity*, 59, 2020.

Flandrin, P. *Time-frequency/time-scale analysis*, volume 10. Academic press, 1998.

Flandrin, P. Time–frequency filtering based on spectrogram zeros. *IEEE Signal Processing Letters*, 22(11):2137–2141, 2015.

Forrester, P. J. Exact integral formulas and asymptotics for the correlations in the $1/r^2$ quantum many-body system. *Phys. Lett. A*, 179(2):127–130, 1993a.

Forrester, P. J. The spectrum edge of random matrix ensembles. *Nuclear Phys. B*, 402(3):709–728, 1993b. ISSN 0550-3213.

Friedli, S. and Velenik, Y. *Statistical mechanics of lattice systems: a concrete mathematical introduction*. Cambridge University Press, 2017.

García-Zelada, D. A large deviation principle for empirical measures on polish spaces: Application to singular gibbs measures on manifolds. *Preprint arXiv:1703.02680. To appear in Annales de l'Institut Henri Poincaré.*, 2018.

García-Zelada, D. Concentration for coulomb gases on compact manifolds. *Electron. Commun. Probab.*, 24(12):18p, 2019.

Gautier, G., Polito, G., Bardenet, R., and Valko, M. Dppy: Sampling dpps with python. *Journal of Machine Learning Research - Machine Learning Open Source Software (JMLR-MLOSS)*, 20(180):1–7, 2019.

Ghosh, S. and Peres, Y. Rigidity and tolerance in point processes: Gaussian zeros and ginibre eigenvalues. *Duke Mathematical Journal*, 166(10):1789–1858, 2017.

Gozlan, N. and Léonard, C. Transport inequalities. A survey. *Markov Process. Related Fields*, 16(4):635–736, 2010.

Gröchenig, K. *Foundations of time-frequency analysis*. Birkhäuser, 2001.

Guionnet, A. and Bodineau, T. About the stationary states of vortex systems. *Ann. Inst. H. Poincaré Probab. Statist.*, 53(2):205–237, 1999.

Hardin, D. P., Michaels, T., and Saff, E. B. A comparison of popular point configurations on $\mathbb{S}^2$. *Dolomites Res. Notes Approx.*, 9:16–49, 2016.

Hiai, F., Petz, D., and Ueda, Y. Free transportation cost inequalities via random matrix approximation. *Probab. Theory Related Fields*, 130(2):199–221, 2004.

Holcomb, D. and Paquette, E. The maximum deviation of the Sine$_\beta$ counting process. *Preprint. ArXiv:1801.08989*, 2018.

Holcomb, D. and Valkó, B. Large deviations for the Sine$_\beta$ and Sch$_\tau$ processes. *Probability Theory and Related Fields*, 163(1-2):339–378, 2015.

Holcomb, D. and Valkó, B. Overcrowding asymptotics for the Sine$_\beta$ process. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*, 53(3):1181–1195, 2017.

Hough, J. B., Krishnapur, M., Peres, Y., and Virág, B. *Zeros of Gaussian analytic functions and determinantal point processes*, volume 51. American Mathematical Society Providence, RI, 2009.

Hurwitz, A. über die erzeugung der invarianten durch integration. *Nachr. Ges. Wiss.Göttingen*, pages 71–90, 1897.

Ismail, M. E. H. *Classical and quantum orthogonal polynomials in one variable*. Cambridge University Press, 2005.

Jiang, T. and Matsumoto, S. Moments of traces of circular beta-ensembles. *Ann. Probab.*, 43(6):3279–3336, 2015.

Johansson, K. On Szegö's asymptotic formula for Toeplitz determinants and generalizations. *Bull. Sci. Math.*, 112(3):257–304, 1988.

Johansson, K. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209 (2):437–476, 2000. ISSN 0010-3616.

Killip, R. and Stoiciu, M. Eigenvalue statistics for cmv matrices: from poisson to clock via random matrix ensembles. *Duke Mathematical Journal*, 146(3):361–399, 2009.

Knowles, A. and Yin, J. Anisotropic local laws for random matrices. *Probab. Theory Relat. Fields*, 169:257–352, 2017.

Kritchevski, E., Valkó, B., and Virág, B. The scaling limit of the critical one-dimensional random schrödinger operator. *Communications in Mathematical Physics*, pages 1–32, 2012.

Kuijlaars, A. and Miña-Díaz, E. Universality for conditional measures of the sine point process. *J. Approx. Theory*, 243:1–24, 2019.

Lambert, G. Mesoscopic central limit theorem for the circular beta-ensembles and applications. *Preprint arXiv:1902.06611*, 2019.

Lambert, G., Ledoux, M., and Webb, C. Quantitative normal approximation of linear statistics of beta-ensembles. *Preprint arXiv:1706.10251*, 2017.

Leblé, T. and Serfaty, S. Large Deviation Principle for Empirical Fields of Log and Riesz Gases. *Inventiones mathematicae*, 210(3), 2017.

