UNIVERSITÉ DE LILLE

THÈSE

pour obtenir le grade de :

DOCTEUR DE L'UNIVERSITÉ DE LILLE

dans la spécialité

« MATHÉMATIQUES PURES ET APPLIQUÉES »

par

Imane Akjouj

Équilibres d'écosystèmes de grande taille via la théorie des matrices aléatoires.

Soutenance le 29 juin 2023 devant le jury composé de :

M.	NICOLAS CHAMPAGNAT	INRIA Nancy Grand Est,	(Rapporteur)
Mme	CATHERINE DONATI-MARTIN	Université de Versailles,	(Rapporteur)
M.	DAVID DEREUDRE	Université de Lille,	(Examinateur)
M.	MAXIME FÉVRIER	Université Paris-Saclay,	(Examinateur)
Mme	VALENTINA ROS	CNRS, Université Paris-Saclay,	(Examinateur)
Mme	MYLÈNE MAÏDA	Université de Lille,	(Directrice de Thèse)
M.	JAMAL NAJIM	CNRS, Université Gustave Eiffel,	(Co-Directeur de Thèse)

Laboratoire Paul Painlevé
BÂTIMENT M2, CITÉ SCIENTIFIQUE, 59655 VILLENEUVE-DASCQ
FRANCE

Résumé

La modélisation mathématique des écosystèmes permet une étude des questions liées à la diversité des espèces et à la complexité de leurs interactions. En biologie et en écologie mathématiques, l'usage de grands systèmes de Lotka-Volterra est courant dans la modélisation de la dynamique des écosystèmes faisant intervenir des espèces qui interagissent entre elles.

Lorsque les écosystèmes, réseaux trophiques ou microbiomes, impliquent de nombreuses espèces, il peut être difficile d'observer ou de mesurer les interactions entre ces espèces et pertinent de considérer les interactions comme aléatoires. Depuis les années 70, certains écologues ont ainsi fait appel aux résultats de la théorie des matrices aléatoires (RMT) dans l'étude des réseaux trophiques. La matrice des interactions est alors une matrice aléatoire.

Dans une première partie de cette thèse se posera la question de l'existence d'un équilibre faisable, c'est-à-dire d'une solution strictement positive du système de Lotka-Volterra, ce qui correspond au scénario où aucune espèce ne disparaît au cours de la dynamique. Par ailleurs, certains modèles en écologie font appel à des matrices creuses, contenant de nombreux zéros ; chaque espèce interagissant avec un petit nombre d'autres espèces. En RMT, l'étude des matrices creuses est assez récente et c'est dans ce contexte que se posent les questions de la faisabilité et de la stabilité de l'équilibre. L'existence asymptotique, lorsque le nombre d'espèces tend vers l'infini, d'un seuil de faisabilité est démontrée pour deux modèles : lorsque la matrice des interactions a une structure par blocs et lorsque le paramètre de sparsité est proportionnel au nombre d'espèces.

La deuxième partie portera sur une toute autre question, celle de la proportion d'espèces survivantes et la mesure empirique du vecteur solution du système de Lotka-Volterra. En particulier, les résultats présentés sont obtenus dans le cas de matrices d'interactions symétriques, appartenant à l'Ensemble Gaussien Orthogonal, ou dans le cas de matrices de Wishart, utilisées pour mesurer notamment la « proximité » des espèces en fonction de leurs traits. Ce chapitre nous permettra de faire le lien entre l'équilibre du système de Lotka-Volterra et la solution du Linear Complementary Problem ainsi que d'appliquer l'algorithme de l'Approximate Message Passing.

L'objectif de la troisième partie sera consacré à l'étude d'un théorème central limite, s'intéressant au comportement asymptotique de la solution d'un système algébrique d'équations couplés dont les coefficients sont aléatoires. La démonstration s'appuiera sur des outils d'analyse combinatoire et de théorie des graphes.

Enfin, dans la dernière partie, divers modèles de matrices aléatoires structurées apparaissant dans la littérature en écologie et en biomathématiques seront présentés. L'accent sera mis sur leur utilisation en écologie théorique tout en faisant le lien avec des résultats mathématiques et des questions ouvertes.

Abstract

Mathematical modeling of ecosystems offers an approach to study issues related to the diversity of species and the complexity of their interactions. In mathematical ecology and biology, it is common to use large systems of Lotka-Volterra (LV) to model the dynamics of ecosystems with interacting species.

When ecosystems, food webs and microbiomes, involve many species, it becomes difficult to observe or quantify the interactions between those species which makes it relevant to consider the interactions as random. Since the seventies, some ecologists use Random Matrix Theory (RMT) in the study of food webs. The interaction matrix is then random.

In the first part of the thesis, the question of the existence of a feasible equilibrium arises. It corresponds to a positive solution of the Lotka-Volterra system, which is the scenario where no species disappears during the dynamics. Furthermore, some ecological models rely on sparse matrices with many zeros, each species interacting with few others. In RMT, the study of sparse matrices is a recent topic and it is in this context that emerge the issues of feasiblity equilibrium and its stability. When the number of species goes to infinity, the asymptotic existence of a feasibility threshold is proved for two models : for an interaction matrix with a block structure and for a sparsity parameter proportional to the species number.

The second part concerns a totally different matter, the one of the proportion of surviving species and the behavior of the empirical measure of the LV equilibrium solution. The provided results are obtain for symmetric interaction matrices, from the Gaussian Orthogonal Ensemble, and for Wishart matrices, used to evaluate the « proximity » between species according to their features. This chapter provides a link between equilibrium of Lotka-Volterra system and the solution of Linear Complementary Problem and allows for an application of the Approximate Message Passing algorithm.

The goal of the third part is to the study of a central limit theorem, showing the asymptotic behavior of the solution of an algebraic system of coupled equations with random coefficients. The proof relies on combinatorics and graph theory tools.

Finally, the last part will be dedicated to present a variety of structured random matrix models, that appear in ecological and biomathematical literature, emphasizing on their use in theoretical ecology and listing mathematical results and questions of interest.

Table des matières

Introduction

Etymologiquement, le terme écologie provient du grec *oîkos* (« maison, habitat ») et *lógos* (« discours ») : il signifie alors la science de l'habitat.
Inventé en 1866 par Ernst Haeckel, biologiste darwiniste, dans son ouvrage *Morphologie générale des organismes* [Haeckel, 1966], ce terme désignait « la science des relations des organismes avec le monde environnant, c'est-à-dire, dans un sens large, la science des conditions d'existence ».
De nos jours, cette science étudie les êtres vivants (animaux, végétaux, micro-organismes, etc.), leur milieu et les interactions qu'ils entretiennent entre eux et avec leur milieu. Lorsque l'accent est mis sur les interactions entre espèces, il s'agira d'écologie des communautés.

En écologie, une espèce est constituée d'individus pouvant se reproduire entre eux et produire une descendance fertile, formant ainsi des populations. En règle générale, les individus sont mutuellement dépendants les uns des autres pour leur survie.
Une communauté est formée des individus de différentes espèces vivant dans une même région. L'ensemble des êtres vivants dans un environnement donné (par exemple : forêt, intestin, lac, etc.) forme un écosystème [Tansley, 1939].

La compréhension des écosystèmes et de leurs mécanismes de fonctionnement constitue un défi majeur en écologie [MacArthur et Wilson, 1967].

A ce jour, les scientifiques ont répertorié plus de 2 millions d'espèces sur Terre. Certains rares écosystèmes sont ainsi très petits (2-3 espèces), tandis que la plupart ont un très grand nombre d'espèces, comme la forêt amazonienne qui abrite 10^6 espèces ou le microbiome qui recèle entre 10^3 et 10^{18} espèces [Coyte *et al.*, 2015]. Cette grande diversité d'espèces est nécessaire à la survie des êtres vivants.

La compréhension de ces grands écosystèmes pourrait permettre de gérer durablement les populations animales et végétales afin de protéger celles qui sont menacées, ou encore d'avoir une meilleure gestion des antibiotiques sur notre flore intestinale. Cependant, de nombreuses questions restent sans réponse, la compréhension de ces grands écosystèmes demandant de nombreuses données empiriques et les études expérimentales n'étant pas adaptées à leur étude.

De nombreuses études expérimentales sont menées sur des systèmes à peu d'espèces, tandis que la récolte de données à grande échelle connaît rapidement des limites logistiques. Ces dernières années, de nombreux outils technologiques ont été développés en laboratoire pour étudier des systèmes microbiologiques et permettraient de faire des comparaisons avec les études théoriques [Hu *et al.*, 2021]. Ainsi, l'émergence du deep learning amène à l'automatisation de nombreux processus permettant, par exemple, de reconnaître des espèces et de les compter sur des images prises par avion ou drone.

Le manque de données peut aussi être compensé par l'utilisation et l'étude de modèles. Ces derniers n'ont pas pour unique objectif de prédire la dynamique de l'écosystème mais également de comprendre les mécanismes qui permettent une grande diversité.

L'une des grandes questions en écologie concerne la relation entre la diversité et la stabilité d'un écosystème. Depuis longtemps, de nombreux écologues ont ainsi suggéré que la diversité des communautés renforçait la stabilité des écosystèmes [MacArthur, 1955].
Cependant, partant d'un modèle théorique qu'il a introduit dans les années 70, [May, 1972] a remis en question la relation diversité-stabilité en utilisant une analyse de stabilité linéaire sur un modèle de communauté construit de manière aléatoire et a découvert que la diversité seule tend à déstabiliser le système. Cela a conduit au débat sur la diversité-stabilité [May, 1973, Yodzis, 1981, McCann, 2000, Ives et Carpenter, 2007, Jacquet *et al.*, 2016, Landi *et al.*, 2018]. Les enjeux théoriques consistent à trouver, entre autres, les arguments et mécanismes manquants au modèle de May.

Dans cette thèse, il sera question de l'étude de grands écosystèmes dans l'objectif de comprendre l'un des principaux facteurs écologiques affectant leur diversité et leur dynamique : les interactions entre espèces.
En particulier, le système de Lotka-Volterra, utilisé pour décrire la dynamique de la communauté, permettra d'appréhender certains effets des interactions sur les propriétés d'un équilibre telles que l'existence, la diversité, la stabilité.

1 Brève introduction à la théorie des matrices aléatoires

Le statisticien John Wishart introduit la théorie des matrices aléatoires à la fin des années 20'. Il s'intéresse alors à des matrices aléatoires de covariance empirique d'échantillons gaussiens multivariés [Wishart, 1928]. Dans les années 50, [Wigner, 1955] y fait appel pour une toute autre raison : expliquer la distribution des niveaux d'énergie dans les noyaux atomiques. L'approche innovante utilisée par [Wigner, 1967] pour décrire le spectre d'une matrice aléatoire Hermitienne a été reprise par d'autres physiciens pour résoudre notamment des problèmes de physique nucléaire [Dyson, 1962].
Par la suite, de nouvelles structures matricielles ont été étudiées. Ainsi, de nombreux travaux ont été réalisés par [Marčenko et Pastur, 1967] sur les grandes matrices de covariance. [Girko, 1985], [Bai, 1997] et [Silverstein et Choi, 1995, Bai et Silverstein, 2010] se sont, quant à eux, penchés sur l'extension des résultats aux matrices non-Hermitiennes. La théorie des matrices aléatoires s'est depuis lors étendue et une multitude de travaux ont été réalisés dans des domaines très divers des mathématiques tels que la combinatoire, les graphes aléatoires, la théorie des probabilités

libres, la théorie du signal, la théorie des nombres, etc.

La théorie des matrices aléatoires doit sa force, entre autres, au comportement universel du spectre de ces matrices, lorsque leur dimension tend vers l'infini. Autrement dit, l'un des résultats clefs de RMT s'apparente à la loi des grands nombres et concerne le spectre de la matrice ; la distribution des valeurs propres de la matrice revêt un caractère déterministe.

Dès lors, de nombreux résultats de la théorie des matrices aléatoires s'intéressent aux propriétés du spectre des matrices : valeurs propres, vecteurs propres, plus grande valeur propre, etc ; et reposent d'un point de vue technique sur une combinaison de divers domaines des mathématiques tels que l'algèbre linéaire, les probabilités, l'analyse complexe ou encore la combinatoire.

1.1 Premières définitions

Soit $A := (A_{ij})_{1\leq i,j\leq n}$ une matrice carrée de taille n à coefficients complexes ($A_{ij} \in \mathbb{C}$). On note $A^* := \bar{A}^T$.

Definition 1.1. Les *valeurs propres* de A, notées $\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)$, sont les racines de son polynôme caractéristique.
L'ensemble des valeurs propres de A est appelé *spectre* de A et est noté $\mathrm{sp}(A)$.

Definition 1.2. On suppose que

$$|\lambda_1(A)| \leq |\lambda_2(A)| \leq \cdots \leq |\lambda_n(A)| \,.$$

Le *rayon spectral* de la matrice A, noté $\rho(A)$, est défini par

$$\rho(A) = |\lambda_n(A)| = \max_{i\in[n]} |\lambda_i(A)| \,, \qquad \text{où } [n] := [\![1; n]\!] \,.$$

Definition 1.3. Les *valeurs singulières* $\sigma_1(A) \leq \sigma_2(A) \leq \cdots \leq \sigma_n(A)$ de la matrice A sont les racines carrées des valeurs propres de la matrice hermitienne A^*A :

$$\sigma_i(A) := \sqrt{\lambda_i(A^*A)} \,, \quad \forall i \in [n] \,.$$

Definition 1.4. Soit un vecteur $\boldsymbol{x} \in \mathbb{R}^n$, on note $\|\boldsymbol{x}\|_2$ sa *norme euclidienne* :

$$\|\boldsymbol{x}\|_2 = \left(\sum_{k=1}^{n} |x_k|^2 \right)^{\frac{1}{2}} \,.$$

La *norme spectrale* de la matrice A, notée $\|A\|$ est définie par sa plus grande valeur singulière :

$$\|A\| := \max_{\|x\|_2=1} \|Ax\|_2 = \sigma_n(A) \,.$$

Definition 1.5. Soit $E \subset \mathbb{C}$, on note δ_λ la *masse de Dirac* au point λ définie par

$$\delta_\lambda(E) = \mathbb{1}_E(\lambda) = \begin{cases} 1 & \text{si } \lambda \in E \,, \\ 0 & \text{sinon} \,, \end{cases}$$

où $\mathbb{1}_I$ désigne la *fonction indicatrice* de I.

Definition 1.6. Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)$ ses valeurs propres. On définit la *mesure spectrale empirique* de A dans $(\mathbb{C}, \mathcal{B}(\mathbb{C}))$ par

$$\mu_A := \frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda_k(A)} \, .$$

Pour tout sous-ensemble $E \subset \mathbb{C}$, la quantité :

$$\mu_A(E) = \frac{\mathrm{card}\{1 \leq k \leq n \mid \lambda_k(A) \in E\}}{n} \, ,$$

désigne la proportion de valeurs propres de A dans E.

La mesure spectrale vient caractériser le spectre d'une matrice et son comportement. Ainsi, la convergence faible d'une mesure spectrale empirique vers une mesure déterministe décrit de nombreux résultats de RMT.

Definition 1.7. On dit que μ_A *converge faiblement* vers une mesure de probabilité μ, i.e. $\mu_A \xrightarrow[n\to\infty]{d} \mu$, si pour toute fonction f continue et bornée sur $\mathbb{R}$,

$$\int f(u)\mu_A(\mathrm{d}u) = \frac{1}{n} \sum_{k=1}^{n} f(\lambda_k) \xrightarrow[n\to\infty]{} \int f(u)\mu(\mathrm{d}u) \, .$$

Remark 1. Si A est aléatoire, alors μ_A est une probabilité aléatoire (à support fini), cela implique $\int f \mu_A(\mathrm{d}u)$ sont aussi des variables aléatoires. Nous dirons alors que presque sûrement (p.s.) μ_A converge faiblement vers μ

$$\mu_A \xrightarrow[n\to\infty]{d} \mu \quad (p.s.) \, .$$

Definition 1.8. Soit $A \in \mathcal{M}_n(\mathbb{C})$, on appelle *résolvante* de la matrice A la matrice $Q = (Q_{ij})_{1 \leq i,j \leq n}$ définie par

$$Q(z) := (A - zI)^{-1} \, , \quad z \notin \mathrm{sp}(A) \, .$$

1.2 Matrices de Wigner

Definition 1.9. Soit W_n une matrice Hermitienne de taille $n \times n$, $W_n = W_n^*$ telle que ses entrées $(W_{k\ell})_{1 \leq k \leq \ell \leq n}$ sont des variables aléatoires indépendantes avec $\mathbb{E}(W_{k\ell}) = 0$ et $\mathbb{E}\left(|W_{k\ell}|^2\right) < \infty$, $\forall 1 \leq k \leq \ell$.
On suppose de plus que la famille $(W_{k\ell})_{1 \leq k < \ell \leq n}$ (resp. la famille $(W_{kk})_{1 \leq k \leq n}$) est composée de variables aléatoires identiquement distribuées.

$W_n/\sqrt{n}$ est appelée *matrice de Wigner*.

Theorem 1. *(Universalité du théorème de Wigner et de la loi semi-circulaire).*
Soit W_n une matrice de Wigner définie par la famille de variables aléatoires i.i.d $(W_{k\ell})_{1 \leq k \leq \ell \leq n}$, telles que :

1. $\mathbb{E}(W_{k\ell}) = 0$, $\forall 1 \leq k \leq \ell \leq n$,
2. $\mathbb{E}(|W_{k\ell}|^2) = \sigma^2 < \infty$, $\forall 1 \leq k \leq \ell \leq n$ et $\sigma > 0$.

Alors, presque sûrement, la mesure spectrale empirique de $W_n/\sqrt{n}$ converge faiblement vers la loi semi-circulaire :

$$\mu_{W_n/\sqrt{n}} := \frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k(W_n/\sqrt{n})} \xrightarrow[n\to\infty]{d} \mu_{sc} \qquad \text{(p.s.)}\,,$$

où μ_{sc} est définie par

$$\mathrm{d}\mu_{sc}(t) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - t^2}\mathbb{1}_{[-2\sigma,2\sigma]}(t)\mathrm{d}t\,.$$

Remark 2. Les valeurs propres de la matrice $W_n/\sqrt{n}$ sont réelles.

On note

$$\lambda_{\max}(W_n) = \max_{k\in[n]} \lambda_k(W_n) \quad \text{et} \quad \lambda_{\min}(W_n) = \min_{k\in[n]} \lambda_k(W_n)\,.$$

Dans le cas des matrices de Wigner, de nombreux travaux ont été réalisés dans les années 80' concernant le comportement de la plus grande valeur propre de la matrice [Füredi et Komlós, 1981, Bai et Yin, 1988].

Theorem 2. *(Convergence des valeurs propres extrêmes).*
Si $\mathbb{E}\left(|W_{k\ell}|^4\right) < \infty$, $\forall 1 \leq k \leq \ell$, alors

$$\lambda_{\max}\left(W_n/\sqrt{n}\right) \xrightarrow[n\to\infty]{p.s} 2\sigma \quad \text{et} \quad \lambda_{\min}\left(W_n/\sqrt{n}\right) \xrightarrow[n\to\infty]{p.s} -2\sigma\,.$$

En particulier,

$$\left\|\frac{W_n}{\sqrt{n}}\right\| = \max\left\{\left|\lambda_{\max}\left(\frac{W_n}{\sqrt{n}}\right)\right| \;,\; \left|\lambda_{\min}\left(\frac{W_n}{\sqrt{n}}\right)\right|\right\} \xrightarrow[n\to\infty]{p.s.} 2\sigma\,.$$

Si $\mathbb{E}\left(|W_{k\ell}|^4\right) = \infty$, $\forall 1 \leq k \leq \ell$, alors

$$\lambda_{\max}\left(W_n/\sqrt{n}\right) \xrightarrow[n\to\infty]{p.s} +\infty\,.$$

1.3 Matrices non-hermitiennes

L'objet de cette nouvelle partie d'introduction sur les matrices aléatoires concernera les matrices non-hermitiennes.

Soit $Y_n \in \mathcal{M}_n(\mathbb{C})$ une matrice aléatoire carrée de dimension $n\times n$ dont les entrées sont i.i.d. centrées de variance σ^2. Les valeurs propres de Y_n ne sont plus réelles mais complexes.
Le résultat principal concerne la convergence de la mesure spectrale empirique de $Y_n/\sqrt{n}$ vers la loi circulaire dans le plan complexe.

Suite aux travaux de [Ginibre, 1965] sur la formule explicite du spectre , [Mehta, 1967] prouve cette convergence pour la distribution spectrale empirique moyenne dans le cas d'une gaussienne complexe. Par la suite, [Edelman, 1997] a établi la loi circulaire dans le cas de variables aléatoires gaussiennes réelles. [Silverstein et Choi, 1995], quant à eux, ont donné un argument pour passer de la convergence moyenne à la convergence presque sûre. [Girko, 1985] a travaillé sur la version universelle (pour d'autres types de distribution) en fournissant quelques éléments de preuves tels que la technique d'hermétisation. Et c'est finalement [Tao *et al.*, 2010] qui sont parvenus à la démonstration dans le cas général.

Theorem 3. Soit Y_n une matrice aléatoire $\mathcal{M}_n(\mathbb{C})$ telle que $(Y_{k\ell})_{k,\ell\in[n]}$ sont des variables aléatoires i.i.d centrées et de variance $\mathbb{E}(|Y_{k\ell}|^2) = \sigma^2$.
Alors, presque sûrement, la mesure spectrale empirique de $Y_n/\sqrt{n}$ converge faiblement vers la loi circulaire

$$\mu_{Y_n/\sqrt{n}} \xrightarrow[n\to\infty]{d} \mu_c \,,$$

où μ_c est la loi circulaire, i.e. la loi uniforme sur le disque de rayon σ de $\mathbb{C}$ avec comme densité

$$\mathrm{d}\mu_c(z) = \frac{1}{\pi\sigma^2}\mathbb{1}_{z\in\mathbb{C},\ |z|\leq\sigma}\mathrm{d}z \,.$$

Dans le cas de la loi du cercle, la position du rayon spectral revêt une certaine importance. De nombreux travaux ont été réalisés sur ce sujet, en particulier par [Bai et Yin, 1988, Geman, 1986, Geman et Hwang, 1982].

Theorem 4. *(Convergence des valeurs propres extrêmes).*
Si $\mathbb{E}(Y_{k\ell}) = 0$ et $\mathbb{E}(|Y_{k\ell}|^4) < \infty$, $\forall 1 \leq k, \ell \leq n$, alors

$$\left\| \frac{Y_n}{\sqrt{n}} \right\| \xrightarrow[n\to\infty]{p.s.} 2\sigma \qquad \text{et} \qquad \rho\left(\frac{Y_n}{\sqrt{n}}\right) \xrightarrow[n\to\infty]{p.s.} \sigma \,.$$

Plus récemment, [Bordenave *et al.*, 2021] ont montré cette convergence sous des hypothèses de moment plus optimales.

1.4 Le modèle elliptique

Le modèle elliptique est initialement introduit par [Girko, 1986] et a depuis été largement étudié [Girko, 1995, Naumov, 2012, Nguyen et O'Rourke, 2015, O'Rourke et Renfrew, 2014].

Definition 1.10. *(Modèle elliptique aléatoire).*
Soit X_n une matrice aléatoire réelle qui satisfait aux trois conditions suivantes :

1. Les paires $(X_{k\ell}, X_{\ell k})$, $k \neq \ell$ sont des variables aléatoires i.i.d avec
$$\forall k \neq \ell\,,\quad \mathbb{E}(X_{k\ell}) = 0\,,\quad \mathbb{E}(|X_{k\ell}|^2) = 1 \quad \text{et} \quad \mathbb{E}(|X_{k\ell}|^4) < \infty \,.$$
2. Pour $k < \ell$, le vecteur $(X_{k\ell}, X_{\ell k})$ est tiré d'une distribution bivariée avec une covariance $\mathbb{E}(X_{k\ell}X_{\ell k}) = \rho$ avec $|\rho| \leq 1$.
3. $(X_{kk}, 1 \leq k \leq n)$ sont des variables aléatoires i.i.d., indépendantes des entrées hors-diagonales avec $\mathbb{E}(X_{kk}) = 0$ et $\mathbb{E}(|X_{kk}|^2) = 1$.

Pour $\rho \in (-1, 1)$, on définit *l'ellipsoïde*

$$\epsilon_\rho := \left\{ z = x + iy \in \mathbb{C} \;;\; \frac{x^2}{(1+\rho)^2} + \frac{y^2}{(1-\rho)^2} \leq 1 \right\} .$$

Remark 3. 1. Pour $\rho = 1$, ϵ_1 est l'intervalle $[-2; 2]$ sur l'axe réel et pour $\rho = -1$, ϵ_{-1} est l'intervalle $[-2; 2]$ sur l'axe imaginaire.

2. Si $\rho = 1$, X_n est une matrice de Wigner.

3. Si $\rho = 0$ et que X_{11} et X_{12} sont de même loi, alors X_n est une matrice non-Hermitienne (telle que définie par le théorème 3).

Theorem 5. *(Loi elliptique).*
Soit X_n une variable aléatoire elliptique satisfaisant les conditions de la définition 1.10. Alors presque sûrement, la mesure spectrale empirique de $X_n/\sqrt{n}$ converge faiblement vers la loi elliptique :

$$\mu_{\frac{X_n}{\sqrt{n}}} \xrightarrow[n\to\infty]{d} \mu_\rho \qquad \text{(p.s.)}\,,$$

où μ_ρ est la mesure de probabilité uniforme sur l'ellipsoïde ϵ_ρ de densité

$$\mu_\rho(z) = \begin{cases} \frac{1}{\pi(1-\rho^2)} & \text{si } z \in \epsilon_\rho\,, \\ 0 & \text{sinon}\,. \end{cases}$$

Pour $\rho \in \{-1, 1\}$, la mesure spectrale empirique de $X_n/\sqrt{n}$ converge faiblement vers la loi semi-circulaire.

Le résultat de [O'Rourke et Renfrew, 2014, Cor. 2.3] fournit des informations sur le rayon spectral d'une matrice elliptique.

Proposition 6. *(Rayon spectral d'une matrice aléatoire elliptique).*
Soit X_n une matrice aléatoire elliptique définie dans la définition 1.10, alors

$$\rho\left(\frac{X_n}{\sqrt{n}}\right) \xrightarrow[n\to\infty]{p.s.} 1 + |\rho|\,.$$

2 Le système de Lotka-Volterra en écologie

2.1 Brève introduction aux équations différentielles ordinaires

En écologie, la dynamique des populations peut être modélisée en temps continu ou discret. En temps continu, des équations différentielles ordinaires (EDO) sont utilisées pour décrire l'évolution du vecteur des abondances $\boldsymbol{x} = (x_1, \ldots, x_n)$ d'un système à n espèces au fil du temps.

En EDO, le problème de Cauchy se traduit comme suit :
Soit $f : \mathbb{R}_+^{n+1} \to \mathbb{R}_+^{n+1}$ continue par rapport à $(\boldsymbol{x}, t)$,

$$\begin{cases} \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} &= f(\boldsymbol{x}(t)) \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0\,. \end{cases} \tag{1}$$

L'étude de ce système (1) repose sur les questions liées à l'existence et l'unicité des équilibres et leurs propriétés. Un équilibre $\boldsymbol{x}^*$ du système (1) est une solution de l'équation :

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = 0 \qquad \text{c'est-à-dire} \qquad f(\boldsymbol{x}(t)) = 0\,.$$

La fonction f peut être complexe, les solutions de ce système ne sont pas nécessairement triviales et il peut y avoir plusieurs équilibres avec des propriétés différentes. Dans la suite de cette thèse, la fonction f sera prise indépendante de la variable temps t.

Un équilibre $\boldsymbol{x}^*$ est dit *faisable* si toutes ces composantes sont strictement positives

$$\boldsymbol{x}^* > 0 \quad \text{i.e.} \quad x_k^* > 0,\ \forall k \in \{1, \dots, n\}\,,$$

ce qui correspond au scénario où toutes les espèces survivent.

Une propriété majeure d'un point d'équilibre est sa *stabilité* [Takeuchi, 1996, Chapitre 3]. Un équilibre est stable s'il revient à sa valeur d'équilibre après une petite perturbation du vecteur d'abondance $\boldsymbol{x}$. Un équilibre non stable sera appelé équilibre instable.

Autrement dit, un équilibre $\boldsymbol{x}^*$ est stable si pour tout voisinage W de $\boldsymbol{x}^*$, il existe un voisinage V tel que pour tout point initial $\boldsymbol{x}^0 \in V$, l'orbite $\{\boldsymbol{x}^t;\ t \geq 0;\ \boldsymbol{x}^0 \in V\}$ reste dans W. Dans le cas d'une équation différentielle linéaire, l'étude de la stabilité est triviale et dépend des valeurs propres de l'opérateur linéaire. Dans le cas non linéaire, elle est plus complexe. On peut cependant linéariser le système pour obtenir des informations sur la stabilité locale autour de l'équilibre.

Soit $\boldsymbol{x}^*$ un point d'équilibre de (1), on dit que $\boldsymbol{x}^*$ est *asymptotiquement stable* s'il est stable et si l'orbite $\{\boldsymbol{x}^t;\ t \geq 0;\ \boldsymbol{x}^0 \in V\}$ converge vers $\boldsymbol{x}^*$.

Theorem 7. *(Stabilité d'un équilibre, cas non-linéaire)*
Soit $\boldsymbol{x}^*$ un équilibre d'un système différentiel non linéaire autonome où f est différentiable en $\boldsymbol{x}^*$ et soit $Df(\boldsymbol{x}^*) = \mathcal{J}(x^*)$ sa Jacobienne au point d'équilibre. Soit $\mathrm{sp}(\mathcal{J}(x^*))$ le spectre de $\mathcal{J}(x^*)$ et $Re(\lambda)$ la partie réelle de λ.

1. Si $\forall \lambda \in \mathrm{sp}(\mathcal{J}(x^*))$, $Re(\lambda) < 0$, alors $\boldsymbol{x}^*$ est asymptotiquement stable et on a

$$\forall \mu \in]0; \min -Re(\lambda)[,\ \forall \epsilon < 0,\ \exists \delta > 0,\ |\boldsymbol{x}(t_0) - \boldsymbol{x}^*| < \delta \\ \Rightarrow \forall t \geq t_0,\ \boldsymbol{x}(t) \ \text{ existe et } \ |\boldsymbol{x}(t) - \boldsymbol{x}^*| \leq \epsilon e^{-\mu(t-t_0)}\,.$$

2. S'il existe $\lambda \in \mathrm{sp}(\mathcal{J}(x^*))$ tel que $Re(\lambda) > 0$, alors $\boldsymbol{x}^*$ est instable.
3. Si $\forall \lambda \in \mathrm{sp}(\mathcal{J}(x^*))$, $Re(\lambda) \leq 0$ et il y a au moins une valeur propre nulle ou purement imaginaire, alors on ne peut pas conclure.

Finalement, l'équilibre $\boldsymbol{x}^*$ est *globalement (asymptotiquement) stable* s'il est asymptotiquement stable et si le voisinage W peut être pris comme étant $\mathbb{R}^n_{*+}$ i.e. pour tout $\boldsymbol{x}_0 > 0$, la solution de (1) qui commence à $\boldsymbol{x}(0) = \boldsymbol{x}_0$ satisfait

$$\boldsymbol{x}(t) \xrightarrow[t \to \infty]{} \boldsymbol{x}^*\,.$$

Une matrice Y est *(Lyapunov) stable* (ses valeurs propres ont une partie réelle strictement négative) si et seulement s'il existe une matrice définie positive H telle que $HY + Y^{\mathrm{T}}H$ est définie négative. Cette condition est issue des travaux de [Liapounoff, 1907], approfondis par [Barker *et al.*, 1978] et [Logofet, 2005].

Une matrice Y est *Lyapunov diagonalement stable* si et seulement s'il existe une matrice diagonale D à éléments strictement positifs telle que $DY + Y^{\mathrm{T}}D$ est définie négative.

2.2 Émergence des modèles densité-dépendant en écologie

Les modèles basés sur les équations différentielles, comme le modèle densité-dépendant, sont fréquemment utilisés en biologie et en écologie pour décrire un écosystème d'espèces en interaction. Le modèle densité-dépendant se définit par le système suivant :

$$\begin{cases} \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} &= \boldsymbol{x}(t) f(\boldsymbol{x}(t)) \quad \text{i.e.} \quad \frac{\mathrm{d}x_k(t)}{\mathrm{d}t} = x_k(t) f(\boldsymbol{x}(t)) \;\; \forall k \in [n]\,, \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0\,, \end{cases} \tag{2}$$

où f correspond au taux de croissance d'une espèce.
Lorsque $f(\boldsymbol{x}(t)) < 0$ (resp. $f(\boldsymbol{x}(t)) > 0$), la dynamique sera décroissante (resp. croissante).

À la fin du 18$^{\text{ème}}$ et au début du 19$^{\text{ème}}$ siècles, Thomas R. Malthus s'est intéressé à la modélisation des fluctuations d'une population.

$$\begin{cases} \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} &= r\boldsymbol{x}(t)\,, \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0\,, \end{cases}$$

où $r = \textit{taux de naissance} \; - \; \textit{taux de mort}$ et la solution analytique est $\boldsymbol{x}(t) = \boldsymbol{x}_0 e^{rt}$. En l'absence de toute contrainte, il apparaît que l'abondance des populations croît de façon exponentielle ce qui amène à une solution réaliste.

Durant le 19$^{\text{ème}}$ siècle, Pierre François Verhulst s'est intéressé à un modèle dans lequel il suppose que l'abondance des espèces est limitée par une capacité de charge $K > 0$, qui correspond à une "taille maximale" :

$$\begin{cases} \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} &= r\boldsymbol{x}(t)\left(1 - \frac{\boldsymbol{x}(t)}{K}\right)\,, \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0\,, \end{cases}$$

Ce système peut permettre de modéliser, par exemple, la croissance d'une population de zèbres dans la savane limitée par la pénurie de ressources.

Par la suite, les scientifiques se sont davantage intéressés à la modélisation des interactions entre espèces. Historiquement, le premier modèle, encore d'actualité, est le modèle de Lotka-Volterra, du nom des deux scientifiques qui l'ont formulé indépendamment à la fin des années 20' [Lotka, 1925, Volterra, 1926].
Régulièrement présenté sous la forme d'un modèle proie-prédateur à 2 espèces, il a été comparé à des données issues de populations naturelles [Huffaker *et al.*, 1958].

D'un point de vue général et dans des dimensions supérieures, le système de Lotka-Volterra joue un rôle clé dans l'étude de la dynamique des populations dans le temps. Sa polyvalence et son analyse mathématique possible font de ce modèle une pierre angulaire dans le développement des modèles écologiques. Ce modèle a été étudié à la fois en écologie ([Wangersky, 1978, Jansen, 1987, Law et Blackford, 1992]) et en mathématiques ([Goh et Jennings, 1977, Goh, 1977, Taylor, 1988, Hofbauer et Sigmund, 1998, Takeuchi, 1996]).

Le système de Lotka-Volterra, composé de n équations différentielles, permet de décrire la dynamique de la population d'un système à n espèces :

$$\frac{\mathrm{d}x_k(t)}{\mathrm{d}t} = x_k(t)\left(r_k - \theta x_k(t) + \sum_{\ell \in [n]} M_{k\ell} x_\ell(t)\right), \tag{3}$$

où $k \in [n]$. L'abondance de l'espèce k au temps t est représentée par $x_k(t)$ et $\boldsymbol{x} = (x_1, \ldots, x_n)$ est le vecteur des abondances des différentes espèces. Le paramètre θ est le coefficient d'autorégulation ou l'interaction intraspécifique de chaque espèce. Le paramètre r_k correspond au taux de croissance intrinsèque de l'espèce k. Le coefficient $M_{k\ell}$ est l'effet de l'espèce ℓ sur l'espèce k.

La matrice M, appelée *matrice des interactions*, représentant la structure du réseau d'interactions, peut souvent être décomposée sous différentes formes : blocs, cascade, etc. [Cohen et Newman, 1988, Stouffer *et al.*, 2005, Lewinsohn *et al.*, 2006, Pocock *et al.*, 2012]. Ce réseau d'interactions est appelé réseau trophique lorsqu'il représente l'ensemble des interactions d'ordre alimentaire entre les êtres vivants d'un écosystème.

L'objet d'étude de nombreux mathématiciens et écologues consiste en la compréhension du comportement du système en lien avec ces différents paramètres. Ainsi, se retrouvent des questions liées au nombre d'équilibres, leur stabilité et leur faisabilité afin de saisir les implications écologiques qui en découlent.

Dans le cas des grands systèmes, la difficulté à observer, mesurer ou estimer les interactions pousse les chercheurs à remplacer les interactions par des variables aléatoires dont les propriétés statistiques (moyenne, variance, etc.) et la structure (blocs, cascade, etc.) codent certaines des véritables propriétés du réseau trophique. Il apparaît alors comme intéressant de choisir la matrice des interactions M comme étant une matrice aléatoire.

2.3 Le modèle de May : apparition des matrices aléatoires en écologie

Les mathématiques derrière le modèle de May

Inspiré par les travaux de [Gardner et Ashby, 1970], Robert May propose un premier modèle mathématique [May, 1972] qui lie la stabilité d'un écosystème à sa complexité.

Dans ce modèle, le vecteur des abondances $\boldsymbol{x}$ satisfait au système d'équations différentielles non linéaires de premier ordre de la forme :

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = F_i(\boldsymbol{x}). \tag{4}$$

May ne considère pas exactement le système d'équations différentielles présenté par (2).

Le principal intérêt de May concerne l'étude de la stabilité de l'équilibre de tels systèmes. L'existence de l'équilibre $\boldsymbol{x}^* = (x_i^*)_{i \in [n]}$ est supposée vérifiée et l'abondance de l'espèce i au temps t s'écrit $x_i(t) = x_i^* + \varepsilon_i(t)$. Au voisinage de l'équilibre, la stabilité du système (4) se résume à la stabilité du système linéaire

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}}{\mathrm{d}t} = \mathcal{J}(\boldsymbol{x}^*)\boldsymbol{\varepsilon},$$

où $J := \mathcal{J}(\boldsymbol{x}^*)$ est la matrice Jacobienne de taille $n \times n$ avec

$$J_{ij} := \frac{\partial F_i}{\partial x_j}(\boldsymbol{x}^*).$$

En particulier, l'équilibre est Lyapounov-diagonalement stable si et seulement si toutes les valeurs propres de J ont une partie réelle strictement négative. La Jacobienne représente ici la *matrice des communautés* qui décrit l'effet de l'espèce ℓ (colonne) sur l'espèce k (ligne) autour du point d'équilibre.

May, en modélisant la Jacobienne par une matrice aléatoire, fait ainsi appel aux résultats mathématiques de RMT dans son étude. Plus précisément, May choisit de remplacer le coefficient d'interaction interspécifique par -1 et tous les autres coefficients sont choisis indépendants les uns des autres, de sorte que :

$$J = -I + \Gamma \tag{5}$$

où $\Gamma_{ii} = 0$ et pour $i \neq j$, Γ_{ij} sont des variables aléatoires centres i.i.d de variance $V := \mathrm{Var}(\Gamma_{ij})$, et de loi indépendante de n.

Il s'agit alors de déterminer les conditions sur n et sur V pour s'assurer que toutes les valeurs propres ont partie réelle strictement négative. En se basant sur un résultat de [Ginibre, 1965], qui démontre qu'asymptotiquement en n, les valeurs propres de la matrice Γ sont contenues dans un disque centré en $(-1, 0)$ de rayon $\sqrt{nV}$. Cela amène May à établir la transition de phase suivante :

Proposition 8. [May, 1972]
Soit J la matrice donnée par (5), alors avec grande probabilité, l'équilibre est stable si

$$V < \frac{1}{n}$$

et instable avec grande probabilité si

$$V > \frac{1}{n}.$$

En réalité, le résultat de Ginibre n'est pas suffisant pour expliquer la transition de phase mais il est nécessaire de s'intéresser au rayon spectral de la matrice J.

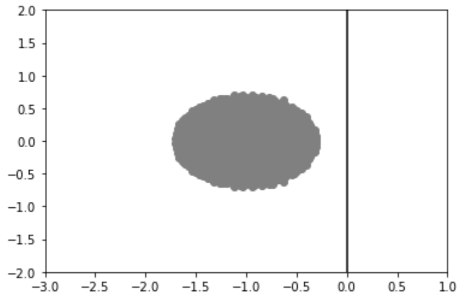
Or, la stabilité étant liée au nombre n d'espèces, pour obtenir rigoureusement le résultat de May, il suffit d'avoir une inégalité de concentration sur le rayon spectral telle que celle obtenue par Bordenave et coauteurs dans le théorème suivant :

Theorem 9. [Bordenave *et al.*, 2018]
Soit $X_n = (X_{ij})$ une matrice aléatoire de taille $n \times n$, à entrées i.i.d. complexes, avec $\mathbb{E}[|X_{11}|^2] \leq 1$. S'il existe $\epsilon > 0$ et $\beta > 0$ tels que $\mathbb{E}[|X_{11}|^{2+\epsilon}] \leq \beta$, alors, pour tout $\delta > 0$, il existe une constante $K := K(\epsilon, \delta, \beta) > 0$, tel que pour tout $n \in \mathbb{N}$:

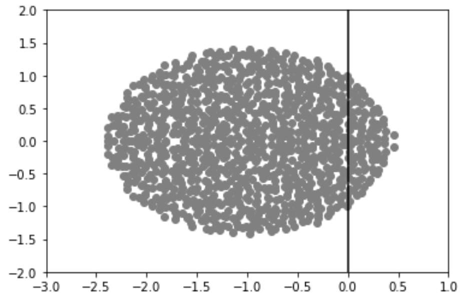
$$\mathbb{P}\left[\rho(X_n) \geq (1+\delta)\sqrt{n}\right] \leq \frac{K}{(\log n)^2}.$$

Cela signifie que dans ce premier modèle, pour n suffisamment grand, avec grande probabilité, il n'y aucune valeur propre de J qui soit en dehors du disque centré en -1 et de rayon $\sqrt{n(V+\delta)}$ pour n assez grand, pour tout $\delta > 0$.

Le spectre de la Jacobienne est illustré par la figure 1.



(a) Stabilité du spectre.



(b) Certaines valeurs propres ont partie réelle positive.

FIGURE 1 – Spectre de la matrice Jacobienne $J = \Gamma - I$, pour $n = 1000$ espèces avec $C = 1$, les entrées Γ_{ij}, pour $i \neq j$ sont des variables aléatoires centrées indépendantes et de variance V, avec $V = \frac{1}{2n} < \frac{1}{n}$ pour (a) et $V = \frac{2}{n} > \frac{1}{n}$ pour (b).

Remark 4. Il y a une légère différence entre le modèle de [Bordenave *et al.*, 2018] et le modèle (5) : pour Bordenave et ses coauteurs, la symétrie des lois des entrées est requise et toutes les entrées de la matrice, y compris les entrées diagonales, sont des variables aléatoires i.i.d tandis que pour le modèle (5), les entrées diagonales sont toutes nulles.

May considère déjà une version "sparse" dans son modèle initial. Chaque interaction a une probabilité C, appelée *connectance*, d'être effective (et une probabilité $1 - C$ d'être égale à 0), indépendamment des autres interactions. Les chapitres 1 et 4 (section 1) permettront d'approfondir cette notion de sparsité.
Dans le modèle de May, cela signifie qu'en moyenne, chaque espèce interagit avec une proportion C de l'ensemble des autres espèces.

La transition de phase s'écrit alors :

Proposition 10. [May, 1972]
Soit C la connectance du modèle, l'équilibre est stable avec grande probabilité si

$$CV < \frac{1}{n}$$

et instable avec grande probabilité si

$$CV > \frac{1}{n}.$$

Mathematiquement, il est possible de faire appel au formalisme suivant. Soit Δ_{ER} la matrice d'adjacence d'un graphe d'Erdös-Rényi (cf chapitre 4 section 2 pour les détails). Chaque entrée de Δ_{ER} est égale à 1 avec probabilité C et à 0 sinon. En d'autres termes, chaque espèce a une probabilité C d'avoir un effet sur une autre espèce. Alors, la matrice Γ introduite par (5) est remplacées par $\widetilde{\Gamma}$ définie par :

$$\widetilde{\Gamma} := \Delta_{\text{ER}} \circ \Gamma = \left([\Delta_{\text{ER}}]_{ij}\Gamma_{ij}\right) .$$

Le paramètre C peut dès lors s'interpréter comme le nombre moyen de voisins pour un sommet donné du graphe.

Dans ce nouveau modèle, la matrice $\widetilde{\Gamma}$ a la même loi que Γ à l'exception du paramètre de variance $\mathrm{Var}(\widetilde{\Gamma}_{ij}) = CV$, pour $i \neq j$, ce qui amène à la proposition 10.

Relation entre le modèle de May et le système de Lotka-Volterra

Supposons qu'il existe un équilibre $\boldsymbol{x}^*$ à l'équation (3) (avec $\theta = 1$). Si $\boldsymbol{x}^* > 0$, il est possible de déterminer la matrice Jacobienne à l'équilibre.
Soit $x_i(t) = x_i^* + \varepsilon_i(t)$. A l'équilibre, puisque $x_i^* > 0$, il vient que

$$r_i = x_i^* - (M\boldsymbol{x}^*)_i \,.$$

Ainsi

$$\begin{aligned}\frac{\mathrm{d}x_i}{\mathrm{d}t} &= x_i(r_i - x_i + (M\boldsymbol{x})_i)\,,\\ &= (x_i^* + \varepsilon_i)(r_i - (x_i^* + \varepsilon_i) + (M(\boldsymbol{x}^* + \boldsymbol{\varepsilon}))_i)\,,\\ &= (x_i^* + \varepsilon_i)(x_i^* - (M\boldsymbol{x}^*)_i - (x_i^* + \varepsilon_i) + (M(\boldsymbol{x}^* + \boldsymbol{\varepsilon}))_i)\,,\\ &= (x_i^* + \varepsilon_i)(-\varepsilon_i + (M\boldsymbol{\varepsilon})_i)\,,\\ &= [\mathrm{diag}(\boldsymbol{x}^*)(-I + M)\boldsymbol{\varepsilon}]_i + o(\boldsymbol{\varepsilon})\,.\end{aligned}$$

et la Jacobienne vaut alors

$$J = \mathcal{J}(\boldsymbol{x}^*) = \mathrm{diag}(\boldsymbol{x}^*)(-I + M)\,.$$

Bien que cette Jacobienne ressemble à celle obtenue par May, il subsiste quelques différences importantes. En effet, dans la formule de la Jacobienne $\mathrm{diag}(\boldsymbol{x}^*)(-I+M)$ apparait le produit entre la matrice diagonale $\mathrm{diag}(\boldsymbol{x}^*)$, et la matrice $-I + M$.
Si M est aléatoire, alors $\mathrm{diag}(\boldsymbol{x}^*)$ et M sont dépendantes puisque $\boldsymbol{x}^* = \boldsymbol{r} + M\boldsymbol{x}^*$ pour un équilibre faisable.

[Stone, 2018] considère une matrice Jacobienne de cette forme tout en simplifiant les hypothèses en considérant que les entrées de $\mathrm{diag}(\boldsymbol{x}^*)$ sont indépendantes de M. Ainsi J est instable si et seulement si M est instable, et le critère de May reste vrai.

Cette hypothèse d'indépendance entre $\mathrm{diag}(\boldsymbol{x}^*)$ et M reste cependant forte puisque le sous-ensemble des espèces survivantes et leur abondance sont une fonction de la matrice des interactions M.

À l'aide de méthodes heuristiques issues de la physique statistique [Bunin, 2017], il a été démontré que le critère de May reste vrai en prenant en compte les corrélations et en remplaçant le nombre total n d'espèces par le nombre d'espèces survivantes $(\mathrm{card}\{i; x_i^* > 0\})$.

Par ailleurs, [Stone, 2018] ne considère pas le régime de normalisation pour lequel l'équilibre $\boldsymbol{x}^*$ est faisable [Bizeul et Najim, 2021]. Il n'est donc pas tout à fait clair que la formule de la Jacobienne est initialement associée à un système de Lotka-Volterra.

Autres modèles

Certains auteurs ont fait appel à d'autres modèles pour la Jacobienne, tel que la matrice elliptique (cf [Tang *et al.*, 2014]) et obtiennent des critères de stabilité similaires concernant les paramètres du modèle.

