EQUILIBRIUM IN A LARGE LOTKA-VOLTERRA SYSTEM WITH PAIRWISE CORRELATED INTERACTIONS

Maxime Clenet, Hafedh El Ferchichi, Jamal Najim

EQUILIBRIUM IN A LARGE LOTKA-VOLTERRA SYSTEM WITH PAIRWISE CORRELATED INTERACTIONS

MAXIME CLENET, HAFEDH EL FERCHICHI, JAMAL NAJIM

Abstract. Consider a Lotka-Volterra (LV) system of coupled differential equations:

$$\dot{x}_k = x_k(r_k - x_k + (B\boldsymbol{x})_k)\,, \quad \boldsymbol{x} = (x_k)\,, \quad 1 \leqslant k \leqslant n\,,$$

where $\boldsymbol{r} = (r_k)$ is a $n \times 1$ vector and B a $n \times n$ matrix. Assume that the interaction matrix B is random and follows the elliptic model:

$$B = \frac{1}{\alpha\sqrt{n}} A + \frac{\mu}{n} \mathbf{1}_n \mathbf{1}_n^{\mathsf{T}}\,,$$

where $A = (A_{ij})$ is a $n \times n$ matrix with $\mathcal{N}(0,1)$ entries satisfying the following dependence structure (i) the entries A_{ij} on and above the diagonal are i.i.d., (ii) for $i < j$ each vector (A_{ij}, A_{ji}) is standard gaussian with covariance ρ, and independent from the other entries; vector $\mathbf{1}_n$ stands for the $n \times 1$ vector of ones. Parameters α, μ are deterministic and may depend on n.

Leveraging on Random Matrix Theory, we analyse this LV system as $n \to \infty$ and study the existence of a positive equilibrium. This question boils down to study the existence of a (componentwise) positive solution to the linear equation:

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

depending on B's parameters (α, μ, ρ), a problem of independent interest in linear algebra.

In the case where no positive equilibrium exists, we provide sufficient conditions for the existence of a unique stable equilibrium (with vanishing components), and following Bunin [9], present a heuristics estimating the number of positive components of the equilibrium and their distribution.

The existence of positive equilibria for large Lotka-Volterra systems has been raised in Dougoud et al. [14], and addressed in various contexts by Najim et al. [1, 7].

Such LV systems are widely used in mathematical biology to model populations with interactions, and the existence of a positive equilibrium known as a *feasible equilibrium* corresponds to the survival of all the species x_k within the system.

1. Introduction

Lotka-Volterra system of coupled differential equations. Lotka-Volterra (LV) systems are widely used in mathematical biology, ecology, chemistry to model populations with interactions or chemical reactions [20, 22, 24, 21]. In the context

Date: May 30, 2022.

2010 *Mathematics Subject Classification.* Primary 15B52, 60G70, Secondary 60B20, 92D40.

Key words and phrases. Linear systems; large random matrices; Gaussian concentration; Lotka-Volterra equations.

Supported by CNRS Project 80 Prime — KARATE..

of theoretical ecology (that we shall adopt hereafter without loss of generality), consider a given foodweb and denote by $\boldsymbol{x}_n^t = (x_k(t))_{1\leqslant k\leqslant n}$ the vector of abundances[1] of the various species at time $t \geqslant 0$. In a LV system, the abundances are connected via the following coupled equations:

$$\frac{dx_k(t)}{dt} = x_k(t)\left(r_k - x_k(t) + \sum_{\ell=1}^{n} B_{k\ell}x_\ell(t)\right) \qquad \text{for} \quad k \in [n] := \{1,\cdots,n\}\,,$$

where $B_n = (B_{k\ell})$ stands for the interaction matrix, and r_k stands for the intrinsic growth of species k. Notice that standard results yield that if the initial condition $\boldsymbol{x}_n^0 = \boldsymbol{x}_n|_{t=0}$ is componentwise positive, then $\boldsymbol{x}_n^t$ remains positive for every $t > 0$.

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:

$$x_k\left(r_k - x_k + \sum_{\ell\in[n]} B_{k\ell}x_\ell\right) = 0 \qquad \text{for} \quad x_k \geqslant 0 \quad \text{and} \quad k \in [n]\,. \tag{1}$$

An important question, which motivated recent developments [14, 7], is the existence of a *feasible* solution $\boldsymbol{x}_n$ to (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) takes the much simpler form:

$$\boldsymbol{x}_n = \boldsymbol{r}_n + B_n\boldsymbol{x}_n\,, \qquad \boldsymbol{r}_n = (r_k)_{k\in[n]}\,.$$

In this article, we will investigate the existence of an equilibrium, potentially feasible, for a large foodweb ($n \to \infty$) whenever the interaction matrix B_n is random. In various models of interest for B_n, Random Matrix Theory (RMT) provides an accurate description of the asymptotic properties of a large random matrix (its spectrum, spectral norm, etc.). We will leverage on RMT to infer the existence of an equilibrium in the case where B_n follows a random elliptic model, to be described hereafter.

To simplify the analysis, we will consider the case where $\boldsymbol{r}_n = \mathbf{1}_n$.

Random elliptic model for the interaction matrix. In the spirit of May[2], we model the interaction matrix B_n as a non-centered random matrix with pairwise correlated entries:

$$B_n = \frac{A_n}{\alpha_n\sqrt{n}} + \frac{\mu}{n}\mathbf{1}_n\mathbf{1}_n^\mathsf{T}\,, \tag{2}$$

where $A_n = (A_{ij})_{i,j\in[n]}$ is a random matrix satisfying the two conditions (i) $(A_{ij}, i \leqslant j)$ are standard Gaussian $\mathcal{N}(0,1)$ independent and identically distributed (i.i.d.) random variables (ii) for $i < j$ the vector (A_{ij}, A_{ji}) is a standard bivariate Gaussian vector, independent from the remaining random variables, with covariance $\mathrm{cov}(A_{ij}, A_{ji}) = \rho$ with $|\rho| \leqslant 1$. The sequence of positive numbers (α_n) is either fixed or goes to infinity. Parameter μ is a fixed real number. As a consequence, the Gaussian entries of the interaction matrix B_n admit the following moments:

$$\mathbb{E}(B_{ij}) = \frac{\mu}{n}\,, \quad \mathrm{var}(B_{ij}) = \frac{1}{\alpha^2 n}\,, \quad \mathrm{cov}(B_{ij}, B_{ji}) = \frac{\rho}{\alpha^2 n}\ (i \neq j)\,.$$

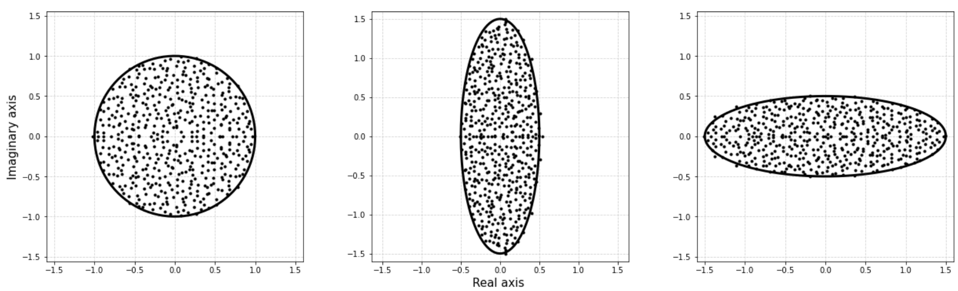
[1] A species abundance is a quantity proportional to the number of individuals for this species.

[2] 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.

Such a matrix model is often called a random elliptic model for $|\rho| < 1$ since the spectrum of matrix $A_n/\sqrt{n}$ is asymptotically an ellipse, see Fig.1, in the sense that the empirical distribution of the eigenvalues of $A_n/\sqrt{n}$ converges towards the uniform distribution on the ellipsoid

$$\mathcal{E}_\rho = \left\{ z \in \mathbb{C},\ \frac{\mathrm{Re}^2(z)}{(1+\rho)^2} + \frac{\mathrm{Im}^2(z)}{(1-\rho)^2} \leqslant 1 \right\} .$$

Originally introduced by Girko [18], this model has since been widely studied [19, 28, 29, 31].



