

# FAISABILITÉ DES GRANDS ÉCOSYSTÈMES DE LOTKA-VOLTERRA À COEFFICIENTS ALÉATOIRES

Imane Akjouj <sup>1</sup>

<sup>1</sup> *Laboratoire Paul Painlevé, Université de Lille, Cité Scientifique, 59655 Villeneuve d'Ascq, imane.akjouj@univ-lille.fr*

**Résumé.** 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 impliquent de nombreuses espèces, il peut être difficile de mesurer les interactions entre les espèces et pertinent de les considérer 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. Par ailleurs, certains modèles en écologie font appel à des matrices creuses, c'est-à-dire contenant de nombreux zéros ; chaque espèce interagissant avec un petit nombre d'autres espèces.

Une question centrale concerne l'existence d'un équilibre faisable, c'est-à-dire une solution strictement positive du système de Lotka-Volterra, ce qui correspond au scénario où aucune espèce ne disparaît au cours de la dynamique. En RMT, l'étude des matrices creuses est assez récente et c'est dans cette optique que l'on s'intéresse à la question de la faisabilité de l'équilibre dans le cadre de matrices d'interactions aléatoires creuses.

**Mots-clés.** Matrices aléatoires, équations de Lotka-Volterra, matrices creuses, systèmes linéaires.

## 1 Introduction

### 1.1 Systèmes de Lotka-Volterra et matrices aléatoires

En écologie mathématique, lors de la modélisation de la dynamique d'écosystèmes faisant intervenir des espèces qui interagissent entre elles, l'usage de grands systèmes de Lotka-Volterra est courant. Ainsi, pour un réseau trophique composé de  $n$  espèces, on note  $\mathbf{x}_n(t) = (x_k(t))_{1 \leq k \leq n}$  le vecteur des abondances des espèces au temps  $t$ , où  $x_k(t)$  désigne l'abondance de l'espèce  $k$  (nombre d'individus, biomasse...) au temps  $t$ . Les abondances sont reliées entre elles par le système d'équations :

$$\frac{dx_k(t)}{dt} = x_k(t) \left( 1 - x_k(t) + \sum_{\ell=1}^n M_{k\ell} x_\ell(t) \right) \quad \text{pour } 1 \leq k \leq n,$$

où  $M_{k\ell}$  désigne l'effet de l'espèce  $\ell$  sur l'espèce  $k$ .

Lorsque les écosystèmes impliquent de nombreuses espèces, il peut être difficile de mesurer les interactions et pertinent de les considérer comme aléatoires.

A l'équilibre  $\frac{d\mathbf{x}_n}{dt} = 0$ , le vecteur des abondances  $\mathbf{x}_n = (x_k)$  est solution du système

$$x_k \left( 1 - x_k + \sum_{\ell=1}^n M_{k\ell} x_\ell \right) = 0 \quad \text{pour } x_k \geq 0 \quad \text{et } 1 \leq k \leq n. \quad (1)$$

La question de l'existence d'un équilibre faisable, c'est-à-dire une solution  $\mathbf{x}_n$  au système (1) dont toutes les composantes  $x_k$  sont strictement positives, se pose. Celle-ci correspond au scénario où aucune espèce ne disparaît au cours de la dynamique.

Dans ce cas, l'équation qui nous intéresse prend la forme plus simple :

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

où  $\mathbf{1}_n = (1, \dots, 1)$ .

Historiquement, en amont de la question de la faisabilité apparaît celle de la stabilité de l'équilibre. En se basant notamment sur la loi circulaire pour de grandes matrices aléatoires à entrées i.i.d., Robert May (1972) fait apparaître un critère reliant la stabilité de l'équilibre à la complexité du système (nombre d'espèces  $n$ , etc) et instaure ainsi l'appel à la théorie des grandes matrices aléatoires dans l'étude des réseaux trophiques (Allesina et Tang (2012)). Plus récemment, Stone (2018) a revisité la relation entre faisabilité et stabilité.