Il est également possible de faire appel à d'autres modèles aléatoires, comme ceux étudiés par [Ben Arous *et al.*, 2021, Fyodorov et Khoruzhenko, 2016] :

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -x_i(t) + f_i(\boldsymbol{x}) \,,$$

où $f_i(\boldsymbol{x})$ est une fonction lisse et modélise la complexité et la non-linéarité des interactions. Autrement dit, la fonction f_i est spécifique à l'espèce i.

2.4 Nature des interactions

Afin de simplifier l'étude du système (3), il est possible de rendre les paramètres r et θ déterministes. ($r_k = 1$, $\forall k \in [n]$ et $\theta = 1$).

Les taux de croissance égaux permettent de réduire le nombre de paramètres et de simplifier grandement les calculs. La compréhension de l'impact de la matrice d'interaction M dans le système de Lotka-Volterra ouvre le champ à de nombreux problèmes ouverts.

Dans le cas $\theta \neq 1$, le système peut être redimensionner pour supprimer le paramètre θ en posant $\tilde{x}_k := \theta x_k$, $\tilde{M}_{k\ell} := M_{k\ell}/\theta$. [Barabás *et al.*, 2017] ont étudié l'importance d'avoir un terme d'autorégulation fort pour que le système ait un équilibre stable.

Les valeurs de la diagonale de la matrice M ne sont pas fixées à 0, mais leurs valeurs microscopiques ont un impact négligeable sur les résultats asymptotiques.

Le système de Lotka-Volterra (3) se réécrit alors

$$\frac{\mathrm{d}x_k(t)}{\mathrm{d}t} = x_k(t)\left(1 - x_k(t) + \sum_{\ell\in[n]} M_{k\ell}x_\ell(t)\right) \,, \qquad \forall k \in [n] \,, \tag{6}$$

autrement dit,

$$\frac{\mathrm{d}x_k}{\mathrm{d}t} = x_k(1 - x_k + (M\boldsymbol{x})_k) \,, \quad \forall k \in [n] \,, \tag{7}$$

où M reste à déterminer.

La compréhension des points d'équilibre du système de Lotka-Volterra (7) et leur stabilité permet de mieux comprendre l'impact du réseau d'interactions, représenté par la matrice d'interaction M, sur l'abondance des espèces. En particulier, le réseau d'interactions a un impact sur la persistance des espèces qui le composent (i.e. le nombre d'espèces survivantes), la faisabilité du système (c'est-à-dire l'existence d'un équilibre pour lequel toutes les espèces survivent) et sa stabilité.

L'enjeu majeur de cette thèse est de comprendre l'impact de la matrice d'interaction M sur la dynamique du modèle Lotka-Volterra.
La matrice M correspond au réseau d'interactions entre les espèces. Les (très) grands réseaux écologiques fortement connectés sont fréquents dans la nature [Dunne *et al.*, 2002, Pimm *et al.*, 1991].

Remark 5. Dans le système de Lotka-Volterra (3), lorsque M est une matrice de Wigner, les interactions entre espèces sont considérées comme étant symétriques :

l'effet d'une espèce k sur une espèce ℓ est identique à l'effet de l'espèce ℓ sur l'espèce k.

Pour les matrices non-hermitiennes, toutes les interactions sont indépendantes.

Cependant, en écologie, les effets réciproques d'une espèce k sur une autre espèce ℓ, ($M_{k\ell}$ et $M_{\ell k}$) sont liés. Mathématiquement, il s'agit d'une corrélation par paire entre les entrées de la matrice. Celle-ci peut être utilisée pour décrire des processus biologiques comme la prédation lorsque le signe des interactions est inversé, i.e que la corrélation est négative.
En RMT, lorsque les interactions par paire sont tirées d'une distribution bivariée, cela permet d'obtenir le modèle elliptique.

Ainsi, la matrice d'interaction M peut être définie par une matrice aléatoire non centrée avec des interactions corrélées deux à deux combinée à une structure de graphe :

$$M = \Delta \circ \left(\frac{A}{\alpha\sqrt{n}} + \frac{\mu}{n}\mathbf{1}_n\mathbf{1}_n^T \right) , \tag{8}$$

où $\circ$ est le produit d'Hadamard, c'est-à-dire $(X \circ Y)_{ij} := X_{ij}Y_{ij}$ et $\mathbf{1}_n$ est le vecteur $(1, \dots, 1)$ de taille $n \times 1$.

La matrice aléatoire $A = (A_{k\ell})_{k,\ell \in [n]}$ satisfait les conditions suivantes :

1. $(A_{k\ell},\ k \leq \ell)$ sont des variables aléatoires indépendantes identiquement distribuées (i.i.d.) et $\mathbb{E}(A_{k\ell}) = 0$, $\mathbb{E}(|A_{k\ell}|^2) = 1$ et $\mathbb{E}(|A_{k\ell}|^4) < \infty$, $\forall\ 1 \leq k \leq \ell$.
2. pour $k < \ell$, le vecteur $(A_{k\ell}, A_{\ell k})$ a une distribution standard bivariée, indépendante des autres variables aléatoires, avec une covariance $\mathrm{cov}(A_{k\ell}, A_{\ell k}) = \mathbb{E}(A_{k\ell}A_{\ell k}) = \rho$ avec $|\rho| \leq 1$.

$\Delta := (\Delta_{k\ell})_{k,\ell \in [n]}$ est la matrice d'adjacence d'un graphe orienté. Chaque espèce est représentée par un sommet et une interaction entre deux espèces par un arête orientée. Ainsi,

$$S_{k\ell} = \begin{cases} 1 & \text{s'il existe un effet de l'espèce } \ell \text{ sur l'espèce } k\,, \\ 0 & \text{sinon}\,. \end{cases}$$

La nature des interactions joue un rôle prépondérant sur les caractéristiques d'un écosystème, et en particulier son (ou ses) équilibre(s). Par ailleurs, le type d'interaction varie en fonction des espèces. Ces paramètres sont gérés par des choix sur les propriétés statistiques des variables aléatoires M_{ij}. Le triplet de paramètres (α, μ, ρ) permet de représenter tout un panel d'interactions.

Le paramètre α permet de déterminer la force des interactions. Une grande valeur de α représente un système avec des interactions faibles tandis qu'une valeur petite pour α permettra d'obtenir des interactions très fortes.

Les paramètres μ et ρ décrivent quant à eux la nature des interactions du système.
Lorsque $\rho < 0$, les interactions entre deux espèces partenaires ont un impact opposé l'une sur l'autre, comme dans les interactions proies-prédateurs (le prédateur a un impact négatif sur l'abondance de sa proie tandis que la présence de la proie affecte positivement celle du prédateur). Au contraire, si $\rho > 0$, les interactions entre partenaires ont un impact similaire les uns sur les autres (interactions mutualistes ou compétitives).

En fonction de son signe, le paramètre d'interaction μ permettra d'augmenter la proportion d'interactions compétitives ou mutualistes.

Étant donnée une interaction par paire $M_{k\ell}/M_{\ell k}$ dans le système, les trois types d'interactions prépondérants sont :

— compétition (relation -/-), qui se produit plus souvent lorsque $\rho > 0$, $\mu < 0$,
— mutualisme (relation +/+), qui se produit plus souvent lorsque $\rho > 0$, $\mu > 0$,
— prédation (relation +/-), qui se produit plus souvent lorsque $\rho < 0$, $\mu \approx 0$ ([Allesina et Tang, 2012]).

D'autres types d'interactions existent tels que le commensalisme ou l'amensalisme [Begon et Townsend, 2021].

2.5 Condition de non-invasibilité

Une caractéristique essentielle pour comprendre la dynamique du système LV (7) est l'existence d'un équilibre $\boldsymbol{x}^* = (x_k^*)_{k\in[n]}$ tel que :

$$\begin{cases} x_k^*(1 - x_k^* + (M\boldsymbol{x}^*)_k) = 0\,, \ \forall k \in [n]\,, \\ x_k^* \geq 0\,. \end{cases} \tag{9}$$

Parmi les premières questions se posent celles de l'existence puis de l'unicité de l'équilibre.
Dans le cas où l'équilibre existe et est unique, se pose la question de la convergence vers cet équilibre, autrement dit la convergence d'une solution $\boldsymbol{x}$ vers l'équilibre $\boldsymbol{x}^*$: $\boldsymbol{x}(t) \xrightarrow[t\to\infty]{} \boldsymbol{x}^*$ si $\boldsymbol{x}(0)$ est suffisamment proche de $\boldsymbol{x}^*$.
Enfin, la description de la stabilité se pose : locale, globale, résilience (i.e. capacité d'un système à retrouver sa structure initiale suite à une perturbation), etc.

L'ensemble $(\mathbb{R}_+^*)^n$ est invariant pour le système de Lotka-Volterra, c'est-à-dire que $\boldsymbol{x}(0) > 0$ (composante par composante) implique $\boldsymbol{x}(t) > 0$ pour tout $t > 0$. Cependant, certaines de ces composantes $x_k(t)$ peuvent converger vers zéro si l'équilibre $\boldsymbol{x}^*$ a des composantes nulles.

Lorsque $\boldsymbol{x}^*$, solution du système (9), a des composantes nulles, une approche naïve et immédiate pour résoudre ce problème est de choisir un sous-ensemble $\mathcal{I} \subset [n]$, définir les composantes correspondantes $\boldsymbol{x}_{\mathcal{I}} = (x_i^*)_{i\in\mathcal{I}}$ à zéro, et de résoudre le système linéaire restant :

$$\boldsymbol{x}_{\mathcal{I}^c} = \mathbf{1}_{|\mathcal{I}^c|} + M_{\mathcal{I}^c}\boldsymbol{x}_{\mathcal{I}^c}\,.$$

S'il existe $\boldsymbol{x}_{\mathcal{I}^c} \geq 0$ qui résout l'équation précédente, alors $\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_{\mathcal{I}} \\ \boldsymbol{x}_{\mathcal{I}^c} \end{pmatrix}$ satisfait (9) et est un équilibre. Le nombre de sous-ensemble $\mathcal{I} \subset [n]$ est 2^n mais surtout croît de manière exponentielle lorsque $n \to \infty$.

Les équations d'équilibre deviennent mal posées car il peut y avoir plusieurs équilibres. Une condition connue en écologie pour les systèmes dynamiques est la condition de non-invasibilité associée à l'équilibre saturé [Law et Morton, 1996, Jansen et Sigmund, 1998]. Un équilibre est saturé s'il est résistant à l'invasion d'une espèce initialement absente. L'étude des équilibres saturés est un sujet de recherche important dans le domaine des systèmes dynamiques [Hofbauer et Sigmund, 1998].

Definition 2.1. *(Équilibre saturé).*
Soit $\mathcal{I}^c$ l'ensemble des espèces persistantes.
Un équilibre $\boldsymbol{x}^*$ est dit saturé si $\forall k \in \mathcal{I}\,;\, 1 - x_k^* + (M\boldsymbol{x}^*)_k \leq 0$.

Lemma 11. S'il existe une solution strictement positive $\boldsymbol{x}(t) > 0$, telle que $\boldsymbol{x}(t) \xrightarrow[t\to\infty]{} \boldsymbol{x}^*$, alors $\boldsymbol{x}^*$ est un équilibre saturé.

En se basant sur [Takeuchi, 1996, Thm 3.2.5], une condition nécessaire pour que l'équilibre $\boldsymbol{x}^*$ soit stable est que pour tout $k \in [n]$.

$$1 - x_k^* + (M\boldsymbol{x}^*)_k \leq 0\,. \tag{10}$$

Ainsi, l'ensemble des conditions suivantes donne un cadre à la recherche de l'équilibre, pour tout $k \in [n]$,

$$\begin{cases} x_k^*\,(1 - x_k^* + (M\boldsymbol{x}^*)_k) & = & 0\,, \\ 1 - x_k^* + (M\boldsymbol{x}^*)_k & \leq & 0\,, \\ x_k^* & \geq & 0\,. \end{cases} \tag{11}$$

Le problème de la recherche d'un équilibre positif entre ainsi dans la classe des problèmes de complémentarité linéaire (LCP), décrit dans la dernière section de cette introduction.

3 Cadre théorique et organisation de la thèse

Comme indiqué par l'équation (7), le système de Lotka-Volterra étudié s'écrit :

$$\frac{\mathrm{d}x_k}{\mathrm{d}t} = x_k(1 - x_k + (M\boldsymbol{x})_k)\ , \quad k \in [n]\,.$$

A la différence du modèle standard de Lotka-Volterra, la présence du paramètre de normalisation $1/\sqrt{n}$ dans la définition (8) de la matrice d'interaction M est due à la raison théorique suivante : limiter l'impact des paramètres d'interaction sur les autres termes tout en gardant une influence asymptotique c'est-à-dire

$$\mathbb{E}\left(\sum_{\ell\in[n]} M_{k\ell}x_\ell(t)\right) \sim O(1) \qquad \text{et} \qquad \|M_n\| \sim O_{\mathbb{P}}(1)\,.$$

D'un point de vue écologique, il semble naturel de considérer que lorsque le nombre d'espèces augmente dans un écosystème et que chaque espèce interagit avec un grand nombre d'autres espèces, l'impact d'une espèce sur l'autre aura tendance à diminuer.

3.1 Faisabilité

Dans les années 70', [Goh et Jennings, 1977] s'intéresse à la question de la faisabilité d'un équilibre $\boldsymbol{x}^* > 0$ pour le modèle Lotka-Volterra. Par la suite, [Logofet, 2018] s'est penché sur ce problème dans le cas d'un système compétitif tandis que [Rossberg, 2013] a également étudié le nombre moyen d'espèces pouvant coexister dans ces communautés compétitives. Plus récemment, [Grilli *et al.*, 2017] ont étudié l'impact des propriétés du réseau d'interactions sur le taux de croissance pour maintenir un équilibre faisable en utilisant les méthodes de stabilité structurelle introduites par [Rohr *et al.*, 2014].

En reprenant le modèle des matrices d'interactions (8), [Bizeul et Najim, 2021] ont étudié mathématiquement le cas d'une matrice d'interaction pleine définie par

$$M_n = \frac{A_n}{\alpha_n \sqrt{n}}\,, \tag{12}$$

où $A_n := (A_{k\ell})_{k,\ell \in [n]}$ est à entrées i.i.d de loi $\mathcal{N}(0,1)$.

En se basant sur les travaux de [Geman et Hwang, 1982], [Dougoud *et al.*, 2018] ont montré que lorsque la matrice d'interaction est aléatoire, si $\alpha_n > 0$ est fixé et indépendant de n alors nécessairement certaines espèces s'éteignent.
[Bizeul et Najim, 2021] obtiennent un seuil de faisabilité de l'équilibre lorsque $\alpha_n^* = \sqrt{2\log(n)}$ et montrent que la faisabilité implique la stabilité. Ce type de résultat avait déjà été observé par [Stone, 2016] ; le seuil de stabilité est d'ailleurs franchi avant le seuil de faisabilité (cf [Clenet *et al.*, 2022b]).

En partant de l'équation (9), si $\boldsymbol{x}^* > 0$, l'ensemble des équations d'équilibre devient une équation linéaire :

$$\boldsymbol{x}^* = \mathbf{1} + M_n \boldsymbol{x}^*\,. \tag{13}$$

Theorem 12. [Bizeul et Najim, 2021]
Soit $\alpha_n \xrightarrow[n\to\infty]{} \infty$ et $\alpha_n^* = \sqrt{2\log(n)}$. Soit $\boldsymbol{x}^* = (x_k^*)_{k\in[n]}$ solution de (13) avec (12) et $A_n = (A_{k\ell})_{k,\ell\in[n]}$ une matrice aléatoire à entrées i.i.d de loi $\mathcal{N}(0,1)$.

1. S'il existe $\epsilon > 0$ tel que $\alpha_n \leq (1-\epsilon)\alpha_n^*$ alors
$$\mathbb{P}\left(\min_{k\in[n]} x_k^* > 0\right) \xrightarrow[n\to\infty]{} 0\,.$$
2. S'il existe $\epsilon > 0$ tel que $\alpha_n \geq (1+\epsilon)\alpha_n^*$ alors
$$\mathbb{P}\left(\min_{k\in[n]} x_k^* > 0\right) \xrightarrow[n\to\infty]{} 1\,.$$

Dans le chapitre 1, il sera de nouveau question de l'existence et de l'unicité d'un équilibre faisable pour le système d'équations de Lotka-Volterra (3). Cette fois-ci, la matrice d'interaction sera définie par

$$M_n = \Delta_n \circ \frac{A_n}{\alpha_n \sqrt{d}}\,,$$

où Δ_n est la matrice d'adjacence d'un graphe d-régulier et $A_n := (A_{k\ell})_{k,\ell\in[n]}$ est à entrées i.i.d de loi $\mathcal{N}(0,1)$.
Plus précisément, on considère un graphe orienté d-régulier à n sommets, c'est-à-dire un graphe tel que chaque sommet i a exactement d arêtes entrantes et d arêtes

sortantes (une boucle étant comptée à la fois dans les arêtes sortantes et dans les arêtes entrantes). Δ_n est définie comme étant la matrice d'adjacence d'un tel graphe, Δ_n a donc d composantes non nulles par ligne et par colonne, soit dn entrées non nulles au total, ces entrées valant toutes 1.

L'évolution du paramètre de normalisation $\sqrt{n}$ en $\sqrt{d}$ sera explicitée par la suite. Dans ce modèle, chaque espèce interagit alors avec d autres espèces et le rapport d/n reflète la connectance du réseau d'interactions.
Deux cas distincts seront étudiés, d'une part lorsque d est proportionnel à n et d'autre part, lorsque $d \geq \log(n)$ et lorsque Δ_n a une structure en blocs.
Dans ce chapitre, il sera aussi question de la stabilité globale de l'équilibre.

3.2 Problème de complémentarité linéaire (LCP)

Le LCP est une classe de problèmes issue de l'optimisation mathématique qui englobe notamment les problèmes de programmation linéaire et quadratique [Murty et Yu, 1988,Cottle *et al.*, 2009].

Soit Y une matrice de taille $n \times n$ et $\boldsymbol{q}$ un vecteur de taille $n \times 1$, le problème de complémentarité linéaire associé, noté $\mathrm{LCP}(Y, \boldsymbol{q})$ consiste à trouver deux vecteurs $\boldsymbol{z}$, $\boldsymbol{w}$ de taille $n \times 1$ satisfaisant les conditions suivantes :

$$\begin{cases} \boldsymbol{z} & \geq \ 0 \quad \text{i.e.} \quad \forall k \in [n], \ z_k \geq 0\,, \\ \boldsymbol{w} = Y\boldsymbol{z} + \boldsymbol{q} & \geq \ 0\,, \\ \boldsymbol{w}^T \boldsymbol{z} & = \ 0\,. \end{cases} \tag{14}$$

Etant donné que $\boldsymbol{w}$ s'obtient directement de $\boldsymbol{z}$, Y et q, on peut noter $\boldsymbol{z} \in \mathrm{LCP}(Y, \boldsymbol{q})$ si $(\boldsymbol{w}, \boldsymbol{z})$ est une solution de (14).

L'étude du LCP remonte aux travaux de [Lemke, 1965] et [Dantzig et Cottle, 1968]. [Lemke et Howson, 1964] ont développé un algorithme basé sur des étapes du pivot pour résoudre le problème (14).

Introduite par [Fiedler et Pták, 1966], la classe des P-matrices est reliée au problème de complémentarité linéaire. En particulier, [Murty, 1972] montre que si Y est une P-matrice alors il existe une unique solution au LCP.

Definition 3.1. Une matrice carrée Y est une *P-matrice* si tous ses mineurs principaux (sous-déterminants) sont strictement positifs, c'est-à-dire

$$\det(Y_{\mathcal{I}}) > 0\,, \quad \forall \mathcal{I} \subset [n] \quad \text{où} \quad Y_{\mathcal{I}} = (Y_{k\ell})_{k,\ell \in \mathcal{I}}\,.$$

De nombreuses propriétés sur les conditions nécessaires et suffisantes pour qu'une matrice réelle soit une P-matrice ont été étudiées par [Rump, 2003] et [Rohn, 2012].

Theorem 13. *(Existence et unicité d'une solution au problème LCP - [Murty, 1972]).*
Une matrice Y est une P-matrice si et seulement si le $\mathrm{LCP}(Y, \boldsymbol{q})$ a une unique solution $(\boldsymbol{w}, \boldsymbol{z})$ pour tout $\boldsymbol{q} \in \mathbb{R}^n$.

Remark 6. En considérant (11), rechercher la solution du système de Lotka-Volterra (6) se ramène à trouver la solution $\boldsymbol{x}^* \in \mathrm{LCP}(I - M, -\mathbf{1})$.

Lemma 14. Soit Y une matrice symétrique de taille $n \times n$.
Y est une $\mathcal{P}-$matrice si et seulement si toutes ses valeurs propres sont strictement positives, i.e. Y est définie positive.

Proposition 15. *[Takeuchi* et al.*, 1978].*
Si Y est Lyapunov diagonalement stable, alors $-Y$ est une P-matrice.

Takeuchi et Adachi (voir par exemple [Takeuchi, 1996, Thm 3.2.1]) fournissent un critère pour l'existence d'un équilibre unique $\boldsymbol{x}^*$ et la stabilité globale du système LV (3).

Theorem 16. *[Takeuchi et Adachi, 1980].*
Si $-\theta I + M$ est Lyapunov diagonalement stable, alors $\mathrm{LCP}(\theta I - M, \boldsymbol{r})$ admet une unique solution.
De plus, pour tout $\boldsymbol{r} \in \mathbb{R}^n$, il y a un unique équilibre $\boldsymbol{x}^*$ à (3), qui est globalement stable dans le sens où pour chaque $\boldsymbol{x}_0 > 0$, la solution de (3) qui démarre à $\boldsymbol{x}(0) = \boldsymbol{x}_0$ satisfait

$$\boldsymbol{x}(t) \xrightarrow[t\to\infty]{} \boldsymbol{x}^* .$$

L'étude du $\mathrm{LCP}(I - M, -r)$ fera l'objet du chapitre 2 dans le cas d'une matrice M issue du GOE ou lorsque M est une matrice de Wishart. L'originalité du chapitre réside dans l'utilisation de l'algorithme dit de l'AMP (Approximate Message Passing).
Ce LCP est aussi à l'origine du théorème limite central étudié dans le chapitre 3.

3.3 Structure des interactions

La structure du réseau d'interactions diffère selon les écosystèmes. En effet, la structure de la matrice d'interaction, définie par (8), peut être affectée par l'existence de communautés, c'est-à-dire de groupes d'espèces qui interagissent préférentiellement entre elles [Thébault et Fontaine, 2010, Allesina *et al.*, 2015]. Dans le système (6), le réseau est représenté par Δ, matrice d'adjacence d'un graphe donné. Plusieurs types de structures, étudiés en écologie et pouvant être modélisés par le graphe Δ, feront l'objet du chapitre 4.

Chapitre 1

Feasibility of sparse large Lotka-Volterra ecosystems

Ce chapitre est tiré de l'article [Akjouj et Najim, 2022], écrit en collaboration avec J. Najim et publié dans *Journal of Mathematical Biology.*

Consider a large ecosystem (foodweb) with n species, where the abundances follow a Lotka-Volterra system of coupled differential equations. We assume that each species interacts with $d = d_n$ other species and that their interaction coefficients are independent random variables. This parameter d reflects the connectance of the foodweb and the sparsity of its interactions especially if d is much smaller that n.

We address the question of feasibility of the foodweb, that is the existence of an equilibrium solution of the Lotka-Volterra system with no vanishing species. We establish that for a given range of d, namely $d \propto n$ or $d \geq \log(n)$ with an extra condition on the sparsity structure, there exists an explicit threshold depending on n and d and reflecting the strength of the interactions, which guarantees the existence of a positive equilibrium as the number of species n gets large.

From a mathematical point of view, the study of feasibility is equivalent to the existence of a positive solution $\boldsymbol{x}_n$ (componentwise) to the equilibrium linear equation :

$$\boldsymbol{x}_n = \mathbf{1}_n + M_n \boldsymbol{x}_n \,,$$

where $\mathbf{1}_n$ is the $n \times 1$ vector with components 1 and M_n is a large sparse random matrix, accounting for the interactions between species. The analysis of such positive solutions essentially relies on large random matrix theory for sparse matrices and gaussian concentration of measure. The stability of the equilibrium is established.

The results in this article extend to a sparse setting the results obtained by [Bizeul et Najim, 2021].

1 Introduction

Lotka-Volterra system of coupled differential equations.

Large Lotka-Volterra (LV) systems are widely used in mathematical biology and ecology to model populations with interactions [Gopalsamy, 1984, Hofbauer et Sigmund, 1998, Kiss et Kovács, 2008]. For a given foodweb, denote by $\boldsymbol{x}_n^t = (x_k(t))_{k\in[n]}$

the vector of abundances of the various species at time $t \geq 0$. In a LV system, the abundances are connected via the following coupled equations :

$$\frac{dx_k(t)}{dt} = x_k(t) \left(r_k - x_k(t) + \sum_{\ell \in [n]} M_{k\ell} x_\ell(t) \right) \qquad \text{for} \quad k \in [n]\,,$$

where $M_n = (M_{k\ell})$ stands for the interaction matrix, and r_k for the intrinsic growth of species k. At the equilibrium $\frac{d\boldsymbol{x}_n}{dt} = 0$, the abundance vector $\boldsymbol{x}_n = (x_k)_{k\in[n]}$ is solution of the system :

$$x_k \left(r_k - x_k + \sum_{\ell \in [n]} M_{k\ell} x_\ell \right) = 0 \qquad \text{for} \quad x_k \geq 0 \quad \text{and} \quad k \in [n]\,. \tag{1.1}$$

An important question, which motivated recent developments [Dougoud *et al.*, 2018, Bizeul et Najim, 2021], is the existence of a *feasible* solution $\boldsymbol{x}_n$ to (1.1), that is a solution where all the x_k's are positive, corresponding to a scenario where no species disappears. Notice that in this latter case, the system (1.1) takes the much simpler form :

$$\boldsymbol{x}_n = \boldsymbol{r}_n + M_n \boldsymbol{x}_n\,,$$

where $\boldsymbol{r}_n = (r_k)$.

Aside from the question of feasibility arises the question of *stability* : for a complex system, how likely a perturbation of the solution $\boldsymbol{x}_n$ at equilibrium will return to the equilibrium ? [Gardner et Ashby, 1970] considered stability issues of complex systems connected at random. Based on the circular law for large random matrices with i.i.d. entries, [May, 1972] provided a complexity/stability criterion and motivated the systematic use of large random matrix theory in the study of foodwebs, see for instance [Allesina et Tang, 2015]. Recently, [Stone, 2018] and [Gibbs *et al.*, 2018] revisited the relation between feasibility and stability.

In the spirit of May[1] and in the absence of any prior information, we shall model the interactions of matrix M_n as random and in order to simplify the analysis, we will consider intrinsic growths $(r_i)_{i\in[n}$ equal to 1, and the equations under study will take the following form in the sequel :

$$\frac{dx_k(t)}{dt} = x_k(t) \left(1 - x_k(t) + \sum_{\ell \in [n]} M_{k\ell} x_\ell(t) \right) \qquad \text{for} \quad k \in [n]\,. \tag{1.2}$$

Sparse foodwebs

One of the most important parameters of the complexity of an ecosystem is its connectance, which is the proportion of interactions between species (see for instance [Pimm, 1984]). This corresponds to the proportion of non-zero entries in the interaction matrix M_n. May's complexity/stability criterion asserts that the instability of an ecosystem increases with the connectance (i.e. the less sparse M_n is, the

1. Beware that May did not consider LV systems but rather used a random matrix model for the jacobian at equilibrium of a generic system of coupled differential equations.

more unstable is the ecosystem equilibrium). More specifically, [Grilli *et al.*, 2017] specifies that the effect of the sparsity depends on the nature of the interactions (random, predator-prey, mutualistic or competitive). In the case of random interactions, [Allesina et Tang, 2012] supports the idea that sparse ecosystems lead to a stable equilibrium. Based on ecological and biological data (see for instance [Dunne *et al.*, 2002]), recent studies [Busiello *et al.*, 2017] suggest that foodwebs can actually be very sparse. In a recent theoretical study, [Marcus *et al.*, 2022] study the properties of sparse ecological communities in relation with the strength of interactions.

To encode this sparsity in a simple parametric way, we first consider a directed d_n-regular graph with n vertices and its associated $n \times n$ adjacency matrix $\Delta_n = (\Delta_{ij})$:

$$\Delta_{ij} = \begin{cases} 1 & \text{if there is an edge pointing from } i \text{ to } j\,, \\ 0 & \text{otherwise.} \end{cases}$$

In the considered graph, each vertex i has d_n edges pointing from a vertex $k \in [n]$ to i, and has d_n other edges pointing from i to a vertex $\ell \in [n]$. An edge pointing from i to i is called a loop. In particular, matrix Δ_n is deterministic, has exactly d_n non-null entries per row and per column, and $n \times d_n$ non-null entries overall.

Denote by A_n a $n \times n$ matrix with independent Gaussian $\mathcal{N}(0,1)$ entries and consider the Hadamard product matrix $\Delta_n \circ A_n = (\Delta_{ij} A_{ij})$. Let $\alpha = \alpha_n$ be a positive sequence. We assume that matrix M_n has the following form

$$M_n = \frac{\Delta_n \circ A_n}{\alpha_n \sqrt{d_n}}\,. \tag{1.3}$$

Let us comment on the normalizing factor $1/(\alpha_n \sqrt{d_n})$. Theoretical results on sparse large random matrices [Bandeira et van Handel, 2016] assert that asymptotically

$$\left\| \frac{\Delta_n \circ A_n}{\sqrt{d_n}} \right\| = \mathcal{O}(1)\,, \quad (n \to \infty)$$

where $\|\cdot\|$ stands for the spectral norm, if the degree d_n of the graph satisfies $d_n \geq \log(n)$, a condition that we will assume in the remaining of the article. In particular, normalization $1/\sqrt{d_n}$ garantees that matrix $\Delta_n \circ A_n/\sqrt{d_n}$ has a macroscopic effect in the LV system, even for large foodwebs (large n).

The extra normalization $1/\alpha_n$ is to be tuned to get a feasible solution.

Denote by $\mathbf{1}_n$ the $n \times 1$ vector of ones and by A^{T} the transpose of matrix A. In the full matrix case $\Delta_n = \mathbf{1}_n \mathbf{1}_n^{\mathrm{T}}$, [Dougoud *et al.*, 2018], based on [Geman et Hwang, 1982], proved that a feasible solution is very unlikely to exist if $\alpha_n \equiv \alpha$ is a constant. We thus consider the regime where $\alpha_n \to \infty$ and will prove that there is a sharp threshold $\alpha_n \sim \sqrt{2\log(n)}$ above which a feasible solution exists (with high probability) and below which does not. This phase transition has already been established in [Bizeul et Najim, 2021] for the full matrix case.

One can notice that, in sparse foodwebs, the interaction coefficients can be stronger than when the interaction matrix is full (i.e. when $d_n = n$) in the sense that $\frac{1}{\sqrt{d_n}} > \frac{1}{\sqrt{n}}$.

Models and feasibility results

The sparse random matrix model under investigation is given in (1.3). Specifying the range of d_n and the structure of the deterministic matrix Δ_n, we introduce hereafter two models amenable to analysis.

Model (A) : Block permutation matrix.

Let $n = d \times m$. Denote by $\mathcal{S}_m$ the group of permutations of $[m] = \{1, \cdots, m\}$. Given $\sigma \in \mathcal{S}_m$, consider the associated permutation matrix

$$P_\sigma = (P_{ij})_{i,j\in[m]} \quad \text{where} \quad P_{ij} = \begin{cases} 1 & \text{if } j = \sigma(i), \\ 0 & \text{else.} \end{cases}$$

Denote by $J_d = \mathbf{1}_d \mathbf{1}_d^{\mathrm{T}}$ the $d \times d$ matrix of ones. Assume that

— matrix M_n is given by (1.3),

— $d = d_n \geq \log(n)$,

— matrix Δ_n introduced in (1.3) is a *block-permutation* adjacency matrix given by

$$\Delta_n = P_\sigma \otimes J_d = (P_{ij} J_d)_{i,j\in[m]} \ , \tag{1.4}$$

where $\otimes$ is the Kronecker matrix product.

Notice that Δ_n still corresponds to the adjacency matrix of a d-regular graph.

Example 1. To illustrate these definitions, we provide an example. Let $n = m \times d$ with $m = 4$ and $\sigma \in \mathcal{S}_4$ defined by

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix} .$$

Matrices P_σ, Δ and $\Delta \circ A$ are respectively given by :

$$P_\sigma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} , \quad \Delta = \begin{pmatrix} J_d & 0 & 0 & 0 \\ 0 & 0 & 0 & J_d \\ 0 & J_d & 0 & 0 \\ 0 & 0 & J_d & 0 \end{pmatrix} , \quad \Delta \circ A = \begin{pmatrix} A^{(1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & A^{(2)} \\ 0 & A^{(3)} & 0 & 0 \\ 0 & 0 & A^{(4)} & 0 \end{pmatrix} ,$$

where $A^{(\mu)}$ ($\mu \in [4]$) is a $d \times d$ matrix with i.i.d. $\mathcal{N}(0,1)$ entries.

Model (B) : d is proportional to n.

Assume that M_n is given by (1.3) and that $d = d_n$ satisfies

$$\lim_{n\to\infty} \frac{d_n}{n} = \beta > 0 . \tag{1.5}$$

We can now state the main result of the article :

Theorem 17. Let A_n be a $n \times n$ matrix with i.i.d. $\mathcal{N}(0,1)$ entries and Δ_n given by Model (A) or (B). Assume that $\alpha_n \xrightarrow[n\to\infty]{} \infty$ and denote by

$$\alpha_n^* = \sqrt{2 \log n} \ .$$

Let $\boldsymbol{x}_n = (x_k)_{k\in[n]}$ be the solution of

$$\boldsymbol{x}_n = \mathbf{1}_n + \frac{1}{\alpha_n\sqrt{d_n}}\left(\Delta_n \circ A_n\right)\boldsymbol{x}_n \ . \tag{1.6}$$

Then

1. If $\exists\,\varepsilon > 0$ such that eventually $\alpha_n \leq (1-\varepsilon)\alpha_n^*$, then $\mathbb{P}\left\{\min_{k\in[n]} x_k > 0\right\} \xrightarrow[n\to\infty]{} 0\,,$
2. If $\exists\,\varepsilon > 0$ such that eventually $\alpha_n \geq (1+\varepsilon)\alpha_n^*$, then $\mathbb{P}\left\{\min_{k\in[n]} x_k > 0\right\} \xrightarrow[n\to\infty]{} 1\,.$

The results of Theorem 17 are illustrated in Fig.1.1.

Remarks

1. By taking $d_n \geq \log(n)$, we guarantee that the spectral norm of matrix $\frac{\Delta_n\circ A_n}{\sqrt{d_n}}$ is of order $\mathcal{O}(1)$, see [Bandeira et van Handel, 2016]. In particular, matrix $\left(I_n - \frac{\Delta_n\circ A_n}{\alpha_n\sqrt{d_n}}\right)$ is invertible and the solution $\boldsymbol{x}_n$ can be represented as :

$$\boldsymbol{x}_n = \left(I_n - \frac{\Delta_n\circ A_n}{\alpha_n\sqrt{d_n}}\right)^{-1}\mathbf{1}_n \ .$$

2. An informal first-order expansion of the solution immediatly explains this phase transition. If we expand the inverse matrix and neglect the remaining terms, we get

$$\boldsymbol{x}_n \ \simeq \ \mathbf{1}_n + \frac{\Delta_n\circ A_n}{\alpha_n\sqrt{d_n}}\mathbf{1}_n \ = \ \mathbf{1} + \frac{\boldsymbol{z}_n}{\alpha_n}$$

where

$$\boldsymbol{z}_n = (z_i) \quad \text{and} \quad z_i = \sum_{j=1}^n \frac{(\Delta_n\circ A_n)_{ij}}{\sqrt{d_n}}\ .$$

Notice that the z_i's remain i.i.d. $\mathcal{N}(0,1)$. Going one step further in the approximation yields

$$\min_{i\in[n]} x_i \ \simeq \ 1 + \frac{\min_{i\in[n]} z_i}{\alpha_n}\,.$$

By standard extreme value results, we have $\min_{i\in[n]} z_i \sim -\sqrt{2\log(n)}$, hence the phase transition.

3. The component-wise positivity of the solution has been studied in the full matrix case, i.e. $\Delta_n = \mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}$ and $d_n = n$, in [Bizeul et Najim, 2021] where the same phase transition phenomenon occurs. Proof of Theorem 17 can be handled as in [Bizeul et Najim, 2021] for Model (B) with non-trivial adaptations that will be specified.

   In the case where $d_n \ll n$, a normalization issue occurs. To say it roughly, the euclidian norm of vector $\mathbf{1}_n/\sqrt{d_n}$ is no longer of order $\mathcal{O}(1)$ but of order $\sqrt{n/d_n}$ and one needs to handle more carefully the sparsity of matrix Δ_n.

In this regard, the block-permutation structure of Model (A) is a technical and simplifying assumption. The problem of the component-wise positivity of $\boldsymbol{x}_n$ for a general adjacency matrix Δ_n of a d-regular graph with $d \geq \log(n)$ remains open.

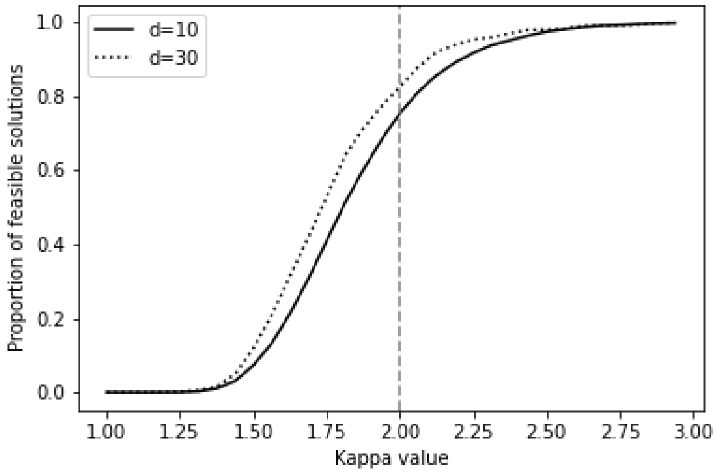


FIGURE 1.1 – Let $n = 15000$ with $\log(n) \simeq 9.61$. For $d = 10$ and $m = 1500$, we first draw at random a permutation $\sigma \in \mathcal{S}_m$ and fix $\Delta_n = P_\sigma \otimes \mathbf{1}_d \mathbf{1}_d^{\mathrm{T}}$ once for all. Each point of the solid line represents the proportion of feasible solutions $\boldsymbol{x}_n$ of (1.6) over 2000 realizations of random matrices A_n for different values of κ, with $\alpha_n = \sqrt{\kappa \log(n)}$. The same simulation is realized with $d = 30$ over 500 realizations of A_n (dotted line).

Stability results

A classical property of (1.2) is the positivity of the orbits[2] : if $\boldsymbol{x}_n^0 \in (\mathbb{R}^{*+})^n$, then $\boldsymbol{x}_n^t \in (\mathbb{R}^{*+})^n$ as well $(t > 0)$.

We first recall definitions related to stability from [Takeuchi, 1996, Chapter 3]. An equilibrium $\boldsymbol{x}_n$ is ***stable*** if for any given neighborhood W of $\boldsymbol{x}_n$, there exists a neighborhood V such that for any initial point $\boldsymbol{x}_n^0 \in V$, the orbit $\{\boldsymbol{x}_n^t;\ t \geq 0;\ \boldsymbol{x}_n^0 \in V\}$ stays in W. In addition, if the equilibrium is stable and the orbit converges to $\boldsymbol{x}_n$, the equilibrium is said ***asymptotically stable***.

In the full matrix case ($\Delta_n = \mathbf{1}_n \mathbf{1}_n^{\mathrm{T}},\ d_n = n$), it has been proved in [Bizeul et Najim, 2021] that in the regime where feasibility occurs, the system is asymptotically stable in the sense that the Jacobian matrix $\mathcal{J}$ of the LV system (1.2) evaluated at $\boldsymbol{x}_n$:

$$\mathcal{J}(\boldsymbol{x}_n) = \mathrm{diag}(\boldsymbol{x}_n)\,(-I_n + M_n) \tag{1.7}$$

has all its eigenvalues with negative real part.

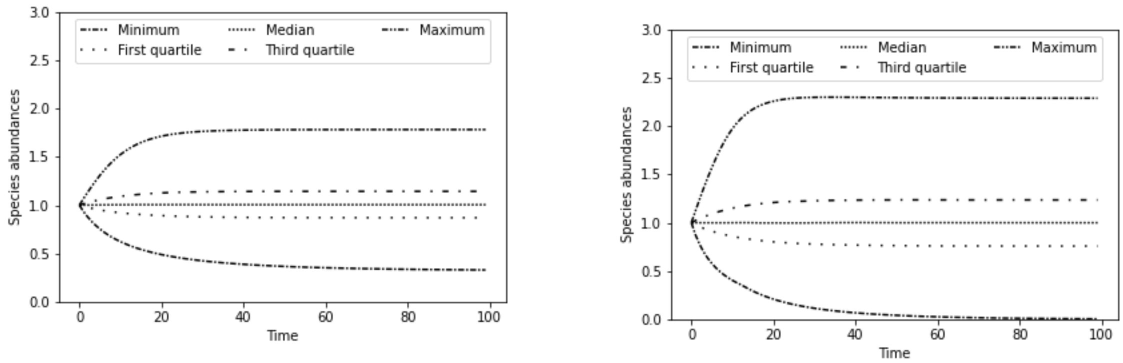
Finally, the equilibrium is ***globally stable*** when it is asymptotically stable and the neighborhood V can be taken as the whole state place $(\mathbb{R}^{*+})^n$.

2. Beware that this property does not prevent some components $x_i(t)$ to converge to zero, hence does not enforce a feasible equilibrium.

We complement Theorem 17 and prove that feasibility and global stability occur simultaneously.

Theorem 18 (Global stability, Takeuchi and Adachi [Takeuchi, 1996, Theorem 3.2.1]). Let $d_n \geq \log(n)$, $\alpha_n \xrightarrow[n\to\infty]{} \infty$, and Δ_n the adjacency matrix of a d_n-regular graph. Then, with probability going to one as $n \to \infty$, Eq. (1.1) admits a unique nonnegative solution $\boldsymbol{x}_n$ which is a globally stable equilibrium.

Beware that in this theorem, the solution, although unique, is no longer (componentwise) positive and may have zero components corresponding to vanishing species. Notice that the assumption over Δ_n covers Models (A) and (B) but is far less restrictive. We illustrate Theorem 18 in Figure 1.2.



(a) Feasible equilibrium for $\alpha_n = \sqrt{3\log(n)}$.

(b) Vanishing species for $\alpha_n = \sqrt{\log(n)}$.

FIGURE 1.2 – LV system with feasible equilibrium (left) and vanishing species (right) : minimum, maximum and mean of the population dynamics $(\boldsymbol{x}_n^t, t > 0)$ solution of (1.2) for $n = 5000$ ($\log(n) \simeq 8.51$), $d = 10$ and Δ_n follows Model (A). In the first figure, $\alpha_n > \sqrt{2\log(n)}$, the minimum abundance remains positive. In the second one, $\alpha_n < \sqrt{2\log(n)}$, the minimum abundance vanishes and the equilibrium is not feasible.

We now specify Theorem 18 in the case of feasibility.

Proposition 19 (Stability and convergence rate). Let $d_n \geq \log(n)$, $\alpha_n \xrightarrow[n\to\infty]{} \infty$, and assume that Δ_n is given by Model (A) or (B). Denote by $\mathcal{S}_n$ the spectrum of the Jacobian matrix $\mathcal{J}(\boldsymbol{x}_n)$ given by (1.7).

Assume that there exists $\varepsilon > 0$ such that eventually $\alpha_n \geq (1+\varepsilon)\alpha_n^*$. Then :

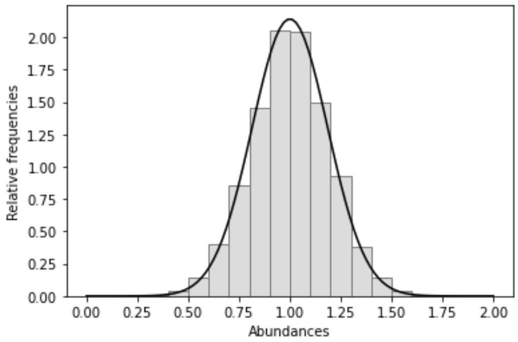
1. The probability that the equilibrium $\boldsymbol{x}_n$ is feasible and globally stable converges to 1,
2. The spectrum $\mathcal{S}_n$ asymptotically coincides with $-\mathrm{diag}(\boldsymbol{x}_n)$ in the sense that :

$$\max_{\lambda \in \mathcal{S}_n} \min_{k \in [n]} |\lambda + x_k| \xrightarrow[n\to\infty]{\mathcal{P}} 0\,,$$

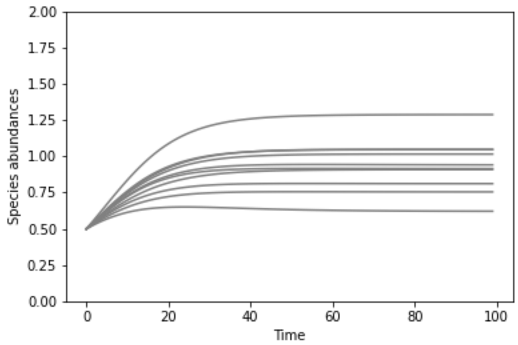
3. Moreover,

$$\max_{\lambda \in \mathcal{S}_n} \mathrm{Re}(\lambda) \leq -(1-\boldsymbol{\ell}^+) + o_P(1) \qquad \text{where} \qquad \boldsymbol{\ell}^+ := \limsup_{n\to\infty} \frac{\alpha_n^*}{\alpha_n} < 1\,. \quad (1.8)$$

As a consequence of (1.8), for any $\boldsymbol{x}_n^0 \in (\mathbb{R}^{+*})^n$, the orbit $\boldsymbol{x}_n^t$ converges to the equilibrium $\boldsymbol{x}_n$ at an exponential convergence rate, see Fig.1.3b.



(a) Histogram of the equilibrium abundances.



(b) Population dynamics with starting abundances equals to $\frac{1}{2}$.

FIGURE 1.3 – Consider the population dynamics $(\boldsymbol{x}_n^t,\, t > 0)$ solution of (1.2) where M is given by (1.3) and Δ_n follows Model (A) with $n = 15000$ species, $m = 1500$ blocks, $d = 10 > \log(n) \simeq 9.62$ and $\alpha_n = \sqrt{3\log(n)}$. On the right, we plot 10 species randomly chosen over 15000 with starting abundances equals to $\frac{1}{2}$. On the left, the histogram of the abundances is represented, and the normal density with mean 1 and variance $\frac{1}{\alpha_n^2}$ is fitted. Notice the substantial spread of the abundances despite the high value of n.

Notations

If $\boldsymbol{v}$ is a vector then $\|\boldsymbol{v}\|$ stands for its Euclidian norm ; if A is a matrix then $\|A\|$ stands for its spectral norm and $\|A\|_F = \sqrt{\sum_{ij} |A_{ij}|^2}$ for its Frobenius norm. Let φ be a function from some space Ψ (usually $\mathbb{R}$) to $\mathbb{R}$ then $\|\varphi\|_\infty = \sup_{x\in\Psi} |\varphi(x)|$. Convergence in probability is denoted by $\xrightarrow{\mathcal{P}}$. When no confusion can occur, we shall drop n and simply denote $A, \Delta, \alpha, d, \boldsymbol{x},$ etc. instead of $A_n, \Delta_n, \alpha_n, d_n, \boldsymbol{x}_n,$ etc.

Organization of the chapter

In Section 2, properties of the spectral norm of a sparse matrix are presented, together with the general strategy of proof. Proof of Theorem 17 is provided in Section 3 for Model (A), and in Section 4 for Model (B). Theorem 18 is proved in Section 5. In Section 6, we conclude and state an open question.

2 Spectral norm of the interaction matrix and strategy of proof

The spectral norm of $\Delta_n \circ A_n/\sqrt{d}$

In the following proposition which proof is based on [Bandeira et van Handel, 2016], we provide an estimate of $\|\Delta \circ A/\sqrt{d}\|$. The fact that A's entries are $\mathcal{N}(0,1)$

and that $d_n \geq \log(n)$ is crucial.