(A) $\rho = 0$ (B) $\rho = -0.5$ (C) $\rho = 0.5$

FIGURE 1. Spectrum of non-Hermitian matrix B_n ($n = 500$) in the centered case ($\mu = 0$) with distinct parameter $\rho \in \{-0.5,\ 0,\ 0.5\}$. The solid line represents the ellipse $\{z = x + \boldsymbol{i}y \in \mathbb{C},\ \frac{x^2}{(1+\rho)^2} + \frac{y^2}{(1-\rho)^2} = 1\}$ which is the boundary of the support of the limiting spectral distribution for an elliptic model.

The spectral norms of A_n and $\mathbf{1}_n\mathbf{1}_n^\mathsf{T}$ satisfy

$$\left\| \frac{A_n}{\sqrt{n}} \right\| = \mathcal{O}(1) \quad \text{and} \quad \left\| \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\mathsf{T} \right\| = 1$$

hence both the random and deterministic parts of the interaction matrix B_n may have an impact as $n \to \infty$.

Presentation of the main results. In this article, we address the following issues.

Feasibility. We first describe the conditions over parameters (ρ, α_n, μ) for which system (1) admits a unique feasible equilibrium. We prove that feasibility is reached whenever $\alpha_n \gg \sqrt{2\log(n)}$ and $\mu < 1$, and that there is no feasibility otherwise, see Theorem 2.1. Notice that the correlation parameter ρ has no influence since the phase transition threshold is the same as in the i.i.d. case [7]: the induced correlations between components x_k's of solution $\boldsymbol{x}_n$ are too weak. Pushing this remark further, we prove that the same phase transition holds if we consider a covariance profile $(\rho_{ij},\ i < j)$ where $\rho_{ij} = \mathrm{cov}(A_{ij}, A_{ji})$ instead of a fixed covariance parameter ρ.

In [7], Bizeul and Najim established the conditions for feasibility in the centered ($\mu = 0$) model with i.i.d interactions (A_{ij}). In [1], Akjouj and Najim studied a sparse model of interactions where only $d_n \geqslant \log(n)$ interactions are non-null in each row and column of A_n. The study of the feasibility for an elliptic model completes this picture.

Stability without feasibility. If α is fixed, Dougoud et al. [14] showed that no feasible solution can arise. Under this assumption, we establish in Proposition 2.3 sufficient conditions for the existence of a unique stable equilibrium to system (1). In this case, some species will vanish (some of the components x_k's of solution $\boldsymbol{x}_n$ are equal to zero). In order to proceed we combine results by Takeuchi [33] on stability of LV systems with Random Matrix Theory (RMT) results.

Estimating the number of surviving species. We finally conclude with an important question: given a set of parameters (ρ, α, μ) which yields to a unique stable equilibrium, is it possible to estimate the proportion of surviving species? From a mathematical point of view, this is an open question. At a physical level of rigor, Bunin [9] (relying on the cavity method) and Galla [15] (relying on generating functionals techniques) provide a closed-form system of equations to compute the proportion of surviving species. We state the open problem, recall Bunin's and Galla's equations and provide simulations.

In [12], equations and simulations are provided in the simpler case where $\rho = 0$, together with heuristics supporting these equations.

Organisation of the article. Feasibility and stability results together with the open question on the estimation of the number of surviving species are presented in Section 2. Section 3 is devoted to the proof of the feasibility result, Theorem 2.1. Proof of the stability result, Proposition 2.3, is provided in Section 4. Simulations were performed in Python. All the figures and the code are available on Github [11].

Notations. If A is a matrix A^T stands for its transpose. We denote by $\log(x)$ the natural logarithm. If $\boldsymbol{x} = (x_i)_{i\in[n]}$ is a vector, we denote by $\boldsymbol{x} > 0$ (resp. $\boldsymbol{x} \geqslant 0$) the componentwise positivity (resp. non-negativity), that is the fact that $x_i > 0$ (resp. $x_i \geqslant 0$) for every $i \in [n]$.

2. Main results: Feasibility, stability and surviving species

2.1. **Feasibility.** To simplify the analysis, we consider the case where $r_k = 1$ ($k \in [n]$). Hence, the LV system takes the following form in the sequel:

$$\frac{dx_k(t)}{dt} = x_k(t)\left(1 - x_k(t) + \sum_{\ell\in[n]} B_{k\ell}x_\ell(t)\right) \qquad \text{for} \quad k \in [n]\,. \tag{3}$$

In the next theorem, we describe the conditions to reach a feasible equilibrium. We either assume that matrix B is given by the elliptic model or has a more general covariance profile.

Theorem 2.1 (Feasibility for the elliptic model)**.** *Assume that matrix B_n is given by the elliptic model (2), or that B_n has a covariance profile, i.e.*

$$B_n = \frac{\tilde{A}_n}{\alpha_n\sqrt{n}} + \frac{\mu}{n}\mathbf{1}_n\mathbf{1}_n^\mathsf{T}\,, \tag{4}$$

where $\tilde{A}_n$ is a $n \times n$ matrix with entries $(\tilde{A}_{ij}, i \leqslant j)$ i.i.d. $\mathcal{N}(0,1)$ and where $(\tilde{A}_{ij}, \tilde{A}_{ji})$ is a standard bivariate gaussian vector for $i < j$, independent from the *remaining random variables, with covariance* $\mathrm{cov}(\tilde{A}_{ij}, \tilde{A}_{ji}) = \rho_{ij}^{(n)}$*, where* $(\rho_{ij}^{(n)}; i < j; n \geqslant 1)$ *is a collection of deterministic real numbers in* $[-1, 1]$*.*

Let $\alpha_n \xrightarrow[n\to\infty]{} \infty$ *and denote by* $\alpha_n^* = \sqrt{2\log n}$*. If* $\mu \neq 1$ *then the following equation*

$$\boldsymbol{x}_n = \mathbf{1}_n + B_n \boldsymbol{x}_n$$

almost surely admits a unique solution $\boldsymbol{x}_n = (x_k)_{k\in[n]}$*.*

1. *(feasibility) If* $\mu < 1$ *and there exists* $\varepsilon > 0$ *such that, for* n *large enough,* $\alpha_n \geqslant (1+\varepsilon)\alpha_n^*$ *then*
$$\mathbb{P}\left\{\min_{k\in[n]} x_k > 0\right\} \xrightarrow[n\to\infty]{} 1\,.$$
2. *If* $\mu > 1$ *or there exists* $\varepsilon > 0$ *such that, for* n *large enough,* $\alpha_n \leqslant (1-\varepsilon)\alpha_n^*$ *then*
$$\mathbb{P}\left\{\min_{k\in[n]} x_k > 0\right\} \xrightarrow[n\to\infty]{} 0\,.$$

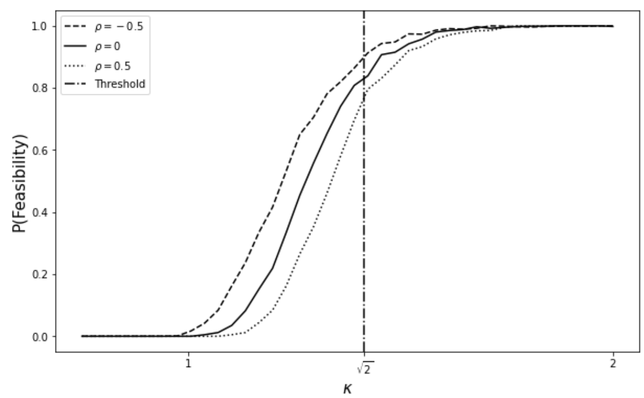


FIGURE 2. Transition towards feasibility for the elliptic model (2). For each κ on the x-axis, we simulate 1000 matrices B_n of size $n = 1000$ and compute the solution x_n of Theorem 2.1 at the scaling $\alpha_n(\kappa) = \kappa\sqrt{\log(n)}$. Each curve represents the proportion of feasible solutions x_n obtained for the 1000 simulations. Three distinct values $\rho \in \{-0.5,\ 0,\ 0.5\}$ are used. The dot-dashed vertical line corresponds to $\kappa = \sqrt{2}$ i.e. the critical scaling $\alpha_n^* = \sqrt{2}\sqrt{\log(n)}$.

