Audition : Université de Paris (LPSM)

Simon Coste

2016 - 2019	thèse, supervision de Charles Bordenave & Justin Salez
2019 - 2021	postdoctorat dans l'équipe de François Baccelli (INRIA)
2021 - 2022	postdoctorat dans l'équipe de Stéphane Mallat (ÉNS Paris)

ENS | PSL

Matrices creuses ~~hermitiennes~~

Matrices creuses ~~hermitiennes~~ : motivations

Mathématiques

- Matrices creuses $\Rightarrow$ autres classes d'universalité [1] !
- Objets poissoniens.
  Liens avec la combinatoire, l'arithmétique.

Modélisation et statistiques

- Énormément de réseaux sont creux [2].
- Énormément de réseaux sont dirigés.
- Trouver des méthodes <u>vraiment</u> non-hermitiennes (ex : non-backtracking).

1. Metz, Neri, Rogers 2018, *Spectral Theory of Sparse Non-Hermitian Random Matrices* (physique)
2. Cf `StanfordLargeNetworkDatasetCollection`

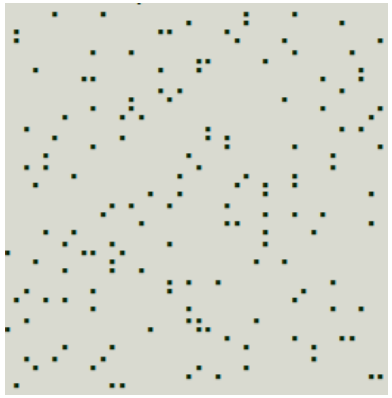
I
Modèles à entrées indépendantes et applications en statistiques

Un modèle naturel : entrées indépendantes

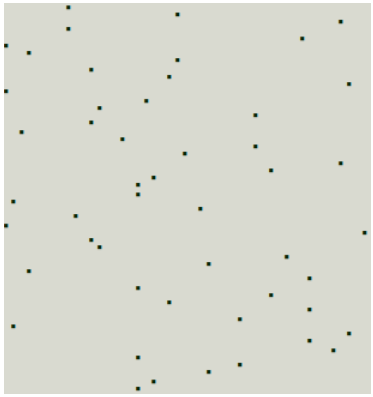
Définition ([Bordenave, C, Nadakuditi, 22, J.FoCM])

matrice de poids $n \times n$	*matrice de masque $n \times n$*	*matrice observée A*
$X_{i,j}$	$\xi_{i,j} \sim \mathrm{Bernoulli}(P_{i,j})$	$A_{i,j} = X_{i,j}\xi_{i,j}$
		(ainsi $\mathbb{E}[A] = P \odot X$*)*

modèle d-creux $\Leftrightarrow \max P_{i,j} \leqslant d/n$

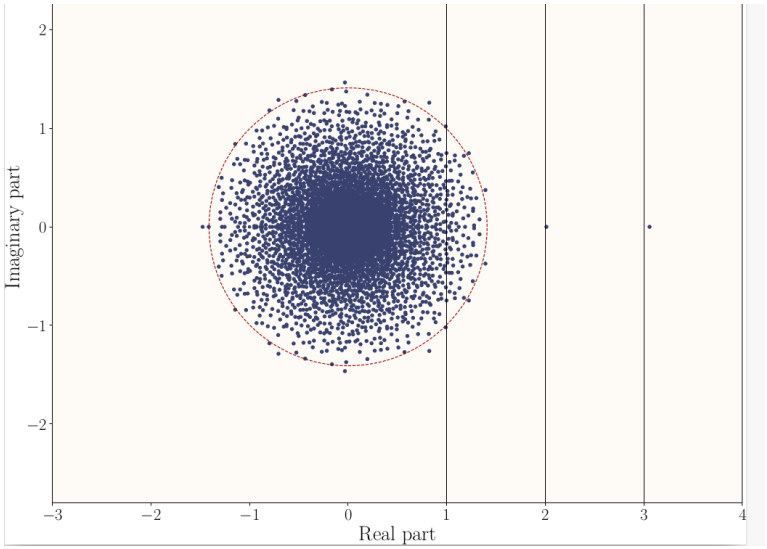


$d = 4$ entrées non nulles par ligne



$d = 1$

Théorie spectrale « extrémale »



$Av_i = \lambda_i v_i \qquad \mathbb{E}[A]u_i = \mu_i u_i$ (valeurs propres, vecteurs propres)

Théorème ([Bordenave, C, Nadakuditi, 22, J.FoCM] et [C, Stephan, 21+])

Il existe un seuil $\vartheta = \sqrt{\|P \odot X \odot X\|_{\mathrm{op}}}$ *tel que si* $|\mu_i| > \vartheta$, *alors* $\lambda_i \sim \mu_i$ *et*

$$\lim_{n\to\infty} |\langle u_i, v_i\rangle| = |\mu_i|^{-1} \times \left(|(\mu_i^2 I - PX^2)^{-1} v_i^2|_1\right)^{-1/2}$$

Toutes les autres valeurs propres de A *sont plus petites que* ϑ.