Proposition 20. [Bandeira et van Handel, 2016, Corollary 3.11]
Let X be the $n \times m$ matrix with $X_{ij} = g_{ij} b_{ij}$, where $\{g_{ij} : 1 \leq i \leq n, 1 \leq j \leq m\}$ are independent $\mathcal{N}(0,1)$ random variables and $\{b_{ij} \ : \ 1 \leq i \leq n, 1 \leq j \leq m\}$ are given scalars. Then,

$$\mathbb{P}\left[\|X\| \geq (1+\epsilon)\left\{\sigma_1 + \sigma_2 + \frac{5}{\sqrt{\log(1+\epsilon)}}\sigma_* \sqrt{\log(n \wedge m)}\right\} + t\right] \leq e^{-t^2/2\sigma_*^2}\,,$$

for any $0 < \epsilon \leq \frac{1}{2}$ and $t \geq 0$, where

$$\sigma_1 := \max_{i\in[n]} \sqrt{\sum_{j=1}^m b_{ij}^2}\,, \qquad \sigma_2 := \max_{j\in[m]} \sqrt{\sum_{i=1}^n b_{ij}^2}\,, \qquad \sigma_* := \max_{i,j} |b_{ij}|\,.$$

Proposition 21. Assume that A is a $n \times n$ matrix with i.i.d. $\mathcal{N}(0,1)$ entries, that Δ is a $n \times n$ adjacency matrix of a d-regular graph, that $d \geq \log(n)$. Then there exists a constant $\kappa > 0$ independent from n (one can take for instance $\kappa = 22$) such that

$$\mathbb{P}\left(\left\|\frac{\Delta \circ A}{\sqrt{d}}\right\| \geq \kappa\right) \xrightarrow[n\to\infty]{} 0\,.$$

In particular, let $\delta \in (0,1)$ be fixed and $\alpha = \alpha(n) \xrightarrow[n\to\infty]{} \infty$. Then

$$\mathbb{P}\left(\left\|\frac{\Delta \circ A}{\alpha\sqrt{d}}\right\| \leq 1-\delta\right) \xrightarrow[n\to\infty]{} 1\,.$$

Démonstration. Applying proposition 20 to $\frac{\Delta\circ A}{\sqrt{d}}$ with $\epsilon = \frac{1}{2}$, $n = m$, i.e. $\sigma_1 = \sigma_2 = 1$ and $\sigma_* = \frac{1}{\sqrt{d}}$, we obtain

$$\mathbb{P}\left(\left\|\frac{\Delta \circ A}{\sqrt{d}}\right\| \geq 3 + \frac{15}{2\sqrt{\log\frac{3}{2}}} \times \frac{\sqrt{\log n}}{\sqrt{d}} + \frac{t}{\sqrt{d}}\right) \leq e^{-\frac{t^2}{2}}\,.$$

Fix $t = \sqrt{\log n}$, then $e^{-\frac{t^2}{2}} = \frac{1}{\sqrt{n}} \xrightarrow[n\to\infty]{} 0$ and $\frac{t}{\sqrt{d}} \leq 1$ by assumption. Furthermore, there exists a rank n_1 such that for all $n \geq n_1$:

$$3 + \frac{15}{2\sqrt{\log(3/2)}} \times \frac{\sqrt{\log n}}{d} + \frac{t}{\sqrt{d}} \quad \leq \quad 4 + \frac{15}{2\sqrt{\log\frac{3}{2}}} \quad < \quad \kappa := 22\,.$$

Thus, $\mathbb{P}\left(\left\|\frac{\Delta\circ A}{\sqrt{d}}\right\| \geq \kappa\right) \to 0$. Since $\alpha \to \infty$, the last part of the proposition immediatly follows. □

Strategy of proof

Based on the previous control of the spectral norm in probability, we reduce the problem of feasibility to the control of the extreme values of high order terms of the resolvent, considered as a Neumann sum, see Lemma 22.

Going back to Eq.(1.6), we can write $\left(I - \frac{\Delta\circ A}{\alpha\sqrt{d}}\right)\boldsymbol{x} = \mathbf{1}$. Introducing the resolvent $Q = \left(I - \frac{\Delta\circ A}{\alpha\sqrt{d}}\right)^{-1}$ which by Proposition 21 exists with probability tending to one, we obtain the representation

$$\boldsymbol{x} = (x_k)_k = \left(I - \frac{\Delta\circ A}{\alpha\sqrt{d}}\right)^{-1}\mathbf{1} = Q\mathbf{1}$$

which holds with growing probability. Denote by $\boldsymbol{e}_k$ the $n\times 1$ k-th canonical vector, then $x_k = \boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{e}_k^{\mathrm{T}}Q\mathbf{1}$. Unfolding the resolvent as a Neumann sum, we obtain

$$x_k \quad = \quad \boldsymbol{e}_k^{\mathrm{T}}Q\mathbf{1} \quad = \quad \sum_{\ell=0}^{\infty}\boldsymbol{e}_k^{\mathrm{T}}\left(\frac{\Delta\circ A}{\alpha\sqrt{d}}\right)^{\ell}\mathbf{1} \quad = \quad 1 + \frac{Z_k}{\alpha} + \frac{R_k}{\alpha^2} \tag{1.9}$$

where

$$Z_k = \boldsymbol{e}_k^{\mathrm{T}}\left(\frac{\Delta\circ A}{\sqrt{d}}\right)\mathbf{1} \qquad \text{and} \qquad R_k = \boldsymbol{e}_k^{\mathrm{T}}\sum_{\ell=2}^{\infty}\frac{1}{\alpha^{\ell-2}}\left(\frac{\Delta\circ A}{\sqrt{d}}\right)^{\ell}\mathbf{1}\,.$$

Notice that the Z_k's are i.i.d. $\mathcal{N}(0,1)$ random variables and denote by $\check{M} = \min_{k\in[n]} Z_k$.

Eq. (1.9) immediatly yields

$$\left\{\begin{array}{lcl} \min_{k\in[n]} x_k & \geq & 1 + \frac{1}{\alpha}\check{M} + \frac{1}{\alpha^2}\min_{k\in[n]} R_k\ , \\[2ex] \min_{k\in[n]} x_k & \leq & 1 + \frac{1}{\alpha}\check{M} + \frac{1}{\alpha^2}\max_{k\in[n]} R_k\ . \end{array}\right. \tag{1.10}$$

Let $\alpha_n^* = \sqrt{2\log n}$, $\beta_n^* = \alpha_n^* - \frac{1}{2\alpha_n^*}\log(4\pi\log n)$ and denote by $G(x) = e^{-e^{-x}}$ the cumulative distribution of a Gumbel distributed random variable. Then it is well-known, see for instance [Leadbetter *et al.*, 1983, Theorem 1.5.3], that

$$\mathbb{P}\left(\alpha_n^*(\check{M}_n + \beta_n^*) \geq x\right) \xrightarrow[n\to\infty]{} G(x)\,. \tag{1.11}$$

By taking into account this convergence, we can rewrite (1.10) as

$$\begin{aligned} 1 + \frac{\alpha_n^*}{\alpha_n}\left(-1 + o_P(1) + \frac{\min_{k\in[n]} R_k}{\alpha_n^*\alpha_n}\right) &\leq \min_{k\in[n]} x_k \\ &\leq 1 + \frac{\alpha_n^*}{\alpha_n}\left(-1 + o_P(1) + \frac{\max_{k\in[n]} R_k}{\alpha_n^*\alpha_n}\right)\,. \end{aligned} \tag{1.12}$$

where we used $(\alpha_n^*)^{-1}(\check{M} + \beta_n^*) = o_P(1)$. Theorem 17 will then follow from the following lemma.

Lemma 22. Under the assumptions of Theorem 17, the following convergence holds

$$\frac{\max_{k\in[n]} R_k}{\alpha_n\sqrt{2\log n}} \xrightarrow[n\to\infty]{\mathcal{P}} 0 \qquad \text{and} \qquad \frac{\min_{k\in[n]} R_k}{\alpha_n\sqrt{2\log n}} \xrightarrow[n\to\infty]{\mathcal{P}} 0\,.$$

Proof of Lemma 22 relies on a careful analysis of the order of magnitude of the extreme values of the remaining term $(R_k)_{k\in[n]}$. The sparse structure of matrix $\Delta\circ A$ (either Model (A) or (B)) requires a specific analysis, substantially different from the one in [Bizeul et Najim, 2021].

3 Proof of the feasibility result for the block model

We assume that Δ_n follows Model (A).

In order to prove Lemma 22, we first take advantage of the fact that $\|\Delta \circ A/\sqrt{d}\|$ is typically lower than κ (see Proposition 21) and replace R_k by a truncated version $\tilde{R}_k$ (step 1). We then prove that $A \mapsto \tilde{R}_k(A)$ is Lipschitz (step 2). The quantity $\tilde{R}_k$ being Lipschitz, its centered version is subgaussian if the matrix entries are Gaussian i.i.d. We finally prove that $\tilde{R}_k(A)$ is uniformily integrable (step 3). The conclusion easily follows. Although the general strategy is similar to the one developed in [Bizeul et Najim, 2021], the proofs are substantially different. In particular, proofs of step 2 and 3 heavily rely on the block permutation structure of the matrices.

Step 1 : Truncation

Toward proving Lemma 22, sub-Gaussiannity is an important property, which follows from Lipschitz properties by standard concentration of measure arguments. Unfortunately $A \mapsto R_k(A)$ fails to be Lipschitz (simply notice that $R_k(A)$ has quadratic and higher order terms). In order to circumvent this issue, we provide a truncated version of R_k.

Let $\kappa > 0$ as in Prop. 21 (one can take $\kappa = 22$), $\eta \in (0,1)$ and $\varphi : \mathbb{R}^+ \to [0,1]$ a smooth function :

$$\varphi(x) = \begin{cases} 1 & \text{if } x \in [0, \kappa + 1 - \eta]\,, \\ 0 & \text{if } x \geq \kappa + 1 \end{cases} \tag{1.13}$$

strictly decreasing from 1 to 0 for $x \in (\kappa + 1 - \eta, \kappa + 1)$. According to Prop. 21,

$$\varphi_d(A) := \varphi\left(\left\|\frac{\Delta \circ A}{\sqrt{d}}\right\|\right)$$

is equal to one with high probability. We introduce the truncated value :

$$\tilde{R}_k(A) = \varphi_d(A) R_k(A)\ .$$

We have

$$\begin{aligned} \mathbb{P}\left(\max_k R_k(A) \neq \max_k \tilde{R}_k(A)\right) \ &\leq\ \mathbb{P}\left(\exists k \in [n],\ R_k(A) \neq \tilde{R}_k(A)\right) \\ &\leq\ \mathbb{P}(\varphi_d(A) < 1)\ \leq\ \mathbb{P}\left(\left\|\frac{\Delta \circ A}{\sqrt{d}}\right\| \geq \kappa\right) \xrightarrow[n\to\infty]{} 0\,, \end{aligned}$$

from which we deduce

$$\frac{\max_{k\in[n]} R_k - \max_{k\in[n]} \tilde{R}_k}{\alpha_n\sqrt{2\log n}} \xrightarrow[n\to\infty]{\mathcal{P}} 0\,. \tag{1.14}$$

It is therefore sufficient to prove

$$\frac{\max_{k\in[n]} \tilde{R}_k}{\alpha_n\sqrt{2\log n}} \xrightarrow[n\to\infty]{\mathcal{P}} 0 \tag{1.15}$$

to establish the first part of Lemma 22. The property of the minimum can be proved similarly.

Step 2 : Lipschitz property for $\tilde{R}_k(A)$

For $\ell \geq 2$, we introduce the following summand terms :

$$\rho_{k,\ell}(A) = \boldsymbol{e}_k^{\mathrm{T}} \frac{1}{\alpha^{\ell-2}} \left(\frac{\Delta \circ A}{\sqrt{d}}\right)^{\ell} \mathbf{1} \qquad \text{and} \qquad \tilde{\rho}_{k,\ell}(A) = \varphi_d(A)\rho_{k,\ell}(A) \;, \tag{1.16}$$

so that $R_k(A) = \sum_{\ell=2}^{\infty} \rho_{k,\ell}(A)$ and $\tilde{R}_k(A) = \sum_{\ell=2}^{\infty} \tilde{\rho}_{k,\ell}(A)$.

The following lemma is the main result of this section.

Lemma 23. Let $\kappa > 0$ as in Proposition 21, $\delta \in (0,1)$ and n_0 such that for all $n \geq n_0$,

$$\frac{\kappa+1}{\alpha_n} \leq 1 - \delta \;.$$

For $\ell \geq 2$ and $n \geq n_0$, the function $\tilde{\rho}_{k,\ell} : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ is K_ℓ-Lipschitz, i.e.

$$|\tilde{\rho}_{k,\ell}(A) - \tilde{\rho}_{k,\ell}(B)| \leq K_\ell \, \|A - B\|_F \;, \tag{1.17}$$

where $K_\ell = K_\ell(\kappa, n_0, \delta) > 0$ is a constant independent from k, d and $n \geq n_0$. Moreover, $K := \sum_{\ell \geq 2} K_\ell < \infty$. In particular, the function $\tilde{R}_k$ is K-Lipschitz :

$$\left|\tilde{R}_k(A) - \tilde{R}_k(B)\right| \leq K \, \|A - B\|_F \;. \tag{1.18}$$

Given a $n \times n$ matrix C, we define its hermitization matrix $\mathcal{H}(C)$ by :

$$\mathcal{H}(C) = \begin{pmatrix} 0 & C \\ C^{\mathrm{T}} & 0 \end{pmatrix} \;.$$

A well-known property of $\mathcal{H}(C)$ is its symmetric spectrum and the fact that the singular values of C are the non-negatives eigenvalues of $\mathcal{H}(C)$. In particular, $\|C\|$ corresponds to the largest eigenvalue of $\mathcal{H}(C)$.

In order to prove Lemma 23, we first consider the case where $\mathcal{H}(\Delta \circ A)$ has a simple spectrum, a sufficient condition for the differentiability of $\|\Delta \circ A\|$, we then prove that the euclidian norm of the gradient of $\tilde{\rho}_{k,\ell}(A)$ is bounded : $\|\nabla \tilde{\rho}_{k,\ell}(A)\| \leq K_\ell$ and finally proceed by approximation to get the general Lipschitz property.

Démonstration. We first consider the case where $\mathcal{H}(\Delta \circ A)$ has a simple spectrum. In this case, $\|\Delta \circ A\|$ is equal to the largest eigenvalue of $\mathcal{H}(\Delta \circ A)$ which has multiplicity 1 and is thus differentiable. Denote by $\partial_{ij} = \frac{\partial}{\partial A_{ij}}$. Notice that if $\Delta_{ij} = 0$, then for any smooth function $f : \mathbb{R}^{n \times n} \to \mathbb{R}$, $\partial_{ij} f(\Delta \circ A) = 0$. If needed, we will take advantage of this property. We have :

$$\|\nabla \tilde{\rho}_{k,\ell}(A)\| = \sqrt{\sum_{i,j=1}^{n} |\partial_{ij} \tilde{\rho}_{k,\ell}(A)|^2} \,,$$

and

$$\begin{aligned} \partial_{ij} \tilde{\rho}_{k,\ell}(A) &= \partial_{ij} \left(\varphi_d(A)\rho_{k,\ell}(A)\right) \\ &= \left(\partial_{ij}\varphi_d(A)\right) \rho_{k,\ell}(A) + \varphi_d(A)\partial_{ij}\rho_{k,\ell}(A) \\ &=: S_{1,ij} + S_{2,ij} \;. \end{aligned}$$

In particular,

$$\sum_{i,j=1}^{n} |\partial_{ij}\tilde{\rho}_{k,\ell}(A)|^2 \leq 2\sum_{i,j=1}^{n} |S_{1,ij}|^2 + 2\sum_{i,j=1}^{n} |S_{2,ij}|^2 \ .$$

We first evaluate $\sum_{ij} |S_{1,ij}|^2$.

Recall that $\|\Delta \circ A\|$ being the maximum eigenvalue of $\mathcal{H}(\Delta \circ A)$ which by assumption is simple, it is differentiable by [Horn et Johnson, 2013, Theorem 6.3.12]. Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be respectively the left and right normalized singular vectors associated to the largest singular value $\|\Delta \circ A\|$ of $\Delta \circ A$. Then

$$\mathcal{H}(\Delta \circ A)\boldsymbol{w} = \|\Delta \circ A\|\,\boldsymbol{w}\,, \quad \text{where} \quad \boldsymbol{w} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} .$$

Notice that $\|\boldsymbol{w}\|^2 = 2$. We have

$$\partial_{ij}\varphi_d(A) = \frac{1}{\sqrt{d}}\varphi'\left(\frac{\|\Delta \circ A\|}{\sqrt{d}}\right)\partial_{ij}\|\Delta \circ A\|$$

and

$$\partial_{ij}\|\Delta \circ A\| = \begin{cases} \frac{1}{\|w\|}(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{e}_i\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{v} + \boldsymbol{v}^{\mathrm{T}}\boldsymbol{e}_j\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{u}) = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{e}_i\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{v} & \text{if } \Delta_{ij} \neq 0\ , \\ 0 & \text{else.} \end{cases} \tag{1.19}$$

Let $i \in [n]$. Denote by

$$\mathcal{I}_i = \{j \in [n],\ \Delta_{ij} = 1\}\,; \tag{1.20}$$

notice that $|\mathcal{I}_i| = d$. We have

$$\begin{aligned} \sum_{i,j\in[n]} |S_{1,ij}|^2 &= \sum_{i\in[n]}\sum_{j\in\mathcal{I}_i}\left|\boldsymbol{u}^{\mathrm{T}}\boldsymbol{e}_i\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{v}\varphi'\left(\left\|\frac{\Delta \circ A}{\sqrt{d}}\right\|\right)\frac{1}{\sqrt{d}}\rho_{k,\ell}(A)\right|^2 , \\ &\leq \left|\varphi'\left(\left\|\frac{\Delta \circ A}{\sqrt{d}}\right\|\right)\frac{1}{\sqrt{d}}\rho_{k,\ell}(A)\right|^2 \sum_{i\in[n]}\left|\boldsymbol{u}^{\mathrm{T}}\boldsymbol{e}_i\right|^2\sum_{j\in[n]}\left|\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{v}\right|^2 , \\ &= \left|\varphi'\left(\left\|\frac{\Delta \circ A}{\sqrt{d}}\right\|\right)\frac{1}{\sqrt{d}}\rho_{k,\ell}(A)\right|^2 . \end{aligned}$$

We now focus on

$$\left|\frac{1}{\sqrt{d}}\rho_{k,\ell}(A)\right|^2 = \left|\boldsymbol{e}_k^{\mathrm{T}}\frac{1}{\alpha^{\ell-2}}\left(\frac{\Delta \circ A}{\sqrt{d}}\right)^{\ell}\frac{\mathbf{1}}{\sqrt{d}}\right|^2 .$$

Notice that $\|\mathbf{1}/\sqrt{d}\| = \sqrt{n/d}$. Since matrix $\Delta \circ A$ follows Model (A), one can notice that $(\Delta \circ A)^{\ell}$ remains a block matrix with only d nonzero terms per row (and per column as well). This property is fundamental for the remaining estimates and fully relies on the Model (A) assumption.

Denote by

$$\mathcal{J}_{k,\ell} = \left\{p \in [n],\ \left[(\Delta \circ A)^{\ell}\right]_{kp} \neq 0\right\} \tag{1.21}$$

and by $\mathbf{1}^{\mathcal{J}_{k,\ell}}$ the $n \times 1$ vector with zero coordinates except those belonging to $\mathcal{J}_{k,\ell}$, set to 1. In particular, $\|\mathbf{1}^{\mathcal{J}_{k,\ell}}\| = \sqrt{d}$. Then

$$\boldsymbol{e}_k^{\mathrm{T}}(\Delta \circ A)^{\ell}\mathbf{1} = \boldsymbol{e}_k^{\mathrm{T}}(\Delta \circ A)^{\ell}\mathbf{1}^{\mathcal{J}_{k,\ell}} \ .$$

We have

$$\begin{aligned}
\left|\frac{1}{\sqrt{d}}\rho_{k,\ell}(A)\right|^2 &= \left|\boldsymbol{e}_k^{\mathrm{T}}\left(\frac{\Delta \circ A}{\alpha\sqrt{d}}\right)^{\ell-2}\left(\frac{\Delta \circ A}{\sqrt{d}}\right)^2 \frac{\mathbf{1}}{\sqrt{d}}\right|^2 , \\
&= \left|\boldsymbol{e}_k^{\mathrm{T}}\left(\frac{\Delta \circ A}{\alpha\sqrt{d}}\right)^{\ell-2}\left(\frac{\Delta \circ A}{\sqrt{d}}\right)^2 \frac{\mathbf{1}^{\mathcal{J}_{k,\ell}}}{\sqrt{d}}\right|^2 , \\
&\leq \|\boldsymbol{e}_k^{\mathrm{T}}\|^2 \left\|\left(\frac{\Delta \circ A}{\alpha\sqrt{d}}\right)^{\ell-2}\right\|^2 \left\|\frac{\Delta \circ A}{\sqrt{d}}\right\|^2 \left\|\frac{\mathbf{1}^{\mathcal{J}_{k,\ell}}}{\sqrt{d}}\right\|^2 , \\
&\leq \left\|\frac{\Delta \circ A}{\alpha\sqrt{d}}\right\|^{2(\ell-2)} \left\|\frac{\Delta \circ A}{\sqrt{d}}\right\|^4 .
\end{aligned}$$

Using the fact that $\varphi'\left(\left\|\frac{\Delta\circ A}{\sqrt{d}}\right\|\right) = 0$ if $\left\|\frac{\Delta\circ A}{\sqrt{d}}\right\| \geq \kappa + 1$, we have

$$\begin{aligned}
\left|\varphi'\left(\left\|\frac{\Delta \circ A}{\sqrt{d}}\right\|\right)\frac{1}{\sqrt{d}}\rho_{k,\ell}(A)\right|^2 &\leq \left|\varphi'\left(\left\|\frac{\Delta \circ A}{\sqrt{d}}\right\|\right)\right|^2 \left\|\frac{\Delta \circ A}{\alpha\sqrt{d}}\right\|^{2(\ell-2)} \left\|\frac{\Delta \circ A}{\sqrt{d}}\right\|^4 , \\
&\leq \|\varphi'\|_\infty^2 (1-\delta)^{2(\ell-2)} (1+\kappa)^4 ,
\end{aligned}$$

and finally

$$\sum_{i,j=1}^{n} |S_{1,ij}|^2 \leq \|\varphi'\|_\infty^2 (1-\delta)^{2(\ell-2)} \times (1+\kappa)^4 . \tag{1.22}$$

We now evaluate $\sum_{i,j=1}^n |S_{2,ij}|^2 = \sum_{i,j\in[n]} |\varphi_d(A)\partial_{ij}\rho_{k,\ell}(A)|^2$.

Recall the definitions of $\mathcal{I}_i$ and $\mathcal{J}_{k\ell}$ introduced in (1.20), (1.21). We have

$$\partial_{ij}\rho_{k,\ell}(A) = \frac{1}{\alpha^{\ell-2}(\sqrt{d})^{\ell}} \sum_{p=0}^{\ell-1} \boldsymbol{e}_k^{\mathrm{T}}(\Delta \circ A)^p \boldsymbol{e}_i \boldsymbol{e}_j^{\mathrm{T}}(\Delta \circ A)^{\ell-1-p}\mathbf{1} \quad \text{if} \quad j \in \mathcal{I}_i$$

and zero else. Then

$$
\begin{aligned}
&\sum_{i\in[n]}\sum_{j\in\mathcal{I}_i} |\partial_{ij}\rho_{k,\ell}(A)|^2 \\
&\leq \frac{\ell}{\alpha^{2(\ell-2)}d^\ell}\left(\sum_{i\in[n]}\sum_{j\in\mathcal{I}_i}\left|\boldsymbol{e}_k^{\mathrm{T}}(\Delta\circ A)^{\ell-1}\boldsymbol{e}_i\boldsymbol{e}_j^{\mathrm{T}}\mathbf{1}\right|^2\right. \\
&\qquad\left. + \sum_{p=0}^{\ell-2}\sum_{i=1}^{n}\sum_{j\in\mathcal{I}_i}\left|\boldsymbol{e}_k^{\mathrm{T}}(\Delta\circ A)^p\boldsymbol{e}_i\boldsymbol{e}_j^{\mathrm{T}}(\Delta\circ A)^{\ell-1-p}\mathbf{1}\right|^2\right), \\
&= \frac{\ell}{\alpha^{2(\ell-2)}d^\ell}\left(d\sum_{i\in[n]}\left|[(\Delta\circ A)^{\ell-1}]_{k,i}\right|^2\right. \\
&\qquad\left. + d\sum_{p=0}^{\ell-2}\sum_{i\in[n]}\sum_{j\in\mathcal{I}_i}\left|[(\Delta\circ A)^p]_{k,i}\,\boldsymbol{e}_j^{\mathrm{T}}(\Delta\circ A)^{\ell-1-p}\frac{\mathbf{1}}{\sqrt{d}}\right|^2\right), \\
&\leq \frac{\ell}{\alpha^{2(\ell-2)}d^{\ell-1}}\left(\left[(\Delta\circ A)^{\ell-1}\left((\Delta\circ A)^{\ell-1}\right)^{\mathrm{T}}\right]_{k,k}\right. \\
&\qquad\left. + \sum_{p=0}^{\ell-2}\sum_{i\in[n]}|[(\Delta\circ A)^p]_{k,i}|^2\sum_{j\in\mathcal{I}_i}\left|\boldsymbol{e}_j^{\mathrm{T}}(\Delta\circ A)^{\ell-1-p}\frac{\mathbf{1}}{\sqrt{d}}\right|^2\right).
\end{aligned}
\tag{1.23}
$$

We concentrate on the term $\Theta = \sum_{j\in\mathcal{I}_i}\left|\boldsymbol{e}_j^{\mathrm{T}}(\Delta\circ A)^{\ell-1-p}\frac{\mathbf{1}}{\sqrt{d}}\right|^2$ and prove that

$$
\Theta \quad\leq\quad \|\Delta\circ A\|^{2(\ell-1-p)}\,. \tag{1.24}
$$

Let $I_{\mathcal{I}_i} = \mathrm{diag}\,\{\mathbf{1}_{\mathcal{I}_i}(k);\, k\in[n]\}$, where $\mathbf{1}_{\mathcal{I}_i}$ is the indicator function of $\mathcal{I}_i$, then

$$
\Theta = \frac{\mathbf{1}^{\mathrm{T}}}{\sqrt{d}}\left[(\Delta\circ A)^{\ell-1-p}\right]^{\mathrm{T}} I_{\mathcal{I}_i}(\Delta\circ A)^{\ell-1-p}\frac{\mathbf{1}}{\sqrt{d}}\,.
$$

Notice that $(\Delta\circ A)^{\ell-1-p}$ has the form $(P_\tau\otimes\mathbf{1}_d\mathbf{1}_d^{\mathrm{T}})\circ B$ for some $\tau\in\mathcal{S}_m$ and some $n\times n$ matrix B. In particular, taking into account the matching between the indices of $I_{\mathcal{I}_i}$ and $(\Delta\circ A)^{\ell-1-p}$'s blocs, there exists a $d\times d$ bloc of matrix $(\Delta\circ A)^{\ell-p-1}$ say B_i such that matrix

$$
\left[(\Delta\circ A)^{\ell-1-p}\right]^{\mathrm{T}} I_{\mathcal{I}_i}(\Delta\circ A)^{\ell-1-p}
$$

is zero except a $d\times d$ block $B_i^{\mathrm{T}}B_i$ on the diagonal and

$$
\Theta = \frac{\mathbf{1}_d^{\mathrm{T}}}{\sqrt{d}}B_i^{\mathrm{T}}B_i\frac{\mathbf{1}_d}{\sqrt{d}} \;\leq\; \|B_i^*B_i\| \;\leq\; \|B_i\|^2 \;\leq\; \left\|(\Delta\circ A)^{\ell-p-1}\right\|^2 \;\leq\; \|\Delta\circ A\|^{2(\ell-p-1)}\,.
$$

Eq.(1.24) is established. Notice in particular that the estimate does not depend on the index i. Plugging this estimate into (1.23) yields

$$\begin{aligned}
&\sum_{i\in[n]}\sum_{j\in\mathcal{I}_i}|\partial_{ij}\rho_{k,\ell}(A)|^2\\
&\leq \frac{\ell}{\alpha^{2(\ell-2)}d^{\ell-1}}\left(\left\|(\Delta\circ A)^{\ell-1}\right\|^2+\sum_{p=0}^{\ell-2}\left[((\Delta\circ A)^p)^*(\Delta\circ A)^p\right]_{kk}\|\Delta\circ A\|^{2(\ell-p-1)}\right),\\
&\leq \frac{\ell}{\alpha^{2(\ell-2)}d^{\ell-1}}\left(\|\Delta\circ A\|^{2(\ell-1)}+\sum_{p=0}^{\ell-2}\|\Delta\circ A\|^{2p}\|\Delta\circ A\|^{2(\ell-p-1)}\right),\\
&= \frac{\ell^2}{\alpha^{2(\ell-2)}d^{\ell-1}}\|\Delta\circ A\|^{2(\ell-1)} \quad = \quad \ell^2\left\|\frac{\Delta\circ A}{\alpha\sqrt{d}}\right\|^{2(\ell-2)}\left\|\frac{\Delta\circ A}{\sqrt{d}}\right\|^2.
\end{aligned}$$

Multiplying by $|\varphi_d(A)|^2$ finally yields the appropriate estimates :

$$\sum_{i,j\in[n]}|S_{2,ij}|^2\leq \ell^2|\varphi_d(A)|^2\left\|\frac{\Delta\circ A}{\alpha\sqrt{d}}\right\|^{2(\ell-2)}\left\|\frac{\Delta\circ A}{\sqrt{d}}\right\|^2 \quad\leq\quad \ell^2(1-\delta)^{2(\ell-2)}(1+\kappa)^2. \tag{1.25}$$

Combining (1.22) and (1.25), we obtain :

$$\|\nabla\tilde{\rho}_{k,\ell}(A)\| \;\leq\; \sqrt{2\sum_{i,j=1}^n|S_{1,ij}|^2+2\sum_{i,j=1}^n|S_{2,ij}|^2} \;\leq\; 2(1-\delta)^{\ell-2}(\kappa+1)^2(\|\varphi'\|_\infty+\ell) \;=:K_\ell. \tag{1.26}$$

where K_ℓ does not depend upon k,n,d and is summable. So far, we have established a local estimate over $\|\nabla\tilde{\rho}_{k,\ell}(A)\|$ for any matrix A such that $\mathcal{H}(\Delta\circ A)$ has a simple spectrum. We first establish the Lipschitz estimate (1.17) for two such matrices A and B.

Let A, B such that $\mathcal{H}(\Delta\circ A)$ and $\mathcal{H}(\Delta\circ B)$ have simple spectrum and consider the interpolation matrix

$$A_t=(1-t)A+tB$$

for $t\in[0;1]$. The continuity of the eigenvalues implies that there exists $\epsilon>0$ sufficiently small such that $\mathcal{H}(\Delta\circ A_t)$ has a simple spectrum for $t\leq\epsilon$ and $t\geq 1-\epsilon$. By an argument in [Kato, 1995, Chapter 2.1], the number of eigenvalues of $\mathcal{H}(\Delta\circ A_t)$ remains constant for $t\in[0,1]$, except maybe for a finite number of points $(t_l; 1\leq l\leq L)$: $t_0=0<t_1<\cdots<t_L<t_{L+1}=1$. Since $\mathcal{H}(\Delta\circ A_t)$ has simple spectrum for $t\in[0;\epsilon)\cup(1-\epsilon;1]$, it has simple spectrum for all $t\notin\{t_l,\ l\in[L]\}$. We can now proceed :

$$\begin{aligned}
|\tilde{\rho}_{k,\ell}(A_{t_1})-\tilde{\rho}_{k,\ell}(A)| &= \left|\lim_{\tau\nearrow t_1}\int_0^\tau\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\rho}_{k,\ell}(A_t)\,\mathrm{d}t\right| = \left|\lim_{\tau\nearrow t_1}\int_0^\tau\nabla\tilde{\rho}_{k,\ell}(A_t)\circ\frac{\mathrm{d}}{\mathrm{d}t}(A_t)\mathrm{d}t\right|,\\
&\leq \lim_{\tau\nearrow t_1}\int_0^\tau\|\nabla\tilde{\rho}_{k,\ell}(A_t)\|\times\|B-A\|_F\,\mathrm{d}t\leq K_\ell t_1\|B-A\|_F.
\end{aligned}$$

By iterating the process over the intervals $(t_{\ell-1}, t_\ell)$, we get

$$\begin{aligned} |\tilde\rho_{k,\ell}(B) - \tilde\rho_{k,\ell}(A)| &\leq \sum_{l=1}^{L+1} \left|\tilde\rho_{k,\ell}(A_{t_l}) - \tilde\rho_{k,\ell}\left(A_{t_{l-1}}\right)\right| , \\ &\leq \sum_{l=1}^{L+1} K_\ell (t_l - t_{l-1}) \|B - A\|_F = K_\ell \|B - A\|_F . \end{aligned}$$

Hence the Lipschitz property along the segment $[A, B]$.

To go beyond, we proceed by density and prove that for a given matrix Δ as in Model (A), the set of matrices $(\Delta \circ A)$ such that $\mathcal{H}(\Delta \circ A)$ has a simple spectrum is dense in the set of matrices $(\Delta \circ A,\ A \in \mathbb{R}^{n\times n})$.

Let P_σ be the permutation matrix used to define Δ in (1.4) and I_d the identity matrix of size d. We define the following $n \times n$ matrices

$$\Pi = P_\sigma \otimes I_d \qquad \text{and} \qquad D_A = (\Delta \circ A)\Pi^{\mathrm{T}} . \tag{1.27}$$

Notice that Π is a $n \times n$ permutation matrix and that D_A is a block diagonal matrix with $d \times d$ blocks on the diagonal. Since $\Pi\Pi^{\mathrm{T}} = \Pi^{\mathrm{T}}\Pi = I_n$, we also have

$$D_A \Pi = \Delta \circ A .$$

In the framework of Example 1, matrices Π and D_A are given by :

$$\Pi = \begin{pmatrix} I_d & 0 & 0 & 0 \\ 0 & 0 & 0 & I_d \\ 0 & I_d & 0 & 0 \\ 0 & 0 & I_d & 0 \end{pmatrix} \qquad \text{and} \qquad D_A = \begin{pmatrix} A^{(1)} & 0 & 0 & 0 \\ 0 & A^{(2)} & 0 & 0 \\ 0 & 0 & A^{(3)} & 0 \\ 0 & 0 & 0 & A^{(4)} \end{pmatrix} .$$

An important feature of D_A is that $\Delta \circ A$ and D_A have the same singular values :

$$D_A D_A^{\mathrm{T}} = (\Delta \circ A)\, \Pi^{\mathrm{T}}\, \Pi\, (\Delta \circ A)^{\mathrm{T}} = (\Delta \circ A)(\Delta \circ A)^{\mathrm{T}} ,$$

hence $\mathcal{H}(\Delta \circ A)$ and $\mathcal{H}(D_A)$ have the same eigenvalues and their spectrum, if simple, is simultaneously simple. Denote by $(A_{(\mu)})_{\mu\in[m]}$ the m diagonal $d \times d$ blocks of matrix D_A and consider their SVD

$$A_{(\mu)} = U_{(\mu)} \Lambda_{(\mu)} V_{(\mu)} .$$

Consider a simultaneous ε-perturbation of the $\Lambda_{(\mu)}$'s into $\Lambda^\varepsilon_{(\mu)}$ so that all the $\Lambda^\varepsilon_{(\mu)}$'s have distinct diagonal elements, ε-close to the $\Lambda_{(\mu)}$'s. Denote by

$$A^\varepsilon_{(\mu)} = U_{(\mu)} \Lambda^\varepsilon_{(\mu)} V_{(\mu)} .$$

and let D^ε_A be the block diagonal matrix with blocks $(A^\varepsilon_{(\mu)})_{\mu\in[m]}$. Then $\mathcal{H}(D^\varepsilon_A)$ is arbitrarily close to $\mathcal{H}(D_A)$ and has a simple spectrum. Note that $D^\varepsilon_A \Pi$ is ε-close to $\Delta \circ A$, is such that $\mathcal{H}(D^\varepsilon_A \Pi)$ has a simple spectrum and has the same pattern as $\Delta \circ A$ in the sense that :

$$\Delta_{ij} = 0 \quad \Rightarrow \quad (D^\varepsilon_A \Pi)_{ij} = 0 .$$

To emphasize this property, we introduce the $n \times n$ matrix A^ε defined as

$$[A^\varepsilon]_{ij} = \begin{cases} [D_A^\varepsilon \Pi]_{ij} & \text{if } \Delta_{ij} = 1\,, \\ A_{ij} & \text{else} \end{cases}$$

so that $D_A^\varepsilon \Pi = \Delta \circ A^\varepsilon$ and

$$\|\Delta \circ A^\varepsilon - \Delta \circ A\|_F = \|A^\varepsilon - A\|_F \xrightarrow[\varepsilon \to 0]{} 0$$

We can now conclude. Let $\Delta \circ A$, $\Delta \circ B$ be given and $D_A^\varepsilon \Pi = \Delta \circ A^\varepsilon$ and $D_B^\varepsilon \Pi = \Delta \circ B^\varepsilon$ constructed as previously ; notice that $C \mapsto \tilde{\rho}_{k,\ell}(C)$ is continuous. Then

$$\begin{aligned} |\tilde{\rho}_{k,\ell}(B) - \tilde{\rho}_{k,\ell}(A)| &\leq |\tilde{\rho}_{k,\ell}(B^\varepsilon) - \tilde{\rho}_{k,\ell}(B)| + K_\ell \|B_\epsilon - A_\epsilon\|_F + |\tilde{\rho}_{k,\ell}(A_\epsilon) - \tilde{\rho}_{k,\ell}(A)|\,, \\ &\xrightarrow[\epsilon \to 0]{} K_\ell \|B - A\|_F\,. \end{aligned}$$

This conclude the proof of the Lipschitz property. □

Step 3 : uniform estimate for $\mathbb{E}\tilde{R}_k(A)$

As a consequence of the Lipschitz property of $\tilde{R}_k$, $\tilde{R}_k(A)$ if centered is sub-Gaussian if A is a $n \times n$ matrix with i.i.d. $\mathcal{N}(0,1)$ entries. The following estimate easily follows using Tsirelson-Ibragimov-Sudakov inequality ([Boucheron *et al.*, 2013, Theorem 5.5]).

Proposition 24. Under the assumptions of Lemma 23, the following estimate holds true :

$$\mathbb{E} \max_{k \in [n]} \left(\tilde{R}_k - \mathbb{E}\tilde{R}_k \right) \leq K\sqrt{2 \log n}\,.$$

For the proof, see [Bizeul et Najim, 2021, Proposition 2.3].

The rest of the section is devoted to the control of $\mathbb{E}\tilde{R}_k(A)$.

Proposition 25. Under the assumptions of Theorem 17, there exists $n_1 \in \mathbb{N}$ and a constant $C > 0$ such that for all $n \geq n_1$,

$$\sup_{k \in [n]} \left| \mathbb{E}\tilde{R}_k(A) \right| \quad \leq \quad C\,.$$

Démonstration. Recall that $n = d \times m$ and that $\Delta \circ A$ is a block permutation matrix with m blocks $(A^{(\mu)})_{\mu \in [m]}$ of size $d \times d$. We choose a given block $A^{(\mu)}$ and denote by $\mu_1, \cdots, \mu_d$ the d indices corresponding to the rows of block $A^{(\mu)}$ in $\Delta \circ A$. By exchangeability, we have

$$\forall k \in [d]\,, \quad \mathbb{E}\tilde{R}_{\mu_k}(A) = \mathbb{E}\tilde{R}_{\mu_1}(A)\,.$$

Denote by $\mathbf{1}^{(\mu)}$ the $n\times 1$ vector with ones for the indices $(\mu_i)_{i\in[d]}$ and zeros elsewhere. We have

$$
\begin{aligned}
\left|\mathbb{E}\tilde{R}_{\mu_k}(A)\right| &= \left|\frac{1}{d}\sum_{i=1}^{d}\mathbb{E}\tilde{R}_{\mu_i}(A)\right| \\
&= \left|\frac{1}{d}\sum_{i=1}^{d}\mathbb{E}\left(\varphi_d(A)\boldsymbol{e}_{\mu_i}^{\mathrm{T}}\left(\frac{\Delta\circ A}{\sqrt{d}}\right)^2\left(I-\frac{\Delta\circ A}{\alpha\sqrt{d}}\right)^{-1}\mathbf{1}\right)\right|, \\
&= \left|\frac{1}{d}\mathbb{E}\left(\varphi_d(A)\mathbf{1}^{(\mu)\mathrm{T}}\left(\frac{\Delta\circ A}{\sqrt{d}}\right)^2\left(I-\frac{\Delta\circ A}{\alpha\sqrt{d}}\right)^{-1}\mathbf{1}\right)\right|, \\
&\leq \mathbb{E}\left|\varphi_d(A)\frac{\mathbf{1}^{(\mu)\mathrm{T}}}{\sqrt{d}}\left(\frac{\Delta\circ A}{\sqrt{d}}\right)^2\left(I-\frac{\Delta\circ A}{\alpha\sqrt{d}}\right)^{-1}\frac{\mathbf{1}}{\sqrt{d}}\right|. \qquad (1.28)
\end{aligned}
$$

We start by expanding $\left(I-\frac{\Delta\circ A}{\alpha\sqrt{d}}\right)^{-1}$:

$$
\left|\varphi_d(A)\frac{\mathbf{1}^{(\mu)\mathrm{T}}}{\sqrt{d}}\left(\frac{\Delta\circ A}{\sqrt{d}}\right)^2\left(I-\frac{\Delta\circ A}{\alpha\sqrt{d}}\right)^{-1}\frac{\mathbf{1}}{\sqrt{d}}\right| = \left|\varphi_d(A)\sum_{\ell=2}^{\infty}\frac{\mathbf{1}^{(\mu)\mathrm{T}}}{\alpha^{\ell-2}\sqrt{d}}\left(\frac{\Delta\circ A}{\sqrt{d}}\right)^{\ell}\frac{\mathbf{1}}{\sqrt{d}}\right|.
$$

Notice that $(\Delta\circ A)^{\ell}$ is a block matrix constituted of m blocks of size $d\times d$. In particular, among the d rows

$$
\left(\left[(\Delta\circ A)^{\ell}\right]_{ij}\right)_{i\in\{\mu_1,\cdots,\mu_d\},\, j\in[n]},
$$

there exist $\nu_1,\cdots,\nu_d$ (consecutive) indices such that the only non-null entries are

$$
\left(\left[(\Delta\circ A)^{\ell}\right]_{ij}\right)_{i\in\{\mu_1,\cdots,\mu_d\},\, j\in\{\nu_1,\cdots,\nu_d\}}.
$$

Denote by $\mathbf{1}^{(\nu)}$ the $n\times 1$ vector of ones for the indices $(\nu_i)_{i\in[d]}$ and zeroes elsewhere. As a consequence of the previous remark,

$$
\mathbf{1}^{(\mu)\mathrm{T}}(\Delta\circ A)^{\ell}\mathbf{1} = \mathbf{1}^{(\mu)\mathrm{T}}(\Delta\circ A)^{\ell}\mathbf{1}^{(\nu)}
$$

and

$$
\begin{aligned}
\frac{1}{\alpha^{\ell-2}}\left|\frac{\mathbf{1}^{(\mu)\mathrm{T}}}{\sqrt{d}}\left(\frac{\Delta\circ A}{\sqrt{d}}\right)^{\ell}\frac{\mathbf{1}}{\sqrt{d}}\right| &= \frac{1}{\alpha^{\ell-2}}\left|\frac{\mathbf{1}^{(\mu)\mathrm{T}}}{\sqrt{d}}\left(\frac{\Delta\circ A}{\sqrt{d}}\right)^{\ell}\frac{\mathbf{1}^{(\nu)}}{\sqrt{d}}\right|, \\
&\leq \frac{1}{\alpha^{\ell-2}}\left\|\frac{\mathbf{1}^{(\mu)}}{\sqrt{d}}\right\|\left\|\frac{\Delta\circ A}{\sqrt{d}}\right\|^{\ell}\left\|\frac{\mathbf{1}^{(\nu)}}{\sqrt{d}}\right\|, \\
&\leq \left\|\frac{\Delta\circ A}{\alpha\sqrt{d}}\right\|^{\ell-2}\left\|\frac{\Delta\circ A}{\sqrt{d}}\right\|^{2}.
\end{aligned}
$$

Let $\kappa > 0$ as in Proposition 21, $\delta \in (0,1), n_0 \in \mathbb{N}$ as in Lemma 23, then

$$\begin{aligned}\sum_{\ell=2}^{\infty} \frac{1}{\alpha^{\ell-2}} \left| \varphi_d(A) \frac{\mathbf{1}^{(\mu)\mathrm{T}}}{\sqrt{d}} \left(\frac{\Delta \circ A}{\sqrt{d}} \right)^{\ell} \frac{\mathbf{1}}{\sqrt{d}} \right| &\leq \varphi_d(A) \sum_{\ell=2}^{\infty} \left\| \frac{\Delta \circ A}{\alpha\sqrt{d}} \right\|^{\ell-2} \left\| \frac{\Delta \circ A}{\sqrt{d}} \right\|^{2}, \\ &= \varphi_d(A) \left\| \frac{\Delta \circ A}{\sqrt{d}} \right\|^{2} \sum_{\ell=0}^{\infty} \left\| \frac{\Delta \circ A}{\alpha\sqrt{d}} \right\|^{\ell}, \\ &\leq (1+\kappa)^2 \sum_{\ell=2}^{\infty} (1-\delta)^{\ell}, \\ &\leq \frac{(1+\kappa)^2}{\delta}.\end{aligned}$$

Plugging this estimate into (1.28) concludes the proof of the estimation of $|\mathbb{E}\tilde{R}_{\mu_k}(A)|$. This estimate being uniform over $\mu_1, \cdots, \mu_d$ and over all the blocks $(A^{(\mu)})$, the proposition is proved.

□

Proof of lemma 22

Combining Lemma 23, Propositions 24 and 25 one can prove Lemma 22 as in [Bizeul et Najim, 2021, Section 2.3] with minor adaptations.

4 Proof of the feasibility result for the $d \propto n$ model

We assume that Δ_n follows Model (B).

The strategy of proof closely follows the one in [Bizeul et Najim, 2021], with one specific issue to handle : the uniform bound on $\mathbb{E}\tilde{R}_k$. An important property exploited in [Bizeul et Najim, 2021] to establish a uniform bound over $\mathbb{E}\tilde{R}_k$ was the exchangeability of the $\tilde{R}_k$'s (or block exchangeability in the case of Model (A)). There is not enough structure in Model (B) to guarantee this exchangeability (which might not hold).

We carefully address this issue hereafter.

A uniform bound over $\mathbb{E}\tilde{R}_k$ for Model (B)

Proposition 26. Under the assumptions of Theorem 17, uniformly in $k \in [n]$,

$$\mathbb{E}\tilde{R}_k = \mathcal{O}\left(\frac{\alpha}{\sqrt{d}} \right).$$

Proof of Proposition 26 relies on two important facts.

— The fact that almost everywhere $\mathcal{H}(\Delta \circ A)$ has a simple spectrum, hence the Lipschitz function $\|\Delta \circ A\|$ is almost surely differentiable with an explicit formula for the partial derivatives, see (1.19). Details are provided in Appendix 7.

— The Gaussian integration by parts (i.b.p.) formula : If $Z \sim \mathcal{N}(0,1)$ then $\mathbb{E}\, Zf(Z) = \mathbb{E} f'(Z)$. Interestingly, this formula holds for f Lipschitz. In this case, f is absolutely continuous hence almost surely differentiable (see for instance [Hartman et Mikusiński, 1961, Chap. 7, Thm. 4]) with linear growth at infinity.

Recall that $\varphi_d(A) = \varphi\left(\frac{\|\Delta \circ A\|}{\sqrt{d}}\right)$, where φ is defined in (1.13).