Proof of Theorem 2.1 is established in Section 3 under the assumption that B_n follows the elliptic model. The adaptations needed to cover the covariance profile case are provided in Appendix A.

2.2. **No feasibility but a unique stable equilibrium.** 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 and Ashby [16] considered stability issues of complex systems connected at random. Based on the circular law for large random matrices with i.i.d. entries, May [27] 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 al. [3]. Recently, Stone [32] and Gibbs et al. [17] revisited the relation between feasibility and stability.

For a generic LV system

$$\frac{d\,y_k(t)}{dt} = y_k(r_k + (C\boldsymbol{y})_k)\,, \quad k \in [n]\,, \tag{5}$$

Takeuchi and Adachi provide a criterion for the existence of a unique equilibrium $\boldsymbol{y}^*$ and the global stability of LV systems, see Theorem 3.2.1 in [33].

Theorem 2.2 (Takeuchi and Adachi 1980)**.** *If there exists a positive diagonal matrix Δ such that $\Delta C + C^{\mathsf{T}}\Delta$ is negative definite, there is a unique non-negative equilibrium $\boldsymbol{y}^*$ to (5), which is globally stable:*

$$\forall \boldsymbol{y}_0 > 0\,, \quad \begin{cases} \boldsymbol{y}(0) = \boldsymbol{y}_0 \\ \boldsymbol{y}(t) \ \textit{satisfies}\ (5) \end{cases} \quad , \quad \boldsymbol{y}(t) \xrightarrow[t\to\infty]{} \boldsymbol{y}^*\,.$$

Combining this result (setting $I - B = -C$) with results from Random Matrix Theory, we can guarantee the existence of a globally stable equilibrium $\boldsymbol{x}^*$ of (3) for a wide range of parameters (ρ, α, μ). Denote by

$$\mathcal{A} = \Bigg\{ (\rho,\alpha,\mu) \in (-1,1) \times (0,\infty) \times \mathbb{R}\,, \\ \alpha > \sqrt{2(1+\rho)}, \quad \mu < \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2(1+\rho)}{\alpha^2}} \Bigg\} \tag{6}$$

the set of admissible parameters.

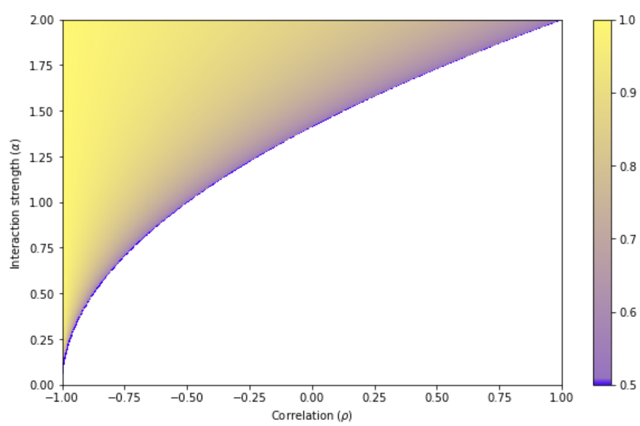


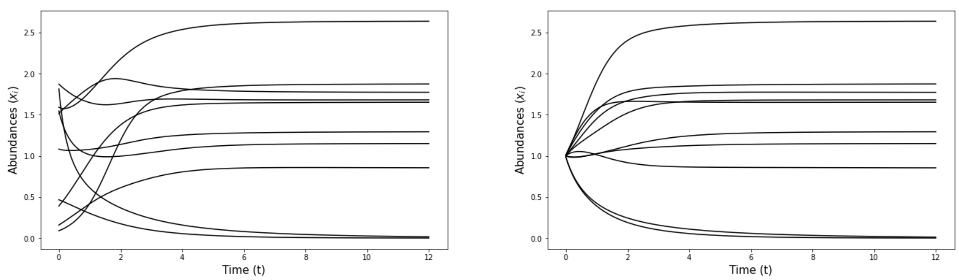
FIGURE 3. Representation of the set of admissible parameters $\mathcal{A}$ by a heat map. The set $\mathcal{A}$ given by (6) yields the existence of a unique (random) globally stable equilibrium $\boldsymbol{x}^*$. The x-axis corresponds to ρ, the y-axis to σ and the intensity of the color μ.

Proposition 2.3. *Let $(\rho, \alpha, \mu) \in \mathcal{A}$, then almost surely, matrix*

$$(I - B) + (I - B)^{\mathsf{T}}$$

is eventually positive definite: with probability one, for a given realization ω, there exists $N(\omega)$ such that for $n \geqslant N(\omega)$, $(I - B^\omega) + (I - B^\omega)^\mathsf{T}$ is positive definite. In particular, there exists a unique globally stable non-negative equilibrium $\boldsymbol{x}^*$.

Proof of Proposition 2.3 is provided in Section 4.



(A) Initial conditions drawn in $(0, 2)$, (B) Initial conditions equal to 1.

FIGURE 4. Representation of the dynamics of a ten-species system. For a fixed matrix of interactions B_{10} with parameters $(\rho = 0, \alpha = 2, \mu = 0) \in \mathcal{A}$, we consider two distinct initial conditions. Simulations show that the abundances converge in both cases toward the unique globally stable equilibrium x^* predicted by Proposition 2.3. Notice that since $\alpha < \sqrt{2\log(10)} \simeq 2.14$, we witness vanishing species.

2.3. **Estimating the number of surviving species: Towards Bunin and Galla's equations.** After giving conditions for the realization of a feasible equilibrium and investigating the existence and uniqueness of a stable sub-population (i.e some species vanish), we address the question of estimating the proportion of surviving species as a function of the model paramaters (ρ, α, μ).

To our knowledge, this question has not received yet an answer at a mathematical level of rigor and remains open. However theoretical physicists/ecologists provided a solution to this problem supported by simulations. Tools from physics to study population dynamics in the context of Lotka-Volterra equations were first introduced by Opper et al. [13, 30]. In 2017, Bunin [9] precisely answers the question of estimating the proportion of surviving species for the model under investigation (non-centered elliptic model B). He uses the dynamical cavity method (a review of which can be found in [6]). The key concept consists of assuming that a unique fixed point exists and introducing a new species with new interactions in the existing system. Provided the coherence of the assumption, an analogy between the properties of the solutions with n and $n+1$ species yields closed-form equations that we present hereafter.

Notice that recently, similar equations were obtained by Galla [15] using generating functional techniques.

The system of equations presented hereafter is a version of Bunin's equations without the carrying capacity. It is similar to the equations obtained by the replicator equations [13, 30]. Notice that we mention but do not discuss the many implicit assumptions yielding the system of equations.

