Introduction to determinantal point processes

DPP and fermions 2022 - ENS Lyon

Mylène MAÏDA - Université de Lille

mylene.maida@univ-lille.fr

Aknowledgements : this course is largely based on a course elaborated with Adrien Hardy (U. Lille, on leave)

Outline of the course

1. (Very) quick historical background
2. What is a determinantal point process (DPP)?
3. Some important examples
4. DPP associated to projections
5. Asymptotics of DPP : global regime
6. Asymptotics of DPP : local regime.
7. Variants of DPP give interesting point processes
8. Conclusion

1. A quick historical background

- 50' Wigner : link between the energy levels of heavy nuclei and eigenvalues of some random matrices

$$f(\lambda_1, \dots, \lambda_N) \propto e^{-\frac{N}{2}\sum_{j=1}^{N} \lambda_j^2} \prod_{i<j} (\lambda_i - \lambda_j)^2$$

- 1962 Dyson : determinantal structure of the correlations

~ 1975 Macchi : study of the determinantal structure of "fermionic systems"

- 00' → DPP arising in many (discrete or continuous) models : OP ensembles, random tilings, random permutations, etc.

- 10' → using DPP's in applied math (machine learning, signal processing ...)

2. What is a determinantal point process (DPP)?

2.1. What is a point process?

- A point process is a random configuration of points.

→ A (simple) configuration of points γ on $\mathbb{X} \subset \mathbb{R}^d$ is a locally finite set

ie $\forall K$ compact, $\mathrm{Card}(\gamma \cap K) < \infty$

Equivalently, one can see γ as an integer-valued Radon measure:

$$\gamma = \sum_{x \in \gamma} \delta_x$$

(so that, $\forall x \in \mathbb{R}^d$, $\gamma(\{x\}) \in \{0,1\}$ (γ is simple))

→ $\boxed{\text{A point process } \mathbb{P} \text{ on } \mathbb{X} \text{ is a probability measure on the set of configurations such that } \mathbb{P}(\gamma \text{ is simple}) = 1.}$

Ex: • if $X_1, \dots, X_n$ r.v. on $\mathcal{X}$ such that $\forall i \neq j$, $\mathbb{P}(X_i = X_j) = 0$

and $\forall \sigma \in \mathfrak{S}_n$, $(X_{\sigma(1)}, \dots, X_{\sigma(n)}) \overset{law}{=} (X_1, \dots, X_n)$,

then $\{X_1, \dots, X_n\}$ $\left(\sum_{i=1}^{n} \delta_{X_i} \right)$ is a point process (p.p.)

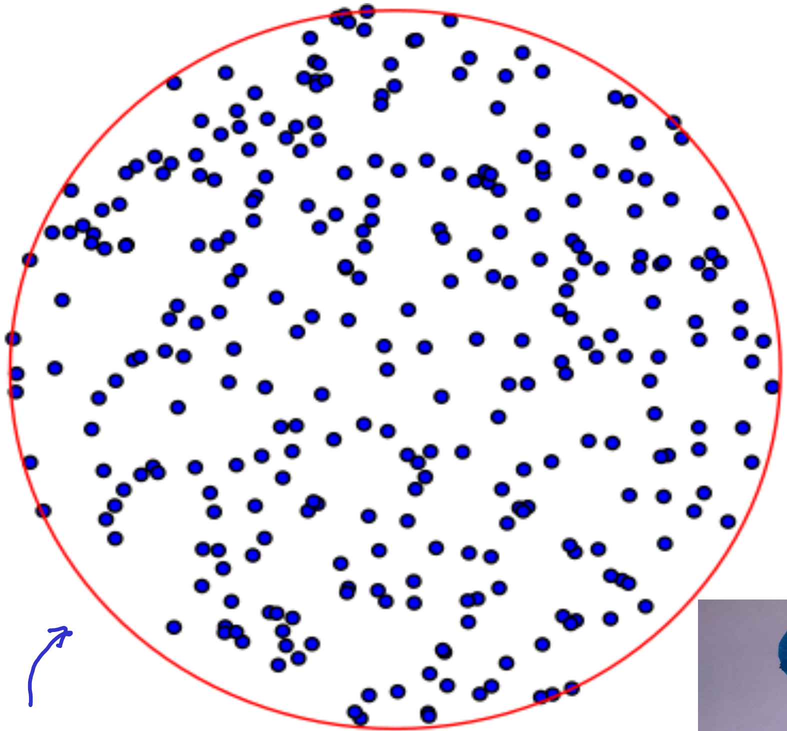
(cf Wigner example)

• Poisson point process ← important as a benchmark

Let f be non-negative and locally integrable on $\mathbb{R}^d$ the intensity measure.
Then there exists a unique p.p. $\mathbb{P}$ such that, under $\mathbb{P}$,