Statistique non supervisée : reconstruction de matrices

Problème statistique : on veut reconstruire X à partir de l'observation de A.

Théorème

1) Séparer aléatoirement $A = A_1 + A_2$.
2) Étudier les vecteurs propres de $A_1 A_2^*$ *plutôt que ceux de* AA^*.

[Bordenave, C, Nadakuditi, 22, J.FoCM] : identification de la limite

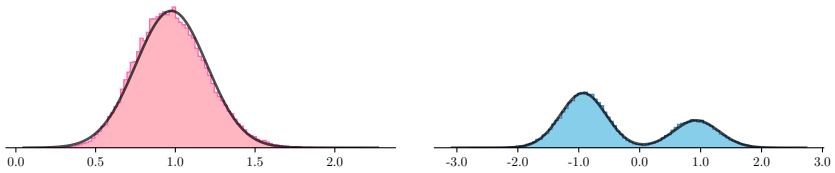
$$RMSE = (nm)^{-0.5} \|X - \hat{X}\|_F$$

Détection des communautés dans un SBM

$$X = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \qquad P = \frac{1}{n} \begin{pmatrix} a & \dots & a & b & \dots & b \\ \vdots & & \vdots & \vdots & & \vdots \\ a & \dots & a & b & \dots & b \\ c & \dots & c & d & \dots & d \\ \vdots & & \vdots & \vdots & & \vdots \\ c & \dots & c & d & \dots & d \end{pmatrix}$$

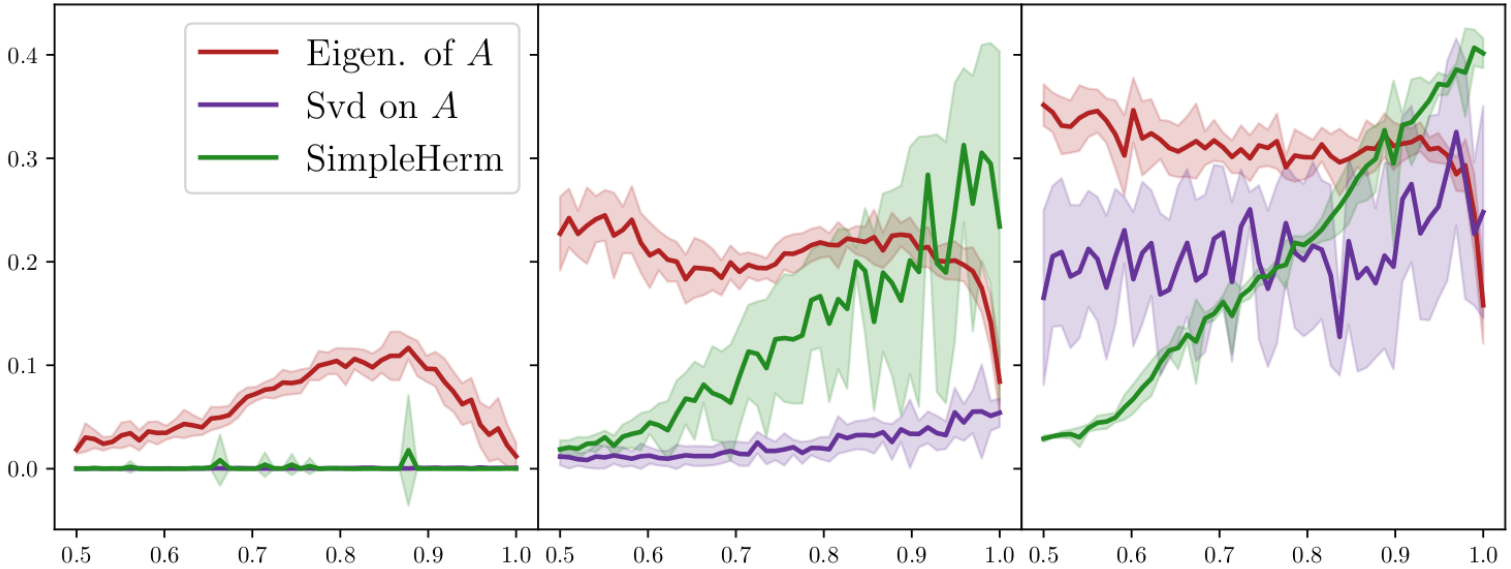
Theorem ([C, Stephan, 21+])