Let $(\rho, \alpha, \mu) \in \mathcal{A}$ and $\boldsymbol{x}^*$ given by Proposition 2.3. We first introduce the following quantities:

$$\phi = \frac{\operatorname{Card}\{x_i^* > 0, i \in [n]\}}{n} \quad , \qquad \langle \boldsymbol{x} \rangle = \frac{1}{n}\sum_{j=1}^{n} x_j^* \ , \qquad \left\langle \boldsymbol{x}^2 \right\rangle = \frac{1}{n}\sum_{j=1}^{n} (x_j^*)^2 \ . \tag{7}$$

Denote by $Z \sim \mathcal{N}(0,1)$ and set

$$\Delta = (1 + \langle \boldsymbol{x} \rangle \mu)\frac{\alpha}{\sqrt{\langle \boldsymbol{x}^2 \rangle}} \, .$$

The following system of 4 equations has 4 unknowns, among which the (supposedly existing) asymptotic limits of $\phi, \langle \boldsymbol{x} \rangle, \left\langle \boldsymbol{x}^2 \right\rangle$, denoted (by abuse of notations) by the same notations. The fourth unknown v is a parameter essentially related to the dynamical cavity method. This system is supposed to admit a unique solution:

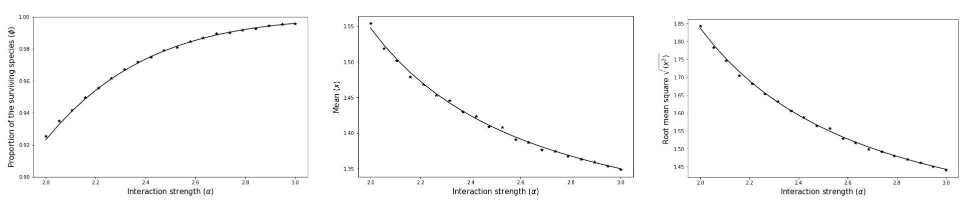
$$\phi = \frac{1}{\sqrt{2\pi}}\int_{-\Delta}^{+\infty} e^{\frac{-z^2}{2}}\, dz \tag{8}$$

$$\langle \boldsymbol{x} \rangle = \frac{\phi}{1 - \frac{\rho v}{\alpha}}\left((1 + \langle \boldsymbol{x} \rangle \mu) + \frac{\sqrt{\langle \boldsymbol{x}^2 \rangle}}{\alpha}\mathbb{E}(Z | Z > -\Delta) \right) \tag{9}$$

$$\begin{aligned} \left\langle \boldsymbol{x}^2 \right\rangle = \left(\frac{\sqrt{\phi}}{1 - \frac{\rho v}{\alpha}} \right)^2 \Bigg(& (1 + \langle \boldsymbol{x} \rangle \mu)^2 + 2(1 + \langle \boldsymbol{x} \rangle \mu)\frac{\sqrt{\langle \boldsymbol{x}^2 \rangle}}{\alpha}\mathbb{E}(Z | Z > -\Delta) \\ & + \frac{\langle \boldsymbol{x}^2 \rangle}{\alpha^2}\mathbb{E}(Z^2 | Z > -\Delta) \Bigg) \end{aligned} \tag{10}$$

$$v = \phi\left(\frac{1}{\alpha - \rho v} \right) \tag{11}$$

The theoretical solutions of system (8)-(11) are compared with the empirical values obtained by Monte-Carlo experiments. As shown in Fig. 5, the matching is remarkable.



(A) ϕ versus α, (B) $\langle x \rangle$ versus α, (C) $\left\langle x^2 \right\rangle$ versus α.

FIGURE 5. Theoretical values of ϕ, $\langle \boldsymbol{x} \rangle$ and $\langle \boldsymbol{x}^2 \rangle$ (solid line) obtained by solving the system (8)-(11) given the parameters $(\mu = 0.2, \rho = 0.5)$, compared to the empirical values (dots) obtained by Monte-Carlo simulations (size of matrix $n = 500$, number of random samples $P = 200$). The x-axis corresponds to the interaction strength α.

The impact of the correlation ρ on the proportion of the surviving species is shown in Figure 6.

Remark 2.4. *From a theoretical ecology point of view, notice that a negative correlation (prey-predator) seems to slow down the decline of the surviving species, whereas a positive correlation (mutualism and competition) reverses the trend. These types of results are similar to Allesina and Tang [2] where they notice that prey-predator interactions seem to stabilize the system.*

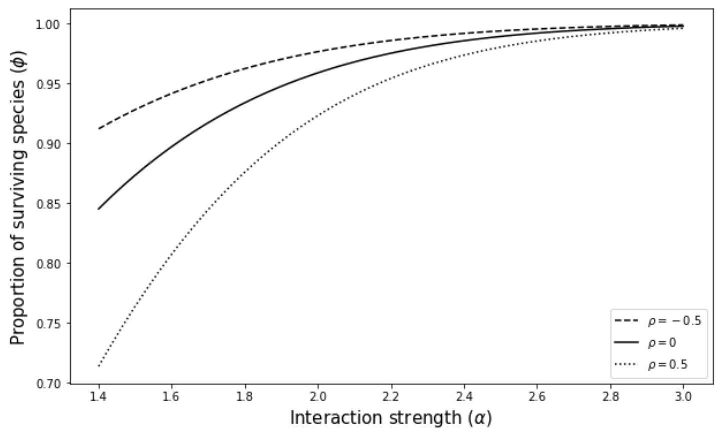


FIGURE 6. Effect of the correlation ρ and the interaction strength α on the proportion of surviving species ϕ. Each curve is plotted by resolving the system (8)-(11) in the centered case $\mu = 0$.

3. Feasibility: Proof of Theorem 2.1

We assume that matrix B_n is given by (2) (elliptic model). The case where matrix B_n is given by (4) (covariance profile model) needs extra arguments which are provided in Appendix A.

3.1. Preliminary results.

Extreme Value Theory (EVT) and the Normal Comparison Lemma. Let $(Z_k)_{k\in\mathbb{N}}$ be a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables and denote:

$$\begin{cases} M_n = \max_{k\in[n]} Z_k \\ \breve{M}_n = \min_{k\in[n]} Z_k \end{cases}, \quad \alpha_n^* = \sqrt{2\log n}\,, \quad \beta_n^* = \alpha_n^* - \frac{1}{2\alpha_n^*}\log(4\pi\log n)\,. \tag{12}$$

Let $G(x) = e^{-e^{-x}}$ be the Gumbel cumulative distribution function, then classical EVT results (see for instance [26, Theorem 1.5.3]) yield that for every $x \in \mathbb{R}$,

$$\mathbb{P}\left\{\alpha_n^*(M_n - \beta_n^*) \leqslant x\right\} \xrightarrow[n\to\infty]{} G(x)\,, \quad \mathbb{P}\left\{\alpha_n^*(\breve{M}_n + \beta_n^*) \geqslant -x\right\} \xrightarrow[n\to\infty]{} G(x)\,. \tag{13}$$

We consider the following dependent framework: Let $(\mathcal{Z}_{k,n})_{k\in[n]}$ be a Gaussian vector whose components are $\mathcal{N}(0,1)$ with covariance

$$\operatorname{cov}(\mathcal{Z}_{k,n};\mathcal{Z}_{\ell,n}) = \frac{\rho}{n}\,, \quad |\rho| \leqslant 1\,, \quad k \neq \ell\,.$$

We are interested in the behaviour of $\mathcal{M}_n = \max_{k\in[n]} \mathcal{Z}_{k,n}$ and $\breve{\mathcal{M}}_n = \min_{k\in[n]} \mathcal{Z}_{k,n}$, and shall prove the counterpart of (13) with the help of the Normal Comparison Lemma (NCL):

Theorem 3.1 (Theorem 4.2.1, [26]). *Suppose that $(\xi_i,\, i \in [n])$ is a gaussian vector where the ξ_i's are standard normal variables, with covariance matrix $\Lambda^1 = \left(\Lambda^1_{ij}\right)$. Similarly, let $(\eta_i,\, i \in [n])$ be a gaussian vector where the η_i's are standard normal, with covariance matrix $\Lambda^0 = \left(\Lambda^0_{ij}\right)$. Denote by $\rho_{ij} = \max\left\{|\Lambda^0_{ij}|, |\Lambda^1_{ij}|\right\}$ and let $(u_i,\, i \in [n])$ be real numbers. Then:*

$$
\begin{aligned}
&\left|P\left\{\xi_j \leqslant u_j\,,\ j \in [n]\right\} - P\left\{\eta_j \leqslant u_j\,,\ j \in [n]\right\}\right| \\
&\qquad\leqslant \frac{1}{2\pi} \sum_{1 \leqslant i < j \leqslant n} \left|\Lambda^1_{ij} - \Lambda^0_{ij}\right| \left(1 - \rho^2_{ij}\right)^{-1/2} \exp\left(-\frac{\frac{1}{2}(u_i^2 + u_j^2)}{1 + \rho_{ij}}\right)\,. \qquad (14)
\end{aligned}
$$

