RMT - DM2

Hugo Simon, Eloi Tanguy

March 14, 2022

Observations préliminaires

1. On se place conditionnellement à q.
On a que $A = \mathbb{E}[A] + (A - \mathbb{E}[A]) =: \mathbb{E}[A] + X$ où X est par construction centrée.
Or $\forall i, j \in [\![1, n]\!]$,

$$\mathbb{E}\left[A_{ij}\right] = \mathbb{E}\left[\mathbb{E}\left[A_{ij}|C_{ab}\right]\right] = \mathbb{E}\left[q_i q_j C_{ab}\right] = q_i q_j (1 + \frac{\mathbb{E}[M_{ab}]}{\sqrt{n}})$$

En notant $J = [j_1, \dots, j_K] \in \mathcal{M}_{n,k}(\mathbb{R})$ où $j_i \in \mathbb{R}^n$ est le vecteur canonique de la classe $\mathcal{C}_i$,
$\bar{M} = \mathbb{E}\left[M\right] = \left(\mathbb{E}\left[M_{ab}\right]\right)_{1 \leq a,b \leq K}$, $Q = \mathrm{diag}(q)$ et $\mathbb{1} \in \mathcal{M}_K(\mathbb{R})$ la matrice pleine de 1, on peut réécrire

$$\boxed{\mathbb{E}[A] = QJ(\mathbb{1} + \frac{1}{\sqrt{n}}\bar{M})J^\top Q}$$

Comme $\mathbb{1} + \frac{1}{\sqrt{n}}\bar{M}$ est carrée de taille K, $\mathbb{E}[A]$ est nécessairement de rang au plus K.

2. De même $\forall i, j \in [\![1, n]\!]$,

$$\mathbb{E}\left[B_{ij}\right] = q_i q_j \frac{\mathbb{E}[M_{ab}]}{\sqrt{n}}$$

i.e. $\mathbb{E}\left[B\right] = \frac{1}{\sqrt{n}} QJ\bar{M}J^\top Q$ de rang au plus K.
D'où

$$\boxed{\frac{B}{\sqrt{n}} = \frac{1}{n} QJ\bar{M}J^\top Q + \frac{X}{\sqrt{n}}}$$

3. On prend $n = 2000$, $K = 3$, et M égale à $\bar{M}$ perturbée d'une variable normale symétrique où $\bar{M} = \mathrm{diag}(80, 112, 160)$. Les 3 cas considérés sont :

(a) $q_i = q_0 = 0.9$ (b) $q_i \sim \mathcal{U}([0.3, 0.9])$ (c) $q_i \sim \mathcal{U}(\{0.1, 0.9\})$

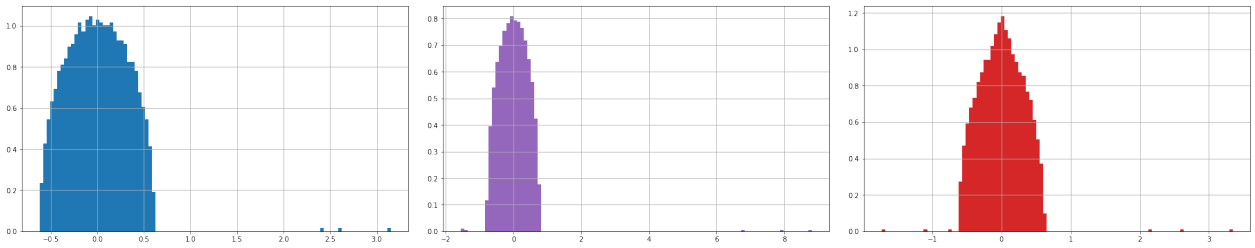


Figure 1: Mesure spectrale de $\frac{B}{\sqrt{n}}$ dans les 3 cas considérés.

De là, on observe que

(a) correspond à un cas matrice de Wigner + spikes. C'est à dire que la mesure spectrale suit une densité du demi-cercle à laquelle s'ajoute K valeurs propres isolées correspondant aux valeurs propres de $\frac{\mathbb{E}[B]}{\sqrt{n}}$.