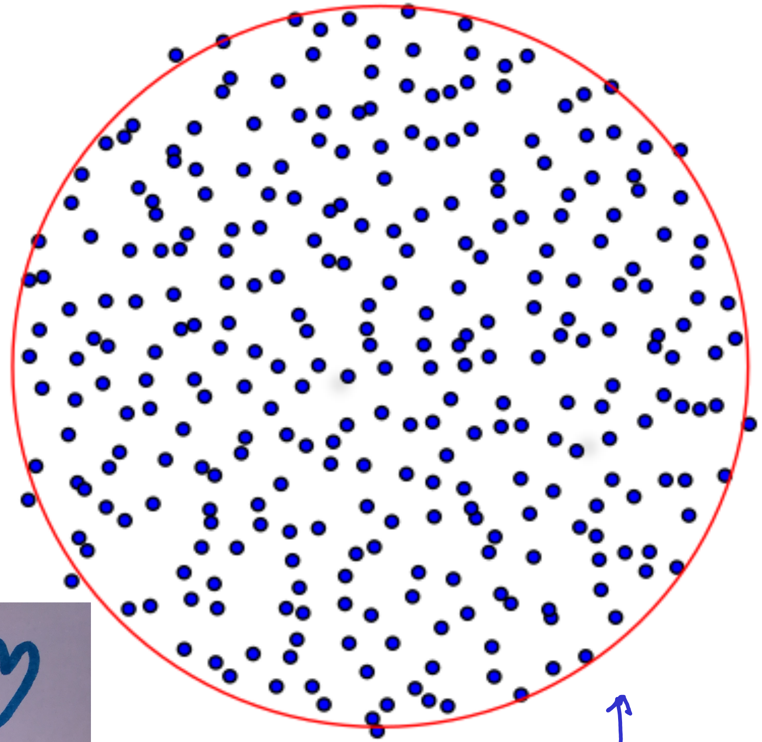
- $\forall A \in \mathcal{B}(\mathbb{R}^d)$ bounded, $\mathrm{Card}(\gamma \cap A) \overset{law}{=} \mathrm{Poi}\left(\int_A f \, d\lambda_d \right)$
- $\forall A_1, \dots, A_k$ disjoints, $\mathrm{Card}(\gamma \cap A_1), \dots, \mathrm{Card}(\gamma \cap A_k)$ are independent.

2.2. How does a DPP looks like ?

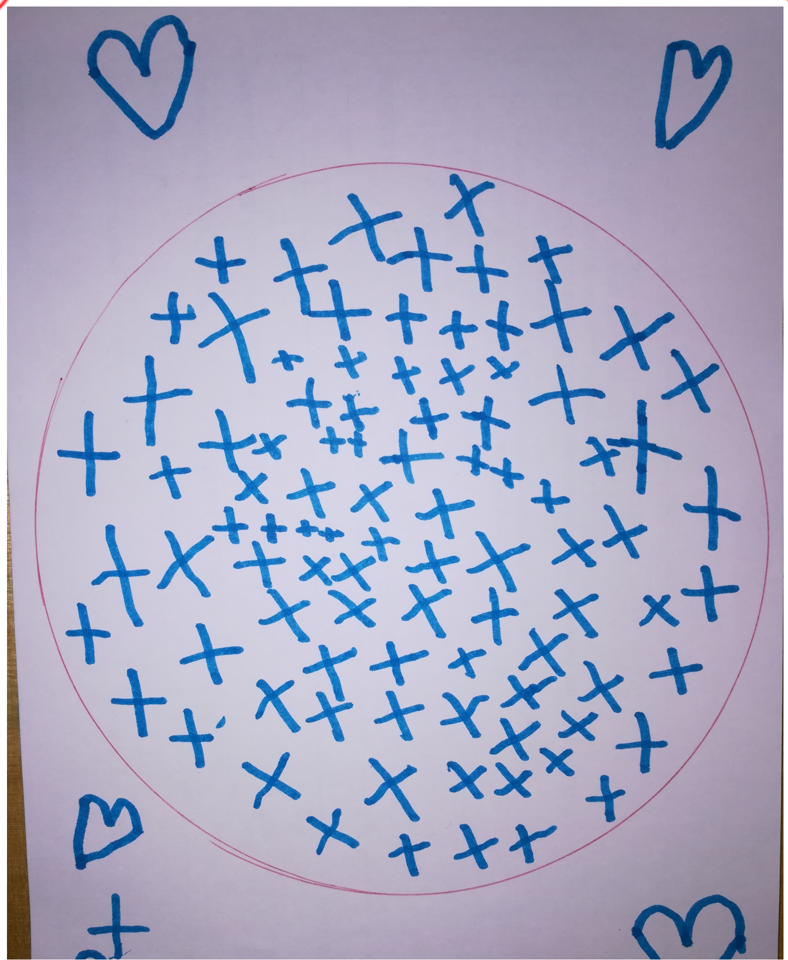


Independent uniform points on the unit disk

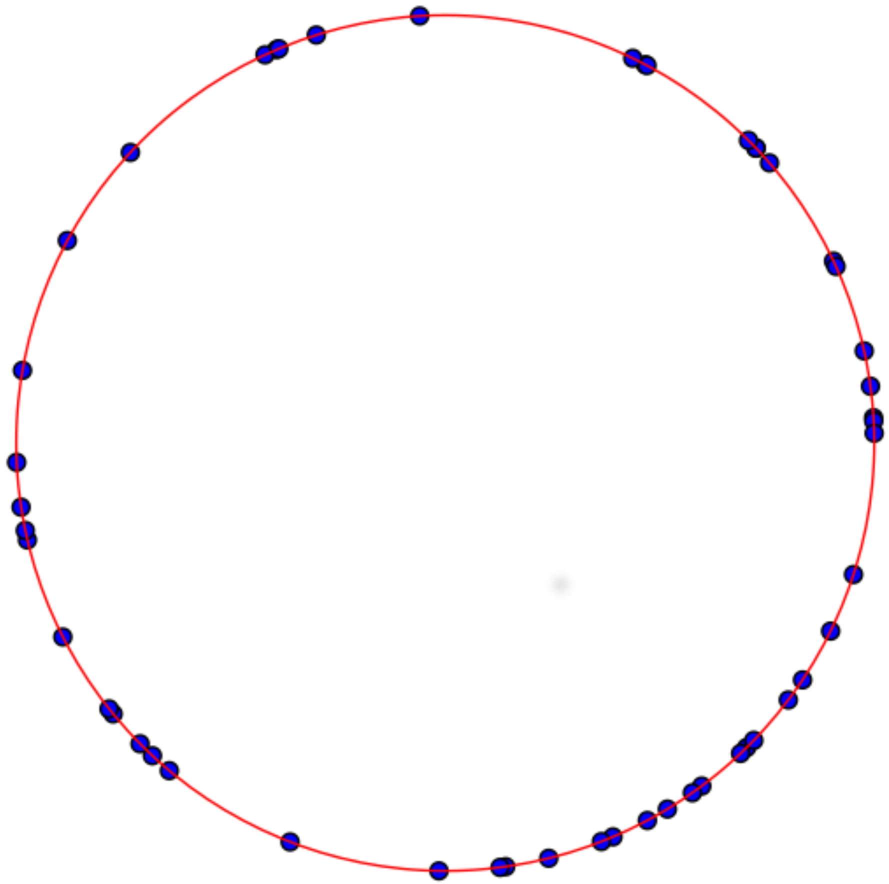
($\approx$ Poisson point process)