Corollary 3.2. *Recall the definition of $(\mathcal{Z}_{k,\ell})_{k\in[n]}$, $\mathcal{M}_n$ and $\breve{\mathcal{M}}_n$ above, then*

$$
\mathbb{P}\left\{\alpha^*_n(\mathcal{M}_n - \beta^*_n) \leqslant x\right\} \xrightarrow[n\to\infty]{} G(x)\,, \quad \mathbb{P}\left\{\alpha^*_n(\breve{\mathcal{M}}_n + \beta^*_n) \geqslant -x\right\} \xrightarrow[n\to\infty]{} G(x)\,. \qquad (15)
$$

Proof. We apply the NCL to $(Z_k)_{k\in[n]}$ and $(\mathcal{Z}_{k,n})_{k\in[n]}$. Let $\rho_{ij} = \frac{|\rho|}{n}$ and $u_n(x) = \frac{x}{\alpha^*_n} + \beta^*_n$, then

$$
\begin{aligned}
&|\mathbb{P}\{\alpha^*_n(M_n - \beta^*_n) \leqslant x\} - \mathbb{P}\{\alpha^*_n(\mathcal{M}_n - \beta^*_n) \leqslant x\}| \\
&= \quad |\mathbb{P}\{Z_j \leqslant u_n(x)\,,\ j \in [n]\} - \mathbb{P}\{\mathcal{Z}_{j,n} \leqslant u_n(x)\,,\ j \in [n]\}|\ , \\
&\leqslant \quad \frac{1}{2\pi}\frac{n(n-1)}{2}\frac{|\rho|}{n}\left(1 - \frac{\rho^2}{n^2}\right)^{-\frac{1}{2}} \exp\left(-\frac{u_n^2(x)}{1 + \frac{|\rho|}{n}}\right) \quad \leqslant \quad K\, n \exp\left(-\frac{u_n^2(x)}{1 + \frac{1}{n}}\right)\,.
\end{aligned}
$$

Now eventually $u_n(x) = \alpha_n(1 + o(1)) \geqslant \kappa\alpha_n$ for any $\kappa < 1$ and eventually

$$
n \exp\left(-\frac{u_n^2(x)}{1 + \frac{1}{n}}\right) \quad \leqslant \quad n \exp\left(-\frac{2\kappa^2 \log(n)}{1 + \frac{1}{n}}\right) \quad = \quad n^{-\left(\frac{2\kappa^2}{1+\rho/n} - 1\right)}\,.
$$

This last term goes to zero as $n \to \infty$ for a well-chosen κ sufficiently close to one. This concludes the proof for $\mathcal{M}_n$. The proof for $\breve{\mathcal{M}}_n$ can be handled similarly with minor modifications. ∎

Random Matrix Theory. Let B_n be given by model (2).

Lemma 3.3. *Let A_n a $n \times n$ matrix with i.i.d. $\mathcal{N}(0,1)$ entries for $i \leqslant j$ and (A_{ij}, A_{ji}) a standard bivariate Gaussian vector with covariance ρ for $i < j$, then the following estimate holds true: almost surely,*

$$
\limsup_{n\to\infty} \left\|\frac{A_n}{\sqrt{n}}\right\| \quad \leqslant \quad \sqrt{2}\left(\sqrt{1+\rho} + \sqrt{1-\rho}\right) \quad \leqslant \quad 2\sqrt{2}\,.
$$

Proof. The proof relies on two arguments: the classical estimate of the asymptotic spectral norm of a Wigner matrix [4, Th. 5.1] and the following decomposition of matrix $A_n/\sqrt{n}$ as linear combination of Hermitian Wigner matrices:

$$
\frac{A_n}{\sqrt{n}} = \frac{A_n + A_n^\mathsf{T}}{2\sqrt{n}} - \boldsymbol{i}\frac{\left[\boldsymbol{i}\left(A_n - A_n^\mathsf{T}\right)\right]}{2\sqrt{n}}\,, \qquad (\boldsymbol{i}^2 = -1)\,. \qquad (16)
$$

Notice that both matrices $W_n^1 = \frac{A_n + A_n^\mathsf{T}}{2\sqrt{n}}$ and $W_n^2 = \frac{\left[\boldsymbol{i}\left(A_n - A_n^\mathsf{T}\right)\right]}{2\sqrt{n}}$ are Wigner matrices, with off-diagonal variances ($i<j$):

$$\mathrm{var}\left(\left[\frac{A_n + A_n^\mathsf{T}}{2}\right]_{ij}\right) = \frac{1+\rho}{2} \qquad \text{and} \qquad \mathrm{var}\left(\left[\frac{\boldsymbol{i}\left(A_n - A_n^\mathsf{T}\right)}{2}\right]_{ij}\right) = \frac{1-\rho}{2}\,.$$

Hence,

$$\limsup_n \left\|\frac{A_n}{\sqrt{n}}\right\| \leqslant \limsup_n \|W_n^1\| + \limsup_n \|W_n^2\| = 2\left(\sqrt{\frac{1+\rho}{2}} + \sqrt{\frac{1-\rho}{2}}\right)$$

An elementary analysis yields $\sqrt{2}(\sqrt{1+\rho}+\sqrt{1-\rho}) \leqslant 2\sqrt{2}$ for $|\rho| \leqslant 1$. □

3.2. **Proof of Theorem 2.1 - the centered case $\mu = 0$.**

Some preparation and strategy of proof. We first prove Theorem 2.1 in the case where $\mu = 0$ and focus on the equation

$$\boldsymbol{x}_n = \mathbf{1}_n + \frac{A_n}{\alpha_n\sqrt{n}}\boldsymbol{x}_n\,. \tag{17}$$

By Lemma 3.3, $\limsup_n \|A_n/\sqrt{n}\|$ is a.s. bounded hence

$$\left\|\frac{A_n}{\alpha_n\sqrt{n}}\right\| \xrightarrow[n\to\infty]{a.s.} 0\ .$$

As a consequence, the resolvent $Q_n = \left(I_n - A_n/(\alpha_n\sqrt{n})\right)^{-1}$ is a.s. eventually well-defined and the solution $\boldsymbol{x}_n = (x_k)_{k\in[n]}$ of (17) writes $\boldsymbol{x}_n = Q_n\mathbf{1}_n$. Denote by $\boldsymbol{e}_k$ the kth canonical vector of $\mathbb{R}^n$. The following representation holds true (we shall often drop index n in the following)

$$\begin{aligned} x_k &= \boldsymbol{e}_k^\mathsf{T}\boldsymbol{x} = \boldsymbol{e}_k Q\mathbf{1} = \sum_{\ell=0}^{\infty} \boldsymbol{e}_k^\mathsf{T}\left(\frac{A}{\alpha\sqrt{n}}\right)^\ell \mathbf{1}\ , \\ &= 1 + \frac{1}{\alpha}\boldsymbol{e}_k^\mathsf{T}\left(\frac{A}{\sqrt{n}}\right)\mathbf{1} + \frac{1}{\alpha^2}\boldsymbol{e}_k^\mathsf{T}\left(\frac{A}{\sqrt{n}}\right)^2 Q\,\mathbf{1}\,. \end{aligned} \tag{18}$$

Denote by

$$Z_{k,n} = \boldsymbol{e}_k^\mathsf{T}\left(\frac{A}{\sqrt{n}}\right)\mathbf{1} = \frac{1}{\sqrt{n}}\sum_i A_{ki} \quad \text{and} \quad R_{k,n}(A) = \boldsymbol{e}_k^\mathsf{T}\left(\frac{A}{\sqrt{n}}\right)^2 Q\,\mathbf{1}\,. \tag{19}$$