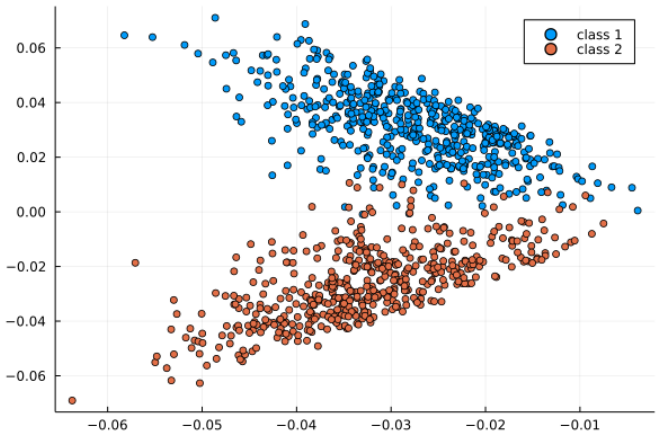
1) Nous avons identifié tous les seuils de Kesten-Stigum dans ce type de modèle.
2) Nous avons « identifié » la limite de la loi empirique des entrées des vecteurs propres associés aux outliers



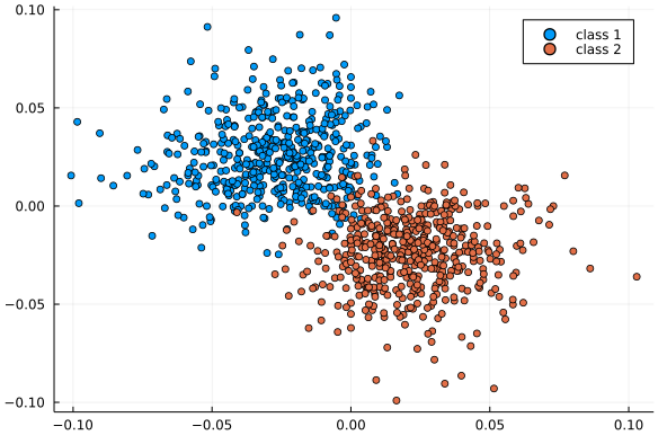
Résultats expérimentaux de [C, Stephan, 21+] : détection de communautés

Performance du clustering à partir de vecteurs propres de matrices non-normales





plongement spectral classique



plongement « non-symétrique »

À faire avec la méthode des traces

Projets

- Markovian Block-Models
  Clustering in block markov chains (Proutière, Sanders, Yun, *Annals of Stat.* 20).
- Alignement de bases de données creuses
  Travaux de Ganassali, Lelarge, Massoulié, Jiaming Xu
- Graphes dynamiques : méthodes tensorielles

II
Modèles de matrices aléatoires avec des dépendances

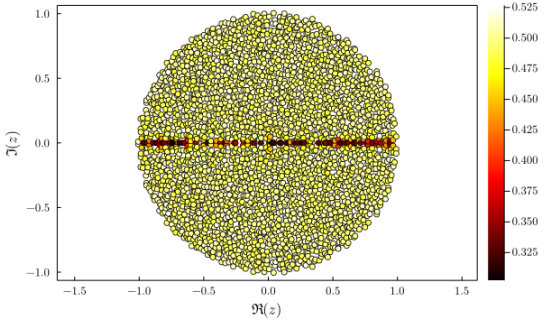
Matrices avec des entrées **dépendantes**

Modèle 1. graphes de configuration [C, 20, Ann. IHP].

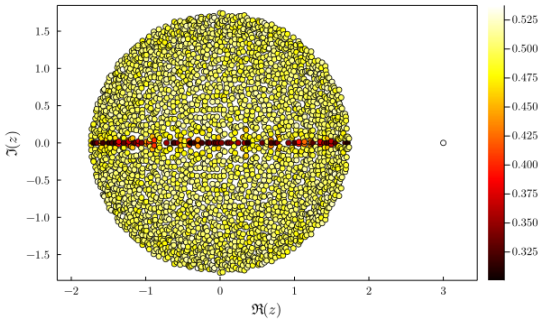
Modèle 2. Soient $\pi_1, \dots, \pi_d$ des permutations de $\{1, \dots, n\}$ aléatoires.