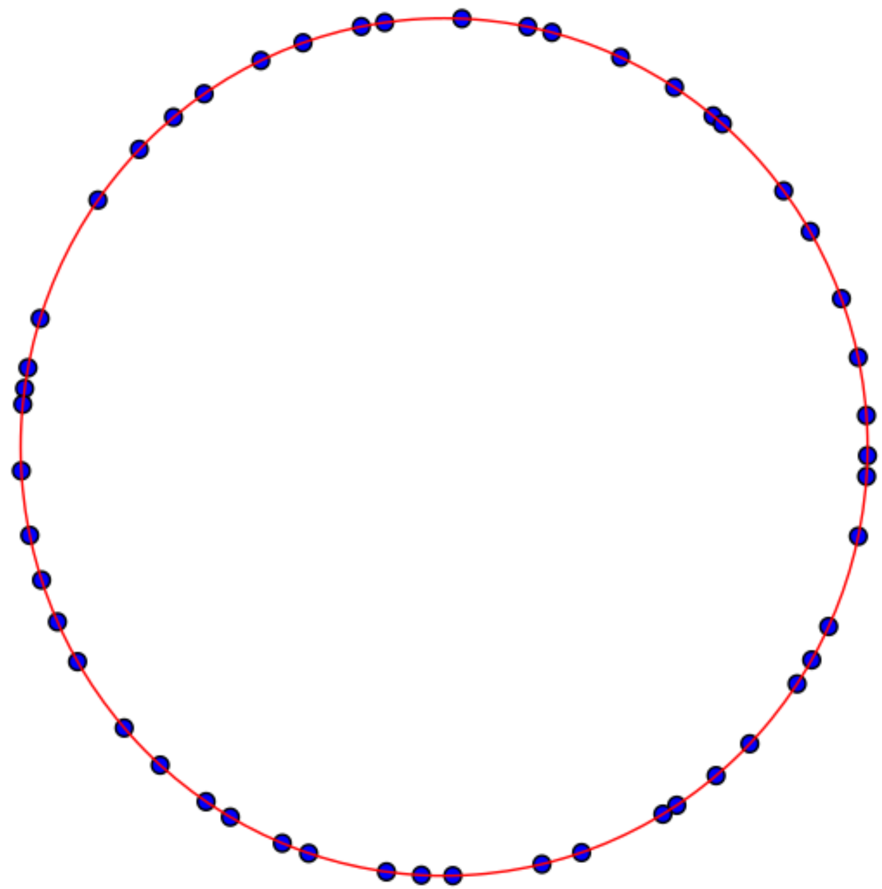
DPP: eigenvalues of a Ginibre matrix



A 5-years old drawing "many random points" on the unit disk.



Independent uniform points on the unit circle



Eigenvalues of a random Haar matrix.

2.3. What is a DPP?

In one sentence: " a DPP is a point process for which all correlation functions can be written as a determinant of minors of the same kernel matrix "

- Correlation functions: $\mathbb{P}$ a point process, μ reference measure

$\forall k \geq 1$, $\mathbb{P}$ has k-point correlation ρ_k iff,

$\forall f : \mathbb{R}^d \to \mathbb{R}$ bounded measurable,

$$\boxed{\mathbb{E}\Big(\sum_{\underbrace{\{X_1,\dots,X_k\}}_{\text{set of } k \text{ distinct points.}} \subset \gamma} f(X_1,\dots,X_k) \Big) = \int f(y_1,\dots,y_k) \rho_k(y_1,\dots,y_k)\, \mu(dy_1)\dots\mu(dy_k)}$$

set of k distinct points.

$\rho_k(y_1,\dots,y_k)$ = density of probability of having a k-tuple of points of the configuration γ at $y_1,\dots,y_k$.

Rq. • if μ is the counting measure over a discrete set :

$$\rho_k(y_1, \dots, y_k) = \mathbb{P}(\{y_1, \dots, y_k\} \subset \gamma)$$

• if μ is the Lebesgue measure on $\mathbb{R}$,

$$\rho_k(y_1, \dots, y_k) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^k} \mathbb{P}(\forall i \leq k, \gamma \text{ has a point in } [y_i, y_i + \varepsilon])$$

Ex 1 : if $\mathbb{P}$ is a PPP with intensity f, then $\rho_k = f^{\otimes k}$, that is $\rho_k(y_1, \dots, y_k) = f(y_1) \times \dots \times f(y_k)$

• ρ_1 is often called the intensity of the process.
= mean density of points.