Notice that the $Z_{k,n}$'s are standard $\mathcal{N}(0,1)$ however they are not independent as

$$\mathrm{cov}(Z_{k,n}, Z_{\ell,n}) = \frac{1}{n}\mathrm{cov}(A_{k\ell}, A_{\ell k}) = \frac{\rho}{n}\,, \quad k \neq \ell\,.$$

Introducing $M_n = \max_{k\in[n]} Z_{k,n}$ and $\breve{M}_n = \min_{k\in[n]} Z_{k,n}$, we proved in Corollary 3.2 that

$$\mathbb{P}\left\{\alpha_n^*(M_n - \beta_n^*) \leqslant x\right\}\ ,\ \mathbb{P}\left\{\alpha_n^*(\breve{M}_n + \beta_n^*) \geqslant -x\right\} \xrightarrow[n\to\infty]{} G(x)\,. \tag{20}$$

In the sequel, we often drop n and simply write $R_k(A)$ instead of $R_{k,n}(A)$. Following the same strategy as in [7], we notice that (18) yields

$$\begin{cases} \min_{k\in[n]} x_k & \geqslant \quad 1 + \frac{1}{\alpha}\breve{M} + \frac{1}{\alpha^2}\min_{k\in[n]} R_k(A) \\ \min_{k\in[n]} x_k & \leqslant \quad 1 + \frac{1}{\alpha}\breve{M} + \frac{1}{\alpha^2}\max_{k\in[n]} R_k(A) \end{cases}\ ,$$

which we can rewrite

$$\begin{aligned}\min_{k\in[n]} x_k &\geqslant 1+\frac{\alpha_n^*}{\alpha_n}\left(\frac{\check{M}+\beta_n^*}{\alpha_n^*}-\frac{\beta_n^*}{\alpha_n^*}+\frac{\min_{k\in[n]}R_k(A)}{\alpha_n^*\alpha_n}\right)\\ &= 1+\frac{\alpha_n^*}{\alpha_n}\left(-1+o_P(1)+\frac{\min_{k\in[n]}R_k(A)}{\alpha_n^*\alpha_n}\right),\end{aligned}$$

where we have use the fact that $\frac{\check{M}+\beta_n}{\alpha_n^*}=o(1)$, cf. (20). Similarly, we have:

$$\min_{k\in[n]} x_k \leqslant 1+\frac{\alpha_n^*}{\alpha_n}\left(-1+o_P(1)+\frac{\max_{k\in[n]}R_k(A)}{\alpha_n^*\alpha_n}\right) \tag{21}$$

The proof in the centered case follows then from the following lemma:

Lemma 3.4. *Let $R_{k,n}(A)$ be defined as in (19) and recall that $\alpha_n \xrightarrow[n\to+\infty]{} +\infty$, then:*

$$\frac{\max_{k\in[n]}R_{k,n}(A)}{\alpha_n\sqrt{2\log n}} \xrightarrow[n\to\infty]{\mathcal{P}} 0 \qquad \textit{and} \qquad \frac{\min_{k\in[n]}R_{k,n}(A)}{\alpha_n\sqrt{2\log n}} \xrightarrow[n\to\infty]{\mathcal{P}} 0\,.$$

The remaining of the section is devoted to the proof of Lemma 3.4.

Lipschitziannity and Gaussian concentration. We first introduce a truncated version of $R_{k,n}(A)$. Let $\eta\in(0,1)$ and $\varphi:\mathbb{R}^+\to[0,1]$ a smooth function satisfying:

$$\varphi(x)=\begin{cases}1 & \text{if } x\in[0,2\sqrt{2}+\eta]\\ 0 & \text{if } x\geqslant 4\end{cases}, \tag{22}$$

decreasing from 1 to 0 gradually as x goes from $2\sqrt{2}+\eta$ to 4. Let

$$\widetilde{R}_{k,n}(A)=\varphi_n R_{k,n}(A) \qquad \text{where} \qquad \varphi_n=\varphi\left(\left\|\frac{A_n}{\sqrt{n}}\right\|\right). \tag{23}$$

Notice that $\widetilde{R}_k(A)$ differs from $R_k(A)$ if $\varphi_n<1$ which happens with vanishing probability as $\mathbb{P}\{\varphi_n<1\}=\mathbb{P}\left\{s_n>2\sqrt{2}+\eta\right\}\xrightarrow[n\to\infty]{}0$ by Lemma 3.3. The following lemma is a first step towards Gaussian concentration.

Lemma 3.5. *Let $\widetilde{R}_k$ defined by (23) and M an $n\times n$ matrix. Then the function*

$$M\mapsto \widetilde{R}_k(M)=\boldsymbol{e}_k^{\mathsf{T}}\left(\frac{M}{\sqrt{n}}\right)^2\left(I-\frac{M}{\alpha\sqrt{n}}\right)^{-1}\mathbf{1}$$

is K-Lipschitz, i.e.

$$\left|\widetilde{R}_k(M)-\widetilde{R}_k(N)\right|\leqslant K\,\|M-N\|_F \tag{24}$$

where M,N are $n\times n$ matrices, $\|M\|_F=\sqrt{\sum_{ij}|M_{ij}|^2}$ is the Frobenius norm and K a constant independent from k and n.

The second step is to notice that $\widetilde{R}_k(A)$ (where A has Gaussian entries but with off-diagonal pairwise correlations) can be in fact expressed as a Lipschitz function of i.i.d. $\mathcal{N}(0,1)$ entries.

Lemma 3.6. *Consider the linear function $\Gamma:\mathbb{R}^{n\times n}\to\mathbb{R}^{n\times n}$ defined by*

$$\Gamma_{ii}(X)=X_{ii} \qquad \textit{and} \qquad \begin{cases}\Gamma_{ij}(X)=\sqrt{\frac{1+\rho}{2}}X_{ij}+\sqrt{\frac{1-\rho}{2}}X_{ji} & (i<j)\,,\\ \Gamma_{ji}(X)=\sqrt{\frac{1+\rho}{2}}X_{ij}-\sqrt{\frac{1-\rho}{2}}X_{ji} & (i<j)\,.\end{cases}$$

Then

(1) We have $\|\Gamma(X)\|_F \leqslant K_\rho \|X\|_F$ where $K_\rho = 2\sqrt{1+|\rho|}$ hence Γ is K_ρ-Lipschitz.
(2) If matrix $X_n = (X_{ij})$ has i.i.d. $\mathcal{N}(0,1)$ entries, then $A_n = \Gamma(X_n)$ has i.i.d. $\mathcal{N}(0,1)$ entries on and above the diagonal ($i \leqslant j$) and each vector (A_{ij}, A_{ji}) is a standard bivariate Gaussian vector with covariance ρ for $i < j$.

The proof is straightforward and is thus omitted.

A consequence of this lemma is that $\widetilde{R}_k(A) = \widetilde{R}_k(\Gamma(X))$ is $K \times K_\rho$-Lipschitz. Applying Tsirelson-Ibragimov-Sudakov inequality [8, Theorem 5.5] finally yields:

Proposition 3.7. *Let K the Lipschitz constant of Lemma 3.5 and $K_\rho = 2\sqrt{1+|\rho|}$. Then*

$$\mathbb{E} \max_{k\in[n]} \left(\widetilde{R}_k(A) - \mathbb{E}\widetilde{R}_k(A) \right) \leqslant 2\, K_\rho\, K \sqrt{\log n}\,.$$

Details of the proof are similar to those in [7] and are thus omitted.

Remark 3.8. *Notice that $\varphi_n \leqslant 1$ and that $\varphi_n = 0$ if $\|A/\sqrt{n}\| \geqslant 4$. In particular,*

$$\varphi_n \left\| \frac{A}{\sqrt{n}} \right\| \leqslant 4 \qquad \text{and} \qquad \varphi_n \|Q\| \leqslant \frac{1}{1-4\alpha^{-1}} \leqslant 2$$

for n large enough. For the latter estimate, write $Q = \left(I - \frac{A}{\alpha\sqrt{n}}\right)^{-1}$, $Q^{-1}Q = I$ and $Q = I + \frac{A}{\alpha\sqrt{n}} Q$, then apply the triangular inequality.