## 1.2 Réseaux trophiques et connectance

L'un des paramètres intervenant dans la complexité d'un écosystème est la connectance, qui correspond à la proportion d'interactions non nulles entre les espèces (proportion d'entrées non nulles de la matrice  $M_n$ ) ; chaque espèce interagissant avec un petit nombre d'autres espèces. En effet, en se basant sur des données écologiques et biologiques, Busiello et al. (2017) suggèrent que la connectance des réseaux trophiques est très faible.

Pour ce faire, une première approche est de considérer le produit de Hadamard  $\Delta_n \circ A_n = (\Delta_{ij} A_{ij})$  entre une matrice d'adjacence  $\Delta_n = (\Delta_{ij})$  de taille  $n$  associée à un graphe déterministe  $d_n$ -régulier et  $A_n$  une matrice aléatoire de taille  $n$  à entrées i.i.d de loi Gaussienne  $\mathcal{N}(0, 1)$ . La matrice des interactions  $M_n$  est alors de la forme :

$$M_n = \frac{\Delta_n \circ A_n}{\alpha_n \sqrt{d_n}}, \quad (2)$$

où  $\alpha = (\alpha_n)$  est une suite de réels strictement positifs. La matrice  $M_n$  a alors précisément  $d_n$  entrées non nulles par lignes et par colonnes, ces entrées sont des variables aléatoires i.i.d de loi  $\mathcal{N}\left(0, \frac{1}{d_n \alpha_n^2}\right)$ .

Le facteur de normalisation  $1/(\alpha_n\sqrt{d_n})$  apparaît comme un élément déterminant dans l'étude de la faisabilité de ce modèle. En particulier, la normalisation en  $\sqrt{d_n}$  permet aux entrées de la matrice  $\Delta_n \circ A_n/\sqrt{d_n}$  d'avoir un effet macroscopique sur le système de Lotka-Volterra, même dans le cas de grands réseaux trophiques ( $n$  grand).

Par ailleurs, si  $\alpha_n \equiv \alpha$ , autrement dit si  $\alpha_n$  est une constante, Dougoud et al. (2018) ont montré qu'il est très improbable qu'une solution faisable existe suivant le TCL de Geman et Hwang (1982). Dès lors, le régime d'intérêt est celui d'une suite  $(\alpha_n)$  de réels strictement positifs telle que  $\alpha_n \rightarrow \infty$ .

## 2 Modèles et résultat

### 2.1 Modèles étudiés

#### 2.1.1 Modèle A : Matrice $d_n$ -régulière par blocs.

Soit  $n = d_n \times m$ .

Supposons que  $d_n \geq \log(n)$ . Cette condition sera explicitée dans la section 2.3.1.

Soit  $\sigma$  une permutation du groupe symétrique  $\mathcal{S}_m$ . La matrice de permutations  $P_\sigma = (P_{ij})_{1 \leq i, j \leq m}$  associée à  $\sigma$  est définie par :  $P_{ij} = \begin{cases} 1 & \text{si } j = \sigma(i), \\ 0 & \text{sinon.} \end{cases}$

Dans ce premier modèle, la matrice  $\Delta_n$ , introduite par l'équation (2), est la matrice par blocs définie par :

$$\Delta_n = P_\sigma \otimes J_d = (P_{ij} J_d)_{1 \leq i, j \leq m},$$

où  $J_d = \mathbf{1}_d \mathbf{1}_d^*$  est la matrice de taille  $d$  composée de 1 et où  $\otimes$  est le produit de Kronecker.

**Remarque 1.** La matrice  $\Delta_n$  correspond bien à la matrice d'adjacence d'un graphe  $d_n$ -régulier.

**Exemple.** Soient  $m = 4$  et  $\sigma = (1)(243)$ .

Les matrices  $P_\sigma$ ,  $\Delta$  et  $\Delta \circ A$  sont respectivement données par :

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

où  $A^{(\mu)}$  ( $\mu \in [4]$ ) est une matrice aléatoire de taille  $d$  composée d'entrées i.i.d.  $\mathcal{N}(0, 1)$ .