Take $f = \mathbb{1}_A$,

$$\mathbb{E}\Big(\sum_{X \in \gamma} \mathbb{1}_A(X)\Big) = \mathbb{E}\big(\mathrm{Card}(\gamma \cap A)\big) = \int_A \rho_1(x)\, \mu(dx)$$

- ρ_2 is often called the <u>pair - correlation function</u>,

it is related to the variance of the linear statistics

$$\mathrm{Var}\left(\mathrm{Card}(\gamma \cap A)\right) = \underbrace{\int_A \rho_1(x)\,\mu(dx) - \left(\int_A \rho_1(x)\,\mu(dx)\right)^2}_{} + \int_{A\times A} \rho_2(x,y)\,\mu(dx)\,\mu(dy)$$

- <u>determinantal point process</u>

→ $K : \mathbb{X} \times \mathbb{X} \longrightarrow \mathbb{C}$ a semidefinite positive kernel :

$$\forall k \geq 1,\ \forall x_1, \dots, x_k \in \mathbb{X},\ \det\left(K(x_i, x_j)\right)_{i,j \leq k} \geqslant 0$$

→ μ reference measure.

def : $\mathbb{P}$ a point process is a DPP with kernel K and reference measure μ iff :

$$\forall k \geq 1,\ x_1, \dots, x_k \in \mathbb{X},\quad \boxed{\rho_k(x_1, \dots, x_k) = \det\left(K(x_i, x_j)\right)_{i,j \leq k}}$$

- In most examples, the kernel K is hermitian: $K(y,x) = \overline{K(x,y)}$
  or real symmetric: $K(y,x) = K(x,y)$.

Let us go back to the computation of expectation and variance of linear statistics.

$$\mathbb{E}\left(\sum_{X \in \gamma} f(X) \right) = \int_{\mathcal{X}} f(x) \underbrace{K(x,x)}_{\rho_1(x)} dx$$

$$\mathrm{Var}\left(\sum_{X \in \gamma} f(X) \right) = \int f^2(x) K(x,x)\, dx - \int f(x) f(y) K(x,y) K(y,x)\, dx\, dy$$

$$\overset{\text{si } K \text{ hermitien}}{=} \int f^2(x) K(x,x)\, dx - \int f(x) f(y) |K(x,y)|^2\, dx\, dy.$$

Rq: a DPP is characterized by its ref. measure μ and its kernel K, we denote it by $\underline{DPP(K,\mu)}$.

But K is not unique!

If h does not vanish, $\tilde{K}(x,y) := K(x,y) \dfrac{h(x)}{h(y)}$,

$$DPP(\tilde{K}, \mu) = DPP(K, \mu)$$

3. Some important examples of DPP

3.1. Eigenvalues of some specific random matrix ensembles

Here "ensemble" = probability measure on a set of matrices.

3.1.1. Ginibre ensemble

$$G := \frac{X + iY}{\sqrt{2}} \quad , \quad X, Y \text{ indep real } \mathcal{N}(0,1) \text{ random variables}$$

$(G_{ij})_{1 \leq i,j \leq n}$ iid copies of G

$$G_n = \left(\frac{G_{ij}}{\sqrt{n}} \right)_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{C})$$

The law g_n of G_n is the so-called Ginibre ensemble :

its density w.r.t Lebesgue measure of $\mathcal{M}_n(\mathbb{C})$ is $M \longmapsto \left(\frac{n}{\pi}\right)^{n^2} e^{-n \operatorname{Tr}(MM^*)}$

One can check that almost surely, G_n is diagonalisable, with distinct eigenvalues and we consider the point process $\mathbb{P}_{\mathrm{Gin}_n}$ of its eigenvalues.

Prop: if we denote by $(\lambda_1, .., \lambda_n)$ the eigenvalues in decreasing modulus, then the random vector $(\lambda_1, .., \lambda_n)$ has density $\mathbb{1}_{\Delta_n} f_n$, w.r.t. Lebesgue measure on $\mathbb{R}^n$, with

$$f_n(z_1, .., z_n) := \frac{n^{\frac{n(n+1)}{2}}}{\pi^n \prod_{j=1}^{n} j!} \prod_{i<j} |z_i - z_j|^2 \, e^{-n \sum_{j=1}^{n} |z_j|^2}$$

et $\Delta_n = \{ (z_1, .., z_n) \in \mathbb{C}^n \,/\, |z_1| \geq .... \geq |z_n| \}$

One can check that:

$$f_n(z_1, .., z_n) = \frac{1}{n!} \det \left(\sum_{k=0}^{n-1} \varphi_k^n(z_i) \overline{\varphi_k^n(z_j)} \right)_{1 \leq i,j \leq n},$$

with $$\varphi_k^n(z) = \frac{n^{\frac{k+1}{2}}}{\sqrt{\pi k!}} z^k e^{-\frac{n}{2}|z|^2}$$

We denote by:

$$\boxed{K_{Gin_n}(z,w) := \sum_{k=0}^{n-1} \varphi_k^n(z) \overline{\varphi_k^n(w)} = \frac{n}{\pi} e^{-\frac{n}{2}(|z|^2 + |w|^2)} \sum_{k=0}^{n-1} \frac{(n z \overline{w})^k}{k!}}$$

Important remark : K_{Gin_n} is the kernel on a projection operator on a subspace of $L^2(\mathbb{C})$ of finite dimension.

↪ we will go back in more details to this aspect

In practice, we have that for any $z_i, z_j \in \mathbb{C}$

$$\int_{\mathbb{C}} K(z_i, w) K(w, z_j) \, dw = K(z_i, z_j)$$

From there, one can check that, for all $1 \leq k \leq n$,

$$\boxed{\rho_k^{Gin}(z_1, \dots, z_k) = \det\left(K_{Gin_n}(z_i, z_j)\right)_{1 \leq i,j \leq k}.}$$

Otherwise stated,

$$\boxed{\mathbb{P}_{Gin_n} = DPP\left(K_{Gin_n}, d_2\right)}$$

Lebesgue measure on $\mathbb{C}$.