Démonstration. In order to get an asymptotic bound over $\mathbb{E}\tilde{R}_k(A)$, we expand its expression :

$$\begin{aligned}
\mathbb{E}\tilde{R}_k(A) &= \mathbb{E}\left[\varphi_d(A)\boldsymbol{e}_k^{\mathrm{T}}\left(\frac{\Delta \circ A}{\sqrt{d}}\right)^2 Q\mathbf{1}\right], \\
&= \frac{1}{d}\sum_{i\in\mathcal{I}_k}\sum_{j\in[n]} \mathbb{E}\left[\varphi_d(A)(\Delta \circ A)_{ki}\left((\Delta \circ A)Q\right)_{ij}\right], \\
&= \frac{\alpha}{\sqrt{d}}\sum_{i\in\mathcal{I}_k}\sum_{j\in[n]} \mathbb{E}\left[\varphi_d(A)(\Delta \circ A)_{ki}\left(-\delta_{ij} + Q_{ij}\right)\right], \\
&= -\frac{\alpha}{\sqrt{d}}\sum_{i\in\mathcal{I}_k} \mathbb{E}\left[\varphi_d(A)(\Delta \circ A)_{ki}\right] + \frac{\alpha}{\sqrt{d}}\sum_{i\in\mathcal{I}_k}\sum_{j\in[n]} \mathbb{E}\left[\varphi_d(A)(\Delta \circ A)_{ki}Q_{ij}\right].
\end{aligned}$$

At this point, we use the Gaussian i.b.p. formula applied to $A \mapsto \varphi_d(A)$ which is Lipschitz and a.s. differentiable with explicit derivative (see (1.19)).

$$\begin{aligned}
\mathbb{E}\tilde{R}_k(A) &= -\frac{\alpha}{\sqrt{d}}\sum_{i\in\mathcal{I}_k} \mathbb{E}\left[\partial_{ki}\varphi_d(A)\right] + \frac{\alpha}{\sqrt{d}}\sum_{i\in\mathcal{I}_k}\sum_{j\in[n]} \mathbb{E}\left[\partial_{ki}\left(\varphi_d(A)Q_{ij}\right)\right], \\
&= -\frac{\alpha}{d}\sum_{i\in\mathcal{I}_k} \mathbb{E}\left[u_k v_i \varphi'\left(\frac{\|\Delta \circ A\|}{\sqrt{d}}\right)\right] + \frac{\alpha}{d}\sum_{i\in\mathcal{I}_k}\sum_{j\in[n]} \mathbb{E}\left[u_k v_i \varphi'\left(\frac{\|\Delta \circ A\|}{\sqrt{d}}\right) Q_{ij}\right] \\
&\qquad + \frac{\alpha}{\sqrt{d}}\sum_{i\in\mathcal{I}_k}\sum_{j\in[n]} \mathbb{E}\left[\varphi_d(A)\left(\partial_{ki}Q_{ij}\right)\right], \\
&=: T_1 + T_2 + T_3 .
\end{aligned}$$

We first handle the term T_1 by Cauchy-Schwarz inequality :

$$|T_1| \le \frac{\alpha}{d}\mathbb{E}\left|u_k \sum_{i\in\mathcal{I}_k} v_i \varphi'\left(\frac{\|\Delta \circ A\|}{\sqrt{d}}\right)\right| \le \frac{\alpha}{d}\mathbb{E}\left[\sqrt{d}\,\|\boldsymbol{v}\|\left|\varphi'\left(\frac{\|\Delta \circ A\|}{\sqrt{d}}\right)\right|\right] = \mathcal{O}\left(\frac{\alpha}{\sqrt{d}}\right).$$

We now handle the term T_2 :

$$\begin{aligned}
|T_2| &\le \frac{\alpha}{d}\mathbb{E}\left[|\varphi'|\left|\sum_i\sum_j v_i Q_{ij}\right|\right], \\
&\le \frac{\alpha}{d}\mathbb{E}\left[|\varphi'| \cdot |\boldsymbol{v}^* Q\mathbf{1}|\right] \le \frac{\alpha}{\sqrt{d}}\sqrt{\frac{n}{d}}\mathbb{E}\left[|\varphi'|\,\|Q\|\right] = \mathcal{O}\left(\frac{\alpha}{\sqrt{d}}\right).
\end{aligned}$$

We finally handle the term T_3. Notice that $\partial_{ki}Q_{ij} = \frac{1}{\alpha\sqrt{d}}Q_{ik}Q_{ij}$ and denote by

$\boldsymbol{\omega} := (Q_{ik} 1_{\mathcal{I}_k^{(i)}})_{i\in[n]}$. Notice that $\|\boldsymbol{\omega}\|^2 \leq \boldsymbol{e}_k^* Q^* Q \boldsymbol{e}_k$ hence $\|\boldsymbol{\omega}\| \leq \|Q\|$ and

$$
\begin{aligned}
|T_3| &= \frac{1}{d} \left| \mathbb{E}\left[\varphi_d(A) \boldsymbol{\omega}^* Q \mathbf{1}\right]\right| \\
&\leq \frac{1}{d} \mathbb{E}\left[\varphi_d(A) \|\boldsymbol{\omega}\| \, \|Q\| \, \|\mathbf{1}\|\right] \;\leq\; \frac{\sqrt{n}}{d} \mathbb{E}\left[\varphi_d(A) \, \|Q\|^2\right] \;=\; \mathcal{O}\left(\frac{1}{\sqrt{d}}\right).
\end{aligned}
$$

Combining these asymptotic notations finally yields :

$$
\mathbb{E}\tilde{R}_k(A) = \mathcal{O}\left(\frac{\alpha}{\sqrt{d}}\right).
$$

□

Notice that even if the bound obtained in Proposition 26 is weaker than the one obtained in Proposition 25 or in [Bizeul et Najim, 2021, Prop. 2.4], it is still sufficient to establish the feasibility under Model (B).

5 Proofs of the stability results

Proof of Theorem 18

The proof is a combination of Takeuchi and Adachi's theorem [Takeuchi, 1996, Theorem 3.2.1] and Proposition 21.

We first recall the definition of Volterra-Liapunov stability, see for instance [Takeuchi, 1996, Section 3.2] : Let B be a $n \times n$ real matrix. B is ***Volterra-Liapunov stable*** if there exists a $n \times n$ positive definite diagonal matrix D such that $DB + B^{\mathrm{T}} D$ is negative definite.

Going back to Eq. (1.2), according to Takeuchi and Adachi's theorem [Takeuchi, 1996, Th. 3.2.1], this LV system has a unique nonnegative and globally stable equilibrium if $M_n - I_n$ is Volterra-Liapunov stable.

We now rely on the asymptotic spectral properties of M_n to study the Volterra-Liapunov stability of $M_n - I_n$. We drop the subscript n in the sequel. Take $D = I$ then

$$
D(M-I) + (M-I)^{\mathrm{T}} D = M + M^{\mathrm{T}} - 2I
$$

is an hermitian matrix. This matrix is negative definite if all its eigenvalues are negative. Given that $M + M^{\mathrm{T}}$ is also hermitian, we just have to check that the spectral radius $\rho\left(M + M^{\mathrm{T}}\right) < 2$. According to Proposition 21 :

$$
\mathbb{P}\left(\rho\left(M + M^{\mathrm{T}}\right) < 2\right) \quad \geq \quad \mathbb{P}\left(\|M\| < 1\right) \quad \xrightarrow[n\to+\infty]{} \quad 1.
$$

Thus, the probability that $M - I$ is Volterra-Liapunov stable converges to 1 as $n \to \infty$. By [Takeuchi, 1996, Th. 3.2.1], this implies that the probability that the LV system (1.2) has a unique nonnegative and globally stable equilibrium converges to 1 as $n \to \infty$.

Proof of Proposition 19

We first prove the first part of the proposition. By Theorem 18, there exists a unique nonnegative globally stable equilibrium to (1.2). If there exists $\epsilon > 0$ such that eventually $\alpha_n \geq (1+\epsilon)\alpha_n^*$ where $\alpha_n^* = \sqrt{2\log n}$, then this equilibrium $\boldsymbol{x}_n$ is positive by Theorem 17 with overwhelming probability as $n \to \infty$.

The rest of the proof closely follows the proof of [Bizeul et Najim, 2021, Corollary 1.4] and is omitted.

6 Conclusion

In this article we study the feasibility and stability of sparse large ecosystems modelled by a large Lotka-Volterra system of coupled differential equations :

$$\frac{d\,\boldsymbol{x}_n}{dt} = \boldsymbol{x}_n(\mathbf{1}_n - \boldsymbol{x}_n + M_n\boldsymbol{x}_n)\,.$$

Our work is motivated by recent research [Busiello *et al.*, 2017] which suggests that in the light of many ecological and biological datasets living networks are often sparse. It also illustrates the interest to study feasibility in relation with the normalization of the interaction matrix's entries beyond the non-sparse full i.i.d. models, and opens perspectives to study models with more structure such as elliptic interactions or patch models.

In the model under investigation, the interaction matrix M_n is a sparse random matrix, where the sparsity is encoded by a patterned matrix Δ_n based on an underlying d_n-regular graph, and the randomness by i.i.d. random variables (matrix A_n) for non-null entries. The single parameter d_n of the regular graph provides an easy one-dimensional parametrization of the connectance of the foodweb.

Our main conclusion is that beyond the standard normalization $1/\sqrt{d_n}$ of the interaction matrix $\Delta \circ A$, which guarantees a bounded norm

$$\left\| \frac{\Delta \circ A}{\sqrt{d}} \right\| = \mathcal{O}(1)\ ,$$

an extra factor $1/\alpha_n$ with $\alpha_n \to \infty$ is needed to reach feasibility. The interaction matrix finally writes

$$M_n = \frac{\Delta_n \circ A_n}{\alpha_n\sqrt{d_n}}$$

and a sharp phase transition occurs at $\alpha_n^* = \sqrt{2\log(n)}$. Interestingly, the same phase transition as in the non-sparse case occurs. In the very sparse setting $\log(n) \leq d_n \ll n$, we rely on an extra block-structure assumption over matrix Δ_n, namely Model (A), to establish the feasibility and the phase transition. Our method of proof crucially relies on this technical assumption which somehow concentrates the non-null entries of the sparse interaction matrix (and its powers) into localized blocks.

However simulations (cf. Fig 1.4) suggest that this block structure assumption is not necessary and could be relaxed. Hence the following :

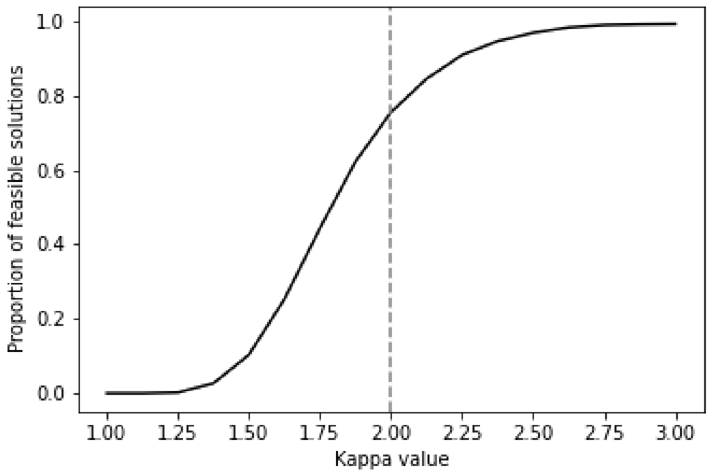


FIGURE 1.4 – Let $n = 15000$, $d = 10$ (notice that $d \geq \log(n) \simeq 9.61$). Matrix Δ_n is drawn at random once for all among the adjacency matrices of d-regular graphs (and a priori does not follow Model (A)). Each point of the curve represents the proportion of feasible solutions $\boldsymbol{x}_n$ of Eq. (1.6) over 1500 realizations of random matrices A_n for different values of κ, with $\alpha_n = \sqrt{\kappa \log(n)}$. The phase transition resemble those of Figure 1.1.

Open question 27. Let Δ_n the adjacency matrix of a deterministic $d_n$regular graph, with $d_n \geq \log(n)$, and A_n a random matrix with i.i.d. $\mathcal{N}(0,1)$ entries. Consider the equation

$$\boldsymbol{x}_n = \mathbf{1}_n + \frac{\Delta_n \circ A_n}{\alpha_n \sqrt{d_n}} \boldsymbol{x}_n \,, \quad \alpha_n \to \infty \,.$$

Is it true that the same phase transition as in Theorem 17 holds ?

7 Appendix : With probability one, the singular values of a sparse random matrix are distinct

We establish hereafter that with probability one the singular values of matrix $\Delta \circ A$ are distinct, a key argument in the proof of Proposition 26 to compute the partial derivatives of $A \mapsto \|\Delta \circ A\|$. The lemma below and its proof are inspired by Nick Cook [Cook, 2021]. We start by mentioning Birkhoff's theorem ([Horn et Johnson, 2013, Theorem 8.7.2]), used in the proof of Cook's lemma 29.

Theorem 28 (Birkhoff [Horn et Johnson, 2013]). A matrix A of size $n \times n$ is doubly stochastic if and only if there are permutation matrices $P_1, \dots, P_N$ of size $n \times n$ and real positive scalars $t_1, \dots, t_N$ such that $t_1 + \cdots + t_N = 1$ and

$$A = t_1 P_1 + \cdots + t_N P_N$$

Moreover, $N \leq n^2 - n + 1$.

Lemma 29 (Cook [Cook, 2021]). Let $n \geq 1$, A_n a $n \times n$ matrix with i.i.d. $\mathcal{N}(0,1)$ entries and Δ_n the adjacency matrix of a d-regular graph. Then with probability one, all the singular values of $\Delta_n \circ A_n$ are distinct.

Remark 7. The original statement of Cook is slightly more general : matrix A_n entries only need a distribution with positive density, and the deterministic matrix Δ_n only needs a generalized diagonal, i.e. $(\Delta_{i\sigma(i)}; \, i \in [n])$ for some $\sigma \in \mathcal{S}_n$, with $n-1$ non null entries.

Démonstration. Let $\mathcal{E}_\Delta$ be the set of matrices with entries supported on the nonzero entries of Δ,

$$\mathcal{E}_\Delta = \left\{ \Delta \circ X \; ; \; X = (X_{ij}) \in \mathbb{R}^{n\times n} \right\} .$$

Thus, $\mathcal{E}_\Delta$ is the support of the law of $\Delta \circ A$. Besides, $\mathcal{E}_\Delta$ is a manifold as a subspace of $\mathbb{R}^{n\times n}$.

Let $\mathcal{R}$ denote the set of matrices with a repeated singular value. It is the set of $n \times n$ matrices X for which the characteristic polynomial p of $X^{\mathrm{T}}X$ has zero discriminant (ρ), see for instance [Cohen, 1993, Section 3.3.2].

$$\mathcal{R} \; = \; \left\{ X \in \mathbb{R}^{n\times n} \; ; \; \rho\left(p\left(X^{\mathrm{T}}X\right)\right) = 0 \right\} \; = \; \left\{ X \in \mathbb{R}^{n\times n} \; ; \; P\left(X\right) = 0 \right\} \; ,$$

where $P : \mathbb{R}^{n\times n} \to \mathbb{R}$ defined by $P(X) = \rho(p(X^{\mathrm{T}}X))$ is a polynomial in the entries of X. It follows that $\mathcal{R}$ is an algebraic manifold in $\mathbb{R}^{n\times n}$.

Hence, $\mathcal{E}_\Delta \cap \mathcal{R}$ is either equal to $\mathcal{E}_\Delta$, or a submanifold. of $\mathcal{E}_\Delta$ of zero Lebesgue measure (under the product measure on $\mathcal{E}_\Delta$).

For the claim, it suffices to show that $\mathcal{E}_\Delta \not\subset \mathcal{R}$ hence to exhibit $Y \in \mathcal{E}_\Delta$ with distinct singular values. By Birkhoff's theorem [Horn et Johnson, 2013, Theorem 8.7.2], the doubly stochastic matrix Δ/d writes

$$\frac{\Delta}{d} = \sum_{\sigma \in \mathcal{S}_n} a_\sigma P_\sigma \; , \quad a_\sigma \geq 0 \quad \text{and} \quad \sum_{\sigma \in \mathcal{S}_n} a_\sigma = 1 \, ,$$

where P_σ is the permutation matrix associated to $\sigma \in \mathcal{S}_n$. There exists in particular σ^* with $a_{\sigma^*} > 0$ and $P_{\sigma^*} \in \mathcal{E}_\Delta$. Let $P_{\sigma^*} = (P_{ij})_{i,j\in[n]}$ then matrix $Y = (iP_{ij})_{i,j\in[n]}$ has distinct singular values $(1, \cdots, n)$. This completes the proof.

□

Chapitre 2

Equilibria of large random Lotka-Volterra systems with vanishing species : a mathematical approach

Ce chapitre est tiré de l'article [Akjouj *et al.*, 2023], écrit en collaboration avec W. Hachem, M. Maïda et J. Najim.

Ecosystems that consist in a large number of species are often modelled as Lotka-Volterra dynamical systems built around a large random interaction matrix. Under some known conditions, global equilibria exist for such dynamical systems. This paper is devoted towards studying rigorously the asymptotic behavior of the distribution of the elements of a global equilibrium vector in the regime of large dimensions. Such a vector is known to be the solution of a so-called Linear Complementarity Problem. It is shown here that the large dimensional distribution of such a solution can be estimated with the help of an Approximate Message Passing (AMP) approach, a technique that has recently aroused an intense research effort in the fields of statistical physics, Machine Learning, or communication theory. Interaction matrices taken from the Gaussian Orthogonal Ensemble, or following a Wishart distribution are considered. Beyond these models, the AMP approach developed in this paper has the potential to address more involved interaction matrix models for solving the problem of the asymptotic distribution of the equilibria.

1 Introduction

In the field of mathematical ecology, the Lotka-Volterra (LV) multi-dimensional differential equations are widely used to model the time evolution of the abundances of n interacting species within an ecosystem [Takeuchi, 1996]. Such systems take the form

$$\dot{x}(t) = x(t) \odot (r + (M - I)\, x(t))\,, \quad x(0) \in (0, \infty)^n,$$

where the vector function $x : [0, \infty) \to \mathbb{R}_+^n = [0, \infty)^n$ represents the abundances of the n species, $\odot$ is the componentwise (or Hadamard) product, $r \in \mathbb{R}_+^n$ is the so-called vector of intrinsic growth rates of the species, and the matrix $M \in \mathbb{R}^{n\times n}$

quantifies the strengths of the food interactions among these species.

When the number of these species is large, the interaction matrix M is often modelled in the ecological literature as a random matrix, rendering the ecological system a large disordered system. Such systems have aroused an important amount of research in the fields of mathematical ecology, borrowing tools from statistical physics, high dimensional probability, or random matrix theory [Akjouj *et al.*, 2022].

In this paper, we shall be interested in the situation where the dynamical system induced by the LV differential equation is well-defined on $\mathbb{R}_+$ and has a globally stable equilibrium vector $x^* = \left[x_i^*\right]_{i=1}^n$. In these conditions, it is of interest to understand the behavior of the distribution of the elements of this vector in the large dimensional regime. Specifically, our purpose is to evaluate the behavior of the probability measure

$$\mu^{x^*} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\delta}_{x_i^*} ,$$

where $\boldsymbol{\delta}_a$ stands for the Dirac measure at a, for large n, which is a random measure defined on the probability space of r and M (we assume in most generality that the intrinsic growth rates are also random). We note here that in general, the equilibirum vector x^* lies at the boundary of $\mathbb{R}_+^n$, i.e. may have vanishing components. In these conditions, it is of particular interest to evaluate the asymptotic behavior of $\|x^*\|_0/n$, which is the proportion of surviving species at the equilibrium (here $\|x\|_0 = \mathrm{card}\,\{i;\ x_i^* > 0\}$ is the number of non-zero elements of the vector x).

Among the “canonical” interaction matrix models considered in the literature devoted to large LV systems are the Gaussian Orthogonal Ensemble (GOE) model, the so-called Ginibre model, or the so-called elliptical model, which can be seen as an interpolation between the first two [Allesina et Tang, 2012]. The asymptotics of μ^{x^*} for these models were recently considered in the physics literature. In [Bunin, 2017], Bunin studied a non-centered elliptical model, and evaluated the large-n behavior of μ^{x^*} with the help of the dynamical cavity method. A similar result was obtained by Galla in [Galla, 2018] by means of generating functionals techniques. Older results related with this problem date back to [Opper et Diederich, 1992, Tokita, 2004]. Heuristic estimates were also proposed in [Clenet *et al.*, 2022a, Clenet *et al.*, 2022b] in the context of the elliptical model.

Many insights are provided by these techniques from a physicist point of view. However, up to our knowledge, no rigorous method for determining the asymptotics of μ^{x^*} can be found in the literature so far. The purpose of this paper is to address this question. The principle of our approach stands as follows. When it exists, the globally stable equilibrium x^* of the LV equation above is known to be the solution of a so-called Linear Complementarity Problem (LCP), see for instance [Takeuchi, 1996, Chap. 3], which consists in finding a real vector that satisfies a system of inequalities involving the matrix M and the vector r (the reader is referred to Section 2.2 below for a quick overview, and to [Cottle *et al.*, 2009] for a comprehensive exposition of the LCP theory). With this at hand, the idea we develop in this paper is that the distribution μ^{x^*} can be estimated for large n with the help of an Approximate

Message Passing (AMP) approach.
The AMP algorithms have been widely popular in the fields of statistical physics, Machine Learning, high-dimensional statistics, or communication theory, as shown in the recent tutorial [Feng *et al.*, 2022] and the references therein. An AMP algorithm produces a sequence of $\mathbb{R}^n$–valued random vectors, say ξ^ℓ, which are iteratively built around a $n \times n$ random so-called measurement matrix. This algorithm is conceived in such a way that for any finite collection $\xi^1, \cdots, \xi^d$ of these vectors, the following joint empirical distribution :

$$\frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\delta}_{\xi_i^1, \cdots, \xi_i^d}$$

converges as $n \to \infty$ to a Gaussian distribution on $\mathbb{R}^d$ that can be characterized by the so-called Density Evolution (DE) equations. In the context of our LV equilibrium problem, it turns out that an AMP algorithm can be designed in such a way that the AMP iterates approximate our LCP solution after an adequate transformation. Thanks to this approximation, the asymptotics of μ^{x^*} can be deduced from the DE equations.

Regarding the statistical model for M, we shall consider in this paper the GOE model [Allesina et Tang, 2012], and the so-called Wishart model. The latter is a particular case of a kernel matrix, which is considered when the interaction between two species depends on a distance between the values of some functional traits attached to these species, see [Akjouj *et al.*, 2022, §4.6] and the references therein, or the recent paper [Rozas *et al.*, 2023].

We also advocate that the LCP/AMP approach for studying μ^{x^*} can be generalized and applied to more involved models for the interaction matrix M. For instance, the recent results of Fan [Fan, 2022] might be used to cover the general rotationally invariant case. In another direction, our approach can be generalized to non-Gaussian interaction matrices with the help of the recent universality results of, *e.g.*, [Bayati *et al.*, 2015, Wang *et al.*, 2022]. Matrices with a variance profile that can be sparse can also be considered. These generalizations are currently under investigation.

This paper is organized as follows. Section 2 is devoted to the problem statement and solution in the GOE case. The main result of this section, Theorem 30, is proven in Section 2.2, where the proof revolves around the LCP formalism for the global equilibrium (recalled in Section 2.2) and a properly designed AMP algorithm (Section 2.2). Section 3 deals with the Wishart case.

2 The GOE case : problem statement, assumptions, and the results

We start by rigorously stating our problem. Let $(A_n)_{n\geq 1}$ be a sequence of random matrices such that A_n is a $n \times n$ GOE matrix. Namely, considering that X_n is a real $n \times n$ matrix with independent $\mathcal{N}(0,1)$ elements,

$$A_n = \frac{X_n + X_n^{\mathrm{T}}}{\sqrt{2}} ,$$

where Y^{T} stands for the transpose of matrix Y. Let $(r_n)_{n=1,2,\ldots}$ be a sequence of random vectors such that $r_n \in [0,\infty)^n$ (notation $r_n \succcurlyeq 0$). We shall assume that r_n and A_n are independent for each n. Let κ be a real non-zero number and denote by M_n the matrix

$$M_n = \frac{A_n}{\kappa\sqrt{n}}\,.$$

For each n, consider the LV dynamical system with trajectories $x_n(t) \succcurlyeq 0$ represented on $t \in \mathbb{R}_+$ by the Ordinary Differential Equation (ODE)

$$\dot{x}_n(t) = x_n(t) \odot (r_n + (M_n - I_n)\, x_n(t))\,, \quad x_n(0) \in (0,\infty)^n, \tag{2.1}$$

where $u \odot v = (u_i v_i)_{1\leq i\leq n}$ for $u = (u_i)$ and $v = (v_i)$ $n \times 1$ vectors. We shall consider the following assumption :

Assumption 1. $\kappa > 2$.

One of the most well-known results of the large random matrix theory is that the spectral measure of the matrix $A_n/\sqrt{n}$ weakly converges in the almost sure sense to the so-called semi circle law (supported by the interval $[-2,2]$), and that the spectral norm $\|A_n/\sqrt{n}\|$ almost surely (a.s.) converges to 2, the edge of the semi circle law, as $n \to \infty$ [Pastur et Shcherbina, 2011].

Independently of the struture of M_n, it is moreover known that when the condition $\|M_n\| < 1$ is satisfied, the ODE (2.1) admits an unique solution on $\mathbb{R}_+$, with a bounded trajectory, for an arbitrary initial value $x_n(0) \in (0,\infty)^n$ [Li *et al.*, 2009]. Since Assumption 1 ensures that with probability one, $\|M_n\| < 1$ for all large n, we obtain that the ODE (2.1) has an unique solution on $\mathbb{R}_+$ for these values of n.

The existence of the solution being settled, we now focus on the possible equilibrium points of the ODE (2.1). It is known, as recalled with more detail in the next section, that when $\|M_n\| < 1$, the ODE (2.1) has a globally stable equilibrium point x_n^* in the classical sense of the Lyapounov theory [Takeuchi, 1996, Chapter 3]. Our purpose is to study the distribution of the elements of x_n^* for large n.

Let us introduce some notions intended to make precise what we mean by the asymptotic distribution of the elements of x_n^*. Given an integer $d > 0$, the notation $\mathcal{P}_2(\mathbb{R}^d)$ refers to the Wasserstein space of the probability measures on $\mathbb{R}^d$ with finite 2$^{\text{nd}}$ moment. We recall that the convergence in $\mathcal{P}_2(\mathbb{R}^d)$, that we denote as $\mu_n \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R}^d)} \boldsymbol{\mu}$, amounts to the weak convergence complemented with the convergence of the second moment of μ_n to $\boldsymbol{\mu}$. An equivalent characterization of the convergence in $\mathcal{P}_2(\mathbb{R}^d)$ will be useful in this paper. A function $\varphi : \mathbb{R}^d \to \mathbb{R}$ is said a *pseudo-Lipschitz* function with Lipschitz constant $L > 0$ and degree 2 when it satisfies for all $x, y \in \mathbb{R}^d$ the inequality

$$|\varphi(x) - \varphi(y)| \leq L\|x - y\|(1 + \|x\| + \|y\|)$$

(here, $\|\cdot\|$ is the Euclidean norm). We denote as $\mathrm{PL}_d(2)$ (or $\mathrm{PL}(2)$ when $d = 1$) the set of the $\mathbb{R}^d \to \mathbb{R}$ pseudo-Lipschitz functions with degree 2. The convergence

$\mu_n \xrightarrow{\mathcal{P}_2(\mathbb{R}^d)} \boldsymbol{\mu}$ amounts to

$$\int \varphi d\mu_n \xrightarrow[n\to\infty]{} \int \varphi d\boldsymbol{\mu}$$

for each function $\varphi \in \mathrm{PL}_d(2)$ [Feng *et al.*, 2022, Villani, 2009].

Further, given a d–uple of vectors $(x_n^1, \cdots, x_n^d) \in (\mathbb{R}^n)^d$ with $x_n^k = \left[x_{1,n}^k, \cdots, x_{n,n}^k\right]^{\mathrm{T}}$, we denote as $\mu^{x_n^1,\cdots,x_n^d} \in \mathcal{P}_2(\mathbb{R}^d)$ the joint distribution of their elements, defined as

$$\mu^{x_n^1,\cdots,x_n^d} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\delta}_{(x_{i,n}^1,\cdots,x_{i,n}^d)},$$

where $\boldsymbol{\delta}_{(a_1,\cdots a_d)}$ is the Dirac measure at $(a_1, \cdots, a_d)$. Our purpose is to characterize the asymptotic behavior of the random probability measure $\mu^{x_n^*}$ in the space $\mathcal{P}_2(\mathbb{R})$.

To this end, we need an assumption on the distributions of the vectors r_n :

Assumption 2. There exists a measure $\bar{\mu} \in \mathcal{P}_2(\mathbb{R})$ such that $\bar{\mu} \neq \boldsymbol{\delta}_0$, and

$$(a.s.) \quad \mu^{r_n} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \bar{\mu}\ .$$

An equivalent formulation for this assumption is that

$$\int \varphi d\mu^{r_n} = \frac{1}{n}\sum_{i=1}^n \varphi(r_{i,n}) \to \int \varphi d\bar{\mu}$$

for each function $\varphi \in \mathrm{PL}(2)$, where $r_n = [r_{1,n}, \ldots, r_{n,n}]^{\mathrm{T}}$. Note that $\bar{\mu}$ is supported by $\mathbb{R}_+$, similarly to the μ^{r_n}.

The main result of this section is stated as follows :

Theorem 30. Let Assumption 1 hold true. Then, $\|M_n\| < 1$ for all large n with probability one. For each of these n, the ODE (2.1) is defined on $\mathbb{R}_+$ and has a globally stable equilibrium x_n^*. For the other n, put $x_n^* = 0$. The distribution $\mu^{x_n^*}$ is a $\mathcal{P}_2(\mathbb{R})$–valued random variable on the probability space where A_n and r_n are defined.

Let Assumption 2 hold true. Let $\bar{r} \geq 0$ be a random variable such that $\bar{\mu} = \mathcal{L}(\bar{r})$. Let $\bar{Z}$ be a $\mathcal{N}(0,1)$ random variable independent of $\bar{r}$. Then, the system of equations

$$\kappa = \delta + \gamma/\delta, \tag{2.2a}$$

$$\sigma^2 = \frac{1}{\delta^2}\mathbb{E}_{\bar{r}}\mathbb{E}_{\bar{Z}}\left(\sigma\bar{Z} + \bar{r}\right)_+^2, \tag{2.2b}$$

$$\gamma = \mathbb{E}_{\bar{r}}\mathbb{P}_{\bar{Z}}\left[\sigma\bar{Z} + \bar{r} > 0\right], \tag{2.2c}$$

admits an unique solution $(\delta, \sigma^2, \gamma)$ in $(1/\sqrt{2}, \infty) \times (0, \infty) \times (0, 1)$.

Let Assumptions 1 and 2 hold true. Then, the convergence

$$\mu^{x_n^*} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}\left(\left(1 + \gamma/\delta^2\right)\left(\sigma\bar{Z} + \bar{r}\right)_+\right) \quad \text{almost surely} \tag{2.3}$$

holds true.

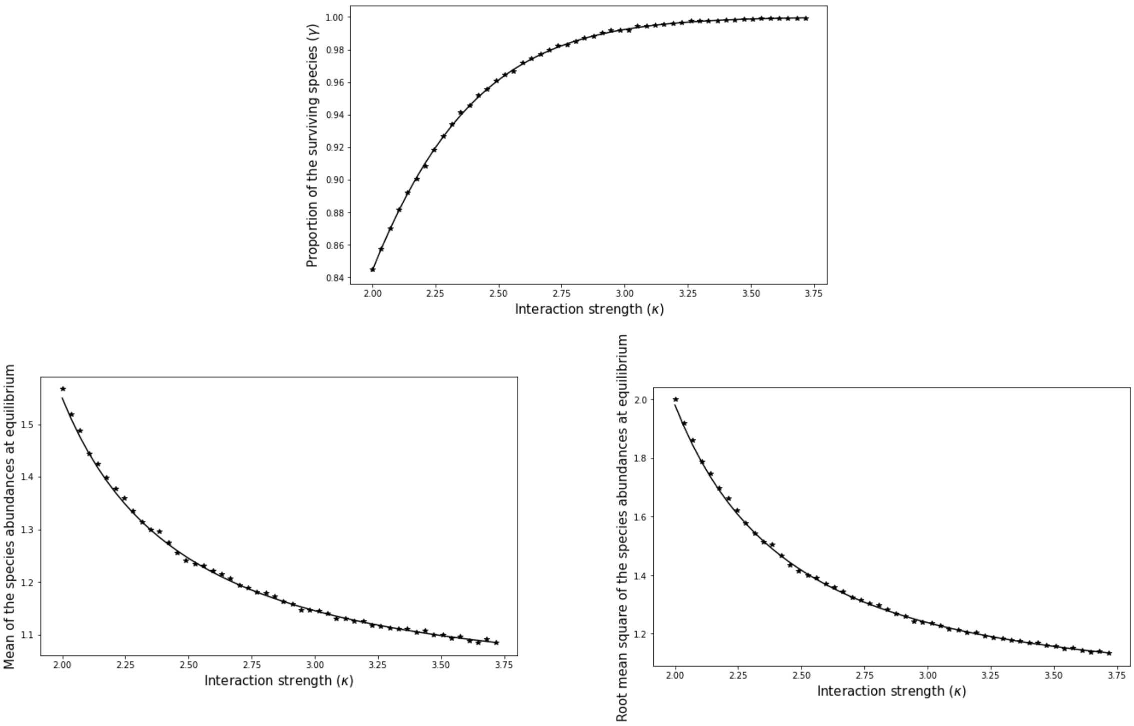


FIGURE 2.1 – The plots represent a comparison between the theoretical solutions γ, the mean and the root mean square at equilibrium of the abundances arising from system 2.2 and their empirical Monte Carlo counterpart (the star marker) as functions of the interaction strength κ. Matrix M_n has size $n = 500$ and the number of Monte Carlo experiments is 100. When interaction κ^{-1} increases, the proportion of surviving species γ decrease but their mean and variance increase.

There is a strong matching between the parameters obtained by solving the system 2.2 and their empirical counterparts obtained by Monte-Carlo simulations. This is illustrated in Fig. 2.1 by comparing γ, the mean and the root mean square of the abundances species at equilibrium.

The theorem calls for some remarks.

Remark 8. The result of Theorem 30 can be obtained at a physical level of rigor by the dynamical cavity method, see [Bunin, 2017] that considers the particular case where $\bar{r}$ is a constant. A similar remark can be made regarding the reference [Galla, 2018].

Remark 9. Write $x_n^* = \left[x_{1,n}^*, \dots, x_{n,n}^*\right]^{\mathrm{T}}$. Inspecting (2.2c) and (2.3), the parameter γ can be seen as "the limit proportion of surviving species", in the sense that

$$\sup_{\varphi} \operatorname*{aslim}_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \varphi(x_{i,n}^*) = \gamma,$$

where $\sup_\varphi$ is taken on the set of functions $\{\varphi : \mathbb{R} \to [0,1] \text{ continuous}, \ \varphi(0) = 0\}$. Note that since the function $\mathbb{1}_{x>0}$ is not continuous, the convergence (2.3) does not imply that $\|x_n^*\|_0/n$ converges to γ, whatever is the sense of this convergence.
Up to our knoweldge, there is no rigorous approach in the literature that solves the problem of the asymptotic behavior of $\|x_n^*\|_0/n$.

	Bunin ([Bunin, 2017])	Clenet, Massol and Najim ([Clenet *et al.*, 2022b])	GOE model
Number of species	S	n	n
Abundances of species i	N_i	x_i	x_i
Carrying capacity	K_i	1	1
Intrinsec growth rate	r_i	r_i	r_i
Proportion of surviving species	Φ	p^*	γ

TABLE 2.1 – Correspondence between the initial notations of the Lotka-Volterra model of each of the three papers.

Remark 10 (Behavior of surviving species proportion)**.** Recalling that $\bar{r} \geq 0$ and $\mathbb{P}[\bar{r} > 0] > 0$, we easily see from Equation (2.2c) $\gamma > 1/2$: More than half the species survive.

Furthermore, an easy calculation involving Equations (2.2b) and (2.2c) shows that γ does not change if we replace $\bar{r}$ with $c\bar{r}$ where $c > 0$ is an arbitrary constant. In particular, γ does not depend on $\mathbb{E}_{\bar{r}}$. In practice, this implies that in the framework of Model (2.1), the empirical mean of the free growth rate vector r_n has little influence on the surviving species proportion.

2.1 Link with Bunin's article

In his paper [Bunin, 2017], Bunin studies the effects of migration and species interactions for a community, with population dynamics described by a Lotka-Volterra system of equations. In particular, diversity and species abundance distributions are two topics of common concern.

For that reason, the aim of this section is to explain how Bunin's model fits with the following mathematical models : the GOE model and the one given by [Clenet *et al.*, 2022b], which corresponds to a model where the interactions are i.i.d centred random variables.

The Lotka-Volterra system of S species studied by Bunin is :

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \frac{r_i}{K_i} N_i \left(K_i - N_i - \sum_{j \neq i} \alpha_{ij} N_j \right) \ , \quad \forall i \in \{1, \dots, S\} \, .$$

where N_i is the abundance of species i, $\alpha_{ij} = \frac{\mu}{S} + \sigma a_{ij}$, for $i \neq j$, encode interspecies interactions, r_i are the intrinsic growth rates, K_i are the carrying capacities.

To simplify the interpretation, one puts $r_i = 1$ and $K_i = 1$.

Table 2.1 matches the initial parameters of the Lotka-Volterra system of each of the three papers. This will make easier to the reader the transition between the models.

Besides, Bunin chooses the variables a_{ij} as centred ($\mathbb{E}(a_{ij}) = 0$), with variance equal to $\frac{1}{S}$ and correlation-factor $\mathbb{E}\left(a_{ij} a_{ji}\right) = \frac{\gamma}{S}$.

At equilibrium, this system is rewritten :

$$n_i\left(\lambda_i - un_i - \sum_{j\neq i} a_{ij}n_j + h\right) = 0\,,\quad \forall i\in\{1,\dots,S\}\,, \tag{2.4}$$

where

$$n_i = \frac{N_i}{<N_j>}\,,\quad u = \frac{1-\mu/S}{\sigma}\,,\quad \lambda_i = \frac{K_i - 1}{\sigma < N_j >}\,,$$
$$h = -\frac{\mu - 1/<N_j>}{\sigma} \quad \text{and} \quad <N_j> = \frac{1}{S}\sum_j N_j\,.$$

As $K_i = 1$, it implies that $\lambda_i = 0$ and $\sigma_\lambda^2 =< \lambda_i^2 >= 0$.

In the dynamical cavity method, it is common to add a new species to the existing system and comparing the properties of the solution with S species to that with $S+1$ species.

Starting from the previous system (2.4), one assumes that the normalized abundances n_i of the species in the pool $i \in \{1,\dots,S\}$ are known. Then, a new species with interactions $(a_{0i}, a_{i0})_{i=1,\dots,S}$ is introduced. A small perturbations ξ_i is added to λ_i in the system for each species. The response to this perturbation is defined by :

$$v_{ij} = \left(\frac{\partial n_i}{\partial \xi_j}\right)_{\xi_j=0} \qquad \text{and} \qquad v =< v_{ii} > \ .$$

Denote by $Z \sim \mathcal{N}(0,1)$ a standard Gaussian random variable, ρ_Z the cumulative Gaussian distribution function of Z, $\Delta = h/\sqrt{q+\sigma_\lambda^2}$ and $\omega_k(\Delta) = \int_{-\Delta}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}(z+\Delta)^k \mathring{d}z$. We can now specify the system (10) of [Bunin, 2017] :

$$\Phi = \omega_0(\Delta) = \mathbb{P}\left(Z > -\Delta\right)\,, \tag{2.5}$$

$$u - \gamma v = \sqrt{q+\sigma_\lambda^2}\,\omega_1(\Delta) = \sqrt{q+\sigma_\lambda^2}\,\mathbb{E}\left[(Z+\Delta)_+\right]\,, \tag{2.6}$$

$$(u-\gamma v)^2 = \left(1+\frac{\sigma_\lambda^2}{q}\right)\omega_2(\Delta) = \left(1+\frac{\sigma_\lambda^2}{q}\right)\mathbb{E}\left[(Z+\Delta)_+^2\right]\,. \tag{2.7}$$

From Bunin to Clenet-Massol-Najim

The model proposed by [Clenet *et al.*, 2022b] corresponds to Bunin's model with γ the correlation factor put to 0.

Henceforth, we use table 2.2 to rewrite Bunin's system (2.4) with the notations used by Clenet and coauthors.

The first equation (2.5) of Bunin becomes :

$$p^* = \mathbb{P}\left(Z > -\delta^*\right) = 1 - \rho_Z(-\delta^*)\,,\quad \text{i.e}\quad \rho_Z^{-1}(1-p^*) = -\delta^*\,.$$

Thus,

$$\begin{aligned} 0 &= \rho_Z^{-1}(1-p^*) + \delta^*\,,\\ &= \rho_Z^{-1}(1-p^*) + \frac{\alpha}{\sigma^*}\left(\frac{1}{\sqrt{p^*}} + \mu m\sqrt{p^*}\right)\,,\\ &= \sigma^*\sqrt{p^*}\rho_Z^{-1}(1-p^*) + \alpha\left(1+\mu m p^*\right)\,, \end{aligned}$$

Bunin ([Bunin, 2017])	Clenet, Massol and Najim ([Clenet *et al.*, 2022b])
$\mathbb{E}(\alpha_{ij}) = \mu$	$-\mu$
$u = \frac{1-\mu/S}{\sigma}$	$\alpha\left(1+\frac{\mu}{n}\right)$
$< N_i >$	p^*m^*
$h = \frac{\frac{1}{<N_i>}-\mu}{\sigma}$	$\alpha\left(\frac{1}{p^*m^*}+\mu\right)$
$< N_i^2 >$	$p^*\sigma^{*2}$
$q =< n^2 >$	$\frac{\sigma^{*2}}{p^*m^{*2}}$
$\Delta = \frac{h}{\sqrt{q+\sigma_\lambda^2}}$	$\delta^* = \frac{\alpha}{\sigma^*}\left(\frac{1}{\sqrt{p^*}}+\mu m\sqrt{p^*}\right)$

TABLE 2.2 – For Bunin's model with correlation between the random variables a_{ij} and a_{ji} put to 0 ($\gamma = 0$), this table matches the parameters involved in system (2.4) with parameters of Clenet, Massol and Najim articles.

which is exactly [Clenet *et al.*, 2022b, Equation (14)]).

The second equation (2.6) gives :

$$\alpha\left(1+\frac{\mu}{n}\right) = \frac{\sigma^*}{\sqrt{p^*m^*}}\mathbb{E}\left[(Z+\delta^*)_+\right] . \tag{2.8}$$

We, first, study the following conditional expectation :

$$\mathbb{E}\left(Z|Z>-\delta^*\right) = \frac{\mathbb{E}\left(Z\mathbb{1}_{Z>-\delta^*}\right)}{\mathbb{P}\left(Z>-\delta^*\right)} \quad \text{i.e.} \quad \mathbb{E}\left(Z\mathbb{1}_{Z>-\delta^*}\right) = p^*\mathbb{E}\left(Z|Z>-\delta^*\right) ,$$

and

$$\mathbb{E}\left(Z\mathbb{1}_{Z>-\delta^*}\right) = \mathbb{E}\left((Z+\delta^*)_+\right) - \delta^*\mathbb{P}\left(Z>-\delta^*\right)$$
$$\text{i.e.} \quad \mathbb{E}\left((Z+\delta^*)_+\right) = p^*\left(\mathbb{E}\left(Z|Z>-\delta^*\right)+\delta^*\right) .$$

Equation (2.8) is rewritten :

$$\begin{aligned}\left(1+\frac{\mu}{n}\right)m^* &= \frac{\sigma^*}{\alpha\sqrt{p^*}}\left(\mathbb{E}\left(Z\mathbb{1}_{Z>-\delta^*}\right)+\delta^*p^*\right) ,\\ &= \frac{\sigma^*}{\alpha\sqrt{p^*}}\left(p^*\mathbb{E}\left(Z|Z>-\delta^*\right)+\delta^*p^*\right) ,\\ &= \frac{\sigma^*\sqrt{p^*}}{\alpha}\mathbb{E}\left(Z|Z>-\delta^*\right)+\frac{\sigma^*\sqrt{p^*}}{\alpha}\delta^* ,\\ &= \frac{\sigma^*\sqrt{p^*}}{\alpha}\mathbb{E}\left(Z|Z>-\delta^*\right)+1+\mu m p^* .\end{aligned}$$

The equations of Clenet, Massol and Najim are obtained at equilibrium, i.e. when n goes to ∞. From there, since factor $1+\mu/n = 1+o(1)$, we get [Clenet *et al.*, 2022b, Equation (15)].

Rewrite the last one (2.7) gives :

$$\alpha^2\left(1+\frac{\mu}{n}\right)^2 = \mathbb{E}\left[(Z+\delta)_+^2\right] . \tag{2.9}$$

Bunin ([Bunin, 2017])	GOE model
$\mathbb{E}(\alpha_{ij}) = \mu$	0
$\sqrt{S\text{Var}(\alpha_{ij})} = \sigma$	$\frac{1}{\kappa}$
$u = \frac{1-\mu/S}{\sigma}$	κ
$< N_i^2 >$	$\left(\delta + \frac{\gamma}{\delta}\right)^2 \sigma^2$
$\Delta = \frac{h}{\sqrt{q+\sigma_\lambda^2}}$	$\frac{\kappa}{(\delta+\gamma/\delta)\sigma} = \frac{1}{\sigma}$

TABLE 2.3 – For Bunin's model with correlation between the random variables a_{ij} and a_{ji} put to 1 ($\gamma = 1$), this table matches the parameters involved in system (2.4) with parameters of the GOE model.

To express this last expectation differently, one uses the conditional expectation :

$$\begin{aligned}
\mathbb{E}\left(Z^2|Z > -\delta^*\right) &= \frac{1}{p^*}\mathbb{E}\left(\left[(Z+\delta^*)_+^2 - 2\delta(Z+\delta^*) + \delta^{*2}\right]\mathbb{1}_{Z>-\delta^*}\right) , \\
&= \frac{1}{p^*}\mathbb{E}\left((Z+\delta^*)_+^2\right) - 2\frac{\delta^*}{p^*}\mathbb{E}\left((Z+\delta^*)_+\right) + \delta^{*2} .
\end{aligned}$$

Yet, equation (2.9) can be rewritten :

$$\begin{aligned}
\alpha^2\left(1+\frac{\mu}{n}\right)^2 &= p^*\mathbb{E}\left(Z^2|Z > -\delta^*\right) + 2\delta^*\mathbb{E}\left((Z+\delta^*)_+\right) - \delta^{*2}p^* , \\
&= p^*\mathbb{E}\left(Z^2|Z > -\delta^*\right) + 2\delta^*p^*\left(\mathbb{E}\left(Z|Z > -\delta^*\right) + \delta^*\right) - \delta^{*2}p^* , \\
&= p^*\mathbb{E}\left(Z^2|Z > -\delta^*\right) + 2\delta^*p^*\mathbb{E}\left(Z|Z > -\delta^*\right) + \delta^{*2}p^* , \\
&= p^*\mathbb{E}\left(Z^2|Z > -\delta^*\right) + 2\frac{\alpha p^*}{\sigma^*}\left(\frac{1}{\sqrt{p^*}} + \mu m\sqrt{p^*}\right)\mathbb{E}\left(Z|Z > -\delta^*\right) \\
&\qquad + \frac{\alpha^2}{\sigma^{*2}}\left(1+\mu m p^*\right)^2 .
\end{aligned}$$

In other words,

$$\begin{aligned}
&\sigma^{*2}\left(1+\frac{\mu}{n}\right)^2 \\
&= \frac{\sigma^{*2}p^*}{\alpha^2}\mathbb{E}\left(Z^2|Z > -\delta^*\right) + 2\frac{\sigma^*p^*}{\alpha}\left(\frac{1}{\sqrt{p^*}} + \mu m\sqrt{p^*}\right)\mathbb{E}\left(Z|Z > -\delta^*\right) + \left(1+\mu m p^*\right)^2 , \\
&= \frac{\sigma^{*2}p^*}{\alpha^2}\mathbb{E}\left(Z^2|Z > -\delta^*\right) + 2\frac{\sigma^*\sqrt{p^*}}{\alpha}\left(1+\mu m p^*\right)\mathbb{E}\left(Z|Z > -\delta^*\right) + \left(1+\mu m p^*\right)^2 ,
\end{aligned}$$

which corresponds asymptotically (where $\mu/n = o(1)$) to Equation (16) of Clenet and coauthors.