(b) correspond à un cas homogène bruité. La mesure spectrale ressemble à un cas $q_0 = \mathbb{E}\left[q_i\right] = 0.6$ à laquelle s'ajoute de faibles spikes parasites.

(c) s'éloigne du modèle Wigner + spikes. Le bulk de la mesure spectrale ressemble peu à la densité du demi-cercle et les valeurs propres isolées se dédoublent du fait des 2 régimes de connectivité $q^{(1)} = 0.1$ et $q^{(2)} = 0.9$.

4.

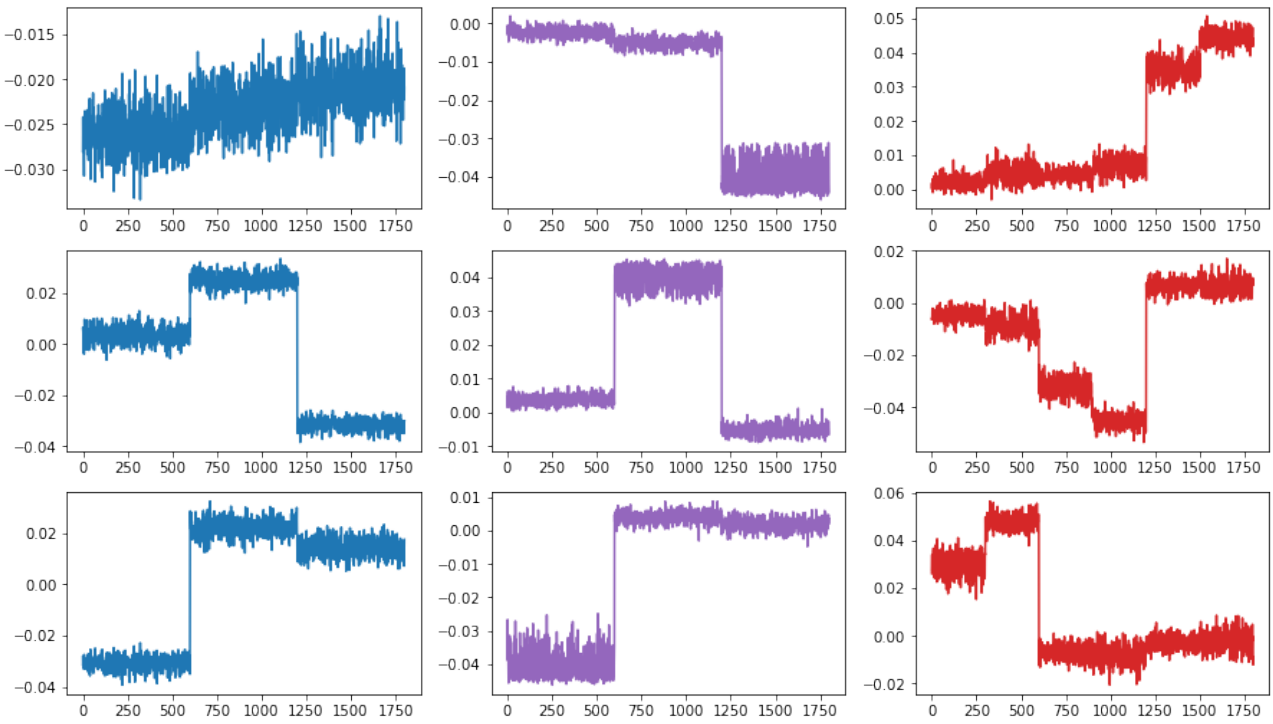


Figure 2: K vecteurs propres extrémaux de $\frac{B}{\sqrt{n}}$ dans les 3 cas considérés.

On observe que les vecteurs propres, sans pour autant converger vers des indicatrices des classes, permettent dans tous les cas de distinguer les différentes classes. On note que le cas (b) est davantage bruité, et on observe également dans le cas (c) les 2 régimes de connectivité $q^{(1)}$ et $q^{(2)}$.