3.1.2. Similar ensembles : CUE, GUE, LUE

Ensemble	Joint law of e.v.	Reference measure	Expression of the kernel.
CUE_n "=" Haar measure on $\mathcal{U}_n(\mathbb{C})$	$\frac{1}{n!}\prod_{j<k}\lvert e^{i\theta_j}-e^{i\theta_k}\rvert^2$	$\frac{d\theta}{2\pi}$ sur $]-\pi,\pi]$	$K_{CUE_n}(\theta,\eta)=\sum_{k=0}^{n-1}\frac{e^{ik(\theta-\eta)}}{2\pi}$
GUE_n : $W_n=\frac{G_n+G_n^*}{\sqrt{2}}$ $\Updownarrow$ $C_n e^{-\frac{n}{2}Tr(M^2)}dM$	$\tilde{C}_n\prod_{i<j}\lvert x_i-x_j\rvert^2\, e^{-\frac{n}{2}\sum_{j=1}^{n}x_j^2}$	Leb sur $\mathbb{R}$	$K_{GUE_n}(x,y)=\sum_{k=0}^{n-1}\varphi_k^n(x)\varphi_k^n(y)$ avec $\varphi_k^n(x)=\left(\frac{n}{2\pi}\right)^{1/4}\frac{1}{\sqrt{k!}}H_k(\sqrt{n}\,x)\,e^{-\frac{nx^2}{4}}$; $k^{\text{ième}}$ polynôme d'Hermite
$LUE_{n,m}$: $W_{n,m}=X_{n,m}X_{n,m}^*$ avec $(X_{n,m})_{ij}=\frac{G_{ij}}{\sqrt{m}}$, $1\leq i\leq n,\ 1\leq j\leq m$	$\tilde{C}_{n,m}\prod_{i<j\leq n}(x_i-x_j)^2\prod_{i=1}^{n}x_i^{m-n}$ $\times\, e^{-m\sum_{j=1}^{n}x_j}$	Leb sur $\mathbb{R}_+^*$	$K_{LUE_{n,m}}(x,y)=\sum_{k=0}^{n-1}\psi_k^m(x)\psi_k^m(y)$, avec $\psi_k^m(x)=c'_{n,m,k}\,x^{\frac{m-n}{2}}\,e^{-\frac{m}{2}x}\,L_k^{(m-n)}(x)$; $k^{\text{ième}}$ pol. de Laguerre de param. $m-n$

3.2. Zero sets of Gaussian analytic functions.

We let $(a_n)_{n\in\mathbb{N}}$ be iid copies of $G = \frac{X+iY}{\sqrt{2}}$.

$$\forall z \in \mathbb{D}(0,1),\ f(z) := \sum_{n=0}^{\infty} a_n z^n \longleftarrow \text{Gaussian analytic function (hyperbolic GAF)}$$

Thm (Peres, Virag 2004): The point process given by the zeros of the hyperbolic GAF is a DPP on $\mathbb{D}(0,1)$, with reference measure $\mu = \frac{1}{\pi}\,\mathrm{Leb}_{|\mathbb{D}(0,1)}$

and kernel $K_H(z,w) = \dfrac{1}{(1-z\overline{w})^2}$.

→ many interesting features will be developed in Subhro Ghosh's lecture.

3.3. Many discrete examples will not been developed here.

↪ uniform spanning trees: if T is a uniform spanning tree on a connected undirected graph G, the set of edges in T form a DPP (Burton-Pemantle)

↪ non-intersecting walks: consider n symmetric r.w. on $\mathbb{Z}$, starting from $i_1 < \dots < i_n$, conditioned to return at $i_1 < \dots < i_n$ and to not intersect. Their positions at time $t/2$ form a DPP on $\mathbb{Z}$. (Karlin-McGregor)

→ <u>Poissonized Plancherel measure:</u> let $n \in \mathbb{N}^*$ be fixed.

A collection of integers $\lambda = (\lambda_1, \dots, \lambda_m)$ such that $\lambda_1 \geq \dots \geq \lambda_m$ and $\lambda_1 + \dots + \lambda_m = n$ can be represented as a Young shape λ (and encodes an irreducible representation of $\mathfrak{S}_n$).

$\dim \lambda$ is the dimension of this rep^n and also the nb of standard fillings of λ.

As $\sum_\lambda \dim(\lambda)^2 = n!$, $\quad \mathbb{P}_n(\lambda) := \dfrac{(\dim \lambda)^2}{n!}$ (Plancherel measure)

defines a probability measure on the partitions of n, $\mathrm{Par}(n)$.

Let $\theta > 0$, we define a new proba. measure $\mathbb{P}^\theta$ on $\bigcup_{n \in \mathbb{N}} \mathrm{Par}(n)$,

$$\mathbb{P}^\theta(\lambda) = e^{-\theta} \frac{\theta^n}{n!} \mathbb{P}_n(\lambda) \quad \text{(Poissonized Plancherel measure)}$$

If we denote by $p_j = \lambda_j - j$ and $q_j = \lambda'_j - j$ (with λ' the conjugate shape),

then $\left\{ p_1 + \frac{1}{2}, \dots, p_{d(\lambda)} + \frac{1}{2}, -q_1 - \frac{1}{2}, \dots, -q_{d(\lambda')} - \frac{1}{2} \right\}$

is a DPP with an explicit kernel (discrete Bessel kernel).

⇝ other similar examples (Schur measures) developed by Borodin and Okounkov.

4. DPP associated to projection kernels

4. 1. Integral operator point of view

We start with an integral operator $K : \mathbb{L}^2(\mu) \longrightarrow \mathbb{L}^2(\mu)$ that we assume to be non-negative and locally trace-class.

$\rightarrow$ integral operator : $Kf\,(x) := \int_{\mathbb{X}} K(x,y)\, f(y)\, \mu(dy)$

such that $\forall f \in \mathbb{L}^2(\mu)$, $Kf \in \mathbb{L}^2(\mu)$ (ok if K is Hilbert-Schmidt : $\int |K(x,y)|^2\, \mu(dx)\, \mu(dy) < \infty$)

$\rightarrow$ trace-class : $Tr(K) = \sum_{p \in \mathbb{N}} \langle K f_p, f_p \rangle$, (f_p) Hilbert-basis.