From Bunin to GOE

The GOE model refers to Bunin's model with high correlation factor $\gamma = 1$. One reminds [Bunin, 2017, Eq.(11)] which concerns the response to the perturbation factor v due to the arrival of a new species by :

$$v = \frac{\Phi}{u - \gamma v}, \qquad \text{which becomes (using table 2.1) : } \; v(\kappa - v) = \gamma \; .$$

The correspondence between the last equality and equation (2.2a) involves either $v = \delta$ or $v = \kappa - \delta$.
While parameter δ seems to be previously an artefact of the AMP method, it is related to the indirect effects of the existing community on the new species.

In the same way than previously, rewriting Bunin's system (equation (2.5) and (2.7)) using this chapter notation (column GOE of table 2.3) allows to recover equations (2.2c) and (2.2b) when $v = \kappa - \delta$. We briefly detail the calculation for (2.7), which becomes using our notations :

$$(\kappa - (\kappa - \delta))^2 = \mathbb{E}\left[\left(Z + \frac{1}{\sigma}\right)_+^2\right],$$

that is equation (2.2b) with $r = 1$,

$$\sigma^2 = \frac{1}{\delta^2}\mathbb{E}\left[(\sigma Z + 1)_+^2\right] .$$

2.2 Proof of Theorem 30

Characterization of x_n^* through a LCP

One reminds that the stable equilibria of the ODE (2.1) can be seen as solutions of LCP's in the theory of mathematical programming [Takeuchi, 1996]. We recall herein some elements of the LCP theory in order to establish the first part of Theorem 30.

Given a matrix $B \in \mathbb{R}^{n\times n}$ and a vector $c \in \mathbb{R}^n$, the LCP problem, denoted as LCP(B, c), consists in finding a couple of vectors $(z, y) \in \mathbb{R}^n \times \mathbb{R}^n$ such that

$$\begin{aligned} & y = Bz + c \succcurlyeq 0, \\ & z \succcurlyeq 0, \\ & \mathrm{supp}(z) \cap \mathrm{supp}(y) = \emptyset, \end{aligned}$$

where supp(z) is the support (set of indices of non-zero elements) of the vector z. When a solution (z, y) exists and is unique, we write $z = \mathrm{LCP}(B, c)$.

Obviously, an equilibrium $\bar{x}$ of our LV dynamical system (2.1) is defined by the system

$$\begin{aligned} & \bar{x} \succcurlyeq 0, \\ & \bar{x} \odot (r_n + (M_n - I_n)\,\bar{x}) = 0. \end{aligned}$$

Furthermore, the supplementary condition

$$r_n + (M_n - I_n)\,\bar{x} \preccurlyeq 0$$

(with the obvious meaning of $\preccurlyeq$) turns out to be a necessary condition for the equilibrium $\bar{x}$ to be stable in the classical sense of Lyapounov theory (see [Takeuchi, 1996, Chapter 3] to recall the different notions of stability, and [Takeuchi, 1996, Theorem 3.2.5] for this result). These three conditions can be rewritten as

$$\begin{aligned} &\bar{w} = (I_n - M_n)\,\bar{x} - r_n \succcurlyeq 0, \\ &\bar{x} \succcurlyeq 0, \\ &\operatorname{supp}(\bar{w}) \cap \operatorname{supp}(\bar{x}) = \emptyset, \end{aligned} \tag{2.10}$$

in other words, the couple $(\bar{x}, \bar{w})$ solves the problem $\mathrm{LCP}(I_n - M_n, -r_n)$.

Theorem 30 is more specific, since it asserts that the ODE (2.1) has a (unique) *globally stable* equilibrium when $\|M_n\| < 1$. Recalling that M_n is a symmetric matrix, this can be obtained from the following result :

Proposition 31 (Lemma 3.2.2 and Theorem 3.2.1 of [Takeuchi, 1996]). Given a symmetric matrix $B \in \mathbb{R}^{n\times n}$ and a vector $c \in \mathbb{R}^n$, consider the LV ODE

$$\dot{x}(t) = x(t) \odot (c + Bx(t))\,, \quad t \geq 0,\ x(0) \in (0,\infty)^n. \tag{2.11}$$

Then, the LCP problem $\mathrm{LCP}(-B, -c)$ has an unique solution for each $c \in \mathbb{R}^n$ if and only if B is negative definite (notation $B < 0$). On the domain where $B < 0, c \in \mathbb{R}^n$, the function $x = \mathrm{LCP}(-B, -c)$ is measurable. Moreover, if $B < 0$, then for each $c \in \mathbb{R}^n$, the ODE (2.11) has a globally stable equilibrium x given as $x = \mathrm{LCP}(-B, -c)$.

It is clear that under Assumption 1, $(I_n - M_n)$ is positive definite with probability one for all large n, and the vector x_n^* defined as

$$x_n^* = \begin{cases} \mathrm{LCP}(I_n - M_n, -r_n) & \text{if } \|M_n\| < 1, \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

satisfies the statement of the first part of Theorem 30.

We now turn to the second part of this theorem, devoted towards the system of equations (2.2).

Existence and uniqueness of the solution of System (2.2)

We begin with the following result. The second part of this lemma will be used in the next section.

Lemma 32. For a given $\delta > 0$, Equation (2.2b) admits a solution σ^2 if and only if $\delta > 1/\sqrt{2}$, and the case being, this solution is unique, and is written $\sigma^2(\delta)$.

Assume $\delta > 1/\sqrt{2}$. Starting with an arbitrary $\sigma_0 \geq 0$, consider the iterates in $t = 0, 1, \ldots$

$$\sigma_{t+1}^2 = \frac{1}{\delta^2}\mathbb{E}_{\bar{r}}\mathbb{E}_{\bar{Z}}\left[\left(\sigma_t\bar{Z} + \bar{r}\right)_+^2\right].$$

Then, $\sigma_t^2 \to_t \sigma^2(\delta)$.

Démonstration. Writing $\mathbb{E} = \mathbb{E}_{\bar{r}}\mathbb{E}_{\bar{Z}}$, consider the function $f(\sigma^2) \triangleq \mathbb{E}(\sigma\bar{Z} + \bar{r})_+^2$. Then, Equation (2.2b) is written

$$\frac{f(\sigma^2)}{\delta^2} = \sigma^2. \tag{2.13}$$

Let us compute $df(\sigma^2)/d\sigma^2$. Fixing $\sigma > 0$ and assuming that the real number h has a small enough absolute value, the function

$$\varphi_h(z,r) \triangleq \frac{((\sigma+h)z+r)_+^2 - (\sigma z + r)_+^2}{h}$$

is easily shown to satisfy $|\varphi_h(z,r)| \leq |z| \left(((\sigma+h)z+r)_+ + (\sigma z + r)_+\right)$. Thus, by Assumption 2 and the Dominated Convergence Theorem, it holds that

$$\begin{aligned}\frac{df(\sigma^2)}{d\sigma^2} &= \frac{1}{2\sigma}\frac{df(\sigma^2)}{d\sigma} = \frac{1}{2\sigma}\mathbb{E}\left(\frac{d(\sigma\bar{Z}+\bar{r})_+^2}{d\sigma}\right) = \frac{1}{\sigma}\mathbb{E}\left[\bar{Z}(\sigma\bar{Z}+\bar{r})_+\right] , \\ &= \mathbb{E}\left[\bar{Z}^2\mathbb{1}_{\bar{Z}\geq -\bar{r}/\sigma}\right] + \frac{\bar{r}}{\sigma}\mathbb{E}\left[\bar{Z}\mathbb{1}_{\bar{Z}\geq -\bar{r}/\sigma}\right].\end{aligned}$$

By doing an Integration by Parts involving the Gaussian law, we obtain that

$$\mathbb{E}\left(\bar{Z}^2\mathbb{1}_{\bar{Z}\geq -\bar{r}/\sigma}\right) = \mathbb{E}_{\bar{r}}\left[\bar{\rho}\Big(-\frac{\bar{r}}{\sigma}\Big) - \frac{\bar{r}}{\sigma}\frac{\exp(-\bar{r}^2/2\sigma^2)}{\sqrt{2\pi}}\right],$$

where

$$\bar{\rho}(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp(-t^2/2)\,dt.$$

We also have

$$\frac{\bar{r}}{\sigma}\mathbb{E}\left(\bar{Z}\mathbb{1}_{\bar{Z}\geq -\bar{r}/\sigma}\right) = \mathbb{E}_{\bar{r}}\left[\frac{\bar{r}}{\sigma}\frac{\exp(-\bar{r}^2/2\sigma^2)}{\sqrt{2\pi}}\right],$$

thus,

$$\frac{df(\sigma^2)}{d\sigma^2} = \mathbb{E}_{\bar{r}}\left(\bar{\rho}\Big(-\frac{\bar{r}}{\sigma}\Big)\right).$$

We observe from this equation that $df(\sigma^2)/d\sigma^2$ decreases from 1 asymptotically to $1/2$ as σ^2 increases from 0 to ∞.

The function

$$g_\delta(\sigma^2) \triangleq \frac{f(\sigma^2)}{\delta^2} - \sigma^2$$

satisfies $g_\delta(0) = \delta^{-2}\mathbb{E}_{\bar{r}}\bar{r}^2 > 0$. If $\delta \leq 1/\sqrt{2}$, then $dg_\delta(\sigma^2)/d\sigma^2 \geq 2df(\sigma^2)/d\sigma^2 - 1 > 0$ for each $\sigma^2 > 0$, therefore, $g_\delta(\cdot)$ is increasing and the equation $g_\delta(\sigma^2) = 0$, *i.e.*, Equation (2.2b), has no solution.

We now consider the case $\delta > 1/\sqrt{2}$. If $\delta > 1$, then, $\delta^{-2}\sup_{\sigma^2} df(\sigma^2)/d\sigma^2 < 1$, and the results of the lemma are consequences of Banach's fixed point theorem.

Let us consider the case where $\delta \in (1/\sqrt{2}, 1]$, or, equivalently, $\delta^{-2} \in [1,2)$. There exists a point $\tilde{\sigma}^2 > 0$ that satisfies the following properties :

1. $\delta^{-2}\sup_{\sigma^2\in[\tilde{\sigma}^2,\infty)} df(\sigma^2)/d\sigma^2 < 1$.
2. $g_\delta(\sigma^2) > 0$ on $[0,\tilde{\sigma}^2]$. To see why this is possible, remember that $g_\delta(0) > 0$ and $dg(\sigma^2)/d\sigma^2 = \delta^{-2}df(\sigma^2)/d\sigma^2 - 1$. Thus the function g_δ starts increasing at $\sigma^2 = 0$ (unless $\delta = 1$), then eventually decreases. The point $\tilde{\sigma}^2$ can be chosen in such a way that g_δ is decreasing at $\tilde{\sigma}^2$ but is still positive.

By this choice of $\tilde{\sigma}^2$, we have

$$\delta^{-2}f([\tilde{\sigma}^2,\infty)) \subset [\tilde{\sigma}^2,\infty). \tag{2.14}$$

Indeed, since $f(\sigma^2)$ is increasing, for each $\sigma^2 \geq \tilde{\sigma}^2$, it holds that $\delta^{-2} f(\sigma^2) \geq \delta^{-2} f(\tilde{\sigma}^2) > \tilde{\sigma}^2$, the last inequality being due to the fact that $g_\delta(\tilde{\sigma}^2) > 0$.

We furthermore have that

$$\min_{\sigma^2 \in [0, \tilde{\sigma}^2]} g_\delta(\sigma^2) > 0 \tag{2.15}$$

since $g_\delta(\sigma^2)$ is continuous and positive on the compact $[0, \tilde{\sigma}^2]$.

By Items 1 and 2 above, Inclusion (2.14), and Banach's fixed point theorem, Equation (2.13) has a unique solution, and moreover, if we start with $\sigma_0^2 \geq \tilde{\sigma}^2$, then σ_t^2 converges to this fixed point.

If $\sigma_0^2 \in [0, \tilde{\sigma}^2)$, then the iterates σ_t^2 will reach the interval $[\tilde{\sigma}^2, \infty)$ after a finite number of steps by (2.15), and we are led back to the first case. □

We can now show that the system (2.2) has an unique solution $(\delta, \sigma^2, \gamma) \in (1/\sqrt{2}, \infty) \times (0, \infty) \times (0, 1)$. To that end, we reuse the notations of the previous proof.

For a given $\delta > 1/\sqrt{2}$, the previous lemma shows that Equation (2.2b) admits a unique solution $\sigma^2(\delta)$. Plugging this value in the right hand side of Equation (2.2c), we can write

$$\gamma = \gamma(\delta) = \mathbb{E}_{\bar{r}}\left[\bar{\rho}\Big(-\frac{\bar{r}}{\sigma(\delta)}\Big)\right] = \left.\frac{df(\sigma^2)}{d\sigma^2}\right|_{\sigma^2 = \sigma^2(\delta)}. \tag{2.16}$$

All what remains to show is that the equation

$$\kappa = \delta + \frac{\gamma(\delta)}{\delta} \tag{2.17}$$

has an unique solution $\delta > 1/\sqrt{2}$. We thus need to study the behavior of $\gamma(\delta)$. In all the remainder, differentiability issues can be easily checked and are skipped.

In the previous proof, $\sigma^2(\delta)$ has been obtained through Banach's fixed point theorem, which implies that

$$\left.\frac{df(\sigma^2)}{d\sigma^2}\right|_{\sigma^2 = \sigma^2(\delta)} < \delta^2. \tag{2.18}$$

Recalling that $df(\sigma^2)/d\sigma^2$ decreases from 1 asymptotically to $1/2$ as σ^2 increases from 0 to ∞, the last inequality has two consequences. First, $\sigma^2(\delta) \to \infty$ as $\delta \downarrow 1/\sqrt{2}$. Second, we obtain by using Equation (2.13) and taking the derivatives with respect to δ that

$$\frac{d\sigma^2(\delta)}{d\delta}\left(1 - \frac{1}{\delta^2}\left.\frac{df(\sigma^2)}{d\sigma^2}\right|_{\sigma^2 = \sigma^2(\delta)}\right) = -\frac{2f(\sigma^2(\delta))}{\delta^3},$$

which shows that $\sigma^2(\delta)$ is a decreasing function. Getting back to Equation (2.16), and observing that $\bar{\rho}(\cdot)$ is a decreasing function, we obtain that $\gamma(\delta)$ is increasing.

We now have all the elements to study Equation (2.17). For $\delta \downarrow 1/\sqrt{2}$, $\sigma^2(\delta) \to \infty$ by what precedes, thus, $\gamma(\delta) \downarrow 1/2$, and $\delta + \gamma(\delta)/\delta \to \sqrt{2} < \kappa$. For δ near infinity, $\delta + \gamma(\delta)/\delta \sim \delta > \kappa$. Consequently, Equation (2.17) has a solution by continuity. To establish uniqueness, we show that the function $\delta \mapsto \delta + \gamma(\delta)/\delta$ is increasing. Indeed,

$$\frac{d}{d\delta}\left(\delta + \frac{\gamma(\delta)}{\delta}\right) = 1 + \frac{\gamma'(\delta)}{\delta} - \frac{\gamma(\delta)}{\delta^2} \geq 1 - \frac{\gamma(\delta)}{\delta^2} > 0$$

as shown by Inequality (2.18), and we are done.

Study of the LCP solution via AMP

We begin with some of the fundamental results of the AMP theory. The now classical form of an AMP iterative algorithm, as formalized in the article [Bayati et Montanari, 2011] of Bayati and Montanari based in part on a result of Bolthausen [Bolthausen, 2014], can be presented as follows. For $t = 0, 1, \ldots$, let $(h^t)_t$ be a sequence of Lipschitz $\mathbb{R}^2 \to \mathbb{R}$ functions. Let $a_n \in \mathbb{R}^n$ be a random vector of so-called auxiliary information. Recall that A_n is the GOE matrix introduced in Section 2. Starting with a vector $u_n^0 \in \mathbb{R}^n$, the AMP recursion reads

$$u_n^{t+1} = \frac{A_n}{\sqrt{n}} h^t(u_n^t, a_n) \ - \ \langle \partial_1 h^t(u_n^t, a_n) \rangle_n \ h^{t-1}(u_n^{t-1}, a_n).$$

Here, writing $u_n^t = \left[u_{i,n}^t\right]_{i=1}^n$ and $a_n = [a_{i,n}]_{i=1}^n$, we set $h^{-1} = 0$ and $h^t(u_n^t, a_n) = \left[h^t(u_{i,n}^t, a_{i,n})\right]_{i=1}^n \in \mathbb{R}^n$. Observing that the derivative $dh^t(u,a)/du$ is defined almost everywhere thanks to the Lipschitz property of h^t, the function $\partial_1 h^t(u,a)$ is any function that coincides with this derivative where it is defined. Finally, $\langle x \rangle_n \triangleq n^{-1} \sum x_i$ for $x = [x_i]_{i=1}^n$.

With this construction, it turns out to be possible to evaluate precisely the asymptotic behavior of the empirical measures of the type $\mu^{a_n, u_n^1, \ldots, u_n^t}$ as $n \to \infty$ for each value of t, leading to the so-called Density Evolution (DE) equations that characterize the limits of these measures. The so-called "Onsager term" $\langle \cdots \rangle_n h^{t-1}(u_n^{t-1}, a_n)$ (equal to zero for $t = 0$) plays a crucial role in making possible this convergence. For a detailed exposition of the AMP theory, along with the description of many of its applications, the reader is referred to the recent tutorial [Feng *et al.*, 2022].

To establish Theorem 30, we shall study the properties of the following AMP algorithm. For each n, let $(u_n^0, a_n) \in \mathbb{R}^n \times \mathbb{R}^n$ be a couple of random vectors independent of A_n, and such that $a_n \succcurlyeq 0$. Assume that there exists a couple of L^2 random variables $(\bar{u}, \bar{a})$, with $\bar{a} \neq 0$, such that

$$\mu^{u_n^0, a_n} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R}^2)} \mathcal{L}((\bar{u}, \bar{a})) \quad \text{almost surely.} \tag{2.19}$$

We can now take u_n^0 as the deterministic vector $u_n^0 = \mathbf{1}_n$; the vector a_n will be precised later. Note that $\bar{a} \geq 0$. Take $h^t \equiv h$ for each integer $t \geq 0$, where

$$h(u, a) = \frac{(u + a)_+}{\delta},$$

and where $\delta > 1/\sqrt{2}$ solves the system (2.2). Our AMP iteration reads then

$$u_n^{t+1} = \frac{A_n}{\delta \sqrt{n}} (u_n^t + a_n)_+ - \frac{\langle \mathbb{1}_{u_n^t + a_n > 0} \rangle_n \, (u_n^{t-1} + a_n)_+}{\delta^2}. \tag{2.20}$$

The DE equations for this algorithm are provided by the following proposition, which is a direct application of, *e.g.*, [Feng *et al.*, 2022, Theorem 2.3] (see also [Bayati et Montanari, 2011, Theorem 4]) :

Proposition 33. Consider the algorithm (2.20). Then, for each $t \geq 1$,

$$\mu^{a_n, u_n^1, \ldots, u_n^t} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R}^{t+1})} \mathcal{L}((\bar{a}, Z^1, \ldots, Z^t)) \quad \text{almost surely,}$$

where $[Z^1, \ldots, Z^t]$ is a centered Gaussian vector, independent of $(\bar{u}, \bar{a})$, which covariance matrix is R^t is recursively defined in t as follows :

$$R^1 = \mathbb{E}\left[(Z^1)^2\right] = \frac{1}{\delta^2}\mathbb{E}\left[(\bar{u}+\bar{a})_+^2\right].$$

Given R^t, only the last row and column of R^{t+1} have to be determined. This is done through the equations

$$\left[R^{t+1}\right]_{t+1,k} = \mathbb{E}\left(Z^{t+1}Z^k\right) = \frac{1}{\delta^2}\begin{cases} \mathbb{E}\left[(Z^t+\bar{a})_+(Z^{k-1}+\bar{a})_+\right] & \text{if} \quad k \in \{2, \ldots, t+1\}, \\ \mathbb{E}\left[(Z^t+\bar{a})_+(\bar{u}+\bar{a})_+\right] & \text{if} \quad k=1. \end{cases},$$

Note that by writing $U^{t+1} = \left[(\bar{u}+\bar{a})_+, (Z^1+\bar{a})_+, \ldots, (Z^t+\bar{a})_+\right]^{\mathrm{T}}$, we see that $R^{t+1} = \mathbb{E}\left(U^{t+1}(U^{t+1})^{\mathrm{T}}\right)$, which imediately shows that R^{t+1} is a positive semidefinite (actually, definite) matrix.

With this result at hand, the idea of the proof of the remainder of Theorem 30 can be summarized as follows :

- If n is large enough then if t is large enough, $(u_n^t + a_n)_+$ will be seen as the solution of a perturbed version of the LCP problem in (2.12). To obtain this result, we need to show that the vectors $(u_n^t + a_n)_+$ and $(u_n^{t-1} + a_n)_+$ tend to be aligned. To this end, we rely on a result of Montanari and Richard in [Montanari et Richard, 2016], see also [Donoho et Montanari, 2016].
- The behavior of $\mu^{(u_n^t+a_n)_+}$ for large n then for large t is described by the DE equations.
- This behavior can be transferred to $\mu^{x_n^*}$ with the help of the LCP perturbation results of Chen and Xiang in [Chen et Xiang, 2007].

AMP iterates in an LCP form, and DE estimates

Write $x_n^t = u_n^t + a_n$, and let $\gamma_n^t = \langle \mathbb{1}_{(u_n^t+a_n)>0}\rangle_n = \langle \mathbb{1}_{x_n^t>0}\rangle_n$. Recall the expression of γ in (2.2c). With these notations, the AMP recursion can be rewritten as

$$\begin{aligned} x_n^{t+1} &= \frac{A_n}{\delta\sqrt{n}}(x_n^t)_+ - \frac{\gamma_n^t}{\delta^2}(x_n^{t-1})_+ + a_n \\ &= \frac{A_n}{\delta\sqrt{n}}(x_n^t)_+ - \frac{\gamma}{\delta^2}(x_n^{t-1})_+ + a_n + \frac{\gamma-\gamma_n^t}{\delta^2}(x_n^{t-1})_+ \\ &= \frac{A_n}{\delta\sqrt{n}}(x_n^t)_+ - \frac{\gamma}{\delta^2}(x_n^t)_+ + a_n + \frac{\gamma-\gamma_n^t}{\delta^2}(x_n^{t-1})_+ + \frac{\gamma}{\delta^2}((x_n^t)_+ - (x_n^{t-1})_+), \end{aligned}$$

thus,

$$x_n^t = \frac{A_n}{\delta\sqrt{n}}(x_n^t)_+ - \frac{\gamma}{\delta^2}(x_n^t)_+ + a_n + \varepsilon_n^t,$$

where

$$\varepsilon_n^t = \frac{\gamma-\gamma_n^t}{\delta^2}(x_n^{t-1})_+ + x_n^t - x_n^{t+1} + \frac{\gamma}{\delta^2}((x_n^t)_+ - (x_n^{t-1})_+). \tag{2.21}$$

We now put the next to last equation under the form of a LCP problem. Specifically, writing $x = x_+ - x_-$ for a vector x, we have

$$-(x_n^t)_- = \frac{A_n}{\delta\sqrt{n}}(x_n^t)_+ - \frac{\gamma}{\delta^2}(x_n^t)_+ - (x_n^t)_+ + a_n + \varepsilon_n^t,$$

in other words,

$$\frac{(x_n^t)_-}{1+\gamma/\delta^2} = \left(I_n - \frac{A_n}{\delta(1+\gamma/\delta^2)\sqrt{n}} \right) (x_n^t)_+ - \frac{a_n + \varepsilon_n^t}{1+\gamma/\delta^2}.$$

Remember that $\kappa = \delta + \gamma/\delta$ and that $M_n = A_n/(\kappa\sqrt{n})$. Thus, when $\|M_n\| < 1$, we have

$$(x_n^t)_+ = \mathrm{LCP}\left(I_n - M_n, -\frac{a_n + \varepsilon_n^t}{1+\gamma/\delta^2} \right). \tag{2.22}$$

By putting

$$a_n = (1+\gamma/\delta^2) r_n,$$

as we shall do hereinafter, this equation can be compared with (2.12). Note that if we set, *e.g.*, $u_n^0 = \mathbf{1}_n$, then the assumption (2.19) on $\mu^{u_n^0, a_n}$ is satisfied thanks to Assumption 2 with $\bar{a} = (1+\gamma/\delta^2)\bar{r}$.

Before bounding ε_n^t in (2.22), let us study the behavior of $\mu^{(x_n^t)_+}$ with the help of Proposition 33. Recall that $\bar{Z} \sim \mathcal{N}(0,1)$. Proposition 33 shows that for $t \geq 2$, $\mu^{u_n^t} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}(Z^t)$, with $Z^t \sim \boldsymbol{\sigma}_t \bar{Z}$, and

$$\boldsymbol{\sigma}_{t+1}^2 = \frac{1}{\delta^2} \mathbb{E}\left[(\boldsymbol{\sigma}_t \bar{Z} + (1+\gamma/\delta^2)\bar{r})_+^2 \right].$$

Since the function $\varphi(u,a) = (u+a)_+$ is Lipschitz, it is clear that $\mu^{(x_n^t)_+} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}((\boldsymbol{\sigma}_t \bar{Z} + (1+\gamma/\delta^2)\bar{r})_+)$. Furthermore, since the distribution function of the law of $\boldsymbol{\sigma}_t \bar{Z} + (1+\gamma/\delta^2)\bar{r}$ has no discontinuity, it holds that

$$\gamma_n^t \to \boldsymbol{\gamma}^t \triangleq \mathbb{P}\left[\boldsymbol{\sigma}_t \bar{Z} + (1+\gamma/\delta^2)\bar{r} > 0 \right] = \mathbb{E}_{\bar{r}} \mathbb{P}_{\bar{Z}}\left[\boldsymbol{\sigma}_t \bar{Z} + (1+\gamma/\delta^2)\bar{r} > 0 \right].$$

Now, by replacing $\bar{r}$ with $(1+\gamma/\delta^2)\bar{r}$ in the statement of Lemma 32, we obtain by this lemma that the sequence $(\boldsymbol{\sigma}_t^2)$ converges to $\boldsymbol{\sigma}_\infty^2$ defined as the unique solution of the equation

$$\boldsymbol{\sigma}_\infty^2 = \frac{1}{\delta^2} \mathbb{E}\left[(\boldsymbol{\sigma}_\infty \bar{Z} + (1+\gamma/\delta^2)\bar{r})_+^2 \right].$$

A consequence is that the sequence $(\boldsymbol{\gamma}^t)$ converges to $\boldsymbol{\gamma}^\infty$ defined as

$$\boldsymbol{\gamma}^\infty = \mathbb{E}_{\bar{r}} \mathbb{P}_{\bar{Z}}\left[\boldsymbol{\sigma}_\infty \bar{Z} + (1+\gamma/\delta^2)\bar{r} > 0 \right].$$

By comparing these two equations with (2.2b) and (2.2c) respectively, we easily see that

$$\boldsymbol{\gamma}^\infty = \gamma, \quad \text{and} \quad \frac{\boldsymbol{\sigma}_\infty^2}{(1+\gamma/\delta^2)^2} = \sigma^2.$$

Since $(x_n^t)_+$ is designed to be an approximation of x_n^*, our purpose is to show that

$$\mu^{x_n^*} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \boldsymbol{\mu}^* \triangleq \mathcal{L}(\boldsymbol{\sigma}_\infty \bar{Z} + (1+\gamma/\delta^2)\bar{r})_+) = \mathcal{L}((1+\gamma/\delta^2)(\sigma\bar{Z} + \bar{r})_+) \tag{2.23}$$

to establish Theorem 30.

Control of the error term ε_n^t

Getting back to the expression (2.21) of ε_n^t, it holds by what precedes that $\|(x_n^{t-1})_+\|^2/n \to \mathbb{E}(\boldsymbol{\sigma}_{t-1}\bar{Z}+(1+\gamma/\delta^2)\bar{r})_+^2 = \boldsymbol{\sigma}_t^2\delta^2$, which is bounded, and furthermore $\lim_t \operatorname{aslim}_{n\to\infty}(\gamma-\gamma_n^t) = 0$. We thus have

$$\lim_{t\to\infty} \operatorname*{aslim}_{n\to\infty} \frac{(\gamma-\gamma_n^t)^2}{\delta^4}\frac{\|(x_n^{t-1})_+\|^2}{n} = 0. \tag{2.24}$$

The main idea to control the two other terms at the right hand side of (2.21) is to show that the correlation coefficient $\mathbb{E}(Z^{t-1}Z^t)/(\boldsymbol{\sigma}_{t-1}\boldsymbol{\sigma}_t)$ converges to 1 as $t\to\infty$. This was done in a similar context in [Montanari et Richard, 2016], see also [Donoho et Montanari, 2016]. For self-containedness, we summarize herein their approach :

Lemma 34. The sequence (Q_t) defined as

$$Q_t \triangleq \frac{\mathbb{E}(Z^{t-1}Z^t)}{\boldsymbol{\sigma}_{t-1}\boldsymbol{\sigma}_t}$$

starting with Q_2 satisfies $Q_t \xrightarrow[t\to\infty]{} 1$.

Démonstration. Let $[X_1,X_2]$ be a centered Gaussian vector such that $\mathbb{E}X_1^2=\mathbb{E}X_2^2=1$, and $\mathbb{E}X_1X_2 = q\in[0,1]$. Let W be a random variable independent of $[X_1,X_2]$, and such that $\mathbb{E}(X_1+W)^2<\infty$. Consider the function $\mathsf{H}:[0,1]\to[0,1]$ defined as

$$q\mapsto \mathsf{H}(q) = \frac{\mathbb{E}\left[(X_1+W)_+(X_2+W)_+\right]}{\mathbb{E}\left[(X_1+W)_+^2\right]}.$$

It is shown in [Montanari et Richard, 2016, Lemma 38 and proof of Lemma 37] that H is a continuous increasing function on $[0,1]$ such that $\mathsf{H}(q)>q$ for all $q\in[0,1)$, and $\mathsf{H}(1)=1$. Thus, starting with any $q\in[0,1)$, the sequence $q_{t+1}=\mathsf{H}(q_t)$ is increasing and is bounded by 1. Since its limit q_∞ satisfies $q_\infty=\mathsf{H}(q_\infty)$, it holds that $q_\infty=1$.

Writing $Z^t=\boldsymbol{\sigma}_t\bar{Z}^t$, we have

$$\begin{aligned}Q_{t+1} = \frac{\mathbb{E}\left(Z^tZ^{t+1}\right)}{\boldsymbol{\sigma}_t\boldsymbol{\sigma}_{t+1}} &= \frac{\mathbb{E}\left[(\boldsymbol{\sigma}_{t-1}\bar{Z}^{t-1}+\bar{a})_+(\boldsymbol{\sigma}_t\bar{Z}^t+\bar{a})_+\right]}{\sqrt{\mathbb{E}\left[(\boldsymbol{\sigma}_{t-1}\bar{Z}^{t-1}+\bar{a})_+^2\right]\mathbb{E}\left[(\boldsymbol{\sigma}_t\bar{Z}^t+\bar{a})_+^2\right]}}, \\ &= \frac{\mathbb{E}\left[(\bar{Z}^{t-1}+\bar{a}/\boldsymbol{\sigma}_{t-1})_+(\bar{Z}^t+\bar{a}/\boldsymbol{\sigma}_t)_+\right]}{\sqrt{\mathbb{E}\left[(\bar{Z}^{t-1}+\bar{a}/\boldsymbol{\sigma}_{t-1})_+^2\right]\mathbb{E}\left[(\bar{Z}^t+\bar{a}/\boldsymbol{\sigma}_t)_+^2\right]}}.\end{aligned}$$

When defining the fonction $\widetilde{H}$ by :

$$\widetilde{H}(q,\sigma_1,\sigma_2) \triangleq \frac{\mathbb{E}\left[(\sigma_1X_1+W)_+(\sigma_2X_2+W)_+\right]}{\sqrt{\mathbb{E}\left[(\sigma_1X_1+W)_+^2\right]\mathbb{E}\left[(\sigma_2X_2+W)_+^2\right]}},$$

with $\sigma_1,\ \sigma_2>0$, one gets

$$Q_{t+1} = \widetilde{H}(Q_t,\boldsymbol{\sigma}_{t-1},\boldsymbol{\sigma}_t).$$

Note that $\widetilde{H}$ is a continuous function. Furthermore, if we set $W = \bar{a}/\boldsymbol{\sigma}_\infty$ in the definition of H above, then we have that $\widetilde{H}(q, \boldsymbol{\sigma}_\infty, \boldsymbol{\sigma}_\infty) = \mathsf{H}(q)$. Also, it is clear from the previous display (and by Cauchy-Schwartz inequality) that $Q_t \in [0, 1]$.

The lemma is established if we prove that $Q_* \triangleq \liminf_t Q_t$ satisfies $Q_* = 1$. Let us first show that $\liminf \mathsf{H}(Q_t) \geq \mathsf{H}(Q_*)$.
If we assume that $Q_* = 0$, then $\forall t$, $Q_t \geq Q_*$ and $\mathsf{H}(Q_t) \geq \mathsf{H}(Q_*) > 0$, hence $\liminf_t \mathsf{H}(Q_t) \geq \mathsf{H}(Q_*) > Q_*$.
It is enough to assume that $Q_* > 0$. For each $\varepsilon > 0$, $Q_t \geq Q_* - \varepsilon$ for all t large enough. Thus, $\mathsf{H}(Q_t) \geq \mathsf{H}(Q_* - \varepsilon)$ for all t large, which implies that $\liminf \mathsf{H}(Q_t) \geq \mathsf{H}(Q_* - \varepsilon)$. Since $\varepsilon > 0$ is arbitrary, we have $\liminf \mathsf{H}(Q_t) \geq \mathsf{H}(Q_*)$. With this, we have

$$\begin{aligned} Q_* &= \liminf_t \widetilde{H}(Q_t, \boldsymbol{\sigma}_{t-1}, \boldsymbol{\sigma}_t) \\ &= \liminf_t \widetilde{H}(Q_t, \boldsymbol{\sigma}_\infty, \boldsymbol{\sigma}_\infty) \\ &= \liminf_t \mathsf{H}(Q_t) \\ &\geq \mathsf{H}(Q_*), \end{aligned}$$

where the second equality is due to the continuity of $\widetilde{H}$. This shows that $Q_* = 1$. □

Let us now deal with the second term at the right hand side of (2.21). Taking $\varphi(x_1, x_2) = (x_1 - x_2)^2 \in \mathrm{PL}_2(2)$ and using Proposition 33, we have

$$\frac{\|x_n^t - x_n^{t+1}\|^2}{n} = \frac{1}{n} \sum_{i=1}^n \varphi(u_{i,n}^t, u_{i,n}^{t+1}) \to \mathbb{E}\left[\left(Z^{t+1} - Z^t\right)^2\right] = \boldsymbol{\sigma}_{t+1}^2 + \boldsymbol{\sigma}_t^2 - 2\boldsymbol{\sigma}_{t+1}\boldsymbol{\sigma}_t Q_{t+1}.$$

Thus, it holds by Lemma 34 that

$$\lim_{t\to\infty} \operatorname*{aslim}_{n\to\infty} \frac{\|x_n^t - x_n^{t+1}\|^2}{n} = 0. \tag{2.25}$$

Regarding the last term, we similarly have $\lim_{t\to\infty} \operatorname*{aslim}_{n\to\infty} \frac{\|(x_n^t)_+ - (x_n^{t-1})_+\|^2}{n} = 0$. Combining this convergence with (2.24) and (2.25), and using that

$$\frac{\|\varepsilon_n^t\|^2}{n} \leq 3\left(\frac{(\gamma - \gamma_n^t)^2}{\delta^4} \frac{\|(x_n^{t-1})_+\|^2}{n} + \frac{\|x_n^t - x_n^{t+1}\|^2}{n} + \frac{\gamma^2}{\delta^4} \frac{\|(x_n^t)_+ - (x_n^{t-1})_+\|^2}{n}\right),$$

we get that

$$\lim_{t\to\infty} \operatorname*{aslim}_{n\to\infty} \frac{\|\varepsilon_n^t\|^2}{n} = 0 \tag{2.26}$$

(that the $\operatorname*{aslim}_{n\to\infty}$ at the left hand side exists can be deduced again from Proposition 33).

Use of a LCP perturbation result

When $\|M_n\| < 1$, we can compare (2.22) with (2.12) by using the LCP perturbation results of [Chen et Xiang, 2007]. Applying [Chen et Xiang, 2007, Theorems 2.7 and 2.8], we have

$$\|x_n^* - (x_n^t)_+\| \leq \left\|\left(I - M_n\right)^{-1}\right\| \frac{\|\varepsilon_n^t\|}{1 + \gamma/\delta^2}.$$

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be an arbitrary function in PL(2). For a given positive integer t, we have

$$\frac{1}{n}\sum_{i=1}^{n}\varphi(x^*_{i,n}) - \int \varphi d\boldsymbol{\mu}^* = \frac{1}{n}\sum_{i=1}^{n}\left(\varphi(x^*_{i,n}) - \varphi((x^t_{i,n})_+)\right) + \frac{1}{n}\sum_{i=1}^{n}\varphi((x^t_{i,n})_+) - \int \varphi d\boldsymbol{\mu}^*$$
$$\triangleq \xi^1_n(t) + \xi^2_n(t).$$

Dealing with $\xi^2_n(t)$, we have from Proposition 33 that

$$\xi^2_n(t) \to \mathbb{E}\varphi(\boldsymbol{\sigma}_t\bar{Z} + \bar{a})_+ - \mathbb{E}\varphi(\boldsymbol{\sigma}_\infty\bar{Z} + \bar{a})_+.$$

Using that the function x_+ is 1–Lipschitz and $|x_+| \le |x|$, the limit satisfies

$$\left|\mathbb{E}\left(\varphi(\boldsymbol{\sigma}_t\bar{Z} + \bar{a})_+ - \varphi(\boldsymbol{\sigma}_\infty\bar{Z} + \bar{a})_+\right)\right| \le L|\boldsymbol{\sigma}_t - \boldsymbol{\sigma}_\infty|\mathbb{E}\left(|\bar{Z}|(1 + |\boldsymbol{\sigma}_t\bar{Z} + \bar{a}| + |\boldsymbol{\sigma}_\infty\bar{Z} + \bar{a}|)\right),$$

which is easily seen to be bounded by a constant $C(t)$ that converges to zero as $t \to \infty$, since $\boldsymbol{\sigma}_t \to \boldsymbol{\sigma}_\infty$.

We now turn to $\xi^1_n(t)$. Denoting as L the Lipschitz constant of φ, we have by the Cauchy-Schwarz inequality

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\left|\varphi(x^*_{i,n}) - \varphi((x^t_{i,n})_+)\right| &\le \frac{L}{n}\sum_{i=1}^{n}\left|x^*_{i,n} - (x^t_{i,n})_+\right|\left(1 + |x^*_{i,n}| + |(x^t_{i,n})_+|\right) \\ &\le \frac{L}{n}\left\|x^*_n - (x^t_n)_+\right\|\left(\sum_i (1 + |x^*_{i,n}| + |(x^t_{i,n})_+|)^2\right)^{1/2} \\ &\le 3L\frac{\|x^*_n - (x^t_n)_+\|}{\sqrt{n}}\left(1 + \frac{\|x^*_n\|}{\sqrt{n}} + \frac{\|(x^t_n)_+\|}{\sqrt{n}}\right).\end{aligned}$$

Writing $B_n \triangleq \|(I - A_n/(\kappa\sqrt{n}))^{-1}\|/(1 + \gamma/\delta^2)$, we obtain that

$$|\xi^1_n(t)| \le 3LB_n\frac{\|\varepsilon^t_n\|}{\sqrt{n}}\left(1 + 2\frac{\|(x^t_n)_+\|}{\sqrt{n}} + B_n\frac{\|\varepsilon^t_n\|}{\sqrt{n}}\right).$$

We know from Assumption 1 that B_n converges almost surely to a positive constant. From Proposition 33, we furthermore have

$$\frac{\|(x^t_n)_+\|}{\sqrt{n}} \to \left(\mathbb{E}(\boldsymbol{\sigma}_t\bar{Z} + \bar{a})^2_+\right)^{1/2},$$

which is bounded in t. Thus, using (2.26), we obtain that $\limsup_n |\xi^1_n(t)|$ is bounded with probability one by a constant $C_1(t)$ that converges to zero as $t \to \infty$.

We finally obtain that

$$\limsup_n \left|\frac{1}{n}\sum_{i=1}^{n}\varphi(x^*_{i,n}) - \int \varphi d\boldsymbol{\mu}^*\right| \le C(t) + C_1(t) \quad \text{with probability one.}$$

Since $C(t) + C_1(t)$ can be made arbitrarily small, we have

$$\text{a.s.} \quad \frac{1}{n}\sum_{i=1}^{n}\varphi(x^*_{i,n}) \to \int \varphi d\boldsymbol{\mu}^*,$$

which terminates the proof of Theorem 30.

3 Wishart model

3.1 Problem statement, assumptions and results

Beyond this GOE model, part of the literature on ecological networks considers the interaction matrix as a Wishart matrix (see for instance [Akjouj *et al.*, 2022, § 4.6] or [Rozas *et al.*, 2023]). Indeed, the interactions between two species are seen as dependent on a distance between their respective values of some functional traits.

Let's take (A_n) a sequence of matrices such that $A_n \in \mathcal{M}_{p\times n}(\mathbb{R})$ has independent Gaussian $\mathcal{N}(0,1)$ entries. Let $(r_n)_{n=1,2,\dots}$ be a sequence of random vectors such that $r_n \succcurlyeq 0$. We shall assume that $p = p(n)$ where $n/p \to c \in (0,\infty)$ when $n \to \infty$, and that r_n and A_n are independent for each n.

For a given foodweb of n species, the i^{th} column of A_n is a vector modelling the traits of species i. Related to this model, the LV differential equation is :

$$\dot{\boldsymbol{x}}_n^t = \boldsymbol{x}_n^t \odot \left(r_n + (M_n - I_n)\, \boldsymbol{x}_n^t \right), \quad \boldsymbol{x}_n^0 \in (0,\infty)^n, \tag{2.27}$$

where the vector function $\boldsymbol{x}_n : [0,\infty) \to \mathbb{R}_+^n = [0,\infty)^n$ represents the abundances of the n species, $\odot$ is the componentwise (or Hadamard) product, $r \in \mathbb{R}_+^n$ is the so-called vector of intrinsic growth rates of the species, and the matrix $M_n \in \mathbb{R}^{p\times p}$ quantifies the strengths of the food interactions among these species.

$$M_n = \frac{A_n^{\mathrm{T}} A_n}{\kappa p}$$

for κ a real positive number.

Assumption 3. $\kappa > (1+\sqrt{c})^2$,

A well-known result of the large random matrix theory is that the largest eigenvalue (and the spectral norm) of the matrix $\frac{A_n^* A_n}{p}$ almost surely (a.s.) tends to $(1+\sqrt{c})^2$ as $n \to \infty$ [Bai et Silverstein, 2010, Thm 5.8].

We briefly remind assumption 2 :

Assumption (Ass. 2)**.** There exists a measure $\bar{\mu} \in \mathcal{P}_2(\mathbb{R})$ such that $\bar{\mu} \neq \boldsymbol{\delta}_0$, and

$$(a.s.) \quad \mu^{r_n} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \bar{\mu} \ .$$

The result of this section reads as follows :

Theorem 35. Let Assumption 3 hold true. Then, $\|M_n\| < 1$ for alla large n with probability one. For each of these n, the LV ODE solution is defined on $\mathbb{R}_+$ and has a globally stable equilibrium $\boldsymbol{x}_n^*$. For the other n, put $\boldsymbol{x}_n^* = 0$. The distribution $\mu^{\boldsymbol{x}_n^*}$ is a $\mathcal{P}_2(\mathbb{R})$–valued random variable on the probability space where A_n and r_n are defined.

Let Assumption 2 hold true. Let $\bar{r} \geq 0$ be a random variable such that $\bar{\mu} = \mathcal{L}(\bar{r})$. Let $\bar{Z}$ be a $\mathcal{N}(0,1)$ random variable independent of $\bar{r}$. Then, the system of equations

$$\kappa = (\delta + c\gamma)\left(1 + \frac{1}{\delta}\right), \tag{2.28a}$$

$$\tau^2 = \frac{c}{\delta^2} \mathbb{E}_{\bar{r}} \mathbb{E}_{\bar{Z}} \left[\left(\tau \bar{Z} + \bar{r}\right)_+^2 \right], \tag{2.28b}$$

$$\gamma = \mathbb{E}_{\bar{r}} \mathbb{P}_{\bar{Z}} \left[\tau \bar{Z} + \bar{r} > 0 \right], \tag{2.28c}$$

admits an unique solution (δ, τ^2, γ) in $(\sqrt{c/2}, \infty) \times (0, \infty) \times (0, 1)$.

With Assumptions 3 and 2 hold true, the convergence

$$\mu^{x_n^*} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}\left((1 + 1/\delta)\left(\tau \bar{Z} + \bar{r}\right)_+\right) \quad \text{almost surely} \tag{2.29}$$

holds true.

There is a strong matching between the parameters obtained by solving the system 2.28 and their empirical counterparts obtained by Monte-Carlo simulations. This is illustrated in Fig. 2.2 by comparing γ, the mean and the root mean square of the abundances species at equilibrium.

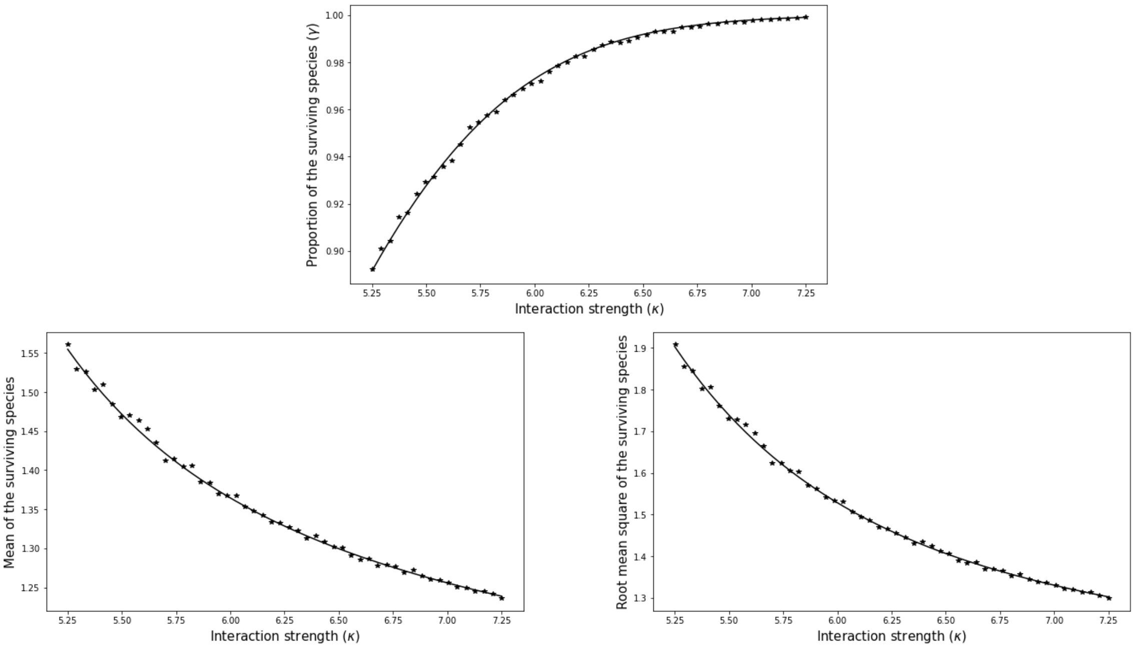


FIGURE 2.2 – The plots represent a comparison between the theoretical solutions γ, the mean and the root mean square at equilibrium of the abundances arising from system 2.28 and their empirical Monte Carlo counterpart (the star marker) as functions of the interaction strength κ. Matrix M_n has size $n = 500$ and $p = 300$. The number of Monte Carlo experiments is 50. When interaction κ^{-1} increases, the proportion of surviving species γ decrease but their mean and variance increase.

The proof of this theorem relies on an asymmetric version of the AMP algorithm. We provide herein the main elements of this proof.

3.2 Proof of Theorem 35

Existence and uniqueness of the solution of System (2.28)

We begin by an adaptation of lemma 32 to the system (2.28).