Proposition 3.9. *The following estimate $\mathbb{E}\widetilde{R}_k(A_n) = \mathcal{O}(1)$ holds true, uniformly for $k \in [n]$.*

Proof. We shall prove that the variables $\widetilde{R}_k$ have a common distribution for $k \in [n]$, which in particular implies that

$$\mathbb{E}\widetilde{R}_k = \mathbb{E}\widetilde{R}_i\,, \quad \forall k, i \in [n] \qquad \text{and} \qquad \mathbb{E}\widetilde{R}_k = \frac{1}{n}\sum_{i\in[n]} \mathbb{E}\widetilde{R}_i\,. \tag{25}$$

Once this fact is established, the proof is straightforward:

$$\left|\mathbb{E}\widetilde{R}_k\right| = \left| \frac{1}{n}\sum_{i\in[n]} \mathbb{E}\widetilde{R}_i \right| = \left| \frac{1}{n}\mathbb{E}\varphi_n \mathbf{1}^\top \left(\frac{A}{\sqrt{n}}\right)^2 Q\mathbf{1} \right| \leqslant \left\| \frac{\mathbf{1}}{\sqrt{n}} \right\|^2 \mathbb{E}\varphi_n \left\| \frac{A}{\sqrt{n}} \right\|^2 \|Q\| = \mathcal{O}(1)\,,$$

where the last equality follows from the arguments developed in Remark 3.8.

Let us now establish (25).

Denote by Δ_σ the matrix associated to the permutation $\sigma : [n] \mapsto [n]$ and defined by

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

Notice in particular that $\Delta_\sigma \boldsymbol{e}_i = \boldsymbol{e}_{\sigma(i)}$, $\Delta_\sigma\Delta_\tau = \Delta_{\sigma\tau}$ for σ, τ two permutations and $\Delta_{\sigma^{-1}} = \Delta_\sigma^\top$. Denote by (ij) the transposition swapping i and j, i.e. $(ij)i = j$, $(ij)j = i$ and $(ij)\ell = \ell$ otherwise. We consider $\breve{A} = \Delta_{(ij)} A \Delta_{(ij)}$, that is $\breve{A}$ is obtained by swapping A's ith and jth column, then the ith and jth row. Observe that A and $\breve{A}$ have the same distribution and so is the case for $R_k(A)$ and $R_k(\breve{A})$.

We have $\Delta_{(ij)}^2 = I_n$, implying that $\breve{A}^k = \Delta_{(ij)} A^k \Delta_{(ij)}$ and then

$$\begin{aligned} R_i(\breve{A}) &= \boldsymbol{e}_i^\top \sum_{k\geqslant 2} \left(\frac{\breve{A}}{\alpha\sqrt{n}}\right)^k \mathbf{1} = \boldsymbol{e}_i^\top \Delta_{(ij)} \sum_{k\geqslant 2} \left(\frac{A}{\alpha\sqrt{n}}\right)^k \Delta_{(ij)}\mathbf{1} = \boldsymbol{e}_j^\top \sum_{k\geqslant 2} \left(\frac{A}{\alpha\sqrt{n}}\right)^k \mathbf{1} \\ &= R_j(A)\,. \end{aligned}$$

This proves that $R_i(A), R_i(\breve{A}), R_j(A)$ have the same law, hence the same expectation. Eq.(25) is established, which concludes the proof. □

We are now in position to prove Lemma 3.4.

Proof of lemma 3.4. Recall that $\mathbb{E}\widetilde{R}_k(A) = \mathbb{E}\widetilde{R}_1$. Since $\max_{k\in[n]} \widetilde{R}_k(A) - \widetilde{R}_1(A) \geqslant 0$, Markov inequality yields:

$$\begin{aligned} \mathbb{P}\left\{\frac{\max_{k\in[n]} \widetilde{R}_k(A) - \widetilde{R}_1(A)}{\alpha\sqrt{2\log n}} \geqslant \varepsilon\right\} &\leqslant \frac{\mathbb{E}\left(\max_{k\in[n]} \widetilde{R}_k(A) - \widetilde{R}_1(A)\right)}{\varepsilon\alpha\sqrt{2\log n}}\,, \\ &= \frac{\mathbb{E}\left(\max_{k\in[n]} \widetilde{R}_k\,(A) - \mathbb{E}\widetilde{R}_k(A)\right)}{\varepsilon\alpha\sqrt{2\log n}}\,, \\ &= \frac{\mathbb{E}\left(\max_{k\in[n]} \left(\widetilde{R}_k\,(A) - \mathbb{E}\widetilde{R}_k(A)\right)\right)}{\varepsilon\alpha\sqrt{2\log n}}\,, \\ &\leqslant \frac{\sqrt{2}K \times K_\rho}{\varepsilon\alpha}\,, \end{aligned}$$

where the last inequality follows from Proposition 3.7.
This implies that

$$\frac{\max_{k\in[n]} \widetilde{R}_k\,(A) - \widetilde{R}_1\,(A)}{\alpha\sqrt{2\log n}} \xrightarrow[n\to\infty]{\mathcal{P}} 0\,.$$

It remains to prove that

$$\frac{\widetilde{R}_1\,(A)}{\alpha\sqrt{2\log n}} \xrightarrow[n\to\infty]{\mathcal{P}} 0 \qquad \text{and} \qquad \frac{\max_{k\in[n]} R_k\,(A)}{\alpha\sqrt{2\log n}} \xrightarrow[n\to\infty]{\mathcal{P}} 0\,.$$

The arguments are similar to those in [7, Section 2.3]. Proof of the second assertion of Lemma 3.4 can be done similarly. This concludes the proof. □

3.3. **Proof of Theorem 2.1 - the non centered case.** Recall that $\alpha \to \infty$ as $n \to \infty$. Denote by $\boldsymbol{u}_n = \frac{1}{\sqrt{n}}\mathbf{1}_n$ and notice that the spectrum of $I_n - \mu \boldsymbol{u}_n \boldsymbol{u}_n^\top$ is $\{1-\mu, 1\}$, the eigenvalue 1 with multiplicity $n-1$. Notice in particular that if $\mu \neq 1$, then $I - \mu\boldsymbol{u}\boldsymbol{u}^\top$ is invertible. So is (eventually) $I - \frac{A}{\alpha\sqrt{n}} - \mu\boldsymbol{u}\boldsymbol{u}^\top$ as $\|A/(\alpha\sqrt{n})\| \to 0$ a.s. We shall also rely on the fact that $\|Q - I\| \xrightarrow[n\to\infty]{} 0$ a.s. As a consequence,

$$\boldsymbol{u}^\top Q \boldsymbol{u} \xrightarrow[n\to\infty]{a.s.} 1\,.$$

Denote by $\tilde{\boldsymbol{x}}$ and $\boldsymbol{x}$ the vectors solutions of the equations:

$$\tilde{\boldsymbol{x}} = \mathbf{1} + B\tilde{\boldsymbol{x}} = \mathbf{1} + \left(\frac{A}{\alpha\sqrt{n}} + \mu\boldsymbol{u}\boldsymbol{u}^\top\right)\tilde{\boldsymbol{x}} \qquad \text{and} \qquad \boldsymbol{x} = \mathbf{1} + \frac{A}{\alpha\sqrt{n}}\boldsymbol{x}\,.$$

The following representations hold:

$$\tilde{\boldsymbol{x}} = (I - B)^{-1}\,\mathbf{1} \qquad \text{and} \qquad \boldsymbol{x} = \left(I - \frac{A}{\alpha\sqrt{n}}\right)^{-1}\mathbf{1}\,.$$