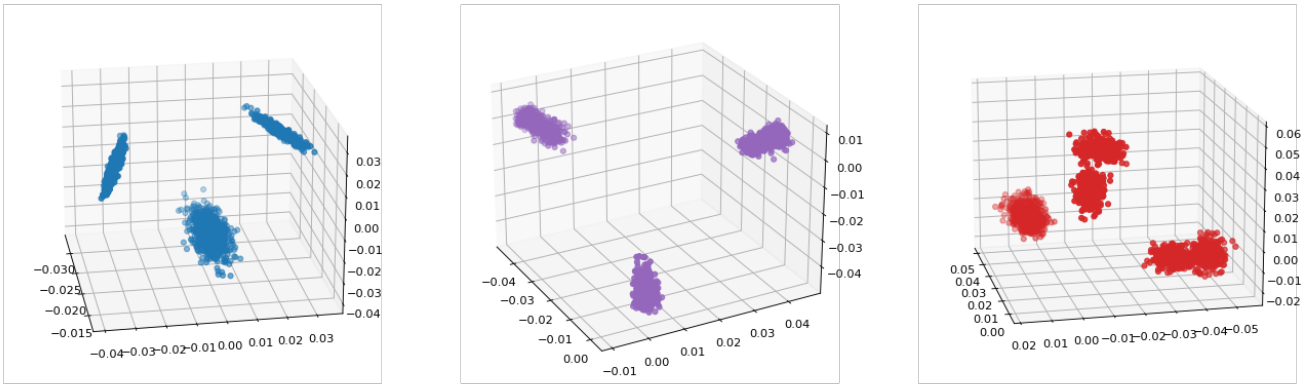


Figure 3: Plongements de dimension K issus des vecteurs propres extrémaux dans les 3 cas considérés.

Observons ces K vecteurs propres comme des plongements de dimension K des n noeuds. On constate encore une fois une nette distinction entre les classes qui peut nous permettre ici d'effectuer un partitionnement. On propose donc l'algorithme de partitionnement spectral suivant :

Algorithm 1 Partitionnement spectral

Input: Matrice d'adjacence A, un entier K

1. Effectuer une décomposition spectrale de A
2. Approximer le spectre de B en retirant au spectre de A sa plus grande valeur propre
3. Considerer les K valeurs propres restantes extrémales
4. Appliquer K-`means` au plongement de dimension K induit par les K vecteurs propres associés

Output: Un partitionnement en K classes des noeuds

Notons que lorsque les q_i ne sont pas égaux, les classes se dédoublent sous les différents régimes de connectivité et cela peut constituer un bruit trop important pour permettre le partitionnement des plongements.

Cas Homogène

Dans le cadre de cette partie, on a $\frac{B}{\sqrt{n}} = \frac{X}{\sqrt{n}} + \frac{q_0^2}{n} JDJ^T =: \frac{X}{\sqrt{n}} + P$
où $J = (j_1, \cdots j_K) \in \mathcal{M}_{n,K}(\mathbb{R})$ les vecteurs canoniques des classes, X est de Wigner,
et $D = \text{diag}(m_1, \cdots m_K)$, avec $m_k := \mathbb{E}[M_{k,k}]$
On posera de plus Q_B et Q_X les résolvantes respectives de $\frac{B}{\sqrt{n}}, \frac{X}{\sqrt{n}}$.