$$A = \pi_1 + \cdots + \pi_d \quad \in \mathscr{M}_{n,n}(\{0,1\})$$

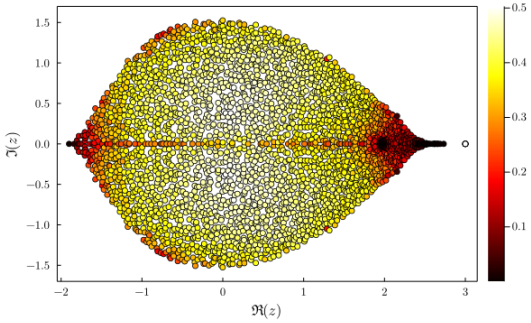
Exemples de valeurs propres de A



$d = 50$



uniforme, $d = 3$

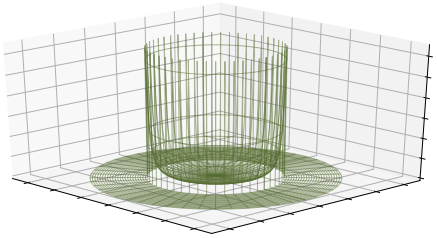


Ewens$(500), d = 3$

Conjecture « Oriented Kesten-McKay » (difficile)

$$\mathrm{ESD} = \mu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i}$$

<u>Conjecture</u> : μ_n converge vers $\frac{d^2(d-1)}{\pi}\frac{\mathbf{1}_{|z|<\sqrt{d}}}{(d^2-|z|^2)^2}dz$.



Quelles sont les autres lois limites des matrices creuses ?

1. Loi limite pour Erdos-Rényi dirigé ?
   Mesure de Brown, [Haagerup et Larsen 2000]
2. Équivalent unitaire (libre) de la mesure de Ewens ?

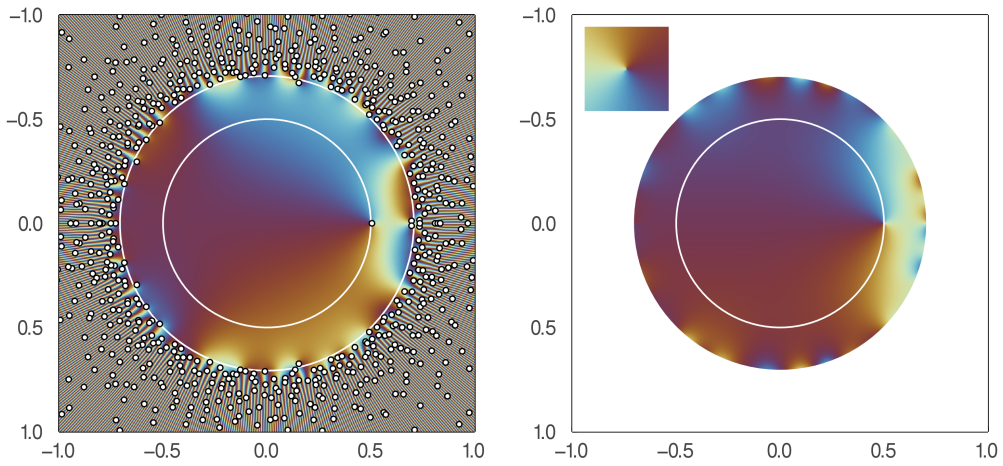
Convergence du polynôme caractéristique

Théorème ([C, 21+] et [C, Lambert, Zhu, 22+])

A_n = *au choix,*

- *la matrice d'adjacence d'un graphe d'Erdős-Rényi dirigé* $G(n, d/n)$
- *la somme de* d *matrices de permutations uniformes (si* d *est entier),*

Alors $det(I_n - zA_n)$ *converge vers* $e^{-f(z)}$ *lorsque* $|z| < \sqrt{d}$.



Extensions vraiment faciles : ajout de poids sur les arêtes
Moins évident : modèles d'attachement préférentiel

Un champs poissonnien log-corrélé

Soit $Y_\ell \sim \text{Poisson}(d^\ell/\ell)$, indépendantes.