Lemma 36. For a given $\delta > 0$, Equation (2.28b) admits a solution τ^2 if and only if $\delta > \sqrt{c/2}$, and the case being, this solution is unique, and is written $\tau^2(\delta)$.

Assume $\delta > 1/\sqrt{c/2}$. Starting with an arbitrary $\tau_0 \geq 0$, consider the iterates in $t = 0, 1, \ldots$

$$\tau_{t+1}^2 = \frac{c}{\delta^2} \mathbb{E}_{\bar{r}} \mathbb{E}_{\bar{Z}} \left[\left(\tau_t \bar{Z} + \bar{r} \right)_+^2 \right].$$

Then, $\tau_t^2 \underset{t}{\to} \tau^2(\delta)$.

Démonstration. Writing $\mathbb{E} = \mathbb{E}_{\bar{r}} \mathbb{E}_{\bar{Z}}$, consider the function $f(\tau^2) \triangleq \mathbb{E}\left[(\tau \bar{Z} + \bar{r})_+^2 \right]$. Then, Equation (2.28b) is written

$$\frac{c}{\delta^2} f(\tau^2) = \tau^2. \tag{2.30}$$

Let us compute $df(\tau^2)/d\tau^2$. Fixing $\tau > 0$ and assuming that the real number h has a small enough absolute value, the function

$$\varphi_h(z, r) \triangleq \frac{((\tau + h)z + r)_+^2 - (\tau z + r)_+^2}{h}$$

is easily shown to satisfy $|\varphi_h(z, r)| \leq |z| \left[((\tau + h)z + r)_+ + (\tau z + r)_+ \right]$. Thus, by Assumption 2 and the Dominated Convergence Theorem, it holds that

$$\begin{aligned} \frac{df(\tau^2)}{d\tau^2} &= \frac{1}{2\tau} \frac{df(\tau^2)}{d\tau} = \frac{1}{2\tau} \mathbb{E}\left[\frac{d(\tau \bar{Z} + \bar{r})_+^2}{d\tau} \right] = \frac{1}{\tau} \mathbb{E}\left[\bar{Z} (\tau \bar{Z} + \bar{r})_+ \right], \\ &= \mathbb{E}\left[\bar{Z}^2 \mathbb{1}_{\bar{Z} \geq -\bar{r}/\tau} \right] + \frac{\bar{r}}{\tau} \mathbb{E}\left[\bar{Z} \mathbb{1}_{\bar{Z} \geq -\bar{r}/\tau} \right]. \end{aligned}$$

By doing an Integration by Parts involving the Gaussian law, we obtain that

$$\mathbb{E}\left[\bar{Z}^2 \mathbb{1}_{\bar{Z} \geq -\bar{r}/\tau} \right] = \mathbb{E}_{\bar{r}} \left[\bar{\rho}\Big(-\frac{\bar{r}}{\tau} \Big) - \frac{\bar{r}}{\tau} \frac{\exp(-\bar{r}^2/2\tau^2)}{\sqrt{2\pi}} \right],$$

where

$$\bar{\rho}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-t^2/2)\, dt.$$

We also have

$$\frac{\bar{r}}{\tau} \mathbb{E}\left[\bar{Z} \mathbb{1}_{\bar{Z} \geq -\bar{r}/\tau} \right] = \mathbb{E}_{\bar{r}} \left[\frac{\bar{r}}{\tau} \frac{\exp(-\bar{r}^2/2\tau^2)}{\sqrt{2\pi}} \right],$$

thus,

$$\frac{df(\tau^2)}{d\tau^2} = \mathbb{E}_{\bar{r}} \left[\bar{\rho}\Big(-\frac{\bar{r}}{\tau} \Big) \right].$$

We observe from this equation that $df(\tau^2)/d\tau^2$ decreases from 1 asymptotically to $1/2$ as τ^2 increases from 0 to ∞.

The function

$$g_\delta(\tau^2) \triangleq \frac{c}{\delta^2} f(\tau^2) - \tau^2$$

satisfies $g_\delta(0) = \frac{c}{\delta^2} \mathbb{E}_{\bar{r}}(\bar{r}^2) > 0$.
If $\delta \leq \sqrt{c/2}$, then $dg_\delta(\tau^2)/d\tau^2 \geq 2df(\tau^2)/d\tau^2 - 1 > 0$ for each $\tau^2 > 0$, therefore, $g_\delta(\cdot)$ is increasing and the equation $g_\delta(\tau^2) = 0$, *i.e.*, Equation (2.28b), has no solution.

We now consider the case $\delta > \sqrt{c/2}$.

If $\delta^2/c > 1$, then, $\frac{c}{\delta^2} \sup_{\tau^2} df(\tau^2)/d\tau^2 < 1$, and the results of the lemma are consequences of Banach's fixed point theorem.

Let us consider the case where $\delta \in (\sqrt{c/2}, \sqrt{c}]$, or, equivalently, $\delta^{-2} \in [1/c, 2/c)$. There exists a point $\tilde{\tau}^2 > 0$ that satisfies the following properties :

1. $\frac{c}{\delta^2} \sup_{\tau^2 \in [\tilde{\tau}^2, \infty)} df(\tau^2)/d\tau^2 < 1$.
2. $g_\delta(\tau^2) > 0$ on $[0, \tilde{\tau}^2]$. To see why this is possible, remember that $g_\delta(0) > 0$ and $dg(\tau^2)/d\tau^2 = \frac{c}{\delta^2} df(\tau^2)/d\tau^2 - 1$. Thus the function g_δ starts increasing at $\tau^2 = 0$ (unless $\delta = \sqrt{c}$), then eventually decreases. The point $\tilde{\tau}^2$ can be chosen in such a way that g_δ is decreasing at $\tilde{\tau}^2$ but is still positive.

By this choice of $\tilde{\tau}^2$, we have

$$\frac{c}{\delta^2} f([\tilde{\tau}^2, \infty)) \subset [\tilde{\tau}^2, \infty). \tag{2.31}$$

Indeed, since $f(\tau^2)$ is increasing, for each $\tau^2 \geq \tilde{\tau}^2$, it holds that $\frac{c}{\delta^2} f(\tau^2) \geq \frac{c}{\delta^2} f(\tilde{\tau}^2) > \tilde{\tau}^2$, the last inequality being due to the fact that $g_\delta(\tilde{\tau}^2) > 0$.

We furthermore have that

$$\min_{\tau^2 \in [0, \tilde{\tau}^2]} g_\delta(\tau^2) > 0 \tag{2.32}$$

since $g_\delta(\tau^2)$ is continuous and positive on the compact $[0, \tilde{\tau}^2]$.

By Items 1 and 2 above, Inclusion (2.31), and Banach's fixed point theorem, Equation (2.30) has an unique solution, and moreover, if we start with $\tau_0^2 \geq \tilde{\tau}^2$, then τ_t^2 converges to this fixed point.

If $\tau_0^2 \in [0, \tilde{\tau}^2)$, then the iterates τ_t^2 will reach the interval $[\tilde{\tau}^2, \infty)$ after a finite number of steps by (2.32), and we are led back to the first case. □

We can now show that the system (2.28) has an unique solution $(\delta, \tau^2, \gamma) \in (\sqrt{c/2}, \infty) \times (0, \infty) \times (0, 1)$. To that end, we reuse the notations of the previous proof.

For a given $\delta > \sqrt{c/2}$, the previous lemma shows that Equation (2.28b) admits an unique solution $\tau^2(\delta)$. Plugging this value in the right hand side of Equation (2.28c), we can write

$$\gamma = \gamma(\delta) = \mathbb{E}_{\bar{r}} \left[\bar{\rho}\left(-\frac{\bar{r}}{\tau(\delta)} \right) \right] = \left. \frac{df(\tau^2)}{d\tau^2} \right|_{\tau^2 = \tau^2(\delta)}. \tag{2.33}$$

All what remains to show is that the equation

$$\kappa = (1 + \delta) \left(1 + \frac{c\gamma(\delta)}{\delta} \right) \tag{2.34}$$

has an unique solution $\delta > \sqrt{c/2}$. We thus need to study the behavior of $\gamma(\delta)$. In all the remainder, differentiability issues can be easily checked and are skipped.

In the previous proof, $\tau^2(\delta)$ has been obtained through Banach's fixed point theorem, which implies that

$$\left. \frac{df(\tau^2)}{d\tau^2} \right|_{\tau^2 = \tau^2(\delta)} < \frac{\delta^2}{c}. \tag{2.35}$$

Recalling that $df(\tau^2)/d\tau^2$ decreases from 1 asymptotically to $1/2$ as τ^2 increases from 0 to ∞, the last inequality has two consequences. First, $\tau^2(\delta) \to \infty$ as $\delta \downarrow \sqrt{c/2}$.

Second, we obtain by using Equation (2.30) and taking the derivatives with respect to δ that

$$\frac{d\tau^2(\delta)}{d\delta}\left(1 - \frac{c}{\delta^2}\left.\frac{df(\tau^2)}{d\tau^2}\right|_{\tau^2=\tau^2(\delta)}\right) = -\frac{2cf(\tau^2(\delta))}{\delta^3},$$

which shows that $\tau^2(\delta)$ is a decreasing function. Getting back to Equation (2.33), and observing that $\bar{\rho}(\cdot)$ is a decreasing function, we obtain that $\gamma(\delta)$ is increasing.

We now have all the elements to study Equation (2.34).
For δ decreasing to $\sqrt{c/2}$, $\tau^2(\delta)$ goes to ∞ by what precedes, thus, $\gamma(\delta)$ decreases to $1/2$, and $(1+\delta)\left(1+\frac{c\gamma(\delta)}{\delta}\right) \to \left(1+\sqrt{c/2}\right)\left(1+\frac{c/2}{\sqrt{c/2}}\right) < \kappa$.
For δ near infinity, $(1+\delta)\left(1+\frac{c\gamma(\delta)}{\delta}\right) \sim \delta > \kappa$. Consequently, Equation (2.34) has a solution by continuity. To establish uniqueness, we show that the function $\delta \mapsto (1+\delta)\left(1+\frac{c\gamma(\delta)}{\delta}\right)$ is increasing. Indeed,

$$\frac{d}{d\delta}\left[(1+\delta)\left(1+\frac{c\gamma(\delta)}{\delta}\right)\right] = 1 + \frac{c\gamma'(\delta)}{\delta} - \frac{c\gamma(\delta)}{\delta^2} + c\gamma'(\delta) \geq 1 - \frac{c\gamma(\delta)}{\delta^2} > 0$$

as shown by Inequality (2.35), and we are done.

Study of the LCP solution via asymetric AMP

Let (n) be a sequence of integers growing to infinity, and let $(p = p(n))$ be another sequence of integers growing to infinity in such a way that $0 < \inf p/n \leq \sup p/n < \infty$.

For each couple (p, n), let $A_n \in \mathbb{R}^{p\times n}$ be a random matrix consisting in pn independent $\mathcal{N}(0,1)$ elements.
Let $a_n \in \mathbb{R}^n$, $b_p \in \mathbb{R}^p$ be the random vectors of auxiliary information.
For $t = 0, 1, \ldots$, let $(f^{t+1})_t$ and $(g^t)_t$ be two sequences of Lipschitz $\mathbb{R}^2 \to \mathbb{R}$ functions.
Starting with $g^{-1} := 0$, $\langle \partial_1 f^0(u_n^0, a_n)\rangle_p \in \mathbb{R}$ and $f^0(u_n^0, a_n) \in \mathbb{R}^n$, the asymmetric AMP recursion is given by

$$u_n^{t+1} := \frac{A_n^*}{\sqrt{p}} g^t(v_p^t, b_p) \; - \; \langle \partial_1 g^t(v_p^t, b_p)\rangle_p \; f^t(u_n^t, a_n)\,, \tag{2.36}$$

$$v_p^t := \frac{A_n}{\sqrt{p}} f^t(u_n^t, a_n) \; - \; \langle \partial_1 f^t(u_n^t, a_n)\rangle_p \; g^{t-1}(v_p^{t-1}, b_p)\,. \tag{2.37}$$

Here, writing $u_n^t = \left[u_{i,n}^t\right]_{i=1}^n$, $v_p^t = \left[v_{i,p}^t\right]_{i=1}^p$, $a_n = [a_{i,n}]_{i=1}^n$ and $b_p = [b_{i,p}]_{i=1}^p$, we set

$$f^t(u_n^t, a_n) = \left[f^t(u_{i,n}^t, a_{i,n})\right]_{i=1}^n \in \mathbb{R}^n \;\text{ and }\; g^t(v_p^t, b_p) = \left[g^t(v_{i,p}^t, b_{i,p})\right]_{i=1}^p \in \mathbb{R}^p\,.$$

Observing that the derivatives $df^t(u,a)/du$ and $dg^t(v,b)/dv$ are defined almost everywhere thanks to the Lipschitz property of f^t and g^t, the functions $\partial_1 f^t(u,a)$ and $\partial_1 g^t(v,b)$ are any functions that coincide with those derivatives where they are defined. Finally,

$$\langle \partial_1 f^t(u^t, a)\rangle_p = \frac{1}{p}\sum_{j=1}^n \partial_1 f^t(u_j^t, a_j) \;\text{ and }\; \langle \partial_1 g^t(v^t, b)\rangle_p = \frac{1}{p}\sum_{j=1}^p \partial_1 g^t(v_j^t, b_j)\,.$$

With this construction, it turns out to be possible to evaluate precisely the asymptotic behavior of the empirical measures of the type $\mu^{a_n,u_n^1,\ldots,u_n^t}$ and $\mu^{b_p,v_p^0,v_p^1,\ldots,v_p^t}$ as $n \to \infty$ for each value of t, leading to the so-called Density Evolution (DE) equations that characterize the limits of these measures. The so-called "Onsager terms" $\langle \cdots \rangle_p f^t(u_n^t, a_n)$ and $\langle \cdots \rangle_p g^{t-1}(v_p^{t-1}, b_p)$ (equal to zero for $t = 0$) play a crucial role in making possible those convergences. For a detailed exposition of the AMP theory, along with the description of many of its applications, the reader is referred to the recent tutorial [Feng *et al.*, 2022].

To establish Theorem 35, we shall study the properties of the following AMP algorithm. For each n, let $(u_n^0, a_n) \in \mathbb{R}^n \times \mathbb{R}^n$ be a couple of random vectors independent of A_n, and such that $a_n \succcurlyeq 0$. Assume that there exists a couple of L^2 random variables $(\bar{u}, \bar{a})$, with $\bar{a} \neq 0$, such that

$$\mu^{u_n^0,a_n} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R}^2)} \mathcal{L}((\bar{u},\bar{a})) \quad \text{almost surely.} \tag{2.38}$$

We can now take u_n^0 as the deterministic vector $u_n^0 = \mathbf{1}_n$ and $b_p = 0_p$; the vector a_n will be precised later. Note that $\bar{a} \geq 0$. Take $f^t \equiv f$ and $h^t \equiv h$ for each integer $t \geq 0$, where

$$f(u,a) = \frac{(u+a)_+}{\delta} \quad \text{and} \quad g(v,b) = v - b\,,$$

and where $\delta > \sqrt{c/2}$ solves the system (2.28). Our AMP iteration reads then

$$\begin{aligned} u_n^{t+1} &= \frac{A_n^*}{\sqrt{p}} v_p^t - \frac{(u_n^t + a_n)_+}{\delta}\,, \\ v_p^t &= \frac{A_n}{\delta\sqrt{p}}(u_n^t + a_n)_+ - \frac{\langle \mathbb{1}_{u_n^t + a_n > 0}\rangle_p}{\delta} v_p^{t-1}\,. \end{aligned} \tag{2.39}$$

The DE equations for this algorithm are provided by the following proposition, which is a direct application of, *e.g.*, [Feng *et al.*, 2022, Theorem 2.5] :

Proposition 37. Consider the algorithm (2.39). Then, for each $t \geq 1$,

$$\mu^{a_n,u_n^1,\ldots,u_n^t} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R}^{t+1})} \mathcal{L}((\bar{a}, Z_\tau^1, \ldots, Z_\tau^t)) \quad \text{almost surely,}$$

and

$$\mu^{v_p^0,v_p^1,\ldots,v_p^t} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R}^{t+1})} \mathcal{L}((Z_\sigma^0, Z_\sigma^1, \ldots, Z_\sigma^t)) \quad \text{almost surely,}$$

where $(Z_\tau^1, \ldots, Z_\tau^t)$ is a centered Gaussian vector, independent of $(\bar{u}, \bar{a})$, which covariance matrix is $\bar{T}^{[t]}$ and $(Z_\sigma^0, \ldots, Z_\sigma^t)$ is a centered Gaussian vector with covariance matrix $\bar{\Sigma}^{[t]}$.

$\bar{T}^{[t]}$ and $\bar{\Sigma}^{[t]}$ are recursively defined in t as follows :

$$\bar{T}^{[1]} = \bar{T}_{1,1} = \mathbb{E}(Z_\tau^1)^2\,, \quad \bar{\Sigma}^{[1]} = \bar{\Sigma}_{0,0}.$$

Given $\bar{T}^{[t]}$ and $\bar{\Sigma}^{[t]}$, only the last row and column of $\bar{T}^{[t+1]}$ and of $\bar{\Sigma}^{[t+1]}$ have to be determined. This is done through the equations

$$\left[\bar{T}^{[t+1]}\right]_{t+1,k} = \bar{T}_{t+1,k} = \bar{T}_{k,t+1} \triangleq \operatorname{Cov}\left(Z_\tau^k,\ Z_\tau^{t+1}\right) = \mathbb{E}\left(Z_\tau^{t+1} Z_\tau^k\right) = \mathbb{E}\left(Z_\sigma^t Z_\sigma^{k-1}\right)\,,$$

and

$$\left[\bar{\Sigma}^{[t+1]}\right]_{t+1,k} = \bar{\Sigma}_{t,k-1} = \bar{\Sigma}_{k-1,t} \triangleq \mathrm{Cov}\left(Z_\sigma^{k-1},\ Z_\sigma^t\right)$$
$$= \frac{c}{\delta^2}\begin{cases} \mathbb{E}\left[(Z_\tau^t+\bar{a})_+(Z_\tau^{k-1}+\bar{a})_+\right] & \text{if} \quad k\in\{2,\dots,t+1\}, \\ \mathbb{E}\left[(Z_\tau^t+\bar{a})_+(\bar{u}+\bar{a})_+\right] & \text{if} \quad k=1. \end{cases},$$

AMP iterates in an LCP form, and DE estimates

Write $x_n^t = u_n^t + a_n$, $y_p^t = v_p^t$, and let $\tilde{\gamma}_n^t = \langle \mathbb{1}_{(u_n^t+a_n)>0}\rangle_p = \langle \mathbb{1}_{x_n^t>0}\rangle_p$ and $\gamma_n^t = \tilde{\gamma}_n^t/c$. Recall the expression of γ in (2.28c). With these notations, the AMP recursion can be rewritten as

$$\begin{aligned} y_p^t &= \frac{A_n}{\delta\sqrt{p}}(x_n^t)_+ - \frac{\tilde{\gamma}_n^t}{\delta}y_p^{t-1} \\ &= \frac{A_n}{\delta\sqrt{p}}(x_n^t)_+ - \frac{c\gamma}{\delta}y_p^{t-1} + \frac{c\gamma - c\gamma_n^t}{\delta}y_p^{t-1}, \\ &= \frac{A_n}{\delta\sqrt{p}}(x_n^t)_+ - \frac{c\gamma}{\delta}y_p^{t} + \frac{c\gamma - c\gamma_n^t}{\delta}y_p^{t-1} + \frac{c\gamma}{\delta}\left(y_p^t - y_p^{t-1}\right), \end{aligned}$$

which yields

$$\left(1+\frac{c\gamma}{\delta}\right)y_p^t = \frac{A_n}{\delta\sqrt{p}}(x_n^t)_+ + \frac{c\gamma - c\gamma_n^t}{\delta}y_p^{t-1} + \frac{c\gamma}{\delta}\left(y_p^t - y_p^{t-1}\right),$$

i.e.

$$y_p^t = \frac{A_n}{\left(1+\frac{c\gamma}{\delta}\right)\delta\sqrt{p}}(x_n^t)_+ + \frac{1}{1+\frac{c\gamma}{\delta}}\left(\frac{c\gamma - c\gamma_n^t}{\delta}y_p^{t-1} + \frac{c\gamma}{\delta}\left(y_p^t - y_p^{t-1}\right)\right).$$

In this way, x_n^{t+1} can be rewritten :

$$x_n^{t+1} = \frac{A_n^* A_n}{\left(1+\frac{c\gamma}{\delta}\right)\delta p}(x_n^t)_+ + \frac{A_n^*}{\left(1+\frac{c\gamma}{\delta}\right)\sqrt{p}}\left(c\frac{\gamma-\gamma_n^t}{\delta}y_p^{t-1} + \frac{c\gamma}{\delta}\left(y_p^t - y_p^{t-1}\right)\right) - \frac{1}{\delta}(x_n^t)_+ + a_n,$$

hence,

$$x_n^t = \frac{A_n^* A_n}{\left(1+\frac{c\gamma}{\delta}\right)\delta p}(x_n^t)_+ - \frac{1}{\delta}(x_n^t)_+ + a_n + \varepsilon_n^t,$$

where

$$\varepsilon_n^t = \frac{A_n^*}{\left(1+\frac{c\gamma}{\delta}\right)\sqrt{p}}\left(c\frac{\gamma-\gamma_n^t}{\delta}y_p^{t-1} + \frac{c\gamma}{\delta}\left(y_p^t - y_p^{t-1}\right)\right) + x_n^t - x_n^{t+1}. \tag{2.40}$$

We now put the next to last equation under the form of a LCP problem. Specifically, writing $x = x_+ - x_-$ for a vector x, we have

$$-(x_n^t)_- = \frac{A_n^* A_n}{\left(1+\frac{c\gamma}{\delta}\right)\delta p}(x_n^t)_+ - \left(1+\frac{1}{\delta}\right)(x_n^t)_+ + a_n + \varepsilon_n^t,$$

in other words,

$$\frac{(x_n^t)_-}{1+1/\delta} = \left(I_n - \frac{A_n^* A_n}{(1+\delta)(1+c\gamma/\delta)p}\right)(x_n^t)_+ - \frac{a_n + \varepsilon_n^t}{1+1/\delta}.$$

Remember that $\kappa = (1+\delta)(1+c\gamma/\delta) = (c\gamma+\delta)(1+1/\delta)$ and that $M_n = A_n^* A_n/(\kappa p)$. Thus, when $\|M_n\| < 1$, we have

$$(x_n^t)_+ = \text{LCP}\left(I_n - M_n,\ -\frac{a_n + \varepsilon_n^t}{1+1/\delta}\right). \tag{2.41}$$

By putting

$$a_n = (1+1/\delta) r_n,$$

as we shall do hereinafter, this equation can be compared with (2.12). Note that if we set, *e.g.*, $u_n^0 = \mathbf{1}_n$, then the assumption (2.38) on $\mu^{u_n^0, a_n}$ is satisfied thanks to Assumption 2 with $\bar{a} = (1+1/\delta)\bar{r}$.

Before bounding ε_n^t in (2.41), let us study the behavior of $\mu^{(x_n^t)_+}$ with the help of Proposition 37. Recall that $\bar{Z} \sim \mathcal{N}(0,1)$. Proposition 37 shows that for $t \geq 0$, $\mu^{u_n^{t+1}} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}(Z_\tau^{t+1})$ and $\mu^{v_n^t} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}(Z_\sigma^t)$, with $Z_\tau^{t+1} \sim \boldsymbol{\tau}_{t+1}\bar{Z}$ and $Z_\sigma^{t+1} \sim \boldsymbol{\sigma}_t \bar{Z}$, where

$$\boldsymbol{\tau}_{t+1}^2 := \mathbb{E}\left(g(Z_\sigma^t, \bar{b})^2\right) = \mathbb{E}\left((Z_\sigma^t)^2\right) = \boldsymbol{\sigma}_t^2\,,$$

and

$$\boldsymbol{\sigma}_{t+1}^2 := c\mathbb{E}\left(f(Z_\tau^t, \bar{a})^2\right) = \frac{c}{\delta^2}\mathbb{E}\left((\boldsymbol{\tau}_{t+1}\bar{Z} + (1+1/\delta)\bar{r})_+^2\right),$$

which give :

$$\boldsymbol{\tau}_{t+1}^2 = \frac{c}{\delta^2}\mathbb{E}\left((\boldsymbol{\tau}_t\bar{Z} + (1+1/\delta)\bar{r})_+^2\right) \quad \text{and} \quad \boldsymbol{\sigma}_{t+1}^2 = \frac{c}{\delta^2}\mathbb{E}\left((\boldsymbol{\sigma}_t\bar{Z} + (1+1/\delta)\bar{r})_+^2\right).$$

Since the function $(u,a) \mapsto (u+a)_+$ is Lipschitz, it is clear that $\mu^{(x_n^t)_+} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}((\boldsymbol{\tau}_t\bar{Z} + (1+1/\delta)\bar{r})_+)$. Furthermore, since the distribution function of the law of $\boldsymbol{\tau}_t\bar{Z} + (1+1/\delta)\bar{r}$ has no discontinuity, it holds that

$$\gamma_n^t \to \boldsymbol{\gamma}^t \triangleq \mathbb{P}\left[\boldsymbol{\tau}_t\bar{Z} + (1+1/\delta)\bar{r} > 0\right] = \mathbb{E}_{\bar{r}}\mathbb{P}_{\bar{Z}}\left[\boldsymbol{\tau}_t\bar{Z} + (1+1/\delta)\bar{r} > 0\right].$$

Now, by replacing $\bar{r}$ with $(1+1/\delta)\bar{r}$ in the statement of Lemma 36, we obtain by this lemma that the sequence $(\boldsymbol{\tau}_t^2)$ converges to $\boldsymbol{\tau}_\infty^2$ defined as the unique solution of the equation

$$\boldsymbol{\tau}_\infty^2 = \frac{c}{\delta^2}\mathbb{E}\left((\boldsymbol{\tau}_\infty\bar{Z} + (1+1/\delta)\bar{r})_+^2\right).$$

A consequence is that the sequence $(\boldsymbol{\gamma}^t)$ converges to $\boldsymbol{\gamma}^\infty$ defined as

$$\boldsymbol{\gamma}^\infty = \mathbb{E}_{\bar{r}}\mathbb{P}_{\bar{Z}}\left[\boldsymbol{\tau}_\infty\bar{Z} + (1+1/\delta)\bar{r} > 0\right].$$

By comparing these two equations with (2.28b) and (2.28c) respectively, we easily see that

$$\boldsymbol{\gamma}^\infty = \gamma, \quad \text{and} \quad \frac{\boldsymbol{\tau}_\infty^2}{(1+1/\delta)^2} = \tau^2.$$

Since $(x_n^t)_+$ is designed to be an approximation of x_n^*, our purpose is to show that

$$\mu^{x_n^*} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \boldsymbol{\mu}^* \triangleq \mathcal{L}\left((\boldsymbol{\tau}_\infty\bar{Z} + (1+1/\delta)\bar{r})_+\right) = \mathcal{L}((1+1/\delta)(\tau\bar{Z} + \bar{r})_+) \tag{2.42}$$

to establish Theorem 35.

Control of the error term ε_n^t

Getting back to the expression (2.40) of ε_n^t, it holds by what precedes that

$$\frac{\|y_p^{t-1}\|^2}{n} \to \mathbb{E}\left((\boldsymbol{\sigma}_{t-1}\bar{Z})^2\right) = \boldsymbol{\sigma}_{t-1}^2 = \boldsymbol{\tau}_t^2\,,$$

which is bounded,

$$\underset{n\to\infty}{\text{aslim}} \left\|\frac{A_n^*}{\sqrt{p}}\right\|^2 = \left(1+\sqrt{c}\right)^2\,,$$

and furthermore $\lim_t \text{aslim}_n(\gamma - \gamma_n^t) = 0$. We thus have

$$\lim_{t\to\infty} \underset{n\to\infty}{\text{aslim}} \left\|\frac{A_n^*}{\sqrt{p}}\right\|^2 \frac{c^2(\gamma-\gamma_n^t)^2}{(\delta+\gamma)^2}\frac{\|y_p^{t-1}\|^2}{n} = 0. \tag{2.43}$$

The main idea to control the two other terms at the right hand side of (2.40) is to show that the correlation coefficient $\mathbb{E}(Z_\tau^{t+1}Z_\tau^t)/(\boldsymbol{\tau}_{t+1}\boldsymbol{\tau}_t)$ converges to 1 as $t\to\infty$. This was done in a similar context in [Montanari et Richard, 2016], see also [Donoho et Montanari, 2016]. For self-containedness, we summarize herein their approach :

Lemma 38. The sequences (Q_t) and $(\tilde{Q}_t)$ defined as

$$Q_t \triangleq \frac{\mathbb{E}(Z_\tau^{t+1}Z_\tau^t)}{\boldsymbol{\tau}_{t+1}\boldsymbol{\tau}_t} \quad \text{and} \quad \tilde{Q}_t \triangleq \frac{\mathbb{E}(Z_\sigma^{t+1}Z_\sigma^t)}{\boldsymbol{\sigma}_{t+1}\boldsymbol{\sigma}_t}$$

starting with Q_1 and $\tilde{Q}_1$ satisfy $Q_t \xrightarrow[t\to\infty]{} 1$ and $\tilde{Q}_t \xrightarrow[t\to\infty]{} 1$.

The proof of this lemma can be done following the exact same reasoning as for Lemma 34. For $\tilde{Q}$, one will consider $\sigma_t = \tau_{t+1}$.

Taking $\varphi(x_1,x_2) = (x_1-x_2)^2 \in \text{PL}_2(2)$ and using Proposition 37, we have

$$\frac{\|x_n^t - x_n^{t+1}\|^2}{n} = \frac{1}{n}\sum_{i=1}^n \varphi(u_{i,n}^t, u_{i,n}^{t+1}) \to \mathbb{E}\left[\left(Z_\tau^{t+1}-Z_\tau^t\right)^2\right] = \boldsymbol{\tau}_{t+1}^2 + \boldsymbol{\tau}_t^2 - 2\boldsymbol{\tau}_{t+1}\boldsymbol{\tau}_t Q_t\,,$$

and

$$\frac{\|y_p^t - y_p^{t-1}\|^2}{n} = \frac{c}{p}\sum_{i=1}^p \varphi(v_{i,p}^t, v_{i,p}^{t-1}) \to \mathbb{E}\left[\left(Z_\sigma^t - Z_\sigma^{t-1}\right)^2\right] = \boldsymbol{\sigma}_t^2 + \boldsymbol{\sigma}_{t-1}^2 - 2\boldsymbol{\sigma}_t\boldsymbol{\sigma}_{t-1}\tilde{Q}_{t-1}\,,$$

Then, it holds by Lemma 38 that

$$\lim_{t\to\infty} \underset{n\to\infty}{\text{aslim}} \frac{\|x_n^t - x_n^{t+1}\|^2}{n} = 0 \quad \text{and} \quad \lim_{t\to\infty} \underset{n\to\infty}{\text{aslim}} \frac{\|y_p^t - y_p^{t-1}\|^2}{n} = 0. \tag{2.44}$$

Combining this convergence with (2.43), and using that

$$\frac{\|\varepsilon_n^t\|^2}{n} \leq 3\left[\left\|\frac{A_n^*}{\sqrt{p}}\right\|^2 \left(\frac{c^2(\gamma-\gamma_n^t)^2}{(c\gamma+\delta)^2}\frac{\|(y_p^{t-1})_+\|^2}{n} + \frac{c^2\gamma^2}{(c\gamma+\delta)^2}\frac{\|y_p^t - y_p^{t-1}\|^2}{n}\right) + \frac{\|x_n^t - x_n^{t+1}\|^2}{n}\right],$$

we get that

$$\lim_{t\to\infty} \underset{n\to\infty}{\text{aslim}} \frac{\|\varepsilon_n^t\|^2}{n} = 0 \tag{2.45}$$

(that the $\underset{n\to\infty}{\text{aslim}}$ at the left hand side exists can be deduced again from Proposition 37).

Use of a LCP perturbation result

Once again, when $\|M_n\| < 1$, we can compare (2.41) with (2.12) by using the LCP perturbation results [Chen et Xiang, 2007, Theorems 2.7 and 2.8] :

$$\|x_n^* - (x_n^t)_+\| \leq \frac{\delta}{1+\delta} \left\| \Big(I - M_n\Big)^{-1} \right\| \; \|\varepsilon_n^t\| .$$

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be an arbitrary function in PL(2). For a given positive integer t, we have

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n \varphi(x_{i,n}^*) - \int \varphi d\boldsymbol{\mu}^* &= \frac{1}{n}\sum_{i=1}^n \big(\varphi(x_{i,n}^*) - \varphi((x_{i,n}^t)_+)\big) + \frac{1}{n}\sum_{i=1}^n \varphi((x_{i,n}^t)_+) - \int \varphi d\boldsymbol{\mu}^* \\ &\triangleq \xi_n^1(t) + \xi_n^2(t).\end{aligned}$$

Dealing with $\xi_n^2(t)$, we have from Proposition 37 that

$$\xi_n^2(t) \xrightarrow[n\to+\infty]{} \mathbb{E}\left[\varphi(\boldsymbol{\tau}_t \bar{Z} + \bar{a})_+\right] - \mathbb{E}\left[\varphi(\boldsymbol{\tau}_\infty \bar{Z} + \bar{a})_+\right] .$$

Denoting as L the Lipschitz constant of φ and using that the function $x \mapsto x_+$ is 1–Lipschitz and $|x_+| \leq |x|$, the limit satisfies

$$\left|\mathbb{E}\left(\varphi(\boldsymbol{\tau}_t \bar{Z} + \bar{a})_+ - \varphi(\boldsymbol{\tau}_\infty \bar{Z} + \bar{a})_+\right)\right| \leq L|\boldsymbol{\tau}_t - \boldsymbol{\tau}_\infty| \mathbb{E}\left(|\bar{Z}| \left(1 + |\boldsymbol{\tau}_t \bar{Z} + \bar{a}| + |\boldsymbol{\tau}_\infty \bar{Z} + \bar{a}|\right)\right) ,$$

which is easily seen to be bounded by a constant $C_2(t)$ that converges to zero as $t \to \infty$, since $\boldsymbol{\sigma}_t \to \boldsymbol{\sigma}_\infty$.

We now turn to $\xi_n^1(t)$, we have by the Cauchy-Schwarz inequality

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n \left|\varphi(x_{i,n}^*) - \varphi((x_{i,n}^t)_+)\right| &\leq \frac{L}{n}\sum_{i=1}^n \left|x_{i,n}^* - (x_{i,n}^t)_+\right| \left(1 + |x_{i,n}^*| + |(x_{i,n}^t)_+|\right) \\ &\leq \frac{L}{n}\|x_n^* - (x_n^t)_+\| \left(\sum_{i=1}^n (1 + |x_{i,n}^*| + |(x_{i,n}^t)_+|)^2\right)^{1/2} \\ &\leq 3L \frac{\|x_n^* - (x_n^t)_+\|}{\sqrt{n}} \left(1 + \frac{\|x_n^*\|}{\sqrt{n}} + \frac{\|(x_n^t)_+\|}{\sqrt{n}}\right) .\end{aligned}$$

Writing $B_n \triangleq \frac{\delta}{\delta+1}\|(I -_n)^{-1}\|$, we obtain that

$$|\xi_n^1(t)| \leq 3LB_n \frac{\|\varepsilon_n^t\|}{\sqrt{n}} \left(1 + 2\frac{\|(x_n^t)_+\|}{\sqrt{n}} + B_n \frac{\|\varepsilon_n^t\|}{\sqrt{n}}\right) .$$

We know from Assumption 3 that B_n converges almost surely to a positive constant. From Proposition 37, we furthermore have

$$\frac{\|(x_n^t)_+\|}{\sqrt{n}} \to \left(\mathbb{E}(\boldsymbol{\tau}_t \bar{Z} + \bar{a})_+^2\right)^{1/2} ,$$

which is bounded in t. Thus, using (2.45), we obtain that $\limsup_n |\xi_n^1(t)|$ is bounded with probability one by a constant $C_1(t)$ that converges to zero as $t \to \infty$.

We finally obtain that

$$\limsup_n \left| \frac{1}{n} \sum_{i=1}^{n} \varphi(x^*_{i,n}) - \int \varphi d\boldsymbol{\mu}^* \right| \leq C_1(t) + C_2(t) \quad \text{with probability one.}$$

Since $C_1(t) + C_2(t)$ can be made arbitrarily small, we have

$$\frac{1}{n} \sum_{i=1}^{n} \varphi(x^*_{i,n}) \to \int \varphi d\boldsymbol{\mu}^*,$$

which terminates the proof of Theorem 35.

Chapitre 3

A central limit theorem

Ce chapitre sera consacré à la démonstration d'un théorème central limite issu d'un système algébrique d'équations couplées dont les coefficients sont aléatoires. Il s'agit plus précisément d'une extension d'un résultat obtenu par [Geman et Hwang, 1982].

L'étude de ce théorème central limite est apparue lors d'une première approche du problème LCP, ce dernier ayant par la suite pu être étudié dans les cas GOE et Wishart en faisant appel à la théorie de l'AMP (Chapitre 2).

1 Introduction

We start by explaining what brings to study this CLT.

1.1 From the Linear Complementarity Problem to a fixed point equation

Let Y be a $n \times n$ matrix, q be a vector of size n. The LCP(Y, q) consists in finding two vectors w and z of size n such that :

$$\begin{cases} z & \geq 0\,, \\ w = Yz + q & \geq 0\,, \\ w^{\mathrm{T}} z & = 0\,. \end{cases}$$

Assume that Y is a $P-$matrix (definition 3.1). Then, using a result by [Murty, 1972], the LCP(Y, q) has a unique solution (w, z) for all $q \in \mathbb{R}^n$.

Let x be the vector of size n solution of the fixed point equation :

$$x = B|x| + b\,, \tag{3.1}$$

with

$$|x| = (|x_i|)_i \quad \text{and} \quad \begin{cases} b & = -(I+Y)^{-1} q\,, \\ B & = (I+Y)^{-1}(I-Y)\,. \end{cases}$$

The quantities $\begin{cases} z = |x| + x\,, \\ w = |x| - x\,, \end{cases}$ are the unique solution of the LCP(Y, q). Let

$$f(x) = b + B|x|\,.$$

If $\|Y\|_2 < 1$, by a consequence of [Horn et Johnson, 2013, Corollary 5.6.16] and using Banach fixed-point theorem, the recursive problem $x^{(r)} = f(x^{(r-1)})$ converges to the fixed point solution.
This first approach was promising in what follows.

1.2 From Lotka-Volterra equations to fixed point equation

Equilibrium of Lotka-Volterra equations and LCP

As it has been seen previously, in a LV system, the abundances are connected via the following coupled equations :

$$\frac{d\tilde{x}_k(t)}{dt} = \tilde{x}_k(t)\left(r_k - \tilde{x}_k(t) + \sum_{\ell\in[n]} M_{k\ell}\tilde{x}_\ell(t)\right) \qquad \text{for} \quad k \in [n]\,, \tag{3.2}$$

where $M_n = (M_{k\ell})$ stands for the interaction matrix, and r_k for the intrinsic growth of species k.

As usual, M_n is defined by

$$M_n = \frac{A_n}{\alpha_n\sqrt{n}}\,.$$

At the equilibrium $\frac{d\tilde{\boldsymbol{x}}_n}{dt} = 0$, the abundance vector $\tilde{\boldsymbol{x}}_n = (\tilde{x}_k)_{k\in[n]}$ solution of the system :

$$\tilde{x}_k\left(r_k - \tilde{x}_k + \sum_{\ell\in[n]} M_{k\ell}\tilde{x}_\ell\right) = 0 \qquad \text{for} \quad \tilde{x}_k \geq 0 \quad \text{and} \quad k\in[n]\,.$$

may arise from the LCP$(I_n - M_n, -r_n)$.

Fixed point equation

If $I_n - \frac{A_n}{\alpha_n\sqrt{n}}$ is a $P-$matrix, the fixed point equation (3.1) becomes :

$$x = \left(2I_n - \frac{A_n}{\alpha_n\sqrt{n}}\right)^{-1}(r_n + 2|x|) - |x|\,. \tag{3.3}$$

Lemma 39. Let A_n be a Wigner matrix (i.i.d. centred and reduced entries, with symmetry : $A_{ij} = A_{ji}$).
If $\alpha > 2$, then $I - \frac{A_n}{\alpha_n\sqrt{n}}$ is a $P-$matrix (a.s., for n big).

Remark 11.
- — The number of surviving species of (3.2) corresponds to the number of non-zeros entries of $z = |x| + x$, i.e. the number of positive components of x.
- — This brings us to study the distribution of x to estimate

$$\frac{1}{n}\mathbb{E}\left(\text{card}\{x_i > 0\}\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(\mathbb{1}_{x_i>0}\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{P}(x_i>0) \approx \mathbb{P}(x_1)\,,$$

assuming that the x_i are asymptotically i.i.d.

Studying the asymptotic distribution of x from this point fixe equation is not easy. For the GOE and the Wishart cases, this LCP problem has been solved in the previous chapter.
Before reaching this proof (and knowing about the AMP theory), we had been interested in an other recursive problem.

Alternative recursive problem

Let $g : \mathbb{R}^n \times \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}^n$ be the function :

$$g(x, M) := (2I - M)^{-1} (\mathbf{1}_n + 2|x|) - |x| .$$

Let $(M_\ell)_{\ell=1,2,\ldots}$ be a sequence of independent random matrices where $M_\ell = (M_{ij}^{(\ell)})_{1\le i,j\le n}$ is a $n \times n$ real random matrix.

The goal of this part is to study the problem :

$$\begin{cases} y_0 & \in \mathbb{R}^n , \\ y_{\ell+1} & = g(y_\ell, M_\ell) . \end{cases}$$

By independence of the matrices M_ℓ, (y_ℓ) is a Markov chain and we initially hope that asymptotically the distribution of y_ℓ converges to the law of a random variable y independent of M solution to the problem $\mathcal{L}(y) = \mathcal{L}(g(y, M))$.

1.3 Central Limit Theorem

Assumption 4. Let $(A_n)_{n=1,2,\ldots}$ be a sequence of random matrices where $A_n = (A_{ij}^{(n)})_{1\le i,j\le n}$ is a $n \times n$ real random matrix.
We assume that the variables $(A_{ij}^{(n)})_{1\le i,j\le n,\ n\ge1}$ are i.i.d with mean 0 and variance σ^2 and that $A_{11}^{(1)}$ has all its moments finite.

Assumption 5. Let $(x_n)_{n=1,2,\ldots}$ be a sequence of real random vectors where $x_n = (x_{i,n})_{i=1,\ldots,n}$.
We assume that the variables $(x_{i,n})_{1\le i\le n,\ n\ge1}$ are i.i.d with mean μ and variance θ^2 and that the moments of $x_{1,1}$ are finite (yet, all the moments of $|x_{1,1}|$ are finite).

The topic of interest is the fixed point of the following function :

$$\begin{aligned} f(x_n) &= \left(2I_n - \frac{A_n}{\alpha\sqrt{n}}\right)^{-1} (\mathbf{1}_n + 2|x_n|) - |x_n| , \\ &= \left(I_n - \frac{A_n}{2\alpha\sqrt{n}}\right)^{-1} \left(\frac{\mathbf{1}_n}{2} + |x_n|\right) - |x_n| . \end{aligned}$$

Let's have

$$m_n := \frac{\mathbf{1}_n}{2} + |x_n| , \quad z_n := \left(I_n - \frac{A_n}{2\alpha\sqrt{n}}\right)^{-1} m_n \quad \text{and} \quad y_n := f(x_n) - \frac{\mathbf{1}_n}{2} = z_n - m_n .$$

Hence, $\forall i \in \{1, \ldots, n\}$

$$z_{i,n} = m_{i,n} + \sum_{k\in\mathbb{N}^*} e_i^* \left(\frac{A_n}{2\alpha\sqrt{n}}\right)^k m_n \qquad \text{and} \qquad y_{i,n} = \sum_{k\in\mathbb{N}^*} e_i^* \left(\frac{A_n}{2\alpha\sqrt{n}}\right)^k m_n .$$

As we have the assumption that all the moments of $|x_{1,n}|$ are finite, it implies that all the moments of $m_{1,n}$ are also finite.
Denote by π the distribution of the variable $m_{1,1}$ and $\Delta := \mathbb{E}\left(m_{1,1}^2\right)$.

Theorem 40. Let α be a positive real number. Let assumptions 4 and 5 hold true.
We assume that $\sigma^2 < \alpha^2$ and that there exists a constant β such that $\forall k \geq 2$, $\mathbb{E}\left(|A_{11}^{(1)}|^k\right) \leq (2\alpha)^k k^{\beta k}$.
Let $q \in \mathbb{N}^*$ and I_q be the identity matrix of size q.
Then

$$(y_{1,n}, y_{2,n}, \ldots, y_{q,n}) \xrightarrow[n\to+\infty]{d} \mathcal{N}(0, \gamma^2 I_q)\,,$$

and

$$\gamma^2 = \frac{\sigma^2 \Delta}{4\alpha^2 - \sigma^2}\,.$$

Besides,

$$(z_{1,n}, z_{2,n}, \ldots, z_{q,n}) \xrightarrow[n\to+\infty]{d} \mathcal{N}(0, \gamma^2 I_q) * \pi^{\otimes q}\,,$$

where $*$ is the convolution of the two distributions.

2 Elements of Graph theory

Before diving into the core result of this chapter, we first recall some combinatorial definitions and useful properties.

We will consider *directed graphs* with possily multiple edges and directed circuits. The vertices are integers between 1 and n. An edge (x, y) is considered to be directed from x to y ; y is called the head and x is called the tail of the edge.
An isolated vertex is a vertex that is neither the head nor the tail of an edge.

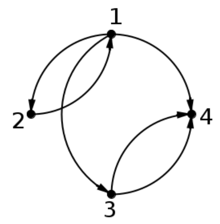
For a directed graph G, we denote V_G the set of its (non-isolated) vertices, $V_G \subset [\![1; n]\!]$, and E_G the set of its edges.
The *multiplicity* of an edge $(x, y) \subset V_G \times V_G$ is the number of edges in the graph G starting by the same vertex x and ending by the same vertex y. Some of the vertices are said *distinguished*.
We denote by v_G the number of non distinguished vertices and e_G the number of edges of G, counted without multiplicity.

Two directed graphs G and H are said to be *equivalent* (see figure 3.1 for an example), if and only if there exists a permutation $\sigma \in \mathcal{S}_n$ such that

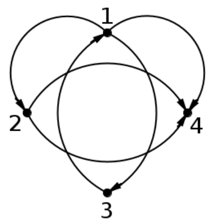
— $x \in V_G$ if and only if $\sigma(x) \in V_H$;
— the distinguished vertices of G are fixed points of σ ;
— if $x, y \in V_G$, the multiplicity of the edge (x, y) in E_G is the same as the multiplicity of the edge $(\sigma(x), \sigma(y))$ in E_H.

From this equivalence relation, we define the *equivalence class of a graph* G, that we will denote $\widehat{G}$ and we call G a *representative* of the class $\widehat{G}$.

A *directed path* of *length* $k \geq 1$ is a (finite or infinite) sequence of edges (i_1, j_1), $(i_2, j_2), \ldots, (i_k, j_k)$ with $i_{\ell+1} = j_\ell$.



(a) Graphe G



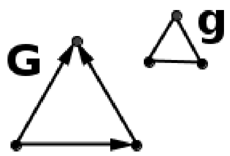
(b) Graphe H

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}$$

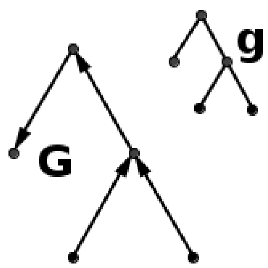
FIGURE 3.1 – Example of equivalent graphs G and H.

A *directed trail* is a directed path in which all edges are distinct.
A *directed circuit* is a non-empty finite directed trail in which the first and last vertices are the same.
A *directed acyclic graph* is a directed graph with no directed circuit.

We now consider a directed graph G, we will denote g the *associated undirected graph* : the edge $\{i, j\} \in E_g$ if and only if $(i, j) \in E_G$ or $(j, i) \in E_G$.
We say that two vertices are in the same *weakly connected component* of G if and only if they are in the same connected component of the undirected graph g.
In particular, if G is a directed acyclic graph and g a forest with q trees, we will say that G is a *directed forest* with q *directed trees* (see figure 3.2 and section 9.1 in [Deo, 2017]).



(a) Directed acyclic graph that is not a tree.