Avant d'étudier le comportement de $\frac{B}{\sqrt{n}}$, étudions $\frac{X}{\sqrt{n}}$, où l'on rappelle que $X = A - \mathbb{E}[A]$. Evidemment les X_{ij} sont centrées, et elles sont indépendantes par hypothèse.
On a $X_{ij} \sim \mathcal{B}(q_0^2 C_{ab})$, ainsi $\mathbb{V}[X_{ij}] = \mathbb{E}[q_0^2 C_{ab}(1 - q_0^2 C_{ab})]$. Mais comme comme $C_{ab} \xrightarrow[n \longrightarrow +\infty]{} 1$ (à une vitesse indépendante de ij et presque-sûrement), on a la convergence (uniforme par rapport à ij) $\mathbb{V}[X_{ij}] \xrightarrow[n \longrightarrow +\infty]{} \sigma^2$, où $\sigma := \sqrt{q_0^2(1 - q_0^2)}$.
On en déduit par le théorème de Wigner que la mesure spectrale de $\frac{X}{\sqrt{n}}$ tend vers $\mathbb{P}_{sc,\sigma}$.
On utilisera également que $Q_X(z) \xrightarrow[n \longrightarrow +\infty]{p.s.} \frac{1}{\sigma} g(z/\sigma)$.

1. *Etape 1: condition sur λ.*
Afin de déterminer une condition d'existence asymptotique de valeurs propres isolées de $\frac{B}{\sqrt{n}}$, on se donne $\lambda \in \mathbb{R} \setminus [-2\sigma, 2\sigma]$ (c'est-à-dire isolé du demi-cercle) qui soit tout de même valeur propre de $\frac{B}{\sqrt{n}}$. On a alors en particulier:
$0 = \det\left(\frac{B}{\sqrt{n}} - \lambda I_n\right) = \det\left(\frac{X}{\sqrt{n}} - \lambda I_n + P\right)$. Or par hypothèse, la matrice $\frac{X}{\sqrt{n}} - \lambda I_n$ est inversible (asymptotiquement, en vertu du théorème "no eigenvalues outside the support"), ce qui permet de factoriser par cette dernière, on a donc:
$0 = \det\left(I_n + P\left(\frac{X}{\sqrt{n}} - \lambda I_n\right)^{-1}\right) = \det\left(I_n + P Q_X(\lambda)\right)$. Or par le théorème de Wigner isotrope, on a l'approximation presque sûre lorsque $n \longrightarrow +\infty$ $\left(\frac{X}{\sqrt{n}} - \lambda I_n\right)^{-1} = g(\lambda/\sigma)/\sigma I_n + o(n)$. Nos considérations asymptotiques nous permettent d'omettre le $o(n)$, quitte à écrire $o(1) = \det(\cdots)$, calculer et passer à la limite.

Ainsi on a la condition $0 = \det\left(I_n + g(\lambda/\sigma)/\sigma P\right)$, ce qui entraîne $0 = \det\left(P + \frac{\sigma}{g(\lambda/\sigma)} I_n\right)$.

Or $P = \frac{q_0^2}{n} J D J^T$, donc d'après l'identité de Sylvester, $0 = \det\left(\frac{q_0^2}{n} D J^T J + \frac{\sigma}{g(\lambda/\sigma)} I_K\right)$.

Puis comme $J^T J = \begin{pmatrix} j_1^T \\ \vdots \\ j_K^T \end{pmatrix} (j_1, \cdots j_K) = \mathrm{diag}(\#\mathcal{C}_1, \cdots \#\mathcal{C}_K)$,

On a pour tout $k \in [\![1, K]\!]$, $\frac{q_0^2 m_k \#\mathcal{C}_k}{n} + \frac{\sigma}{g(\lambda/\sigma)} = 0$. On remplace $\frac{\#\mathcal{C}_k}{n}$ par sa limite c_k: $\boxed{g(\lambda/\sigma) = -\frac{\sigma}{q_0^2 m_k c_k}}$

Remarquons que dans le cas $m_k = 0$, il n'y a pas de valeur propre λ satisfaisant notre condition nécessaire, on suppose donc $m_k \neq 0$.