$$f(z) = \sum_{k=0}^{\infty} z^k \left(\sum_{\ell|k} \ell Y_\ell \right)$$

- fonction méromorphe sur $D(0, 1/\sqrt{d})$ avec un unique pôle en $1/d$.
- $\text{cov}(f(z), f(w)) = -\log(1 - z\bar{w}) + O_d(1)$ (« log-corrélé »)
- $\lim_{d\to\infty} f_d(z/\sqrt{d}) =$ processus gaussien log-corrélé.
- Expression remarquable :

$$e^{-f(z)} = \prod_{k=0}^{\infty} (1 - z^\ell)^{Y_\ell}$$

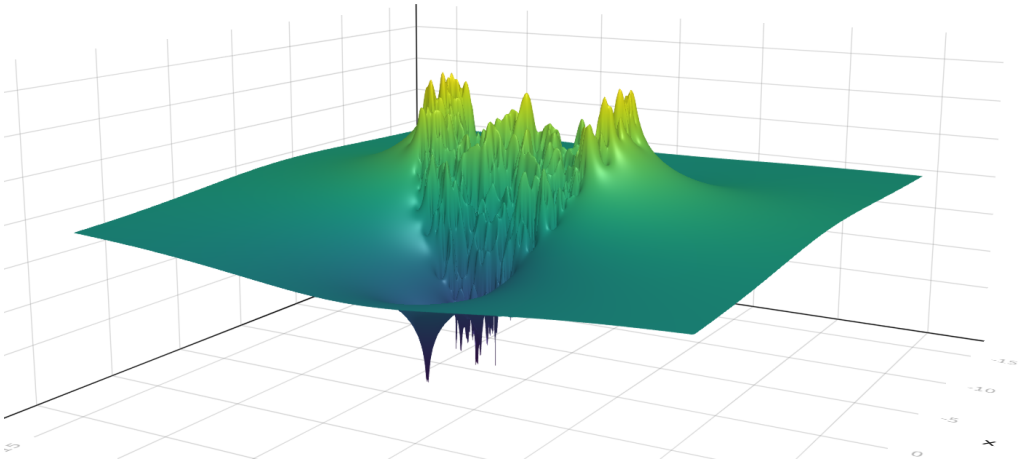
- Fonction génératrice des moments de $F = e^{-f}$ connue [3]

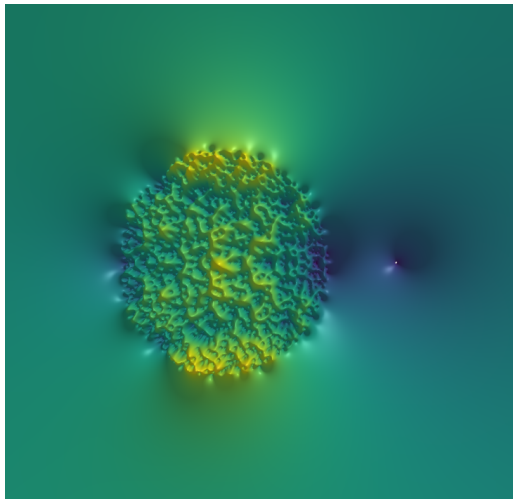
$$\mathbb{E}[F(z_1)\ldots F(z_m)] = \frac{\prod_{S\subset[m] \text{ odd}} (1 - d\prod_{s\in S} z_s)}{\prod_{S\subset[m] \text{ even}} (1 - d\prod_{s\in S} z_s)}$$

...mais pas d'interprétation combinatoire.

3. [C, 21+]section 2.3

Fluctuations et champs poissonniens log-corrélés





Chaos holomorphe Poissonien défini comme une distribution de Schwartz sur $\mathbb{T}$:

$$(\mathrm{PHC}_d, \varphi) = \lim_{r \to d^{-1/2}} \int_{\mathbb{T}} e^{-f(re^{i\theta})} \varphi(e^{i\theta}) d\theta$$

Régularité Sobolev [C, 21+] : $\mathrm{PHC}_d \in H^{\varepsilon}(\mathbb{T})$ pour tout $\varepsilon < -1/2$.

Question dans l'esprit de [Rider et Virag 2006]

$$\max \log |\det(I - zA)| \sim \frac{\text{constante}}{\log n} ?$$

Partie non-harmonique des fluctuations (polynômes en z et $\bar{z}$) ?

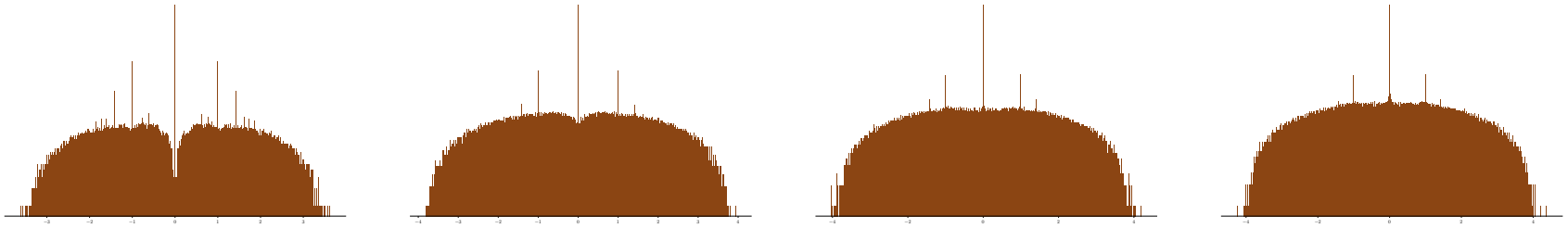
III

Matrices creuses **Hermitiennes**

Matrices creuses **Hermitiennes**

Quid des matrices d'adjacence des graphes d'Erdős-Rényi $G(n, d/n)$?