(b) Directed tree.

FIGURE 3.2 – Examples of skeleton of directed acyclic graphs.

We have the following lemma :

Lemma 41. Let H be a directed graph with q weakly connected components. Then, the number of vertices of H, written $\#V_H$, verifies :

$$\#V_H \leq e_H + q\,,$$

with equality if and only if H is a directed forest with q directed trees.

Démonstration. A proof by induction on the number of vertices will give the inequality.

Base case : Let H be the simpliest graph, made up of a single edge :
— H can have only one vertex (if its edge is a directed circuit), then : $\#V_H = 1 \leq e_H + q = 1 + 1 = 2$;
— or H has two vertices, then $\#V_H = 2 \leq e_H + q = 1 + 1 = 2$.

The property is clearly true in that case.

Induction step : Assume that H is a directed graph with q weakly connected components, such that $\#V_H \leq e_H + q$.
We add a new vertex to H. As we consider only non-isolated vertices, there are several paths we can take :

— we can link this new vertex to an existant weakly connected component of H. In that case, we have improve simultaneously the number of vertices $\#V_G$ and the number of edges without multiplicity e_H and the inequality stays true ;

— or we can add a directed circuit to the new vertex. In that case, we have created a new weakly connected components, so we improve all at once the number of vertices $\#V_H$, the number of edges without multiplicity e_H and the number of weakly connected components q.

— or we can link this new vertex to several existant weakly connected components of H. In that new case, the number of weakly connected components q decreases while the number of vertices $\#V_H$ and the number of edges without multiplicity e_H increase. However, e_H's increment is bigger than q's decrease and the inequality remains true.

For the case of equality, a proof by induction, reasoning on the undirected graph, will give it. □

3 Moments lemma

For all $n \geq 1,\ k \geq 1,\ 1 \leq i \leq n$, let's have

$$\begin{aligned}\nu(i,k,n) &:= e_i^* \left(\frac{A_n}{2\alpha\sqrt{n}}\right)^k m_n\,, \\ &= \left(\frac{1}{2\alpha\sqrt{n}}\right)^k \sum_{i_1,i_2,\ldots,i_k=1}^{n} A^{(n)}_{ii_1} A^{(n)}_{i_1 i_2} \ldots A^{(n)}_{i_{k-1}i_k} m_{i_k,n}\,.\end{aligned} \tag{3.4}$$

Lemma 42 (Moments lemma)**.** Let $(i_1,k_1),(i_2,k_2),\ldots,(i_s,k_s) \in \mathbb{N}^* \times \mathbb{N}$ be distinct pairs. Then :

$$(\nu(i_1,k_1,n),\nu(i_2,k_2,n),\ldots,\nu(i_s,k_s,n)) \xrightarrow[n\to+\infty]{d} (Z_1,Z_2,\ldots,Z_s)\,,$$

where the random variables $(Z_j)_{1\leq j\leq s}$ are independent and satisfy

$$Z_j \sim \begin{cases} \pi & \text{if } \ k_j = 0\,, \\ \mathcal{N}\left(0, \left(\frac{\sigma}{2\alpha}\right)^{2k_j}\Delta\right) & \text{else}\,. \end{cases}$$

Démonstration. **Strategy :** Let $n_1,\ldots,n_s$ be a family of positive integers. We will show that

$$\begin{aligned}&\mathbb{E}\left(\nu(i_1,k_1,n)^{n_1}\ldots\nu(i_s,k_s,n)^{n_s}\right) \\ &\quad\xrightarrow[n\to+\infty]{} \prod_{j=1}^{s}\left(\frac{\sigma}{2\alpha}\right)^{k_j n_j}\left[\Delta^{\frac{n_j}{2}}(n_j-1)!!\mathbb{1}_{k_j\geq 1}\mathbb{1}_{n_j\equiv 0[2]} + \mathbb{E}\left(m_{1,1}^{n_j}\right)\chi_{k_j=0}\right]\,,\end{aligned} \tag{3.5}$$

where $\nu(i,k,n)$ is defined by equation (3.4), $(n_j-1)!! = (n_j-1)\times(n_j-3)\times\cdots\times 3\times 1$ and $\mathbb{1}_E$ is the indicator function of the subset E.

In the sequel, we simply denote A (resp. m) instead of A_n (resp. m_n) and A_{ij} (resp. m_i) the entry $A_{ij}^{(n)}$ (resp. the i^{th} coordinate $m_{i,n}$) of this matrix A_n (resp. vector m_n).

The strategy of proof is divided in three steps. First, to understand the independence between the variables Z_j of the lemma 41, we study the case $s=2$ and $n_1=n_2=1$, i.e. $\mathbb{E}\left(\nu(i,k,n)\nu(j,\ell,n)\right)$. Then, we take a look at $\mathbb{E}\left(\nu(i,k,n)^p\right)$, which brings to the factor $(n_j-1)!!$ in equation (3.5). Finaly, we combine these two steps to get the convergence (3.5).

First part : Let's take a look at $\mathbb{E}\left(\nu(i,k,n)\nu(j,\ell,n)\right)$.

We start by the case where $k,\ell>0$.

For $I=(i,i_1,i_2,\dots,i_k)$, $J=(j,j_1,\dots,j_\ell)$, we denote $A_I := A_{ii_1}A_{i_1i_2}\dots A_{i_{k-1}i_k}$ and $m_I = m_{i_k}$. As well, $A_J := A_{jj_1}A_{j_1j_2}\dots A_{j_{\ell-1}j_\ell}$ and $m_J = m_{j_\ell}$.

For $i\in[\![1;n]\!]$, let $\mathcal{I}_{i,k}$ be the set of $(k+1)-$tuples of integers in $[\![1;n]\!]$ starting by i. By independence of A and m, we have

$$\mathbb{E}\left(\nu(i,k,n)\nu(j,\ell,n)\right) = \left(\frac{1}{2\alpha\sqrt{n}}\right)^{k+\ell} \sum_{\substack{I\in\mathcal{I}_{i,k}\\ J\in\mathcal{I}_{j,\ell}}} \mathbb{E}\left(A_IA_J\right)\mathbb{E}\left(m_Im_J\right).$$

We set $i_0=i$. Let G_I be the directed path with vertices $V_{G_I}=\{i,i_1,i_2,\dots,i_k\}$ and edges $E_{G_I}=\{(i_m,i_{m+1}) \mid m\in\{0,\dots,k-1\}\}$.

We can notice that G_I can also be seen as a directed path starting by vertex i of length k (the graph can contain directed circuits).

The vertices i and j are distinguished vertices.

For I and J given, we write $C_{IJ} := \mathbb{E}\left(A_IA_J\right)\mathbb{E}\left(m_Im_J\right)$.

If there is I' and J' such that $G_{I'}\cup G_{J'}\in\widehat{G_I\cup G_J}$, then

$$C_{IJ} = C_{I'J'} =: C_{\widehat{G_I\cup G_J}} < \infty\,,$$

i.e. all the elements in the same equivalence class have the same contribution.

Let define $\mathcal{G}_{\{(i,k),(j,\ell)\}}$ the set of equivalence classes of graphs composed by the union of a directed walk starting by vertex i of length k and a directed walk starting by vertex j of length ℓ.

Hence,

$$\begin{aligned}\mathbb{E}\left(\nu(i,k,n)\nu(j,\ell,n)\right) &= \left(\frac{1}{2\alpha\sqrt{n}}\right)^{k+\ell} \sum_{\widehat{G}\in\mathcal{G}_{\{(i,k),(j,\ell)\}}} \sum_{\substack{I,J \text{ s.t.}\\ G_I\cup G_J\in\widehat{G}}} C_{\widehat{G}}\,,\\ &= \left(\frac{1}{2\alpha\sqrt{n}}\right)^{k+\ell} \sum_{\widehat{G}\in\mathcal{G}_{\{(i,k),(j,\ell)\}}} C_{\widehat{G}} \sum_{\substack{I,J \text{ s.t.}\\ G_I\cup G_J\in\widehat{G}}} 1\,.\end{aligned}$$

Moreover, for a given equivalence class $\widehat{G}$, all graphs have the same number of distinct edges and we call it $v_{\widehat{G}}$. Then,

$$\sum_{\substack{I,J \text{ s.t.} \\ G_I \cup G_J \in \widehat{G}}} 1 = n \times (n-1) \times \cdots \times (n - v_{\widehat{G}} + 1)\,,$$
$$= n^{v_{\widehat{G}}}(1 + o(1))\,.$$

Hence

$$\mathbb{E}\left(\nu(i,k,n)\nu(j,\ell,n)\right) = \left(\frac{1}{2\alpha\sqrt{n}}\right)^{k+\ell} \sum_{\widehat{G} \in \mathcal{G}_{\{(i,k),(j,\ell)\}}} n^{v_{\widehat{G}}}(1+o(1))C_{\widehat{G}}\,.$$

For i, j, k and ℓ fixed, the number of equivalence classes, i.e. the number of elements of $\mathcal{G}_{\{(i,k),(j,\ell)\}}$, does not depend on n.
We are interested by the equivalence classes that will bring an asymptotic contribution, i.e. such that $v_{\widehat{G}} \geq \frac{k+\ell}{2}$ and $C_{\widehat{G}} \neq 0$. First, it is the skeleton of the contributive classes (which is the same for all the graphs of a given class) that interests us. Then, we will take a look to the number of elements in each contributive class.

Let $\widehat{H} \in \mathcal{G}_{\{(i,k),(j,\ell)\}}$.
By definition, there exist $I = (i, i_1, i_2, \ldots, i_k)$ a $(k+1)-$tuples of integers in $[\![1;n]\!]$ starting by i and $J = (j, j_1, \ldots, j_\ell)$ $(\ell+1)-$tuples of integers in $[\![1;n]\!]$ starting by j and $G = G_I \cup G_J$ such that $G \in \widehat{G_I \cup G_J} = \widehat{H}$.
G is a representative of the equivalence class $\widehat{G} = \widehat{H} \in \mathcal{G}_{\{(i,k),(j,\ell)\}}$.

$$C_{\widehat{G}} = \mathbb{E}\left(A_I A_J\right) \mathbb{E}\left(m_I m_J\right) = \mathbb{E}\left(A_{ii_1} \ldots A_{i_{k-1}i_k} A_{jj_1} \ldots A_{j_{\ell-1}j_\ell}\right) \mathbb{E}\left(m_{i_k} m_{j_\ell}\right)\,.$$

In order to get $C_{\widehat{G}} \neq 0$, the multiplicity of each directed edge of G has to be at least 2.
Thus,

$$e_{\widehat{G}} \leq \frac{k+\ell}{2}\,.$$

Following lemma 41, we have these inequalities :

- If $i = j$, G has only one distinguished vertex and only one weakly connected component, so

$$v_G = \#V_G - 1 \leq e_G\,, \text{with equality if } G \text{ is a directed tree.}$$

- If $i \neq j$, G has two distingueshed vertices and at most two weakly connected components

$$\#V_G \leq e_G + \#\{\text{weakly connected components of } G\}\,,\ \text{ so } v_G = \#V_G - 2 \leq e_G\,,$$

  with equality if G is a directed forest with two distinct directed trees.

Hence,

$$v_{\widehat{G}} \leq e_{\widehat{G}} \leq \frac{k+\ell}{2}\,.$$

The equivalence class $\widehat{G}$ of G asymptotically contributes if and only if

$$v_{\widehat{G}} = e_{\widehat{G}} = \frac{k+\ell}{2}\,, \tag{3.6}$$

the first equality is verified if G is a directed tree when $i = j$ (resp. a directed forest when $i \neq j$) and the second one is verified if the multiplicity of each directed edge is exactly 2 (each edge is crossed exactly twice).

We now want to show that if G contributes asymptotically, then necessarily $I = J$ and all the elements of I are distinct, such that only one equivalence class contributes asymptotically.

We assume that the equivalence class $\widehat{G}$ of G contributes asymptotically, i.e. $\widehat{G}$ verifies the equalities (3.6).

We reason by absurdity and suppose that $i \neq j$:

— First case : G has two weakly connected components G_I and G_J.
As $E_{G_I} \cap E_{G_J} = \emptyset$, I must contain a directed circuit in order to allow G to contribute in the expectation (each edge must be crossed at least twice).
So G_I is not a tree and G doesn't contribute asymptotically, which leads to a contradiction.

— Second case : G has only one weakly connected component.
In this case,

$$v_G + 2 = \#V_G \leq e_G + 1 \text{ i.e. } v_G \leq e_G - 1\,,$$

which contradicts (3.6).

From this first step, we deduce that necessarily $i = j$ and G is a directed tree.

— So that G is a directed tree, the vertex i should appear only in the edges (i, i_1) and (i, j_1) (else, we would have a directed circuit).
Yet, in order to get all the edges crossed twice, the edges (i, i_1) and (i, j_1) have to be the same.
So $i_1 = j_1$.

— By iteration of this arguments, it appears that $\forall m \in [\min(k, \ell)]$, $i_m = j_m$.

— We assume that $k \neq \ell$.
Let's take $k < \ell$, then the path $\{(i_k, j_{k+1}), (j_{k+1}, j_{k+2}), \ldots, (j_{\ell-1}, j_\ell)\}$ withdraws so each edge is crossed exactly twice.
G would have a directed circuit and will no more be a tree, which is absurd.
So $k = \ell$.

To conclude : the equivalence class $\widehat{G}$ of G contributes asymptotically if $G = G_I \cup G_J$ is a directed tree such that $I = J$ and all the elements of I are distinct.

In this case

$$\begin{aligned} C_{\widehat{G}} &= \mathbb{E}\left(A_I^2\right)\mathbb{E}\left(m_I^2\right) = \mathbb{E}\left(A_{ii_1}^2\right)\mathbb{E}\left(A_{i_1i_2}^2\right)\ldots\mathbb{E}\left(A_{i_{k-1}i_k}^2\right)\mathbb{E}\left(m_{i_k}^2\right) \quad \text{by independence,}\\ &= \sigma^{2k}\Delta\,. \end{aligned}$$

We finally have the following convergence :

$$\mathbb{E}\left(\nu(i,k,n)\nu(j,\ell,n)\right) \xrightarrow[n\to+\infty]{} \begin{cases} 0 & \text{if } i \neq j \text{ or } k \neq \ell\,,\\ \left(\frac{\sigma}{2\alpha}\right)^{2k}\Delta & \text{if } i = j \text{ and } k = \ell\,. \end{cases}$$

Remark : we can now take a look to the case $k \geq 0$ and $\ell = 0$:

$$\mathbb{E}\left(\nu(i,k,n)\nu(j,0,n)\right) = \left(\frac{1}{2\alpha\sqrt{n}}\right)^k \sum_{\substack{I\in\mathcal{I}_{i,k}\\ J\in\mathcal{I}_{j,0}}} \mathbb{E}\left(A_I\right)\mathbb{E}\left(m_I m_J\right) .$$

By definition, $\mathcal{I}_{j,0}$ contains only the 1-tuple $J := (j)$. Hence,

$$\begin{aligned}\mathbb{E}\left(\nu(i,k,n)\nu(j,0,n)\right) &= \left(\frac{1}{2\alpha\sqrt{n}}\right)^k \sum_{\widehat{G}\in\mathcal{G}_{\{(i,k),(j,0)\}}} \sum_{\substack{I \text{ s.t.}\\ G_I\cup G_J\in\widehat{G}}} C_{\widehat{G}} ,\\ &= \left(\frac{1}{2\alpha\sqrt{n}}\right)^k \sum_{\widehat{G}\in\mathcal{G}_{\{(i,k),(j,0)\}}} n^{v_{\widehat{G}}}(1+o(1))C_{\widehat{G}} ,\end{aligned}$$

where

$$C_{\widehat{G}} = \mathbb{E}\left(A_I\right)\mathbb{E}\left(m_I m_j\right),$$

with $G = G_I \cup G_J$ a representative of this equivalence class.

One can see that if k is odd, then $\mathbb{E}\left(A_I\right) = 0$, which implies $C_{\widehat{G}} = 0$. So that the graph G contributes, k must be even.
For the same reasons than previously, one gets

$$v_{\widehat{G}} \leq e_{\widehat{G}} \leq \frac{k}{2} .$$

The equivalence class $\widehat{G}$ of G asymptotically contributes if and only if

$$v_{\widehat{G}} = e_{\widehat{G}} = \frac{k}{2} .$$

The first equality is verified when G_I is a directed tree when the second one is verified as soon as the multiplicity of each edge of G_I is 2.

Let's assume that $k \geq 2$, to get the edge (i, i_1) crossed twice, there exists $\ell \in \{1,\ldots,k-1\}$ such that $i_\ell = i$. Which implies that G_I contains a directed circuit, i.e G_I is not a directed tree, which is absurd.

So $k = 0$ and

$$\mathbb{E}\left(\nu(i,0,n)\nu(j,0,n)\right) = \begin{cases} \mathbb{E}(m_1)^2 & \text{if } \ i \neq j ,\\ \Delta & \text{if } \ i = j .\end{cases}$$

Finally,

$$\mathbb{E}\left(\nu(i,k,n)\nu(j,\ell,n)\right) \xrightarrow[n\to+\infty]{} \begin{cases} \left(\frac{\sigma}{2\alpha}\right)^{2k}\Delta & \text{if } \ i = j \text{ and } k = \ell \neq 0 ,\\ \mathbb{E}(m_1)^2 & \text{if } \ i \neq j \text{ and } k = \ell = 0 ,\\ 0 & \text{else.}\end{cases}$$

Second part : We are now interested in $\mathbb{E}\left(\nu(i,k,n)^p\right)$.

In a first time, we assume that $k \neq 0$.
To avoid any ambiguity between the identity matrix I_q and the $(k+1)$-tuples starting by i, we rewrite this last $J = (i, i_1, i_2, \ldots, i_k)$.
For $q \in \{1, \ldots, p\}$, we denote $J_q = (i, i_1^{(q)}, i_2^{(q)}, \ldots, i_k^{(q)})$, $A_{J_q} := A_{ii_1^{(q)}} A_{i_1^{(q)} i_2^{(q)}} \ldots A_{i_{k-1}^{(q)} i_k^{(q)}}$ and $m_{J_q} = m_{i_k^{(q)}}$.
Like previously, the independence between A and m gives

$$\mathbb{E}\left(\nu(i,k,n)^p\right) = \left(\frac{1}{2\alpha\sqrt{n}}\right)^{kp} \sum_{J_1, \ldots, J_p \in \mathcal{I}_{i,k}} \mathbb{E}\left(A_{J_1} A_{J_2} \ldots A_{J_p}\right) \mathbb{E}\left(m_{J_1} m_{J_2} \ldots m_{J_p}\right) .$$

We reuse the notation $\mathcal{G}_{\{(i,k)^p\}}$ for the set of equivalence classes of graphs composed by the union of p directed walks starting by vertex i of length k.

For $\widehat{H} \in \mathcal{G}_{\{(i,k)^p\}}$ given, there exists $J_1, \ldots, J_p$, $(k+1)$-tuples starting by i, such that the graph $G_{J_1} \cup G_{J_2} \cup \cdots \cup G_{J_p}$ is a representative of the equivalence class $\widehat{H}$. So,

$$C_{\widehat{H}} = \mathbb{E}\left(A_{ii_1^{(1)}} A_{i_1^{(1)} i_2^{(1)}} \ldots A_{i_{k-1}^{(1)} i_k^{(1)}} A_{ii_1^{(2)}} \ldots A_{i_{k-1}^{(p)} i_k^{(p)}}\right) \mathbb{E}\left(m_{i_k^{(1)}} m_{i_k^{(2)}} \ldots m_{i_k^{(p)}}\right) .$$

and

$$\begin{aligned}
\mathbb{E}\left(\nu(i,k,n)^p\right) &= \left(\frac{1}{2\alpha\sqrt{n}}\right)^{kp} \sum_{\widehat{G} \in \mathcal{G}_{\{(i,k)^p\}}} C_{\widehat{G}} \sum_{\substack{J_1, \ldots, J_p \in \mathcal{I}_{i,k} \text{ s.t.} \\ \widehat{G_{J_1} \cup \cdots \cup G_{J_p}} = \widehat{G}}} 1 \\
&= \left(\frac{1}{2\alpha\sqrt{n}}\right)^{kp} \sum_{\widehat{G} \in \mathcal{G}_{\{(i,k)^p\}}} C_{\widehat{G}}\, n \times (n-1) \times \cdots \times (n - v_{\widehat{G}} + 1) , \\
&= \left(\frac{1}{2\alpha\sqrt{n}}\right)^{kp} \sum_{\widehat{G} \in \mathcal{G}_{\{(i,k)^p\}}} n^{v_{\widehat{G}}} (1 + o(1)) C_{\widehat{G}} .
\end{aligned}$$

Let G be the graph $G_{J_1} \cup G_{J_2} \cup \cdots \cup G_{J_p}$, representative of the equivalence class $\widehat{H} = \widehat{G}$. Then G has only one weakly connected component and by lemma 41

$$v_G = \#V_G - 1 \leq e_G \;\text{ with equality if } G \text{ is a directed tree.}$$

Besides, so that $C_{\widehat{G}} \neq 0$, the multiplicity of each edge has to be at least 2. So

$$e_G \leq \frac{kp}{2} \quad \text{hence} \quad v_G \leq e_G \leq \frac{kp}{2} .$$

The equivalence class of G contributes if and only if $v_G = e_G = \frac{kp}{2}$, where the first equality if verified if G is a directed tree and the second if the multiplicity of each edge of G is exactly 2.
Thus, we look for the classes that asymptotically contribute in the expectation.

— As G is a directed tree, the vertex i should appear only in the edges $(i, i_1^{(1)})$, $(i, i_1^{(2)})$, ..., $(i, i_1^{(p)})$. If this condition is not respected, there will be a directed cycle and G will no more be a directed tree.

Moreover, for each edge to be crossed exactly twice, one needs

$$\#\{(i,i_1^{(1)}),(i,i_1^{(2)}),\dots,(i,i_1^{(p)})\} = \frac{p}{2}$$

i.e. p is even and $\forall j \in \{1,\dots,p\}$, $\exists! \ell \in \{1,\dots,p\}\backslash\{j\}$ s.t. $i_1^{(j)} = i_1^{(\ell)}$.

— Let $j,\ell \in \{1,\dots,p\}$, with $j \neq \ell$, such that $i_1^{(j)} = i_1^{(\ell)}$ (i.e. edges $(i,i_1^{(j)})$ and $(i,i_1^{(\ell)})$ are the same).
  - — If $i_2^{(j)} \neq i_2^{(\ell)}$ i.e. $(i_1^{(j)},i_2^{(j)}) \neq (i_1^{(\ell)},i_2^{(\ell)})$.
    Then, since edge $(i_1^{(j)},i_2^{(j)})$ is crossed twice, vertex $i_1^{(j)}$ should appear once again and not only in $i_1^{(j)}$ and $i_1^{(\ell)}$.
    And because the multiplicity of edge $(i,i_1^{(\ell)})$ can't be bigger than 2, this implies the existence of a directed circuit, which contradicts the first hypothesis.
  - — By reusing this argument, it comes

$$\forall q \in \{1,\dots,k\}\,, \quad i_q^{(j)} = i_q^{(\ell)}\,.$$

The equivalence class $\widehat{G}$ of G asymptotically contributes if $G = G_{J_1} \cup G_{J_2} \cup \cdots \cup G_{J_p}$ is such that $\forall q \in \{1,\dots,p\}$,

— J_q starts by vertex i and has $k+1$ separate vertices,

— and $\exists! \ell \in \{1,\dots,p\}\backslash\{q\}$ such that $J_q = J_\ell$.

There is as much equivalence classes that make a contribution $\sigma^{kp}\Delta^{\frac{p}{2}}$ than the number of pairings of $\{1,\dots,p\}$.

That finally gives, for $k \neq 0$,

$$\mathbb{E}\left(\nu(i,k,n)^p\right) \xrightarrow[n\to+\infty]{} \begin{cases} 0 & \text{if } p \text{ is odd,} \\ (p-1)!! \left(\frac{\sigma}{2\alpha}\right)^{kp} \Delta^{\frac{p}{2}} & \text{if } p \text{ is even.} \end{cases}$$

On the other hand, if $k=0$, it cames that :

$$\mathbb{E}\left(\nu(i,0,n)^p\right) = \mathbb{E}\left(m_i^p\right) = \mathbb{E}\left(m_1^p\right)\,.$$

Third part : general case.

The goal of this part consists in bringing together the arguments used in the two previous parts.

Let $(i_1,k_1),(i_2,k_2),\dots,(i_s,k_s) \in \mathbb{N}^* \times \mathbb{N}$ be distincts pairs of integers.

Let $n_1,n_2,\dots,n_s \in \mathbb{N}$.

Using the previous rating, it brings that :

$$\begin{aligned}
&\mathbb{E}\left(\nu(i_1,k_1,n)^{n_1}\nu(i_2,k_2,n)^{n_2}\dots\nu(i_s,k_s,n)^{n_s}\right) \\
&= \left(\frac{1}{2\alpha\sqrt{n}}\right)^{\sum_{j=1}^n k_j n_j} \sum_{\substack{I_1^{(1)},\dots,I_{n_1}^{(1)}\in\mathcal{I}_{i_1,k_1} \\ \vdots \\ I_1^{(s)},\dots,I_{n_s}^{(s)}\in\mathcal{I}_{i_s,k_s}}} \mathbb{E}\left(A_{I_1^{(1)}}\dots A_{I_{n_1}^{(1)}}A_{I_1^{(2)}}\dots A_{I_{n_s}^{(s)}}\right)\mathbb{E}\left(m_{I_1^{(1)}}\dots m_{I_{n_s}^{(s)}}\right)\,, \\
&= \left(\frac{1}{2\alpha\sqrt{n}}\right)^{\sum_{j=1}^n k_j n_j} \sum_{\widehat{G}\in\mathcal{G}_{\{(i_1,k_1)^{n_1},\dots,(i_s,k_s)^{n_s}\}}} n^{v_{\widehat{G}}}(1+o(1))C_{\widehat{G}}\,.
\end{aligned}$$

We first look at the case where $k_j \neq 0$ for all $j \leq s$.

To get a non zero asymptotic contribution, it is necessary to match two per two the chains with the same starting vertex i_j and the same length k_j.
So, n_j must be even and there is $\prod_{j=1}^{s}(n_j-1)!!$ possible pairings. Once the pairing is chosen, it determines the graph class, which contributes for

$$\prod_{j=1}^{s}\left(\frac{\sigma}{2\alpha}\right)^{k_j n_j}\Delta^{\frac{n_j}{2}}.$$

Assume now that some k_j are zeros. Let's take a look to the following example :

$$\begin{aligned}
\mathbb{E}\left(\nu(3,1,n)^2\nu(1,0,n)\right) &= \left(\frac{1}{2\alpha\sqrt{n}}\right)^2\sum_{i=1}^{n}\mathbb{E}\left(A_{3i}^2\right)\mathbb{E}(m_i^2 m_1)\,,\\
&= \left(\frac{1}{2\alpha\sqrt{n}}\right)^2\sum_{i=2}^{n}\mathbb{E}\left(A_{3i}^2\right)\mathbb{E}(m_i^2)\mathbb{E}(m_1)\\
&\qquad + \left(\frac{1}{2\alpha\sqrt{n}}\right)^2\left(A_{31}^2\right)\mathbb{E}(m_1^3)\,,\\
&= \left(\frac{1}{2\alpha\sqrt{n}}\right)^2(n-1)\sigma^2\Delta\mathbb{E}(m_1) + \left(\frac{1}{2\alpha\sqrt{n}}\right)^2\sigma^2\mathbb{E}(m_1^3)\,,\\
&\xrightarrow[n\to+\infty]{} \left(\frac{\sigma}{2\alpha}\right)^2\Delta\mathbb{E}(m_1)\,.
\end{aligned}$$

This suggests there is the following relationship to show :

$$\begin{aligned}
&\mathbb{E}\left(\nu(i_1,k_1,n)^{n_1}\nu(i_2,k_2,n)^{n_2}\dots\nu(i_s,k_s,n)^{n_s}\right)\\
&\qquad -\mathbb{E}\left(\prod_{j \text{ s.t. } k_j\neq 0}\nu(i_j,k_j,n)^{n_j}\right)\prod_{j \text{ s.t. } k_j=0}\mathbb{E}\left(m_1^{n_j}\right)\xrightarrow[n\to\infty]{} 0 \quad (3.7)
\end{aligned}$$

which finally gives the lemma.

Now we justify (3.7). One can assume without loss of generality that $k_1 = k_2 = \ldots = k_r = 0$ and for all $r \leq t \leq s, k_t \neq 0$. Then $i_1,\ldots,i_r$ are distinct.
We denote $J := \{i_1,\ldots,i_r\}$. and $\mathcal{I}_{i,k,J}$ the set of all the chains of length k starting by i and ending by an index belonging to J. Hence, one gets the disjoint union $\mathcal{I}_{i,k} = \mathcal{I}_{i,k,J}\bigcup\mathcal{I}_{i,k,J^c}$.
We can verify that the main contribution comes from

$$\left(\frac{1}{2\alpha\sqrt{n}}\right)^{\sum_{j=1}^{n}k_j n_j}\sum_{\substack{I_1^{(1)},\ldots,I_{n_1}^{(1)}\in\mathcal{I}_{i_1,k_1,J^c}\\ \cdots\\ I_1^{(s)},\ldots,I_{n_s}^{(s)}\in\mathcal{I}_{i_s,k_s,J^c}}}\mathbb{E}\left(A_{I_1^{(1)}}\dots A_{I_{n_1}^{(1)}}A_{I_1^{(2)}}\dots A_{I_{n_s}^{(s)}}\right)\mathbb{E}\left(m_{I_1^{(1)}}\dots m_{I_{n_s}^{(s)}}\right)$$

and in this case

$$\mathbb{E}\left(m_{I_1^{(1)}}\dots m_{I_{n_s}^{(s)}}\right) = \prod_{j\leq r}\mathbb{E}\left(m_1^{n_j}\right)\mathbb{E}\left(m_{I_1^{(r+1)}}\dots m_{I_{n_s}^{(s)}}\right).$$

We therefore get

$$\mathbb{E}\left(\nu(i_1,k_1,n)^{n_1}\dots\nu(i_s,k_s,n)^{n_s}\right)$$
$$\xrightarrow[n\to\infty]{} \prod_{j=1}^{s}\left(\frac{\sigma}{2\alpha}\right)^{k_jn_j}\left[\Delta^{\frac{n_j}{2}}(n_j-1)!!\chi_{k_j\geq 1}\chi_{n_j\equiv 0[2]}+\mathbb{E}\left(m_1^{n_j}\right)\chi_{k_j=0}\right],$$

which concludes the proof of the moments lemma 42. □

4 Proof of the Central Limit Theorem

We can now deal with the proof of the CLT 40.
Let $q\in\mathbb{N}^*$ fixed and $\lambda_1,\lambda_2,\dots,\lambda_q\in\mathbb{R}$.

We want to show that

$$\mathbb{E}\left(\exp\left(i\sum_{j=1}^{q}\lambda_j y_{j,n}\right)\right)\xrightarrow[n\to+\infty]{}\exp\left(-\frac{1}{2}\gamma^2\sum_{j=1}^{q}\lambda_j^2\right).$$

Like in the proof of the moments lemma 42, we simply denote y instead of y_n and y_i the i$^{\text{th}}$ coordinate $y_{i,n}$ of this vector.

Let $p\in\mathbb{N}^*$, denote by $y^{(p)}$ the vector

$$y^{(p)}=\sum_{k=1}^{p-1}\left(\frac{A}{2\alpha\sqrt{n}}\right)^k m\,.$$

Hence,

$$\sum_{j=1}^{q}\lambda_j y_j^{(p)}=\sum_{j=1}^{q}\lambda_j\sum_{k=1}^{p-1}\nu(j,k,n)\,.$$

Moments lemma 42 implies

$$\sum_{j=1}^{q}\lambda_j y_j^{(p)}\xrightarrow[n\to+\infty]{}\mathcal{N}\left(0,\ \Delta\sum_{j=1}^{q}\lambda_j^2\sum_{k=1}^{p-1}\left(\frac{\sigma}{2\alpha}\right)^{2k}\right).$$

By writing

$$\tilde{\sigma}:=\frac{\sigma}{2\alpha}\,,\qquad \gamma^2:=\frac{\sigma^2\Delta}{4\alpha^2-\sigma^2}\qquad\text{and}\qquad \gamma_{(p)}^2:=\Delta\frac{\tilde{\sigma}^2-\tilde{\sigma}^{2p}}{1-\tilde{\sigma}^2}\,, \tag{3.8}$$

one gets

$$\Delta\sum_{j=1}^{q}\lambda_j^2\sum_{k=1}^{p-1}\left(\frac{\sigma}{2\alpha}\right)^{2k}=\gamma_{(p)}^2\sum_{j=1}^{q}\lambda_j^2\,.$$

Let denote by $r^{(p)} := y - y^{(p)}$. Yet,

$$\begin{aligned}
r^{(p)} = y - y^{(p)} &= \frac{A}{2\alpha\sqrt{n}} z - y^{(p)}\,, \\
&= \frac{A}{2\alpha\sqrt{n}} y - y^{(p)} + \frac{A}{2\alpha\sqrt{n}} m\,, \\
&= \frac{A}{2\alpha\sqrt{n}} \left(y - y^{(p)}\right) + \frac{A}{2\alpha\sqrt{n}} y^{(p)} - \sum_{k=1}^{p-1} \left(\frac{A}{2\alpha\sqrt{n}}\right)^k m + \frac{A}{2\alpha\sqrt{n}} m\,, \\
&= \frac{A}{2\alpha\sqrt{n}} r^{(p)} + \sum_{k=2}^{p} \left(\frac{A}{2\alpha\sqrt{n}}\right)^k m - \sum_{k=2}^{p-1} \left(\frac{A}{2\alpha\sqrt{n}}\right)^k m\,.
\end{aligned}$$

So

$$r^{(p)} = \frac{A}{2\alpha\sqrt{n}} r^{(p)} + \left(\frac{A}{2\alpha\sqrt{n}}\right)^p m\,. \tag{3.9}$$

For $\delta \in \left]\frac{2\sigma}{\alpha}; 1\right[$, we denote $G_n := \left\{\left\|\frac{A}{2\alpha\sqrt{n}}\right\| < \delta\right\}$, where $\|.\|$ stands for the spectral norm for a matrix and the Euclidean norm for a vector.
According to [Geman, 1980], $\mathbb{1}_{G_n} \xrightarrow[n\to+\infty]{} 1$ almost surely.

We show that $\forall i \in \{1, \ldots, q\}$, $\forall \epsilon > 0$,

$$\limsup_{p\to+\infty} \limsup_{n\to+\infty} \mathbb{P}\left(\left|\mathbb{1}_{G_n} r_i^{(p)}\right| > \epsilon\right) = 0\,.$$

From equation (3.9), one has :

$$\begin{aligned}
\left\|\mathbb{1}_{G_n} r^{(p)}\right\| &\le \left\|\mathbb{1}_{G_n} \frac{A}{2\alpha\sqrt{n}}\right\| \left\|\mathbb{1}_{G_n} r^{(p)}\right\| + \left\|\mathbb{1}_{G_n} \frac{A}{2\alpha\sqrt{n}}\right\|^p \|m\|\ , \\
&\le \delta \left\|\mathbb{1}_{G_n} r^{(p)}\right\| + \delta^p \|m\|\ .
\end{aligned}$$

That is

$$(1-\delta)^2 \mathbb{1}_{G_n} \sum_{i=1}^{n} \left|r_i^{(p)}\right|^2 \le \delta^{2p} \|m\|^2\ .$$

And by symmetry and using the fact that $\mathbb{E}\left(\|m\|^2\right) = n\mathbb{E}\left(m_1^2\right) = n\Delta$,

$$\mathbb{E}\left(\mathbb{1}_{G_n} \left|r_i^{(p)}\right|^2\right) \le \frac{\delta^{2p}\Delta}{(1-\delta)^2}\,.$$

Hence, using Markov's inequality, one gets :

$$\begin{aligned}
\limsup_{p\to+\infty} \limsup_{n\to+\infty} \mathbb{P}\left(\left|\mathbb{1}_{G_n} r_i^{(p)}\right| > \epsilon\right) &\le \limsup_{p\to+\infty} \limsup_{n\to+\infty} \frac{\mathbb{E}\left(\mathbb{1}_{G_n} \left|r_i^{(p)}\right|^2\right)}{\epsilon^2}\,, \\
&\le \limsup_{p\to+\infty} \frac{\delta^{2p}\Delta}{(1-\delta)^2\epsilon^2}\,, \\
&= 0\,,
\end{aligned}$$

which implies that :

$$\forall \epsilon > 0, \quad \limsup_{p\to+\infty} \limsup_{n\to+\infty} \mathbb{P}\left(\left| \sum_{i=1}^{q} \mathbb{1}_{G_n} r_i^{(p)} \right| > \epsilon \right) = 0 \,.$$

Then,

$$\begin{aligned}
\limsup_{n\to+\infty} &\left| \mathbb{E}\left(\exp\left(i \sum_{j=1}^{q} \lambda_j y_j \right) \right) - \exp\left(-\frac{1}{2}\gamma^2 \sum_{j=1}^{q} \lambda_j^2 \right) \right| \\
&\leq \limsup_{n\to+\infty} \left| \mathbb{E}\left(\exp\left(i \sum_j \lambda_j y_j \right) \right) - \mathbb{E}\left(\exp\left(i \sum_j \lambda_j y_j^{(p)} \right) \right) \right| \\
&\quad + \limsup_{n\to+\infty} \left| \mathbb{E}\left(\exp\left(i \sum_j \lambda_j y_j^{(p)} \right) \right) - \exp\left(-\frac{1}{2}\gamma^2 \sum_j \lambda_j^2 \right) \right| , \\
&= \limsup_{n\to+\infty} \left| \mathbb{E}\left(\exp\left(i \sum_j \lambda_j y_j^{(p)} \right) \left(\exp\left(i \sum_j \lambda_j r_j^{(p)} \right) - 1 \right) \right) \right| \\
&\quad + \left| \exp\left(-\frac{1}{2}\gamma_{(p)}^2 \sum_j \lambda_j^2 \right) - \exp\left(-\frac{1}{2}\gamma^2 \sum_j \lambda_j^2 \right) \right| ,
\end{aligned}$$

where $\gamma_{(p)}^2$, $\tilde{\sigma}$ and γ^2 are defined by (3.8). In this way,

$$\limsup_{p\to+\infty} \gamma_{(p)}^2 = \gamma^2 \,. \tag{3.10}$$

So,

$$\limsup_{p} \left| \exp\left(-\frac{1}{2}\gamma_{(p)}^2 \sum_j \lambda_j^2 \right) - \exp\left(-\frac{1}{2}\gamma^2 \sum_j \lambda_j^2 \right) \right| = 0 \,.$$

We, now, take a look to

$$\limsup_{p\to+\infty} \limsup_{n\to+\infty} \left| \mathbb{E}\left(\exp\left(i \sum_j \lambda_j y_j^{(p)} \right) \left(\exp\left(i \sum_j \lambda_j r_j^{(p)} \right) - 1 \right) \right) \right| . \tag{3.11}$$

The following inequality comes

$$\begin{aligned}
&\left| \mathbb{E}\left(\exp\left(i \sum_j \lambda_j y_j^{(p)} \right) \left(\exp\left(i \sum_j \lambda_j r_j^{(p)} \right) - 1 \right) \right) \right| \\
&\qquad \leq \mathbb{E}\left(\left| \exp\left(i \sum_j \lambda_j y_j^{(p)} \right) \right| \left| \exp\left(i \sum_j \lambda_j r_j^{(p)} \right) - 1 \right| \right) , \\
&\qquad \leq \mathbb{E}\left(\left| \exp\left(i \sum_j \lambda_j r_j^{(p)} \right) - 1 \right| \right) .
\end{aligned}$$

Let $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{E}\left(\left| \exp\left(i \sum_j \lambda_j r_j^{(p)} \right) - 1 \right| \right) \\
&\qquad = \mathbb{E}\left(\left| \exp\left(i \sum_j \lambda_j r_j^{(p)} \right) - 1 \right| \left(\mathbb{1}_{\left|\sum_j \lambda_j r_j^{(p)}\right| \leq \epsilon} + \mathbb{1}_{\left|\sum_j \lambda_j r_j^{(p)}\right| > \epsilon} \right) \right) .
\end{aligned}$$

On one side,

$$\limsup_{p\to+\infty}\limsup_{n\to+\infty}\mathbb{E}\left(\left|\exp\left(i\sum_j \lambda_j r_j^{(p)}\right)-1\right|\mathbb{1}_{\left|\sum_j \lambda_j r_j^{(p)}\right|>\epsilon}\right)$$

$$\leq \limsup_{p\to+\infty}\limsup_{n\to+\infty} 2\mathbb{P}\left(\left|\sum_j \lambda_j r_j^{(p)}\right|>\epsilon\right),$$

$$=0\,.$$

On the other side,

$$\limsup_{p\to+\infty}\limsup_{n\to+\infty}\ \mathbb{E}\left(\left|\exp\left(i\sum_j \lambda_j r_j^{(p)}\right)-1\right|\mathbb{1}_{\left|\sum_j \lambda_j r_j^{(p)}\right|\leq\epsilon}\right)$$

$$\leq \limsup_{p\to+\infty}\limsup_{n\to+\infty}\zeta_\epsilon\mathbb{P}\left(\left|\sum_j \lambda_j r_j^{(p)}\right|\leq\epsilon\right),$$

$$=\zeta_\epsilon\,,$$

where $\zeta_\epsilon = |\exp(i\epsilon)-1| \xrightarrow[\epsilon\to 0]{} 0$.

Hence, for all $\epsilon>0$,

$$\limsup_{p\to+\infty}\limsup_{n\to+\infty}\left|\mathbb{E}\left(\exp\left(i\sum_{j=1}^q \lambda_j y_j\right)\right)-\exp\left(-\frac{1}{2}\gamma^2\sum_{j=1}^q\lambda_j^2\right)\right|\leq\zeta_\epsilon\,.$$

Thus

$$\limsup_{p\to+\infty}\limsup_{n\to+\infty}\left|\mathbb{E}\left(\exp\left(i\sum_{j=1}^q \lambda_j y_j\right)\right)-\exp\left(-\frac{1}{2}\gamma^2\sum_{j=1}^q\lambda_j^2\right)\right|=0\,,$$

which concludes the proof of the first part of the theorem 40, the convergence of a finite number of components of y.

Concerning the convergence in distribution of the vector $(z_{1,n}, z_{2,n}, \ldots, z_{q,n})$, one should remember that, asymptotically, the moment lemma 42 implies

$$y_j^{(p)}+m_j \xrightarrow[n\to+\infty]{d}\mathcal{N}\left(0,\Delta\sum_{k=1}^{p-1}\left(\frac{\sigma}{2\alpha}\right)^{2k}\right)*\pi\,.$$

Hence,

$$(z_1^{(p)}, z_2^{(p)},\ldots,z_q^{(p)})\xrightarrow[n\to+\infty]{d}\mathcal{N}(0,\gamma_{(p)}^2 I_q)*\pi^{\otimes q}\,,$$

where $z_j^{(p)}=y_j^{(p)}+m_j$.

For ease of reading, we write $\tilde{z}=(z_1,\ldots,z_q)$, $\tilde{\lambda}=(\lambda_1,\ldots,\lambda_q)$ and the characteristic function is written :

$$\varphi_{\tilde{z}}(\tilde{\lambda})=\varphi_{z_1,\ldots,z_q}(\lambda_1,\ldots,\lambda_q):=\mathbb{E}\left(\exp\left(i\sum_{j=1}^q\lambda_j z_j\right)\right)\,.$$

The aim is to show that :

$$\varphi_{\tilde{z}}(\tilde{\lambda}) \xrightarrow[n\to+\infty]{} \varphi_{\tilde{Y}}(\tilde{\lambda})\varphi_{\tilde{M}}(\tilde{\lambda})\,,$$

where $\tilde{Y} = (Y_1, \dots, Y_q)$ is a random vector of i.i.d variables of law $\mathcal{N}(0, \gamma^2)$ and $\tilde{M} = (M_1, \dots, M_q)$ is a random vector of i.i.d variables of law π independent from the Y_i.
Henceforth,

$$\begin{aligned}
&\limsup_{n\to+\infty} \left|\varphi_{\tilde{z}}(\tilde{\lambda}) - \varphi_{\tilde{Y}}(\tilde{\lambda})\varphi_{\tilde{M}}(\tilde{\lambda})\right| \\
&\quad\leq \limsup_{n\to+\infty} \left|\varphi_{\tilde{z}}(\tilde{\lambda}) - \varphi_{\tilde{z}^{(p)}}(\tilde{\lambda})\right| + \limsup_{n\to+\infty} \left|\varphi_{\tilde{z}^{(p)}}(\tilde{\lambda}) - \varphi_{\tilde{Y}}(\tilde{\lambda})\varphi_{\tilde{M}}(\tilde{\lambda})\right|\,, \\
&\quad= \limsup_{n\to+\infty} \left|\mathbb{E}\left(\exp\left(i\sum_j \lambda_j z_j\right)\right) - \mathbb{E}\left(\exp\left(i\sum_j \lambda_j z_j^{(p)}\right)\right)\right| \\
&\qquad\qquad + \left|\varphi_{\tilde{Y}^{(p)}}(\tilde{\lambda})\varphi_{\tilde{M}}(\tilde{\lambda}) - \varphi_{\tilde{Y}}(\tilde{\lambda})\varphi_{\tilde{M}}(\tilde{\lambda})\right|\,, \\
&\quad= \limsup_{n\to+\infty} \left|\mathbb{E}\left(\exp\left(i\sum_j \lambda_j z_j^{(p)}\right)\left(\exp\left(i\sum_j \lambda_j r_j^{(p)}\right) - 1\right)\right)\right| \\
&\qquad\qquad + \left|\varphi_{\tilde{M}}(\tilde{\lambda})\left(\varphi_{\tilde{Y}^{(p)}}(\tilde{\lambda}) - \varphi_{\tilde{Y}}(\tilde{\lambda})\right)\right|\,,
\end{aligned} \tag{3.12}$$

where $\tilde{Y} = (Y_1^{(p)}, \dots, Y_q^{(p)})$ is a vector of i.i.d variables of law $\mathcal{N}(0, \gamma_{(p)}^2)$ and $\gamma_{(p)}^2$, $\tilde{\sigma}$ and γ^2 are defined by (3.8).

In this way, we deal with the first member of (3.12) by using the exact same reasoning than for (3.11).
To manage the second member of (3.12), we reuse the convergence (3.10), which means that

$$\begin{aligned}
&\limsup_{p} \left|\varphi_{\tilde{M}}(\tilde{\lambda})\ \left(\varphi_{\tilde{Y}^{(p)}}(\tilde{\lambda}) - \varphi_{\tilde{Y}}(\tilde{\lambda})\right)\right| \\
&\qquad\leq \limsup_{p} \left|\varphi_{\tilde{Y}^{(p)}}(\tilde{\lambda}) - \varphi_{\tilde{Y}}(\tilde{\lambda})\right|\,, \\
&\qquad= \limsup_{p} \left|\exp\left(-\frac{1}{2}\gamma_{(p)}^2 \sum_{j=1}^q \lambda_j^2\right) - \exp\left(-\frac{1}{2}\gamma^2 \sum_{j=1}^q \lambda_j^2\right)\right|\,, \\
&\qquad= 0\,.
\end{aligned}$$

To get a control on the first term of equation (3.12), we reuse the same argumentation than the one used for (3.11) and get

$$\limsup_{p\to+\infty} \limsup_{n\to+\infty} \left|\varphi_{\tilde{z}}(\tilde{\lambda}) - \varphi_{\tilde{Y}}(\tilde{\lambda})\varphi_{\tilde{M}}(\tilde{\lambda})\right| = 0\,.$$

Chapitre 4

Structured models

Ce chapitre est tiré de la revue [Akjouj *et al.*, 2022] écrite en collaboration avec M. Barbier, M. Clenet, W. Hachem, M. Maïda, F. Massol, J. Najim et V.C. Tran dans le cadre du projet KARATÉ (*LotKA-VolterRA models – when random maTrix theory meets theoretical Ecology*).
Il s'agit de ma contribution principale à l'article correspondant à la partie 4. Elle développe les modèles impliquant des structures sur la matrice des interactions M ou la matrice des communautés Γ.