#### 2.1.2 Modèle B : $d_n$ proportionnel à $n$ .

La matrice  $M_n$  est donnée par l'équation (2) et on suppose que la suite  $(d_n)$  vérifie

$$\lim_{n \rightarrow \infty} \frac{d_n}{n} = \beta > 0.$$

## 2.2 Seuil de faisabilité

**Théorème.** Soient  $A_n$  une matrice aléatoire de taille  $n$  à entrées i.i.d.  $\mathcal{N}(0, 1)$  et  $\Delta_n$  donnée par le modèle (A) ou le modèle (B).

Supposons que  $\alpha_n \xrightarrow[n \rightarrow \infty]{} \infty$ .

Soit  $\mathbf{x}_n = (x_k)_{1 \leq k \leq n}$  solution de l'équation :

$$\mathbf{x}_n = \mathbf{1}_n + \frac{1}{\alpha_n \sqrt{d_n}} (\Delta_n \circ A_n) \mathbf{x}_n . \quad (3)$$

Alors,

1. si  $\exists \varepsilon > 0$  tel que  $\alpha_n \leq (1 - \varepsilon) \sqrt{2 \log(n)}$  alors  $\mathbb{P} \{ \min_k x_k > 0 \} \xrightarrow[n \rightarrow \infty]{} 0$ ,
2. si  $\exists \varepsilon > 0$  tel que  $\alpha_n \geq (1 + \varepsilon) \sqrt{2 \log(n)}$  alors  $\mathbb{P} \{ \min_k x_k > 0 \} \xrightarrow[n \rightarrow \infty]{} 1$ .

**Remarque 2.** Cette même transition de phase a été précédemment établie par Bizeul et Najim (2021) dans le cas d'une matrice d'interactions pleine (i.e. lorsque  $d_n = n$ ).

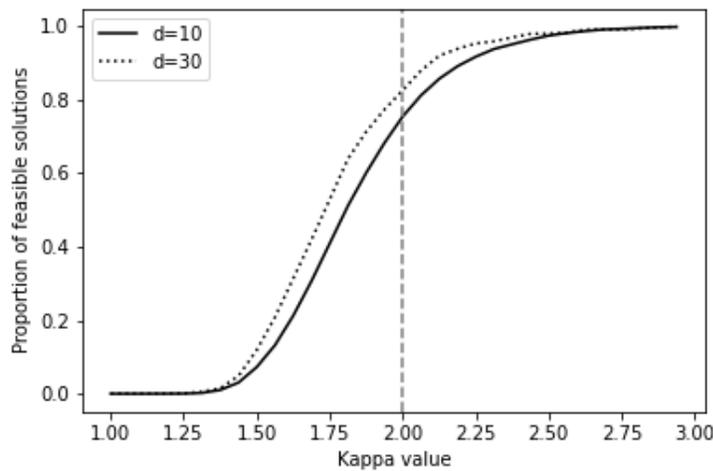


Figure 1: Pour  $m = 1500$  et  $d_n = 10$  ( $d_n > \log(n) \simeq 9.61$ ), suivant le modèle A, la ligne continue représente la proportion de solutions faisables  $\mathbf{x}_n$  pour 2000 réalisations de la matrice aléatoire  $A_n$  par valeur de  $\kappa$  pour 16 valeurs différentes de  $\kappa$ , avec  $\alpha_n = \sqrt{\kappa \log(n)}$ . La même simulation est effectuée pour  $d_n = 30$  pour 500 réalisations de  $A_n$  (ligne pointillée).

## 2.3 Eléments de preuve

### 2.3.1 Gestion de la résolvante

Des résultats théoriques sur les normes des matrices creuses à entrées aléatoires obtenus par Bandeira et van Handel (2016) garantissent que, sous la condition  $d_n \geq \log(n)$ , le rayon spectral de la matrice  $\frac{\Delta_n \circ A_n}{\sqrt{d_n}}$  est d'ordre  $\mathcal{O}_{\mathbb{P}}(1)$ .