Théorème : ESD_n converge vers une mesure borélienne μ_d.



$c = 2$ $c = 2,6$ $c = 2,8$ $c = 3$

Theorem ([C, Salez, 21, Ann. Proba.])

μ_d possède des états étendus en zéro si et seulement si $d > e$.

[C., PhD thesis] : la mesure spectrale limite des arbres aléatoires uniformes n'a pas d'états étendus en zéro.

IV
Modèles probabilistes génératifs et deep learning

Étant donnés des échantillons (x_i) provenant d'une loi inconnue p,

1. quel est le meilleur estimateur $\hat{p}$ de p ?
2. comment générer de nouveaux échantillons de $\hat{p}$?

Théories classiques	Théories « modernes »
estimation non-paramétrique	Energy-based models
MCMC, Langevin	Normalizing Flows, Diffusion models

Idées probabilistes en deep learning

Méthode (Idée des EBM)

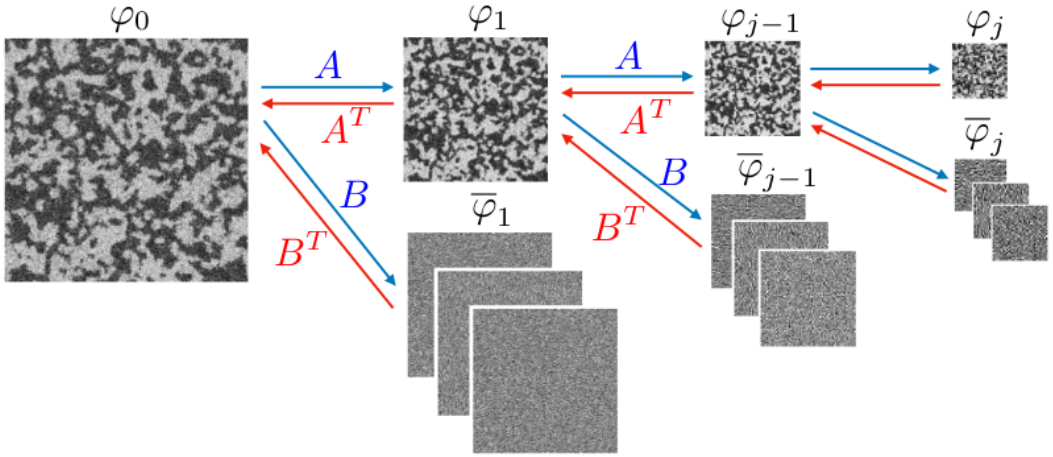
1) On paramétrise une famille d'énergies E_θ *(exemple : un réseau de neurones)*
2) On cherche à trouver θ *qui minimise* $KL(p|p_\theta)$ *où* $p_\theta = e^{-E_\theta}/Z_\theta$.

Problème

L'algorithme de montée de gradient s'écrit

$$\theta_{n+1} - \theta_n = \mathbb{E}_{p_{\theta_n}}[\nabla E_{\theta_n}] - \mathbb{E}_p[\nabla E_{\theta_n}]$$

MCMC à chaque étape, mauvais conditionnement, choix d'une paramétrisation...



$$X_0 \in \mathbb{R}^{2^d} \to X_1 \in \mathbb{R}^{2^{d-1}} \to \cdots \to X_J \in \mathbb{R}^1$$

Idée du Groupe de Renormalisation : paramétriser les lois conditionnelles $p(X_j|X_{j+1})$.

$$p(x_j|x_{j+1}) = \frac{e^{-E_{\theta_j}(x_j,x_{j+1})}}{Z_{\theta_j}(x_{j+1})}$$