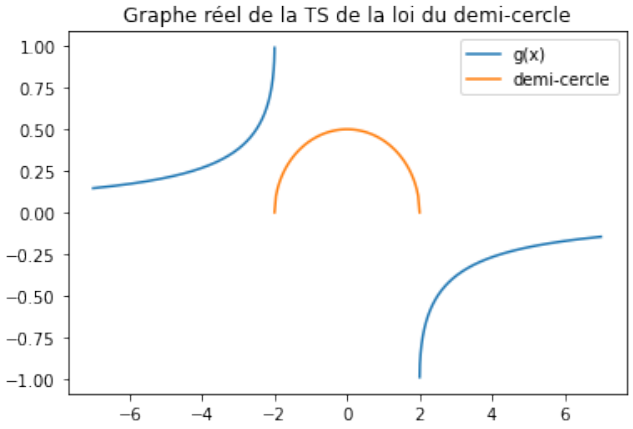
Etape 2: calcul de g sur $\mathbb{R} \setminus [-2, 2]$.

On cherche $g : \mathbb{R} \setminus [-2, 2] \to \mathbb{R}$ caractérisée par les propriétés suivantes:

(1): $\forall x \in \mathbb{R} \setminus [-2, 2], \quad g(x)^2 + x g(x) + 1 = 0$, (2): g est croissante sur $]-\infty, -2[$ et sur $]2, +\infty[$.

La condition (1) entraîne que $\forall x \in \mathbb{R} \setminus [-2, 2], \quad g(x) \in \left\{\frac{-x \pm \sqrt{x^2 - 4}}{2}\right\}$,

Puis la condition (2) impose finalement $\boxed{\forall x \in \mathbb{R} \setminus [-2, 2], \quad g(x) = \frac{-x + \mathrm{sign}(x)\sqrt{x^2 - 4}}{2}}$



Etape 3: condition sur les paramètres

La condition s'écrit $\exists \lambda \in \mathbb{R} \setminus [-2\sigma, 2\sigma] : g(\lambda/\sigma) = -\frac{\sigma}{q_0^2 m_k c_k} \iff -\frac{\sigma}{q_0^2 m_k c_k} \in]-1, 0[\cup]0, 1[$.

La condition est finalement $\boxed{q_0^2 |m_k| c_k > \sigma}$.

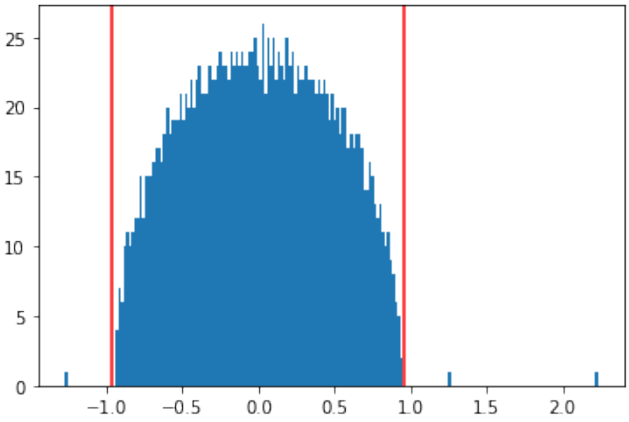
2. Supposons la condition vérifiée, maintenant il s'agit de trouver $\lambda_k \in \mathbb{R} \setminus [-2\sigma, 2\sigma]$ solution de $g(\lambda_k/\sigma) = -\frac{\sigma}{q_0^2 m_k c_k}$. On sait qu'un tel λ_k existe (et est même unique) par le théorème des valeurs intermédiaires, fixons-le et calculons, en écrivant l'équation sous la forme $g(\mu) = \alpha$:

On a $-\mu + \mathrm{sign}(-m_k)\sqrt{\mu^2 - 4} = 2\alpha$ donc $\mu^2 - 4 = (2\alpha + \mu)^2$ puis $\alpha^2 + \alpha\mu + 1 = 0$.

Finalement $\mu = -\frac{1 + \alpha^2}{\alpha}$, puis en remplaçant $\alpha = -\frac{\sigma}{q_0^2 m_k c_k}$ et $\mu = \lambda_k/\sigma$, $\boxed{\lambda_k = \frac{\sigma^2 + (q_0^2 m_k c_k)^2}{q_0^2 m_k c_k}}$.

Pour se rassurer, on constate que $\beta \mapsto \frac{\sigma^2 + \beta^2}{\beta}$ envoie $]\sigma, +\infty[$ sur $]2\sigma, +\infty[$.