Large Random Matrices and RMT play a prominent role in the theoretical study of systems of particles in interactions such as ecosystems, foodwebs, etc. In this thesis and more particularly in the revue, we have studied at large the LV system of differential equations :

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = x_i(r_i - x_i + (M\boldsymbol{x})_i)$$

where matrix M, supposed random, has either i.i.d. entries or follows the elliptic model. Associated to the generic model :

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = x_i\varphi_i(\boldsymbol{x}),$$

another line of research focused on a modelisation of the Jacobian of the system near equilibrium :

$$J = -I + \Gamma \tag{4.1}$$

where Γ is a random matrix, the question being then to understand the relative localisation of the spectrum of Γ with respect to -1 to conclude on the stability of the system. This second approach is historically the first one with the influential paper by [May, 1972] being one of its landmarks.

In order to progress toward a more realistic description of the reality, one is tempted to consider more involved models of random matrices to take into account more properties of the complex systems such as sparsity, existence of underlying structures, randomness beyond independence, etc.

For instance, the question of the effect of the structure of the ecological network on its feasibility and stability already appeared in [Pimm, 1979], following the work of May, where Pimm argues that connectance is not the only parameter that can influence the feasibility and stability of the networks and starts a theoretical study of structured (both deterministic and random) networks.

In this section, we present a variety of random matrix models beyond the i.i.d. and elliptic ones, emphasizing on their use in theoretical ecology and listing mathematical results and questions of interest. Often, mathematical results of interest are not available on the shelf for a direct use and massive simulations remain the main approach to exploit the potentialities of such models.

1 An introduction to sparsity for ecological models

Empirically, in an ecosystem with n species, even if the maximal number of interactions is n^2, the real number L of nonzero interactions is often much smaller. For a LV system, this means that the interaction matrix M has many entries equal to zero. We define the ***connectance*** as $C = L/n^2$.
Before starting the presentation of the models, let us clarify that the notion of sparsity used here is different from the usual one in mathematics, where matrices or networks are said to be *sparse* when the L/n^2 goes to zero with n. We will consider a wider range for the connectance (in particular when C can be of order $O(1)$).

Although the interpretation is more doubtful in the jacobian modelization (4.1), the connectance already appears in [Gardner et Ashby, 1970]'s simulations and in [May, 1972]'s work . It is an important parameter to capture the sparsity of the models of interest.

In more recent works, [Grilli *et al.*, 2017] work explicitely on the interaction matrix of a LV system and study the stability and feasibility of the equilibrium as a function of various parameters, among which the connectance (see also [Marcus *et al.*, 2022, Dunne *et al.*, 2002]). Based on empirical evidence, [Busiello *et al.*, 2017] suggest that foodwebs can actually be very sparse.

Beyond the connectance, it is possible to take into account the structure of the network by setting some of the entries to zero, thus enforcing an absence of interactions. For this, we may use a matrix $\Delta = (\Delta_{ij})$ where Δ_{ij} equals 1 if species j has an effect on species i and 0 otherwise. If one draws the system interactions as a graph, then Δ can be interpreted as the adjacency matrix of this graph and the interaction matrix M or the community matrix Γ can then be represented as proportionnal to $\Delta \circ A$ where $\circ$ represents the Hadamard product of matrices, that is :

$$(\Delta \circ A)_{ij} = \Delta_{ij} A_{ij}$$

and A is random either i.i.d. or elliptic. In such a model, Δ represents the structure of the system and A the (random) intensity of the interactions.

In the following subsections, we consider a number of sparse models, Δ being deterministic or random. We also refer to [Allesina et Tang, 2015] for a presentation of many models in connection with RMT.

2 The simplest model for sparsity for ecological networks : Erdös-Rényi graphs

When all species play the same role in the foodweb and the only parameter of interest is the average number of interactions for a given species,it is natural to choose Δ as the adjacency matrix of an ER graph of size n : each coefficient of the random matrix Δ has probability p to be nonzero, equal to 1, and probability $1-p$ to be put to zero, independently of the others. The average number of edges in the graph is pn^2, hence the connectance C equal to p.

ER in the mathematical literature

ER graphs are reference models in the field of random graphs and their geometric properties have been extensively studied (see e.g. [Bollobás, 2001, Durrett, 2007, der Hofstad, 2017]). The spectral properties of their adjacency matrices have also been studied. In the regime when $C = O(1)$, which is called *dense* by mathematicians, the ER matrix is a rank one deformation of a matrix with centered i.i.d. entries, so that we observe a circular law and one outlier. In fact,

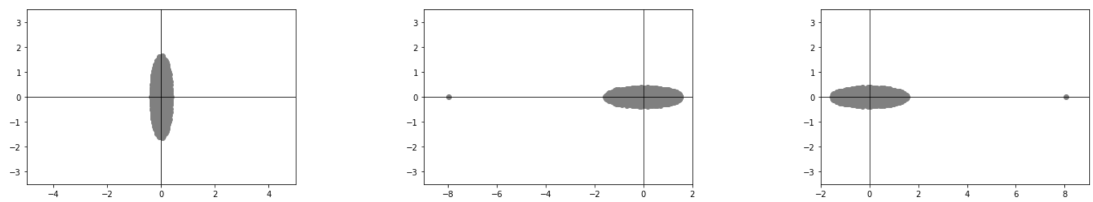
$$\frac{\Delta}{\sqrt{n}} = \frac{1}{\sqrt{n}}(\Delta - \mathbb{E}\Delta) + \frac{1}{\sqrt{n}}\mathbb{E}\Delta \quad \text{with} \quad \left\| \frac{1}{\sqrt{n}}\mathbb{E}\Delta \right\| = \sqrt{n}\, C\,,$$

where $\|.\|$ refers to the spectral norm when applied to a matrix. Notice that the precise understanding of the extreme eigenvalues of $\frac{\Delta}{\sqrt{Cn}}$ in sparse or very sparse regimes is still an active subject in RMT. A concise overview can be found in the introduction of [Alt *et al.*, 2021].

ER in the ecological literature : sparsity increases stability

As developed in the introduction of May's model, the case when $\Gamma_{ij} = \Delta_{ij} A_{ij}$, with Δ the adjacency matrix of a dense ER graph and A has i.i.d. centered entries with variance V has been already considered by May. This sparse model is equivalent to the full model, where the entries have variance CV and in this case sparsity increases stability : in fact, the stability condition $nVC < 1$ is easily satisfied for small C.

The case when Δ is the adjacency matrix of an ER graph but the model for the matrix A is more involved has been studied in particular in [Allesina et Tang, 2012]. They use models for A that are of the same flavour as the elliptic model - for example, $(A_{ij}, A_{ji})_{i<j}$ both positive to model mutualistic systems or with opposite sign to model a prey-predator situation. As illustrated in Figure 4.1, in the mutualistic case, outliers with a large real part may strongly affect the stability. In [Allesina et Tang, 2012], the authors also establish an explicit stability criterion adapted to each case, generalizing May's criterion and emphasize again that sparsity increases stability.



(a) Predator-prey. (b) Competitive. (c) Mutualistic.

FIGURE 4.1 – Spectrum of the interaction matrix for $n = 1000$ species. Δ is ER symmetric with $C = 0.1$. For the competitive (resp. mutualistic) model, $A_{ij} = A_{ji}$ with distribution $-|\mathcal{N}(0,1)|$ (resp. $|\mathcal{N}(0,1)|$) variables. For predator-prey $A_{ij} = -A_{ji}$, distribution $|\mathcal{N}(0,1)|$.

3 Sparsity with a deterministic structure

An alternative to the ER model is the case where matrix Δ is deterministic.

Consider a d-regular oriented graph with n vertices, that is a graph where each vertex i has exactly d oriented edges exiting from i and d edges coming to i. Let Δ be the adjacency matrix of such a graph, then Δ has d non-null entries per row and per column and $L := d \times n$ non-null entries overall. Parameter d which may depend on n accounts for the sparsity of the system and in the framework of a LV system, consider the interaction matrix :

$$M = \frac{1}{\alpha}\frac{\Delta \circ A}{\sqrt{d}} = \frac{1}{\alpha}\left(\frac{\Delta_{ij}A_{ij}}{\sqrt{d}}\right)_{ij} \,,$$

where the A_{ij}'s are i.i.d. and α is an extra normalization which may depend on n. Notice that the normalization is no longer $\sqrt{n}$ but $\sqrt{d}$ accounting for the fact that there are exactly d non-null entries per row. For such a model the connectance C equals :

$$C = \frac{d}{n}$$

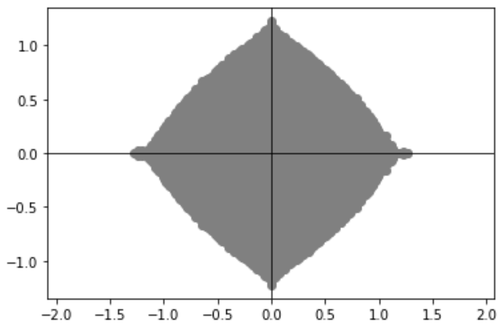
and the interest lies in relatively "small" values of d.

This model has been studied in chapter 1 where specific assumptions on d and Δ are considered, namely either d is proportional to n, or $d \gg \log(n)$ and Δ has a specific block structure. In these models, values of d are slightly large but much smaller than n. In this chapter, it is shown that the same phase transition as in the full model occurs : feasibility and stability hold iff $\alpha = \alpha_n \gg \sqrt{2\log(n)}$.

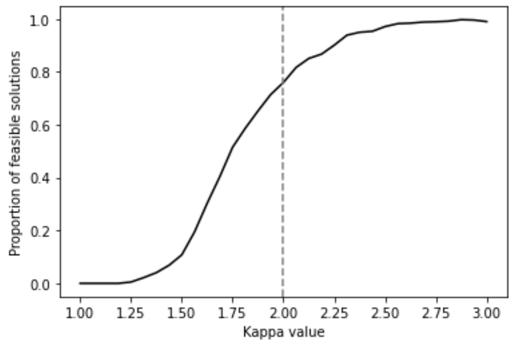
The spectrum of matrix $\frac{\Delta \circ A}{\sqrt{d}}$ together with the proportion of equilibria near the phase transition thershold are plotted in Figure 4.2.

4 Introducing modularity through Stochastic Block Model (SBM)

Beyond the Erdös-Rényi case, when every species equally interacts with any other, it is often more realistic to consider that there exist within the ecosystem *communities* (also called ***modules***), that is groups of species sharing the same connexion patterns. This leads to the celebrated Stochastic Block Model (SBM),



(a) Spectrum of the normalized by $\frac{1}{\sqrt{d}}$ interaction matrix.



(b) Proportion of feasible equilibrium.

FIGURE 4.2 – Deterministic model with $n = 1000$ species and $d = 10$. Each species interacts with $d = 10$ species. The existing interaction are i.i.d. Gaussian $\mathcal{N}(0, 1)$ entries. In figure 4.2b, each point represents the proportion of feasible solutions $\boldsymbol{x}$ over 200 realizations of random matrices M_n for different values of κ, with $\alpha_n = \sqrt{\kappa \log(n)}$.

introduced in [Holland *et al.*, 1983], (see also [Abbe, 2018, Lee et Wilkinson, 2019] for reviews). Let $r \in \mathbb{N}$ be the number of communities. Given

— a vector of positive real numbers $(\pi_1, \cdots, \pi_r)$ such that $\sum_{i=1}^{r} \pi_i = 1$,

— an $r \times r$ matrix P,

the corresponding SBM is a random graph whose vertices are partitioned into r communities $C_1, \cdots, C_r$, where each node belongs to the community C_i, $i \in \{1, \cdots, r\}$, with probability π_i. Then, an edge between a vertex $u \in C_i$ and a vertex $v \in C_j$ exists with probability p_{ij}, independently of all other edges.

SBM in the mathematical literature

There exists a huge literature on the SBM, initially introduced to analyze social networks, and extensively used in machine learning for modelling complex networks and address the community detection problem. The goal there is to design algorithms to cluster the different communities and estimate accurately matrix P, see for example [Matias et Miele, 2017, Baskerville *et al.*, 2011].

Again using the Hadamard product $\Delta \circ A$, the spectrum of the adjacency matrix Δ associated to a SBM can be described, at least in simple cases. Consider for example a SBM with $r = 2$ communities of equal size ($\pi_1 = \pi_2 = 1/2$) and let

$$\begin{pmatrix} p & q \\ q & p \end{pmatrix},$$

with p and q of order $0(1)$ (dense case). Then Δ is a rank-two perturbation of a matrix with centered independent entries. Depending on the values of $\frac{p+q}{2}$ and $\frac{p-q}{2}$, there can be up to two outliers in its spectrum. As in the ER case, sparse cases has also been recently considered, see e.g. [Benaych-Georges *et al.*, 2020].

SBM in the ecological literature : modularity increases stability

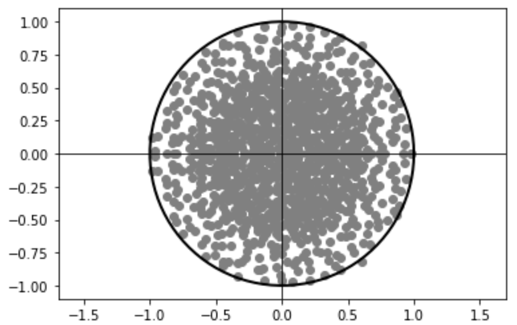
In the seventies, May and Pimm already considered rudimental forms of the SBM into the framework of the Jacobian model (4.1), to take into account some features of ecosystems such as modularity and ***compartmental models***.

[May, 1972] presents a simple occurrence of SBM. He considers a SBM with r modules and a probability vector $(c_1, \dots, c_r)$. This SBM corresponds to modules with no interactions, while within the ith block made of d_i species, the interactions behave like an ER graph with connectance c_i and variance V_i. The stability condition reads :

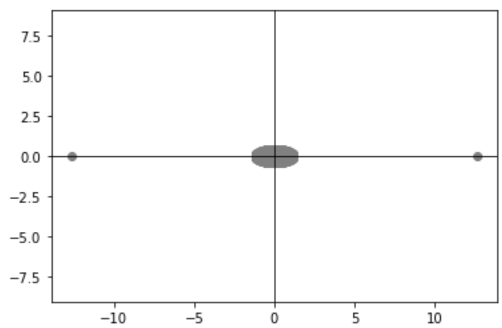
$$\max_{i \in [r]} c_i V_i d_i < 1,$$

hence modularity increases stability. This phenomenon is illustrated in Figure (4.3).

[Pimm, 1979] addresses the following question *"should model systems be organized into compartments of species characterized by strong interactions within compartments, but weak interactions among the compartments ?"* A random version of such a model would correspond to a SBM with a matrix P having large diagonal coefficients and small off-diagonal ones.



(a) Spectrum of the normalized by $\frac{1}{\sqrt{(c_1+q)n}}$ interaction matrix, where $c_1 = 0.5 > c_2 = 0.2 > q = 0.02$. When there exists, the interactions are i.i.d. Gaussian $\mathcal{N}(0,1)$.



(b) Spectrum of the normalized by $\frac{1}{\sqrt{qn}}$ interaction matrix for a bipartite model, where $q = 0.5$ and $c_1 = c_2 = 0$. Δ is symmetric and A has i.i.d $|\mathcal{N}(0,1)|$ entries.

FIGURE 4.3 – SBM with two communities of 500 species, $n = 1000$. The probability of interactions inside the first (resp. second) community is c_1 (resp. c_2) and the probability of interaction with species of the other community is equal to q.

More recently, the effect of modularity on the stability of the networks is extensively explored in [Thébault et Fontaine, 2010] in the framework of a tamed version of LV equations. They evaluate modularity through an index, introduced by [Barber, 2007] and, through simulations, illustrate that persistence, that is the number of surviving species, increases with modularity in trophic networks (see also [Stouffer et Bascompte, 2011]) but decreases with modularity in mutualistic networks. It would be interesting to investigate whether mathematical results on SBM could help to understand their observations.

The question of stability is for example important for plant-pollinator ecosystems. The latter correspond to bipartite mutualistic networks (see e.g. [Thébault et Fontaine, 2010, Billiard *et al.*, 2022]). In [Billiard *et al.*, 2022] in particular, the evolution of abundances can be approximated, when the number of species n tends to infinity, by a kinetic integro-differential equation where the dense graphs are replaced by ***graphons***. The theory of graphons is mathematically well developed but beyond the scope of this review (see for example [Lovász, 2012]).

5 Nested models : a few generalist and many specialist species.

In the Erdös-Rényi model or in SBM, the network is determined by considering the absence or presence of edges for each pair of vertices independently of the others. Other models of random graphs are defined by specifying the degree distributions. For example in the configuration models (also known as Molloy-Reed-Bollobás, see e.g. [Bollobás, 2001, Molloy et Reed, 1995, der Hofstad, 2017]), independent random variables distributed with the target degree distribution are associated to each vertex and edges are formed by pairing at random the half-edges.

By choosing heavy-tailed degree distributions, one can thus create a few vertices with very high degrees (corresponding to generalist species) and a majority of vertices with low degree (corresponding to specialists). Such ecosystems are called ***nested***. They have also been modelled and studied, at least through simulations. This idea has been implemented in [Thébault et Fontaine, 2010] following [Okuyama et Holland, 2008].

Nested ecosystems can also be described through random graphs with given *expected degrees*. This model is known as the *Chung-Lu model* : take a deterministic sequence $w = (w_1, \ldots, w_n)$, that will correspond to the expected degrees and draw an edge between vertex i and vertex j with probability $w_i w_j / \sum_{i=1}^{n} w_i$ independently of all other edges. If we choose all the weights to be equal to pn, we are obviously back to the ER model with connectance p but nested ecosystems can be modeled by choosing a power-law distribution for the weights, that is $w_i = ci^{-\frac{1}{\beta-1}}$, for i greater or equal to some i_0. In this case, the number of species interacting with k others is proportional to $k^{-\beta}$. The spectrum of the adjacency matrix of such a graph has been studied in [Chung *et al.*, 2003] where they point out that a phase transition occurs at $\beta = 2.5$: for $\beta > 2.5$, the largest eigenvalue behaves like $\sqrt{n}$, which is the maximal degree in the graph, whereas for $\beta < 2.5$, the largest eigenvalue behaves like $\bar{d}$, which is the weighted average of the square of the expected degrees. It would be interesting to investigate whether these mathematical results could be effectively used in the study of nested ecosystems.

6 Kernel matrices

Definition of the model.

Part of the literature on ecological networks considers that the interactions between two species depends on a distance between their respective values of some functional traits. The examples that we will present below, fit into the mathematical framework of **kernel matrices**. We have

$$M_{ij} \text{ or } \Gamma_{ij} = f(g(X_i, X_j)),$$

where X_i is a vector modelling the traits of species i, $g : \mathbb{R}^p \times \mathbb{R}^p \longrightarrow \mathbb{R}$ is a symmetric matrix, denoted as the kernel (corresponding to the measure of the distance), and $f : \mathbb{R} \longrightarrow \mathbb{R}$ a function called the envelope. Typical examples are $g(x, y) = x^T y$, or $\|x - y\|^2$ and $f(x) = \exp(ex)$ or $(1 + x)^a$ etc.

Kernel matrices in the mathematical literature.

Among these models, the first interesting and well studied case is the so-called Wishart case[1], when $g(x, y) = x^T y$ and $f(x) = x$. If the entries of the vectors are i.i.d. centered and normalized, then it is well known (see [Marčenko et Pastur, 1967]) that, if the ratio p/n of the number of traits over the number of species converges to τ, the spectrum converges almost surely to the Marcenko-Pastur distribution. The mathematical theory of kernel matrices is developped in the revue [Akjouj *et al.*, 2022, App. E. 2]. The main message is that in the RMT regime, that is when the number of traits is large and of the same order as the number of species, and if g and f are reasonable, they have almost no influence on the spectrum, meaning that, in the models, "any" kernel matrix could be without harm replaced by a Wishart matrix or a Gaussian kernel matrix.

Kernel matrices for ecological networks.

Dieckmann and Doebeli consider a simple co-evolutionary model [Roughgarden, 1979, Dieckmann et Doebeli, 1999, Champagnat *et al.*, 2006]. The birth rate of an individual with trait x is $b(x) = \exp(-x^2/2\sigma_b^2)$, the individual natural death rate is constant $d(x) = d_C$, and the competition between two individuals with traits x and y is $C(x, y) = \eta_c \exp(-(x - y)^2/2\sigma_c^2)$, $\sigma_c > 0$. This would correspond to a Gaussian kernel (also used in [Meszéna *et al.*, 2006]). In [Nuismer *et al.*, 2013], they develop *phenotype matching model*, where the interaction is stronger when the traits of two species are close but also *phenotype difference (or threshold) model*, in which successful interaction depends on the degree to which the trait of the second species surpasses the trait of the first one[2] (see also [Kisdi, 1999, de Andreazzi *et al.*, 2020]). Other models involving thresholds can be found in [Santamaria et Rodriguez-Girones, 2007] and in [Rohr *et al.*, 2016], the models involving a combination of characteristics of the species taken separately and a measure of the similarity between the traits.

We end this paragraph by mentioning the work of [Serván *et al.*, 2020], which lies in the LV framework, with a kernel matrix M for which the distance between the traits, is determined through a distance between species in their phylogenetic tree. He addresses the questions of feasability and stability has been detailed in the revue for the elliptic case and explicitely uses the link with Wishart matrices mentioned above. It would be interesting to investigate whether this point of view can be fruitful in other contexts.

1. That is the empirical covariance matrix of the vectors $X_1, \ldots, X_n$.
2. One can think of fruit and beak sizes in a plants-birds interaction network.

Bibliographie

[Abbe, 2018] Abbe, E. (2018). Community detection and stochastic block models : recent development. *Journal of Machine Learning Research*, 18(177):1–86.

[Akjouj *et al.*, 2022] Akjouj, I., Barbier, M., Clenet, M., Hachem, W., Maïda, M., Massol, F., Najim, J. et Tran, V. C. (2022). Complex systems in ecology : a guided tour with large lotka-volterra models and random matrices.

[Akjouj *et al.*, 2023] Akjouj, I., Hachem, W., Maïda, M. et Najim, J. (2023). Equilibria of large random lotka-volterra systems with vanishing species : a mathematical approach.

[Akjouj et Najim, 2022] Akjouj, I. et Najim, J. (2022). Feasibility of sparse large lotka-volterra ecosystems. *Journal of Mathematical Biology*, 85(66).

[Allesina *et al.*, 2015] Allesina, S., Grilli, J., Barabás, G., Tang, S., Aljadeff, J. et Maritan, A. (2015). Predicting the stability of large structured food webs. *Nature communications*, 6(1):7842.

[Allesina et Tang, 2012] Allesina, S. et Tang, S. (2012). Stability criteria for complex ecosystems. *Nature (London)*, 483(7388):205–208.

[Allesina et Tang, 2015] Allesina, S. et Tang, S. (2015). The stability–complexity relationship at age 40 : a random matrix perspective. *Population Ecology*, 57:63–75.

[Alt *et al.*, 2021] Alt, J., Ducatez, R. et Knowles, A. (2021). Poisson statistics and localization at the spectral edge of sparse erdös–rényi graphs. arxiv :2106.12519.

[Bai et Silverstein, 2010] Bai, Z. et Silverstein, J. (2010). *Spectral Analysis of Large Dimensional Random Matrices*.

[Bai, 1997] Bai, Z. D. (1997). Circular law. *The Annals of Probability*, 25(1):494–529.

[Bai et Yin, 1988] Bai, Z.-D. et Yin, Y.-Q. (1988). Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a wigner matrix. *The Annals of Probability*, pages 1729–1741.

[Bandeira et van Handel, 2016] Bandeira, A. S. et van Handel, R. (2016). Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *Ann. Probab.*, 44(4):2479–2506.

[Barabás *et al.*, 2017] Barabás, G., Michalska-Smith, M. J. et Allesina, S. (2017). Self-regulation and the stability of large ecological networks. *Nature ecology & evolution*, 1(12):1870–1875.

[Barber, 2007] Barber, M. J. (2007). Modularity and community detection in bipartite networks. *Phys. Rev. E*, 76:066102.

[Barker *et al.*, 1978] Barker, G., Berman, A. et Plemmons, R. J. (1978). Positive diagonal solutions to the lyapunov equations. *Linear and Multilinear Algebra*, 5(4):249–256.

[Baskerville *et al.*, 2011] Baskerville, E. B., Dobson, A. P., Bedford, T., Allesina, S., Anderson, T. M. et Pascual, M. (2011). Spatial guilds in the serengeti food web revealed by a bayesian group model. *PLOS Computational Biology*, 7(12):1–11.

[Bayati *et al.*, 2015] Bayati, M., Lelarge, M. et Montanari, A. (2015). Universality in polytope phase transitions and message passing algorithms. *Ann. Appl. Probab.*, 25(2):753–822.

[Bayati et Montanari, 2011] Bayati, M. et Montanari, A. (2011). The Dynamics of Message Passing on Dense Graphs, with Applications to Compressed Sensing. *IEEE Transactions on Information Theory*, pages 764–785.

[Begon et Townsend, 2021] Begon, M. et Townsend, C. R. (2021). *Ecology : from individuals to ecosystems.* John Wiley & Sons.

[Ben Arous *et al.*, 2021] Ben Arous, G., Fyodorov, Y. et Khoruzhenko, B. (2021). Counting equilibria of large complex systems by instability index. *Proceedings of the National Academy of Sciences*, 118(34):e2023719118.

[Benaych-Georges *et al.*, 2020] Benaych-Georges, F., Bordenave, C. et Knowles, A. (2020). Spectral radii of sparse random matrices. *Ann. Inst. H. Poincaré Probab. Statist.*, 56(3):2141–2161.

[Billiard *et al.*, 2022] Billiard, S., Leman, H., Rey, T. et Tran, V. (2022). Continuous limits of large plant-pollinator random networks and some applications. *MathematicS In Action.*

[Bizeul et Najim, 2021] Bizeul, P. et Najim, J. (2021). Positive solutions for large random linear systems. *Proceedings of the American Mathematical Society*, page 1.

[Bollobás, 2001] Bollobás, B. (2001). *Random graphs.* Cambridge University Press, 2 édition.

[Bolthausen, 2014] Bolthausen, E. (2014). An iterative construction of solutions of the TAP equations for the Sherrington–Kirkpatrick model. *Communications in Mathematical Physics*, 325(1):333–366.

[Bordenave *et al.*, 2018] Bordenave, C., Caputo, P., Chafaï, D. et Tikhomirov, K. (2018). On the spectral radius of a random matrix : An upper bound without fourth moment. *The Annals of Probability*, 46(4).

[Bordenave *et al.*, 2021] Bordenave, C., Chafai, D. et García-Zelada, D. (2021). Convergence of the spectral radius of a random matrix through its characteristic polynomial. *Probability Theory and Related Fields.*

[Boucheron *et al.*, 2013] Boucheron, S., Lugosi, G. et Massart, P. (2013). *Concentration inequalities.* Oxford University Press, Oxford. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[Bunin, 2017] Bunin, G. (2017). Ecological communities with Lotka-Volterra dynamics. *Phys. Rev. E*, 95:042414.

[Busiello *et al.*, 2017] Busiello, D. M., Suweis, S., Hidalgo, J. et Maritan, A. (2017). Explorability and the origin of network sparsity in living systems. *Scientific Reports*, 7(1).

[Champagnat *et al.*, 2006] Champagnat, N., Ferrière, R. et Méléard, S. (2006). Unifying evolutionary dynamics : from individual stochastic processes to macroscopic models via timescale separation. *Theoretical Population Biology*, 69:297–321.

[Chen et Xiang, 2007] Chen, X. et Xiang, S. (2007). Perturbation bounds of P-matrix linear complementarity problems. *SIAM J. Optim.*, 18(4):1250–1265.

[Chung *et al.*, 2003] Chung, F., Lu, L. et Vu, V. (2003). Spectra of random graphs with given expected degrees. 100(11):6313–6318.

[Clenet *et al.*, 2022a] Clenet, M., El Ferchichi, H. et Najim, J. (2022a). Equilibrium in a large Lotka–Volterra system with pairwise correlated interactions. *Stochastic Processes and their Applications*, 153:423–444.

[Clenet *et al.*, 2022b] Clenet, M., Najim, J. et Massol, F. (2022b). Equilibrium and surviving species in a large Lotka-Volterra system of differential equations. *ArXiv preprint*, arXiv :2205.00735.

[Cohen, 1993] Cohen, H. (1993). *A course in computational algebraic number theory*, volume 8. Springer-Verlag Berlin.

[Cohen et Newman, 1988] Cohen, J. E. et Newman, C. M. (1988). Dynamic basis of food web organization. *Ecology*, 69(6):1655–1664.

[Cook, 2021] Cook, N. (2021). Multiplicity of singular values for patterned gaussian matrices. *personnal communication.*

[Cottle *et al.*, 2009] Cottle, R. W., Pang, J.-S. et Stone, R. E. (2009). *The linear complementarity problem*, volume 60 de *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. Corrected reprint of the 1992 original [MR1150683].

[Coyte *et al.*, 2015] Coyte, K. Z., Schluter, J. et Foster, K. R. (2015). The ecology of the microbiome : Networks, competition, and stability. *Science*, 350(6261):663–666.

[Dantzig et Cottle, 1968] Dantzig, G. et Cottle, R. (1968). Complementary pivot theory of. mathematical programming. *Mathematics of the decision sciences, part*, 1:115–136.

[de Andreazzi *et al.*, 2020] de Andreazzi, C. S., Astegiano, J. et Guimarães, P. R. (2020). Coevolution by different functional mechanisms modulates the structure and dynamics of antagonistic and mutualistic networks. *Oikos*, 129:224–237.

[Deo, 2017] Deo, N. (2017). *Graph theory with applications to engineering and computer science.* Courier Dover Publications.

[der Hofstad, 2017] der Hofstad, R. V. (2017). *Random Graphs and Complex Networks*, volume 1 de *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, Cambridge.

[Dieckmann et Doebeli, 1999] Dieckmann, U. et Doebeli, M. (1999). On the origin of species by sympatric speciation. *Nature*, 400:354–357.

[Donoho et Montanari, 2016] Donoho, D. et Montanari, A. (2016). High dimensional robust M-estimation : asymptotic variance via approximate message passing. *Probab. Theory Related Fields*, 166(3-4):935–969.

[Dougoud *et al.*, 2018] Dougoud, M., Vinckenbosch, L., Rohr, R. P., Bersier, L.-F. et Mazza, C. (2018). The feasibility of equilibria in large ecosystems : A primary but neglected concept in the complexity-stability debate. *PLOS Computational Biology*, 14(2):1–18.

[Dunne *et al.*, 2002] Dunne, J. A., Williams, R. J. et Martinez, N. D. (2002). Food-web structure and network theory : The role of connectance and size. *Proceedings of the National Academy of Sciences - PNAS*, 99(20):12917–12922.

[Durrett, 2007] Durrett, R. (2007). *Random graph dynamics.* Cambridge University Press, New York.

[Dyson, 1962] Dyson, F. J. (1962). Statistical theory of the energy levels of complex systems. i. *Journal of Mathematical Physics*, 3(1):140–156.

[Edelman, 1997] Edelman, A. (1997). The probability that a random real gaussian matrix haskreal eigenvalues, related distributions, and the circular law. *journal of multivariate analysis*, 60(2):203–232.

[Fan, 2022] Fan, Z. (2022). Approximate message passing algorithms for rotationally invariant matrices. *Ann. Statist.*, 50(1):197–224.

[Feng *et al.*, 2022] Feng, O. Y., Venkataramanan, R., Rush, C. et Samworth, R. J. (2022). A unifying tutorial on approximate message passing. *Foundations and Trends® in Machine Learning*, 15(4):335–536.

[Fiedler et Pták, 1966] Fiedler, M. et Pták, V. (1966). Some generalizations of positive definiteness and monotonicity. *Numerische Mathematik*, 9:163–172.

[Füredi et Komlós, 1981] Füredi, Z. et Komlós, J. (1981). The eigenvalues of random symmetric matrices. *Combinatorica*, 1:233–241.

[Fyodorov et Khoruzhenko, 2016] Fyodorov, Y. et Khoruzhenko, B. (2016). Nonlinear analogue of the May-Wigner instability transition. *Proceedings of the National Academy of Sciences*, 113(25):6827–6832.

[Galla, 2018] Galla, T. (2018). Dynamically evolved community size and stability of random Lotka-Volterra ecosystems. *Europhysics Letters*, 123(4):48004.

[Gardner et Ashby, 1970] Gardner, M. et Ashby, W. (1970). Connectance of large dynamic (cybernetic) systems : Critical values for stability. *Nature*, 228:784.

[Geman, 1980] Geman, S. (1980). A limit theorem for the norm of random matrices. *The Annals of Probability*, 8(2):252 – 261.

[Geman, 1986] Geman, S. (1986). The spectral radius of large random matrices. *Ann. Probab.*, 14(4):1318–1328.

[Geman et Hwang, 1982] Geman, S. et Hwang, C.-R. (1982). A chaos hypothesis for some large systems of random equations. *Z. Wahrsch. Verw. Gebiete*, 60(3): 291–314.

[Gibbs *et al.*, 2018] Gibbs, T., Grilli, J., Rogers, T. et Allesina, S. (2018). Effect of population abundances on the stability of large random ecosystems. *Physical Review E*, 98(2).

[Ginibre, 1965] Ginibre, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *Journal of Mathematical Physics*, 6(3):440–449.

[Girko, 1986] Girko, V. (1986). Elliptic law. *Theory of Probability & Its Applications*, 30(4):677–690.

[Girko, 1995] GIRKO, V. (1995). The elliptic law : ten years later ii.

[Girko, 1985] GIRKO, V. L. (1985). Circular law. *Theory of Probability & Its Applications*, 29(4):694–706.

[Goh et Jennings, 1977] GOH, B. et JENNINGS, L. (1977). Feasibility and stability in randomly assembled lotka-volterra models. *Ecological Modelling*, 3(1):63–71.

[Goh, 1977] GOH, B. S. (1977). Global stability in many-species systems. *The American Naturalist*, 111(977):135–143.

[Gopalsamy, 1984] GOPALSAMY, K. (1984). Global asymptotic stability in volterra's population systems. *Journal of Mathematical Biology*, 19(2):157–168.

[Grilli *et al.*, 2017] GRILLI, J., ADORISIO, M., SUWEIS, S., BARABÁS, G., BANAVAR, J., ALLESINA, S. et MARITAN, A. (2017). Feasibility and coexistence of large ecological communities. *Nature communications*, 8(1).

[Haeckel, 1966] HAECKEL, E. (1966). *Generelle Morphologie der Organismen.*

[Hartman et Mikusiński, 1961] HARTMAN, S. et MIKUSIŃSKI, J. (1961). *The theory of Lebesgue measure and integration.* International Series of Monographs on Pure and Applied Mathematics, Vol. 15. Pergamon Press, New York-Oxford-London-Paris ; Pań-stwowe Wydawnictwo Naukowe, Warsaw. Enlarged ed., translated from Polish by Leo F. Boron.

[Hofbauer et Sigmund, 1998] HOFBAUER, J. et SIGMUND, K. (1998). *Evolutionary games and population dynamics.* Cambridge university press.

[Holland *et al.*, 1983] HOLLAND, P., LASKEY, K. et LEINHARDT, S. (1983). Stochastic blockmodels : some first steps. *Social Networks*, 5:109–137.

[Horn et Johnson, 2013] HORN, R. A. et JOHNSON, C. R. (2013). *Matrix analysis.* Cambridge University Press, Cambridge, second édition.

[Hu *et al.*, 2021] HU, J., AMOR, D. R., BARBIER, M., BUNIN, G. et GORE, J. (2021). Emergent phases of ecological diversity and dynamics mapped in microcosms. *bioRxiv.*

[Huffaker *et al.*, 1958] HUFFAKER, C. *et al.* (1958). Experimental studies on predation : dispersion factors and predator-prey oscillations. *Hilgardia*, 27(14):343–383.

[Ives et Carpenter, 2007] IVES, A. R. et CARPENTER, S. R. (2007). Stability and diversity of ecosystems. *Science*, 317(5834):58–62.

[Jacquet *et al.*, 2016] JACQUET, C., MORITZ, C., MORISSETTE, L., LEGAGNEUX, P., MASSOL, F., ARCHAMBAULT, P. et GRAVEL, D. (2016). No complexity–stability relationship in empirical ecosystems. *Nature Communications*, 7(7): 12573.

[Jansen et Sigmund, 1998] JANSEN, V. A. et SIGMUND, K. (1998). Shaken, not stirred : On permanence in ecological communities.

[Jansen, 1987] JANSEN, W. (1987). A permanence theorem for replicator and lotka-volterra systems. *Journal of Mathematical Biology*, 25(4):411–422.

[Kato, 1995] KATO, T. (1995). *Perturbation theory for linear operators.* Classics in Mathematics. Springer-Verlag, Berlin. Reprint of the 1980 edition.

[Kisdi, 1999] KISDI, E. (1999). Evolutionary branching under asymmetric competition. *J. Theor. Biol.*, 197(2):149–162.

[Kiss et Kovács, 2008] Kiss, K. et Kovács, S. (2008). Qualitative behavior of n-dimensional ratio-dependent predator–prey systems. *Applied Mathematics and Computation*, 199(2):535–546.

[Landi *et al.*, 2018] Landi, P., Minoarivelo, H. O., Brannstrom, A., Hui, C. et Dieckmann, U. (2018). Complexity and stability of ecological networks : a review of the theory. *Population Ecology*, 60(4):319–345.

[Law et Blackford, 1992] Law, R. et Blackford, J. C. (1992). Self-assembling food webs : a global viewpoint of coexistence of species in lotka-volterra communities. *Ecology*, 73(2):567–578.

[Law et Morton, 1996] Law, R. et Morton, R. D. (1996). Permanence and the assembly of ecological communities. *Ecology*, 77(3):762–775.

[Leadbetter *et al.*, 1983] Leadbetter, M. R., Lindgren, G. et Rootzen, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. Springer Verlag.

[Lee et Wilkinson, 2019] Lee, C. et Wilkinson, D. J. (2019). A review of stochastic block models and extensions for graph clustering. *Applied Network Science*, 4.

[Lemke, 1965] Lemke, C. E. (1965). Bimatrix equilibrium points and mathematical programming. *Management science*, 11(7):681–689.

[Lemke et Howson, 1964] Lemke, C. E. et Howson, Jr, J. T. (1964). Equilibrium points of bimatrix games. *Journal of the Society for industrial and Applied Mathematics*, 12(2):413–423.

[Lewinsohn *et al.*, 2006] Lewinsohn, T. M., Inácio Prado, P., Jordano, P., Bascompte, J. et M. Olesen, J. (2006). Structure in plant–animal interaction assemblages. *Oikos*, 113(1):174–184.

[Li *et al.*, 2009] Li, X., Jiang, D. et Mao, X. (2009). Population dynamical behavior of lotka-volterra system under regime switching. *J. Comput. Appl. Math.*, 232(2):427–448.

[Liapounoff, 1907] Liapounoff, A. (1907). Problème général de la stabilité du mouvement. *In Annales de la Faculté des sciences de Toulouse : Mathématiques*, volume 9, pages 203–474.

[Logofet, 2005] Logofet, D. O. (2005). Stronger-than-lyapunov notions of matrix stability, or how “flowers” help solve problems in mathematical ecology. *Linear algebra and its applications*, 398:75–100.

[Logofet, 2018] Logofet, D. O. (2018). *Matrices and graphs : stability problems in mathematical ecology*. CRC press.

[Lotka, 1925] Lotka, A. J. (1925). *Elements of physical biology*. Williams & Wilkins.

[Lovász, 2012] Lovász, L. (2012). *Large networks and graph limits*, volume 60 de *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI.

[MacArthur, 1955] MacArthur, R. (1955). Fluctuations of animal populations and a measure of community stability. *Ecology*, 36(3):533–536.

[MacArthur et Wilson, 1967] MacArthur, R. H. et Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton University Press, rev - revised édition.

[Marcus *et al.*, 2022] Marcus, S., Turner, A. M. et Bunin, G. (2022). Local and collective transitions in sparsely-interacting ecological communities.

[Marčenko et Pastur, 1967] Marčenko, V. A. et Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457.

[Matias et Miele, 2017] Matias, C. et Miele, V. (2017). Statistical clustering of temporal networks through a dynamic stochastic block model. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 79(4):1119–1141.

[May, 1972] May, R. M. (1972). Will a large complex system be stable ? *Nature*, 238:413–414.

[May, 1973] May, R. M. (1973). *Stability and complexity in model ecosystems*. Monographs in population biology,.

[McCann, 2000] McCann, K. S. (2000). The diversity–stability debate. *Nature*, 405(6783):228–233.

[Mehta, 1967] Mehta, M. L. (1967). *Random Matrices and the Statistical Theory of Energy Levels*. Academic Press.

[Meszéna *et al.*, 2006] Meszéna, G., Gyllenberg, M., Pásztor, L. et Metz, J. A. (2006). Competitive exclusion and limiting similarity : A unified theory. *Theoretical population biology*, 69:68–87.

[Molloy et Reed, 1995] Molloy, M. et Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random structures and algorithms*, 6:161–180.

[Montanari et Richard, 2016] Montanari, A. et Richard, E. (2016). Non-negative principal component analysis : message passing algorithms and sharp asymptotics. *IEEE Trans. Inform. Theory*, 62(3):1458–1484.

[Murty, 1972] Murty, K. G. (1972). On the number of solutions to the complementarity problem and spanning properties of complementary cones. *Linear Algebra and its Applications*, 5(1):65–108.

[Murty et Yu, 1988] Murty, K. G. et Yu, F.-T. (1988). *Linear complementarity, linear and nonlinear programming*, volume 3. Citeseer.

[Naumov, 2012] Naumov, A. (2012). Elliptic law for real random matrices. *arXiv preprint arXiv :1201.1639*.

[Nguyen et O'Rourke, 2015] Nguyen, H. H. et O'Rourke, S. (2015). The elliptic law. *International Mathematics Research Notices*, 2015(17):7620–7689.

[Nuismer *et al.*, 2013] Nuismer, S. L., Jordano, P. et Bascompte, J. (2013). Coevolution and the architecture of mutualistic networks. *Evolution*, 67:338–354.

[Okuyama et Holland, 2008] Okuyama, T. et Holland, J. N. (2008). Network structural properties mediate the stability of mutualistic communities. *Ecology letters*, 11(3):208–216.

[Opper et Diederich, 1992] Opper, M. et Diederich, S. (1992). Phase transition and $1/f$ noise in a game dynamical model. *Phys. Rev. Lett.*, 69:1616–1619.

[O'Rourke et Renfrew, 2014] O'Rourke, S. et Renfrew, D. (2014). Low rank perturbations of large elliptic random matrices.

[Pastur et Shcherbina, 2011] Pastur, L. A. et Shcherbina, M. (2011). *Eigenvalue Distribution of Large Random Matrices*, volume 171 de *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI.

[Pimm, 1979] Pimm, S. L. (1979). The structure of food webs. *Theoretical Population Biology*, 16(2):144–158.

[Pimm, 1984] Pimm, S. L. (1984). The complexity and stability of ecosystems. *Nature (London)*, 307(5949):321–326.

[Pimm *et al.*, 1991] Pimm, S. L., Lawton, J. H. et Cohen, J. E. (1991). Food web patterns and their consequences. *Nature*, 350(6320):669–674.

[Pocock *et al.*, 2012] Pocock, M. J., Evans, D. M. et Memmott, J. (2012). The robustness and restoration of a network of ecological networks. *Science*, 335(6071): 973–977.

[Rohn, 2012] Rohn, J. (2012). On rump's characterization of p-matrices. *Optimization Letters*, 6:1017–1020.

[Rohr *et al.*, 2016] Rohr, R. P., Naisbit, R. E., Mazza, C. et Bersier, L.-F. (2016). Matching-centrality decomposition and the forecasting of new links in networks. *Proceedings of the Royal Society. B, Biological sciences*, 283:20152702–.

[Rohr *et al.*, 2014] Rohr, R. P., Saavedra, S. et Bascompte, J. (2014). On the structural stability of mutualistic systems. *Science*, 345(6195):1253497.

[Rossberg, 2013] Rossberg, A. G. (2013). *Food webs and biodiversity : foundations, models, data*. John Wiley & Sons.

[Roughgarden, 1979] Roughgarden, J. (1979). *Theory of population genetics and evolutionary ecology : an introduction*. Macmillan.

[Rozas *et al.*, 2023] Rozas, E., Crumpton, M. J. et Galla, T. (2023). Competitive exclusion and hebbian couplings in random generalised lotka-volterra systems.

[Rump, 2003] Rump, S. M. (2003). On p-matrices. *Linear algebra and its applications*, 363:237–250.

[Santamaria et Rodriguez-Girones, 2007] Santamaria, L. et Rodriguez-Girones, M. A. (2007). Linkage rules for plant-pollinator networks : Trait complementarity or exploitation barriers ? *PLoS biology*, 5:354–362.

[Serván *et al.*, 2020] Serván, C. A., Capitán, J. A., Miller, Z. R. et Allesina, S. (2020). Effects of phylogeny on coexistence in model communities. *bioRxiv*.

[Silverstein et Choi, 1995] Silverstein, J. W. et Choi, S.-I. (1995). Analysis of the limiting spectral distribution of large dimensional random matrices. *Journal of Multivariate Analysis*, 54(2):295–309.

[Stone, 2016] Stone, L. (2016). The google matrix controls the stability of structured ecological and biological networks. *Nature communications*, 7(1):12857.

[Stone, 2018] Stone, L. (2018). The feasibility and stability of large complex biological networks : a random matrix approach. *Scientific reports*, 8(1):1–12.

[Stouffer et Bascompte, 2011] Stouffer, D. B. et Bascompte, J. (2011). Compartmentalization increases food-web persistence. *Proceedings of the National Academy of Sciences - PNAS*, 108(9):3648–3652.

[Stouffer *et al.*, 2005] Stouffer, D. B., Camacho, J., Guimera, R., Ng, C. et Nunes Amaral, L. (2005). Quantitative patterns in the structure of model and empirical food webs. *Ecology*, 86(5):1301–1311.

[Takeuchi, 1996] Takeuchi, Y. (1996). *Global dynamical properties of Lotka-Volterra systems*. World Scientific Publishing Co., Inc., River Edge, NJ.

[Takeuchi et Adachi, 1980] Takeuchi, Y. et Adachi, N. (1980). The existence of globally stable equilibria of ecosystems of the generalized volterra type. *Journal of Mathematical Biology*, 10(4):401–415.

[Takeuchi *et al.*, 1978] Takeuchi, Y., Adachi, N. et Tokumaru, H. (1978). Global stability of ecosystems of the generalized volterra type. *Mathematical Biosciences*, 42(1-2):119–136.

[Tang *et al.*, 2014] Tang, S., Pawar, S. et Allesina, S. (2014). Correlation between interaction strengths drives stability in large ecological networks. *Ecology letters*, 17(9):1094–1100.

[Tansley, 1939] Tansley, A. G. (1939). British ecology during the past quarter-century : The plant community and the ecosystem. *Journal of Ecology*, 27(2):513–530.

[Tao *et al.*, 2010] Tao, T., Vu, V. et Krishnapur, M. (2010). Random matrices : Universality of esds and the circular law.

[Taylor, 1988] Taylor, P. J. (1988). Consistent scaling and parameter choice for linear and generalized lotka-volterra models used in community ecology. *Journal of theoretical biology*, 135(4):543–568.

[Thébault et Fontaine, 2010] Thébault, E. et Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, 329(5993):853–856.

[Tokita, 2004] Tokita, K. (2004). Species abundance patterns in complex evolutionary dynamics. *Phys. Rev. Lett.*, 93:178102.

[Villani, 2009] Villani, C. (2009). *Optimal transport*, volume 338 de *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin. Old and new.

[Volterra, 1926] Volterra, V. (1926). Fluctuations in the abundance of a species considered mathematically. *Nature*, 118(2972):558–560.

[Wang *et al.*, 2022] Wang, T., Zhong, X. et Fan, Z. (2022). Universality of approximate message passing algorithms and tensor networks. *arXiv preprint arXiv :2206.13037*.

[Wangersky, 1978] Wangersky, P. J. (1978). Lotka-volterra population models. *Annual Review of Ecology and Systematics*, 9(1):189–218.

[Wigner, 1955] Wigner, E. P. (1955). Characteristic vectors of bordered matrices with infinite dimensions i. *Annals of Mathematics*, 62:541–545.

[Wigner, 1967] Wigner, E. P. (1967). Random matrices in physics. *SIAM review*, 9(1):1–23.

[Wishart, 1928] WISHART, J. (1928). The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52.

[Yodzis, 1981] YODZIS, P. (1981). The stability of real ecosystems. *Nature*, 289(5799):674–676.