Insertion au LPSM - P7

LPSM
Laboratoire de probabilités
statistique & modélisation

Sandrine Péché
Fanny Augeri

Annexe. Challenge data

Challenge Data

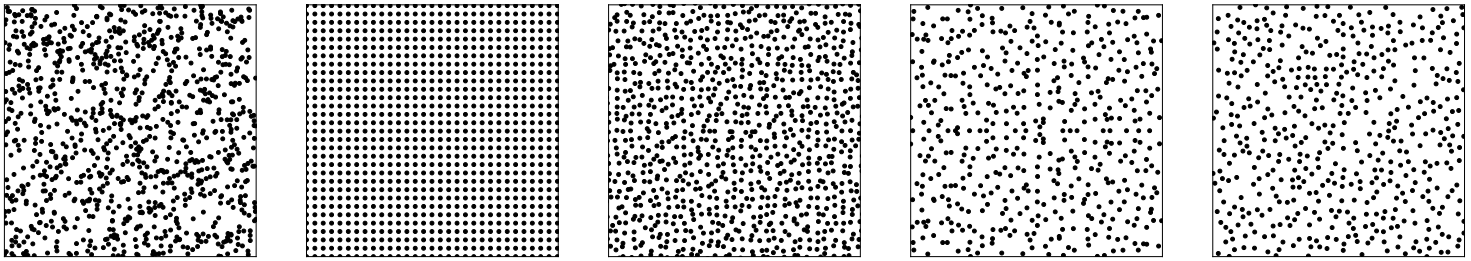
About us

Challenge Data is managed by **the Data team (ENS Paris)**, in partnership with the Collège de France, and the **DataLab** at Institut Louis Bachelier. It is supported by the CFM chair and the PRAIRIE Institute.

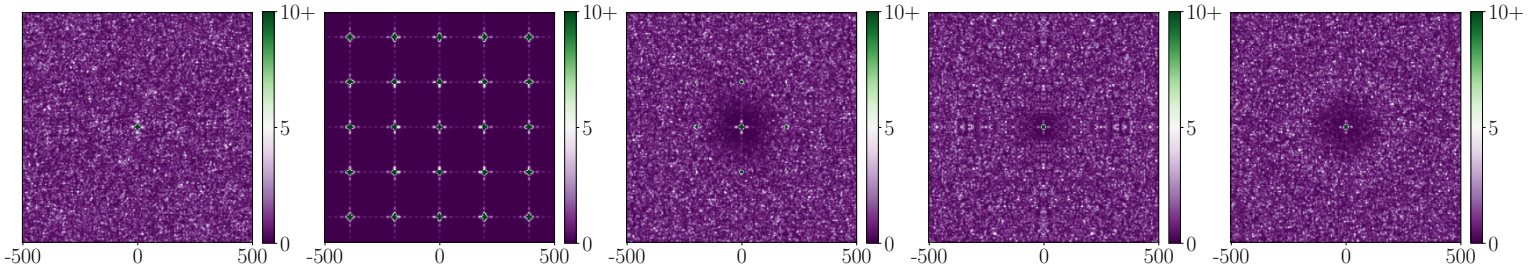
32789	56	9994
Submissions	Challenges currently running	Participants

Annexe. Rigidité des processus ponctuels

Quelques processus avec des rigidités différentes



Poisson | Grid | Perturbed grid | Ginibre | Matern-II



La transformée de Fourier de la fonction de corrélation

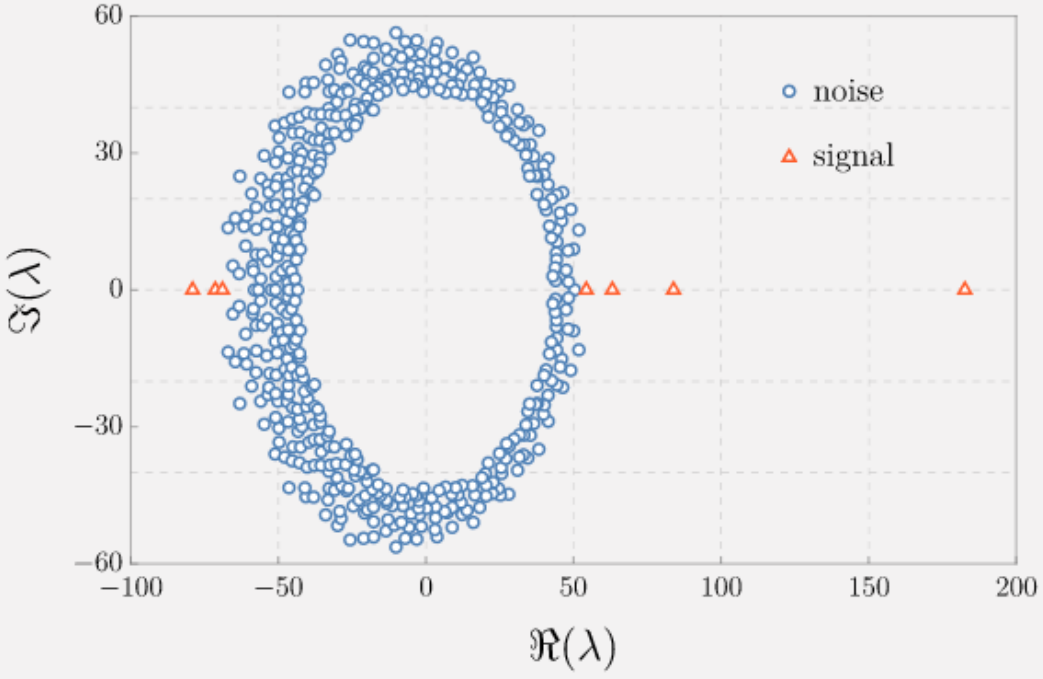
Un survey en cours d'écriture sur les notions de « rigidités » des processus ponctuels : hyperuniformité, number-rigidity, stealthyness...