If K is continuous, $Tr(K) = \int K(x,x)\, \mu(dx) < \infty$

$\rightarrow$ assumed in the sequel

Locally trace-class, $Tr(K_{1_B}) < \infty$, $\forall B$ compact.

Thm : Hermitian locally trace-class integral operator K on $\mathbb{L}^2(\mu)$ defines a DPP iff $0 \leqslant K \leqslant 1$ (ie all eigenvalues are in $[0,1]$)

(Macchi, Soshnikov)

In terms of the Mercer's representation, $K(x,y) = \sum_{p=1}^{\infty} \underset{\in [0,1]}{\lambda_p}\, \varphi_p(x)\, \overline{\varphi_p(y)}$

In many interesting cases, the operator is a projection, that is $K^2 = K$

ie $\displaystyle\int K(x,y)\,K(y,z)\,\mu(dy) = K(x,z) \quad , \forall (x,z) \in X$.

We then have easy formulas for moments of linear statistics:

if we denote by f the operator by multiplication $g \mapsto fg \in \mathbb{L}^2(\mu)$,

we then have:

$$\mathbb{E}\Big(\sum_{X\in\gamma} f(X)\Big) = \mathrm{Tr}(fK) \quad , \quad \mathbb{E}\big(\mathrm{Card}(\gamma\cap B)\big) = \mathrm{Tr}(K\,1_B).$$

$$\mathrm{Var}\Big(\sum_{X\in\gamma} f(X)\Big) = \mathrm{Tr}(f^2K) - \mathrm{Tr}(fKfK)$$

In the case of an Hermitian projection kernel:

$$\mathrm{Var}\Big(\sum_{X\in\gamma} f(X)\Big) = \frac{1}{2}\iint |f(x)-f(y)|^2\,|K(x,y)|^2\,\mu(dx)\,\mu(dy)$$

$\hookrightarrow$ variance may crucially depend on the regularity of f.

4.2. Projection onto a finite-dimensional subspace.

Let V be a subset of dimension n of $\mathbb{L}^2(\mu)$, $f_1,\dots,f_n$ a basis of V

and $g_1,\dots,g_n \in \mathbb{L}^2(\mu)$ such that $\langle f_i, g_j\rangle = \int f_i\,\overline{g_j}\,d\mu = \delta_{ij}$

(biorthogonal family)

We define $K(x,y) := \sum_{p=1}^{m} f_p(x)\,\overline{g_p(y)}$

then the associated operator is a projection operator such that:

- $\mathrm{Im}(K) = V$
- $\mathrm{Ker}(K)^\perp = \mathrm{Span}(g_1, \dots, g_m)$

and the associated DPP has exactly m points.

Rq: an important particular case is when $f_1, \dots, f_m$ is a o.m.b of V and $(g_1, \dots, g_m) = (f_1, \dots, f_m)$.

Examples:

- Gin_m : $V = \mathbb{C}^{m-1}(X)$, $\mu = \frac{1}{\pi} e^{-|z|^2} dz$, $f_p(z) = \frac{z^p}{\sqrt{p!}}$
- GUE_m : $V = \mathbb{R}^{m-1}(X)$, $\mu = e^{-x^2} dx$, $f_p(x) = h_p(x)$

Hermite polynomials.

$\rightarrow$ idem CUE_m, LUE_m etc.

Rq: all these examples are orthogonal polynomial ensembles (OPE).

5. Asymptotic behavior of the DPPs : global regime

Empirical measure of the particles : $\hat{\mu}_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}$

$\forall f$ measurable, $\int f \, d\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} f(x_i)$ random probability measure

Goal : find a probability measure μ_∞ such that,

almost surely, $\forall f$ bounded continuous, $\int f \, d\hat{\mu}_n \xrightarrow[n\to\infty]{} \int f \, d\mu_\infty$

5.1. An easy example : CUE_n

Prop : a.s, for any $f : \mathbb{S}^1 \to \mathbb{C}$ continuous,

$$\int f \, d\hat{\mu}_{CUE_n} \xrightarrow[n\to\infty]{} \int_{\mathbb{S}^1} f \, d\nu = \int_0^{2\pi} f(e^{i\theta}) \frac{d\theta}{2\pi} .$$

Key points of the proof : $K_{CUE_n}(z,w) = \sum_{k=0}^{n-1} z^k \bar{w}^k \quad (z = e^{i\theta}, w = e^{i\eta})$

- moments : $\forall m \in \mathbb{Z}$, $\mathbb{E}\left(\int z^m \hat{\mu}_n(dz) \right) = \delta_{0m}$

$$\left.\begin{aligned} &\mathbb{E}\left(\int z^m \hat{\mu}_n(dz) \right) = \delta_{0m} \\ &\mathrm{Var}\left(\int z^m \hat{\mu}_n(dz) \right) = \frac{1}{n^2} \min(|m|, n) \end{aligned}\right\} \Rightarrow \text{a.s. } \forall m \in \mathbb{Z}, \int z^m \hat{\mu}_n(dz) \to \int z^m \nu(dz)$$

\+ conclude by approximating any continuous functions by polynomials.

5.2. A general result for OPE and consequences for random matrix ensembles

Let $(P_k^n)_{0\leq k\leq n-1}$ be a sequence of polynomials on $\mathbb{R}$ such that, $\forall k,\ell\in\mathbb{N}$

P_k^n is of degree k and $\int P_k^n P_\ell^n \, d\mu_n = \delta_{k\ell}$.

We know that there exits a three-term relation :

$$x\, P_k^n(x) = a_k^n\, P_{k+1}^n(x) + b_k^n\, P_k^n(x) + a_{k-1}^n\, P_{k-1}^n(x)$$

We let $\displaystyle K_n(x,y) = \sum_{k=0}^{n-1} P_k^n(x)\,\overline{P_k^n(y)}$.