Dès lors, la matrice  $\left(I_n - \frac{\Delta_n \circ A_n}{\alpha_n \sqrt{d_n}}\right)$  est inversible et la solution  $\mathbf{x}_n$  de l'équation (3) vérifie l'égalité :

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

### 2.3.2 Transition de phase

Un développement du premier ordre de la résolvante  $\left(I_n - \frac{\Delta_n \circ A_n}{\alpha_n \sqrt{d_n}}\right)^{-1}$  négligeant le reste donne :

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

où

$$\mathbf{z}_n = (z_k)_{1 \leq k \leq n} \quad \text{et} \quad z_k = \sum_{j=1}^n \frac{(\Delta_n \circ A_n)_{kj}}{\sqrt{d_n}} .$$

Les variables aléatoires  $z_i$  sont des variables i.i.d. de loi  $\mathcal{N}(0, 1)$ . De manière approchée, il vient alors:

$$\min_{1 \leq k \leq n} x_k \simeq 1 + \frac{\min_{1 \leq k \leq n} z_k}{\alpha_n} .$$

A l'aide des résultats standards sur les lois d'extremum (loi de Gumbel), on obtient  $\min_k z_k \sim -\sqrt{2 \log(n)}$ , ce qui permet de retrouver la transition de phase.

### 2.3.3 Et le reste

Plus précisément, le développement de la résolvante donne :

$$\mathbf{x}_n = \mathbf{1}_n + \frac{\mathbf{z}_n}{\alpha_n} + \frac{\mathbf{r}_n}{\alpha_n^2} \quad \text{où} \quad \mathbf{r}_n = \sum_{\ell=2}^{\infty} \frac{1}{\alpha_n^{\ell-2}} \left(\frac{\Delta_n \circ A_n}{\sqrt{d_n}}\right)^{\ell} \mathbf{1}_n = (r_k)_{1 \leq k \leq n} .$$

Montrer la faisabilité du modèle revient alors à démontrer le résultat suivant :

**Lemme.**

$$\frac{\max_k r_k}{\alpha_n \sqrt{2 \log n}} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0 \quad \text{et} \quad \frac{\min_k r_k}{\alpha_n \sqrt{2 \log n}} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0 .$$

### 2.3.4 Remarque liée au modèle A

Lorsque  $d_n \ll n$ , apparaît une problématique liée à la norme euclidienne du vecteur  $\mathbf{1}_n/\sqrt{d_n}$ . En effet, celle-ci n'est pas d'ordre  $\mathcal{O}(1)$  mais d'ordre  $\sqrt{n/d_n}$  ce qui rend délicate la gestion de la structure creuse de la matrice  $\Delta_n$ .

A cet effet, la structure par blocs du modèle A est une hypothèse technique et la question de la faisabilité reste ouverte pour  $\Delta_n$  matrice d'adjacence d'un graphe  $d_n$ -régulier avec  $d_n \geq \log(n)$ .

## Bibliographie

- Allesina, S. et Tang, S. (2012). Stability criteria for complex ecosystems, *Nature*, 483, pp. 205–208.
- Bandeira, A. et van Handel, R. (2016). Sharp nonasymptotic bounds on the norm of random matrices with independent entries, *The Annals of Probability*, 44, pp. 2479–2506.
- Bizeul, P. et Najim, J. (2021). Positive solutions for large random linear systems, *Proceedings of the American Mathematical Society*, 149, pp. 2333–2348.
- Busiello, D. et Suweis, S. et Hidalgo, J. et Maritan, A. (2017). Explorability and the origin of network sparsity in living systems, *Scientific Reports*, 7.
- Dougoud, M. et Vinckenbosch, L. et Rohr, R. et 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.
- Geman, S. et Hwang, C-R. (1982). A chaos hypothesis for some large systems of random equations, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 60, pp. 291–314.
- May, R. (1972). Will a Large Complex System be Stable?, *Nature*, 238, pp. 413–414.
- Stone, L. (2018). The feasibility and stability of large complex biological networks: a random matrix approach, *Scientific Reports*, 8.