Pour vérifier, on prend $n = 2000, K = 3, m = (-5, 5, 10), q_0 = 0.8$:



La théorie prévoit des spikes de $(-1.283, 1.283, 2.241)$, et l'on calcule numériquement des spikes $(-1.264, 1.262, 2.234)$, ce qui est acceptable. Ces valeurs numériques ont des fluctuations d'environ ± 0.01 pour différents tirages de B.

Il est important de noter que cette précision dépend de l'ordre de grandeur de m et q (même sous la condition de la question 1), en particulier l'approximation de la variance par σ^2 peut devenir facilement trop fausse.

3. Commençons par quelques notations: on notera $(\hat{u}_i, \hat{\lambda}_i)_{i\in[\![1,n]\!]}$ les éléments spectraux de $\frac{B}{\sqrt{n}}$. On les ordonne de sorte que $\hat{\lambda}_k \xrightarrow[n\to+\infty]{p.s.} \lambda_k$ pour $k \in [\![1, K]\!]$, et λ_i tende p.s. vers une valeur propre du bulk. On fixe $k \in [\![1, K]\!]$ et on se place dans les conditions de la question 1). On suppose les λ_k simples (multiplicité 1), le calcul devenant atroce en général.

On cherche à calculer asymptotiquement l'alignement $\mathcal{A}_k := \left|\frac{j_k}{\sqrt{\#\mathcal{C}_k}} \cdot \hat{u}_k\right|^2$.

Déjà, par définition $Q_B(z) = \left(\frac{B}{\sqrt{n}} - zI_n\right)^{-1} = \sum_{i=1}^{K} \frac{\hat{u}_i\hat{u}_i^T}{\hat{\lambda}_i - z} + \sum_{i=K+1}^{n} \frac{\hat{u}_i\hat{u}_i^T}{\hat{\lambda}_i - z}$ est méromorphe sur $\mathbb{C} \setminus [-2\sigma, 2\sigma]$ de pôles $(\hat{\lambda}_i)_{i\in[\![1,K]\!]}$ et de résidus associés $-\hat{u}_i\hat{u}_i^T$ (entrée par entrée de la matrice).

Soit Γ un contour encerclant λ_k uniquement (par exemple on peut prendre $\gamma(t) = \lambda_k + \varepsilon e^{2i\pi t}$), ce qui est possible pour n suffisamment grand, λ_k étant isolée.

Par le Théorème des Résidus, $-\frac{1}{\#\mathcal{C}_k 2i\pi}\oint_\Gamma j_k^T Q(z) j_k \mathrm{d}z = \left|\frac{j_k}{\sqrt{\#\mathcal{C}_k}} \cdot \hat{u}_k\right|^2$.

Par ailleurs, remanipulons l'intégrande avec $z \in \Gamma$ fixé: $Q(z) = \left(Q_X(z)^{-1} + J\widetilde{D}J^T\right)^{-1}$ avec $\widetilde{D} := \frac{q_0^2}{n}D$.

Par l'identité de Woodbury (cas général): $Q(z) = Q_X(z) - Q_X(z)J\left(\widetilde{D}^{-1} + J^T Q_X(z) J\right)^{-1} J^T Q_X(z)$.

De même que dans 1), par le théorème de Wigner isotrope, on remplace asymptotiquement $Q_X(z)$ par $g(z/\sigma)/\sigma I_n$ (avec l'intention de repasser aux scalaires en applicant $j_k^T[\cdot]j_k$). Dans la suite, on raisonne asymptotiquement et presque sûrement:

$Q(z) = g(z/\sigma)/\sigma I_n - (g(z/\sigma)/\sigma)^2 J\left(\widetilde{D}^{-1} + g(z/\sigma)/\sigma J^T J\right)^{-1} J^T$. Le premier terme étant carrément holomorphe sur un voisinage contenant Γ, son intégrale sur Γ sera nulle.