<u>Thm</u> (Hardy, 18) : Assume that $a_k^n \xrightarrow[n\to\infty]{} a(s)$ and $b_k^n \xrightarrow[n\to\infty]{} b(s)$ as $\frac{k}{n} \xrightarrow[n\to\infty]{} s$, where $a,b : (0,1)\to\mathbb{R}$ are two continuous functions such that $\forall p,q\in\mathbb{N}$, $a^p b^q$ is integrable on $(0,1)$.

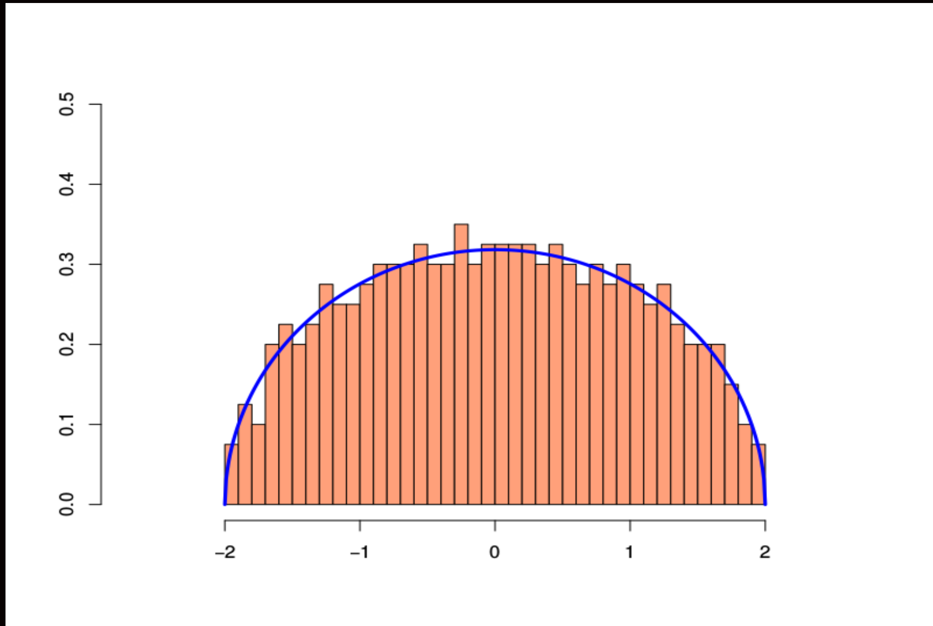
Then, $\forall p\in\mathbb{N}$, $\displaystyle \mathbb{E}\left(\int x^p \,\hat{\mu}_n(dx)\right) \xrightarrow[n\to\infty]{} \int x^p \,\mu_{a,b}(dx)$,

where $\mu_{a,b}$ is the law of $a(U)\,\xi + b(U)$, with U and ξ indep, U uniform on $(0,1)$ and the distribution of ξ is the arcsine law on $[-1,1]$. Moreover, if $\displaystyle\sum_{n\in\mathbb{N}^*}\left(\frac{1}{n}\, a_n^n\right)^2 < \infty$, the convergence holds a.s.

Ex : for GUE_n, $a_k^n = \sqrt{\frac{k}{n}}$ and $b_k^n = 0$ so that $a(s) = \sqrt{s}$ and $b(s) = 0$

in this case, $\mu_{a,b}$ is the semi-circular distribution (Wigner law)

with density $\frac{1}{2\pi} \sqrt{4-x^2}\, \mathbb{1}_{[-2,2]}(x)$



5.3. CLT for linear statistics.

General machinery :

Thm (Soshnikov) : If f is bounded measurable such that

- $\mathrm{Var}\left(n \int f \, d\hat{\mu}_n \right) = \mathrm{Var}\left(\sum_{i=1}^n f(X_i) \right) \xrightarrow[n \to \infty]{} \infty$
- $\exists \delta > 0$ s.t. $\mathbb{E}\left(\sum_{i=1}^n |f(X_i)| \right) = O\left(\mathrm{Var}\left(\sum_{i=1}^n f(X_i) \right) \right)$

then
$$\frac{\dfrac{n\left(\int f\, d\hat{\mu}_n - \mathbb{E}\left(\int f\, d\hat{\mu}_n\right)\right)}{\sqrt{\mathrm{Var}\left(n\int f\, d\hat{\mu}_n\right)}}}{} \xrightarrow[n\to\infty]{} \mathcal{N}(0,1)$$

Claim: the behavior of the variance is completely different for smooth f and for indicator functions.

ex: CUE_n Take $\eta > 0$ and $f = \mathbb{1}_{\underbrace{[1, e^{i\eta}]}_{D}}$ (on $\mathbb{S}^1$)

Then $\mathbb{E}\left(n\int f\, d\hat{\mu}_n\right) = \mathbb{E}\left(\mathrm{Card}(\gamma\cap D)\right) = \frac{n\eta}{2\pi}$

$$\mathrm{Var}\left(\mathrm{Card}(\gamma\cap D)\right) = \frac{1}{\pi^2}\sum_{j=1}^{n-1}\frac{1}{j} - \frac{1}{2\pi^2}\sum_{j=1}^{n-1}\frac{|1-e^{-ij\eta}|^2}{j^2} \sim \frac{1}{\pi^2}\log n$$

In Diaconis-Evans, they show that

if f is $H^{1/2}$, that is $\sum_{j\in\mathbb{N}} j\,|\hat{f}_j|^2 < \infty$,

then $\mathrm{Var}\left(\sum_{i=1}^{n} f(x_i)\right) \xrightarrow[n\to\infty]{} \sum_{j\in\mathbb{N}} j\,|\hat{f}_j|^2$

Soshnikov's result does not apply, but CTL still holds.