<u>Question</u> : le gas de Coulomb en dimension $d \geqslant 2$ est-il hyperuniforme ?

Annexe. Utilisation de [Bordenave, C, Nadakuditi, 22, J.FoCM] dans l'industrie bioinformatique

https://fasma.io

fasma



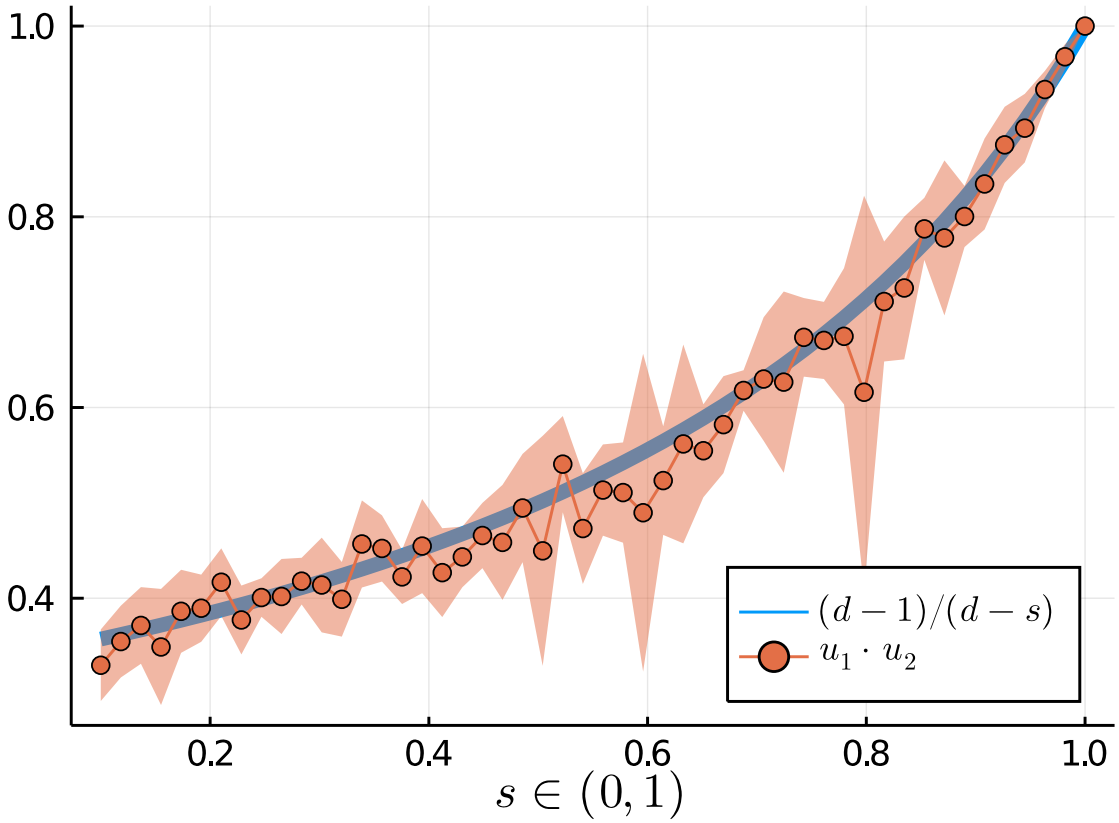
Pow

Becaus
biologic
extreme

We leve
random
unsupe

We war
perform
perform

Annexe. Alignement de base de données / désanonymisation

M_1, M_2 : matrices d'Erdős-Rényi dirigés, avec corrélation des entrées s.
u_1, u_2 : vecteurs propres dominants.

$$\langle u_1, u_2 \rangle \to \frac{1 - 1/d}{1 - s/d}$$