Comme $J^T J = \mathrm{diag}(\#\mathcal{C}_1, \cdots \#\mathcal{C}_K)$, on a $\widetilde{D}^{-1} + g(z/\sigma)/\sigma J^T J = \mathrm{diag}\left(\left(\frac{q_0^2 m_i}{n}\right)^{-1} + g(z/\sigma)/\sigma \#\mathcal{C}_i\right)_{i\in[\![1,K]\!]}$.

Ensuite son inverse $\Omega := \left(\widetilde{D}^{-1} + g(z/\sigma)/\sigma J^T J\right)^{-1}$ est diagonale de terme en $(i,i) \in [\![1,K]\!]^2$: $\frac{\frac{q_0^2 m_i}{n}}{1 + q_0^2 m_i c_i g(z/\sigma)/\sigma}$, où l'on a utilisé $\frac{\#\mathcal{C}_i}{n} \to c_i$.

Mettons tout ensemble: à un terme holomorphe près, l'intégrande (en incluant le facteur $-\frac{1}{\#\mathcal{C}_k}$) s'écrit:

$$\frac{1}{\#\mathcal{C}_k}(g(z/\sigma)/\sigma)^2 j_k^T J\Omega J^T j_k = \frac{1}{\#\mathcal{C}_k}\#\mathcal{C}_k^2 (g(z/\sigma)/\sigma)^2 \Omega_{k,k} = (g(z/\sigma)/\sigma)^2 \frac{q_0^2 m_k c_k}{1 + q_0^2 m_k c_k g(z/\sigma)/\sigma} =: h(z)$$

Il s'agit maintenant de calculer le résidu de h en λ_k, qui est égal à $\frac{1}{2i\pi}\oint_\Gamma h(z)\mathrm{d}z$ et donc à $\mathcal{A}_k$.

Remarquons que par définition, $g(\lambda_k/\sigma)/\sigma = -\frac{1}{q_0^2 m_k c_k}$, ce qui est normal mais aussi très rassurant.

Calculons $\mathcal{A}_k = \lim_{z\to\lambda_k}(z-\lambda_k)\left(g(z/\sigma)/\sigma\right)^2 \frac{q_0^2 m_k c_k}{1+q_0^2 c_k m_k g(z/\sigma)/\sigma}$

Donc $\mathcal{A}_k = \left(g(\lambda_k/\sigma)/\sigma\right)^2 \lim_{z\to\lambda_k}\frac{z-\lambda_k}{g(z/\sigma)/\sigma - g(\lambda_k/\sigma)/\sigma} = g^2(z/\lambda_k)\frac{1}{g'(\lambda_k/\sigma)}$.

Ainsi, en utilisant la formule sur g' donnée dans le sujet, $\mathcal{A}_k = g^2(\lambda_k/\sigma)\frac{1-g^2(\lambda_k/\sigma)}{g^2(\lambda_k/\sigma)} = 1 - g^2(\lambda_k/\sigma)$

Puis finalement $\boxed{\mathcal{A}_k = 1 - \frac{\sigma^2}{\left(q_0^2 m_k c_k\right)^2}}$

Remarquons qu'avec $\beta := q_0^2 m_k c_k$, les hypothèses imposent $|\beta| > \sigma$, donc $1 - \frac{\sigma^2}{\beta^2}$ est bien positif.

4. Nous conservons les mêmes paramètres expérimentaux que dans la question 2.
Les alignements respectifs des spikes à $(-1.283, 1.283, 2.241)$ sont:

- Théorie: $[0.798, 0.798, 0.949]$
- Pratique: $[0.788, 0.804, 0.955]$

5. On se propose d'utiliser l'algorithme 1 de partitionnement spectral introduit précédemment. Afin de l'évaluer, il nous faut une vérité terrain du partitionnement à atteindre.

- Pour commencer, on peut l'appliquer à des jeux de données synthétiques générés par le modèle décrit plus haut.
- Afin d'évaluer la pertinence de ce modèle, on peut appliquer l'algorithme sur des bases de données empiriques et labelisées comme *Zachary's karate club* ou *Cora dataset*.