6. Asymptotic behavior of the DPPs: local behavior

6.1. The sine process.

We consider the DPP on $\mathbb{R}$, with reference measure $\mu = \text{Leb}$.

and kernel
$$K_{\sin}(x,y) = \begin{cases} \dfrac{\sin(\pi(x-y))}{\pi(x-y)} & \text{if } x \neq y \\ 1 & \text{if } x = y. \end{cases}$$

If $\mathcal{F}$ is the Fourier transform on $L^2(\mu)$, one can check that the corresponding operator can be written:

$$K_{\sin} = \mathcal{F}^{-1}\left(\mathbb{1}_{[-\frac{1}{2}, \frac{1}{2}]} \cdot \mathcal{F} \right)$$

From there, it is easy to check that it is a projection operator (onto the subspace of functions f such that $\hat{f}$ is supported on $[-\frac{1}{2}, \frac{1}{2}]$)

$\hookrightarrow$ it plays a very important role in universality results.

Rq: projection kernel with an infinite number of points.

6.2. The example of CUE_n again

Reminder: $\gamma = \{\theta_1, \dots, \theta_n\}$ with kernel $K_n(\theta,\eta) = \sum_{k=0}^{n-1} e^{ik\theta} e^{-ik\eta}$

$$K_n(\theta,\eta) = e^{i\frac{n-1}{2}(\theta-\eta)} \frac{\sin\left(\frac{n}{2}(\theta-\eta)\right)}{\sin\left(\frac{1}{2}(\theta-\eta)\right)} \quad \text{and reference measure} \quad \frac{1}{2\pi}\,\mathrm{Leb}_{|[-\pi,\pi]}$$

We slightly modify the kernel $K'_n(\theta,\eta) = \dfrac{\sin\left(\frac{n}{2}(\theta-\eta)\right)}{\sin\left(\frac{1}{2}(\theta-\eta)\right)}$ defines the same DPP

If we now consider $\frac{n}{2\pi}\gamma = \left\{\frac{n}{2\pi}\theta_1, \dots, \frac{n}{2\pi}\theta_n\right\}$, it is a DPP with kernel

$$\widetilde{K}_n(\theta,\eta) = \frac{1}{n} K'_n\left(\frac{2\pi\theta}{n}, \frac{2\pi\eta}{n}\right) = \frac{\sin(\pi(\theta-\eta))}{n\sin\left(\frac{\pi}{n}(\theta-\eta)\right)} \xrightarrow[n\to\infty]{} K_{\sin}(\theta,\eta)$$

and reference measure $\mathrm{Leb}_{|[-n,n]}$

Rq: all the random matrix ensembles presented above, properly rescaled, converge to the sine process.

6.3. A general result for OPE.

Thm (Lubinsky, 09) Let μ be a finite measure on $[-1,1]$ with density w and $(P_k)_{k\in\mathbb{N}}$ the corresponding orthonormal polynomials with respect to μ.

We denote by γ_n the leading coefficient of P_n : μ is supposed to be regular, that is $\lim_{n\to\infty} \gamma_n^{1/n} = 2$. (true e.g. if $a_n \to \frac{1}{2}$ and $b_n \to 0$)

Let $K_n(x,y) := \sum_{k=0}^{n-1} p_k(x)\, p_k(y) \left(= a_n \frac{P_n(x) P_{n-1}(y) - P_{n-1}(x) P_n(y)}{x-y} \right)$

and $\widetilde{K}_n(x,y) = \sqrt{w(x)}\sqrt{w(y)}\, K_n(x,y)$.

Let $J \subset (-1,1)$ be a compact such that w is positive and continuous on J.

Then, uniformly on $x \in J$, for a, b in a compact set, we have

$$\lim_{n\to\infty} \frac{\widetilde{K}_n\left(x + \frac{a}{\widetilde{K}_n(x,x)},\ x + \frac{b}{\widetilde{K}_n(x,x)}\right)}{\widetilde{K}_n(x,x)} = \frac{\sin(\pi(a-b))}{\pi(a-b)}.$$

Rq : this involves the convergence of all correlation functions for the DPP $\widetilde{K}_n(x,x)(\gamma - x)$

7. Variants of DPPs give point processes of interest

Ginibre $\subset$ Coulomb gases

$\beta = 2 \quad V = |\cdot|^2$

$$\frac{1}{Z_n^\beta} \prod_{i<j} |z_i - z_j|^\beta e^{-n \sum_{i=1}^{n} V(z_i)}$$

(idem CBE_n, GBE_n, $Schur_\beta$ ...)

Hyperbolic GAF $\subset$ GAFs

$\gamma_n = 1$

a_n iid $\mathcal{N}(0,1)$,

zero set of

$$f(z) = \sum_{n=0}^{\infty} \gamma_n a_n z^n$$

k-DPPs.

def : Let $\mathbb{X}$ be a discrete set.

A k-DPP is a point process obtained by conditioning a DPP to have exactly k points a.s.

$\hookrightarrow$ interesting for applications

DPP $\subset$ α-determinantal processes.

$\alpha = -1$

A point process is α-determinantal iff its corr. functions can be written :

$$\rho_k(x_1, \dots, x_k) = \det_\alpha(K(x_i, x_j))$$

$$= \sum_{\pi \in \mathfrak{S}_k} \alpha^{k - c(\pi)} \prod_{i=1}^{k} K(x_i, x_{\pi(i)})$$

$\hookrightarrow$ permanental (bosons), iid copies of DPP's ...

8. Conclusion

→ DPP have a rigid structure that allows many "explicit" computations

→ important examples of point processes from different frameworks are DPP

→ useful for modelling in physics, signal processing etc.

→ but this structure is very sensitive to change of parameters: there are mathematical gems, or benchmarks ...

Some references:

- Hough, Krishnapur, Peres, Virag. Determinantal processes and independance. 2006
- Soshnikov. Determinantal Random Point Fields. 2000
- Macchi. The coincidence approach to stochastic point processes. 1975.
- Lyons. Determinantal probability measures. 2003.