Ledoux, M. and Popescu, I. Mass transportation proofs of free functional inequalities, and free Poincaré inequalities. *J. Funct. Anal.*, 257(4):1175–1221, 2009.

Lu, Y. and Mattingly, J. C. Geometric ergodicity of *l*angevin dynamics with *c*oulomb interactions. *Nonlinearity*, 33(2), 2019.

Maïda, M. and Maurel-Segala, É. Free transport-entropy inequalities for non-convex potentials and application to concentration for random matrices. *Probab. Theory Related Fields*, 159(1-2):329–356, 2014.

Marcenko, V. A. and Pastur, L. A. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.

Marton, K. A simple proof of the blowing-up lemma. *IEEE Trans. Inform. Theory*, 32(3):445–446, 1986.

Meckes, E. S. and Meckes, M. W. Concentration and convergence rates for spectral measures of random matrices. *Probab. Theory Relat. Fields*, 156:145–164, 2013.

Molag, L. and Stevens, M. Universality for conditional measures of the Bessel point process. *Preprint arXiv:1904.04349*, 2019.

Moore, M. A. and Pérez-Garrido. Absence of a finite-temperature melting transition in the classical two-dimensional one-component plasma. *Phys. Rev. Lett.*, 82:4078–4081, 1999.

Nakano, F. Level statistics for one-dimensional schrödinger operators and gaussian beta ensemble. *Journal of Statistical Physics*, 156(1):66–93, 2014.

Nakano, F. and Trinh, K. D. Gaussian beta ensembles at high temperature: eigenvalue fluctuations and bulk statistics. *J. Stat. Phys.*, 173(2):296–321, 2018.

Onatski, A. The Tracy-Widom limit for the largest eigenvalues of singular complex Wishart matrices. *Ann. Appl. Probab.*, 18(2):470–490, 2008. ISSN 1050-5164.

Pakzad, C. Poisson statistics at the edge of gaussian beta-ensembles at high temperature. *Preprint arXiv:1804.08214*, 2018.

Peres, Y. and Virág, B. Zeros of the i.i.d. Gaussian power series: a conformally invariant determinantal process. *Acta Mathematica*, 194(1):1–35, 2005.

Popescu, I. Local functional inequalities in one-dimensional free probability. *J. Funct. Anal.*, 264(6):1456–1479, 2013.

Ross, N. Fundamentals of stein's method. *Probab. Surveys*, 8:210–293, 2011.

Rougerie, N. and Serfaty, S. Higher-dimensional Coulomb gases and renormalized energy functionals. *Comm. Pure Appl. Math.*, 69(3):519–605, 2016.

Sandier, E. and Serfaty, S. From the Ginzburg-Landau model to vortex lattice problems. *Comm. Math. Phys.*, 313(3):635–743, 2012.

Serfaty, S. *Coulomb gases and Ginzburg-Landau vortices*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2015.

Serfaty, S. Gaussian fluctuations and free energy expansion for 2d and 3d coulomb gases at any temperature. *Preprint arXiv:2003.11704*, 2020.

Shcherbina, M. Change of variables as a method to study general β-models: *b*ulk universality. *J. Math. Phy.*, 55(4):043504, 2014.

Shub, M. and Smale, S. Complexity of Bezout's theorem. III. Condition number and packing. *J. Complexity*, 9(1):4–14, 1993. Festschrift for Joseph F. Traub, Part I.

Silverstein, J. W. and Choi, S. Analysis of the limiting spectral distribution of large-dimensional random matrices. *J. Multivariate Anal.*, 54(2):295–309, 1995. ISSN 0047-259X.

Smale, S. Mathematical problems for the next century. *Mathematics: Frontiers and Perspectives*, pages 271–294, 2000.

Strandburg, K. J. Two-dimensional melting. *Rev. Mod. Phys.*, 60:161–207, 1988.

Torquato, S. Hyperuniformity and its generalizations. *Phys. Rev. E*, 94(022122), 2016.

Trinh, K. D. Global spectrum fluctuations for gaussian beta ensembles: a martingale approach. *J. Theor. Probab.*, pages 1–18, 2017.

Valkó, B. and Virág, B. Continuum limits of random matrices and the Brownian carousel. *Inventiones mathematicae*, 177(3):463–508, 2009.

Valkó, B. and Virág, B. Large gaps between random eigenvalues. *Ann. Probab.*, 38 (3):1263–1279, 2010.

Valkó, B. and Virág, B. The Sine$_\beta$ operator. *Inventiones mathematicae*, 209(1): 275–327, 2017.

Valkó, B. and Virág, B. Operator limit of the circular beta ensemble. *Ann. Probab.*, 48(3):1286–1316, 2020.

Venker, M. Particle systems with repulsion exponent β and random matrices. *Electron. Commun. Probab.*, 18(83):1–12, 2013.

Vershynin, R. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.

Villani, C. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003. ISBN 0-8218-3312-X.