Des métriques de classification multi-classe telle que la précision, le score F1 ou l'aire sous la courbe d'efficacité récepteur (AUC-ROC) permettent de quantifier les performances.

Cas Hétérogène

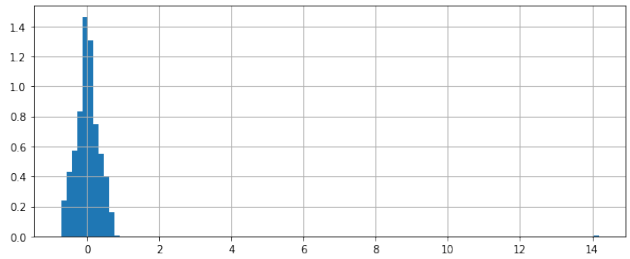
1. On prend comme précédemment $n = 2000$, $K = 3$, avec cette fois ci M égale à $\bar{M}$ perturbée d'une variable normale symétrique où $\bar{M} = \mathrm{diag}(2.5, 3.5, 5)$. On se place dans le cas non-homogène $q_i \sim \mathcal{U}([0.1, 0.8])$.

On observe alors que la variance supplémentaire induite par la non-homogénéité ne permet de distinguer plusx² qu'une seule valeur propre en dehors du bulk, voir figure 4a.

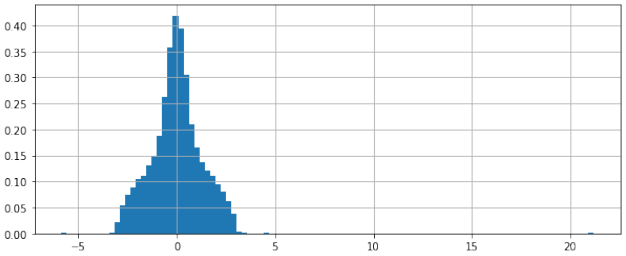
2. On se propose de corriger la non-homogénéitée en renormalisant la matrice d'adjacence A par les degrées des noeuds, i.e. $\tilde{A} = D^{-1}A$ où $D = \mathrm{diag}(\sum_{i=1}^{n} A_{i,j})$.

En figure 4b, on observe 3 valeurs propres de $\tilde{A}$ en dehors du bulk ce qui permet d'appliquer efficacement notre algorithme de partitionnement spectral.

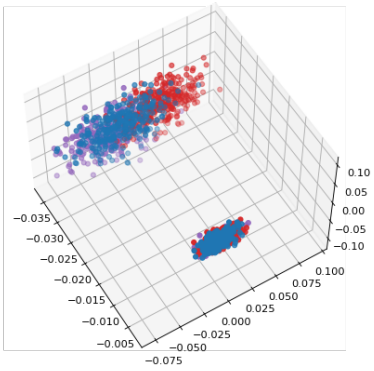
(a) Valeurs propres de $\frac{A}{\sqrt{n}}$



(b) Valeurs propres de $\frac{\tilde{A}}{\sqrt{n}}$



Cela est confirmé par la visualisation des plongements de dimension 3 issus des 3 vecteurs propres extrémaux. Alors que sans renormalisation, le plongement ne permet de distinguer que les régimes de connectivité $q^{(1)} = 0.1$ et $q^{(2)} = 0.8$, la renormalisation par les degrés permet de plus de clairement distinguer les classes.



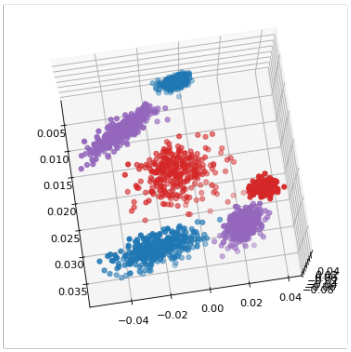


Figure 5: Plongements de dimension K issus des vecteurs propres extrémaux de $\frac{A}{\sqrt{n}}$ et $\frac{\tilde{A}}{\sqrt{n}}$