Recall that $Q = (I - A/(\alpha\sqrt{n}))^{-1}$. By rank one perturbation identity (Woodbury), we have:

$$(I - B)^{-1} = Q + \frac{Q\boldsymbol{u}\boldsymbol{u}^{\mathsf{T}}Q}{1 - \mu\boldsymbol{u}^{\mathsf{T}}Q\boldsymbol{u}}$$

and

$$\tilde{\boldsymbol{x}} \quad = \quad \frac{Q\mathbf{1}(1 - \mu\boldsymbol{u}^{\mathsf{T}}Q\boldsymbol{u}) + \mu Q\boldsymbol{u}\boldsymbol{u}^{\mathsf{T}}Q\mathbf{1}}{1 - \mu\boldsymbol{u}\mathsf{T}Q\boldsymbol{u}} \quad = \quad \frac{\boldsymbol{x}}{1 - \mu\boldsymbol{u}^{\mathsf{T}}Q\boldsymbol{u}}\,.$$

If $\mu < 1$ and $\alpha \geqslant (1+\varepsilon)\alpha^*$ then eventually, $\tilde{\boldsymbol{x}}$ has positive components. This is no longer the case if $\mu > 1$ or $\alpha \leqslant (1-\varepsilon)\alpha^*$. This concludes the proof of Theorem 2.1.

4. Stability: Proof of Proposition 2.3

Proof. We have

$$I - B + I - B^{\mathsf{T}} = 2I - (B + B^{\mathsf{T}}) = 2I - \left(\frac{A + A^{\mathsf{T}}}{\alpha\sqrt{n}} + \frac{2\mu}{n}\mathbf{1}\mathbf{1}^{\mathsf{T}}\right)\,.$$

We will rely on the following condition:

$$2I - (B + B^{\mathsf{T}}) \text{ is positive definite} \quad \Leftrightarrow \quad \lambda_{\max}(B + B^{\mathsf{T}}) < 2\,. \tag{26}$$

Notice that $(A + A^{\mathsf{T}})/\alpha$ is a symmetric matrix with independent $\mathcal{N}(0, 2(1+\rho)/\alpha^2)$ entries above the diagonal (the diagonal entries have a different distribution from the off-diagonal entries, with no asymptotic effect). In this case, it is well known that the largest eigenvalue of the normalized matrix (or equivalently its spectral norm since the matrix is symmetric) almost surely converges to the right edge of the support of the semi-circle law (see [5, Theorem 5.2]):

$$\lambda_{\max}\left(\frac{A + A^{\mathsf{T}}}{\alpha\sqrt{n}}\right) \xrightarrow[n\to\infty]{a.s.} \frac{2\sqrt{2(1+\rho)}}{\alpha}\,. \tag{27}$$

Suppose that $(\rho, \alpha, \mu) \in \mathcal{A}$. Notice that in this case,

$$\frac{\sqrt{1+\rho}}{\alpha\sqrt{2}} \quad < \quad \frac{1}{2} \quad < \quad \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2(1+\rho)}{\alpha^2}}\,.$$

We consider three subcases

(i) $\mu = 0$,
(ii) $\mu \leqslant \frac{\sqrt{1+\rho}}{\alpha\sqrt{2}}$,
(iii) $\mu \in \left(\frac{\sqrt{1+\rho}}{\alpha\sqrt{2}}\,,\, \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2(1+\rho)}{\alpha^2}}\right)$.

In the centered case (i), condition (26) asymptotically occurs whenever $\alpha > \sqrt{2(1+\rho)}$.

Before studying subcases (ii) and (iii), we recall a result on small rank perturbations of large random matrices.

Notice that the rank-one perturbation matrix $P = \frac{2\mu}{n}\mathbf{1}\mathbf{1}^{\mathsf{T}}$ admits a unique non zero eigenvalue 2μ. Denote by $\check{A} = \frac{A + A^{\mathsf{T}}}{\alpha\sqrt{n}}$. We are concerned with the top eigenvalue of the symmetric matrix $\check{A}+P$. Based on a result by Capitaine et al. [10, Theorem 2.1], we have:

$$\lambda_{\max}(\check{A}+P) \xrightarrow[n\to\infty]{a.s} \begin{cases} 2\mu + \frac{1+\rho}{\alpha^2\mu} & \text{if } \mu > \frac{\sqrt{1+\rho}}{\sqrt{2}\alpha}\,, \\ \frac{2\sqrt{2(1+\rho)}}{\alpha} & \text{else.} \end{cases}$$

Consider now subcase (ii), then $\lambda_{\max}(\check{A}+P) \xrightarrow[n\to\infty]{a.s.} \frac{2\sqrt{2(1+\rho)}}{\alpha}$, which is strictly lower than 2 since $(\rho,\alpha,\mu)\in\mathcal{A}$. Hence $\lambda_{\max}(\check{A}+P)$ is eventually strictly lower than 2 in this case.

We finally consider subcase (iii). In this case,

$$\lambda_{\max}(\check{A}+P) \xrightarrow[n\to\infty]{a.s.} 2\mu + \frac{1+\rho}{\alpha^2\mu}\,.$$

We shall prove that $2\mu + \frac{1+\rho}{\alpha^2\mu} < 2$ or equivalently

$$2\alpha^2\mu^2 - 2\alpha^2\mu + 1 + \rho < 0\,. \tag{28}$$

An elementary study of the polynomial $\Omega(X) = 2\alpha^2X^2 - 2\alpha^2X + 1 + \rho$ yields that Ω's discriminant is positive if $\alpha > \sqrt{2(1+\rho)}$ and Ω's roots are given by

$$\Omega(\mu^\pm) = 0 \quad \Leftrightarrow \quad \mu^\pm = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{2(1+\rho)}{\alpha^2}}\,.$$

Also remark that $\Omega\left(\frac{\sqrt{1+\rho}}{\alpha\sqrt{2}}\right) < 0$, so that $\frac{\sqrt{1+\rho}}{\alpha\sqrt{2}} \in (\mu^-,\mu^+)$. In particular condition (28) is fulfilled for $\mu \in \left(\frac{\sqrt{1+\rho}}{\alpha\sqrt{2}},\mu^+\right)$, which is precisely subcase (iii). Hence a.s. $\limsup_{n\to\infty}\lambda_{\max}(\check{A}+P) < 2$. We can then rely on Theorem 2.2 to conclude. $\square$

References

[1] I. Akjouj and J. Najim. Feasibility of sparse large lotka-volterra ecosystems. *arXiv preprint arXiv:2111.11247*, 2021.

[2] S. Allesina and S. Tang. Stability criteria for complex ecosystems. *Nature*, 483(7388):205, 2012.

[3] S. Allesina and S. Tang. The stability–complexity relationship at age 40: a random matrix perspective. *Population Ecology*, 57(1):63–75, 2015.

[4] Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010.

[5] Z. D. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010.

[6] M. Barbier and J.-F. Arnoldi. The cavity method for community ecology. preprint, Ecology, June 2017.

[7] P. Bizeul and J. Najim. Positive solutions for large random linear systems. *Proceedings of the American Mathematical Society*, 149(6):2333–2348, 2021.

[8] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

[9] G. Bunin. Ecological communities with lotka-volterra dynamics. *Physical Review E*, 95(4):042414, 2017.

[10] M. Capitaine, C. Donati-Martin, and D. Féral. The largest eigenvalues of finite rank deformation of large Wigner matrices: Convergence and nonuniversality of the fluctuations. *The Annals of Probability*, 37(1), Jan. 2009.

[11] M. Clenet. Feasibility in a large lotka-volterra system with pairwise correlated interactions. *https://github.com/maxime-clenet/Feasibility-in-a-large-Lotka-Volterra-system-with-pairwise-correlated-interactions*, 2022.

[12] M. Clenet, F. Massol, and J. Najim. Equilibrium and surviving species in a large lotka-volterra system of differential equations. *Submitted. arxiv:2205.00735*, 2022.
[13] S. Diederich and M. Opper. Replicators with random interactions: A solvable model. *Physical Review A*, 39(8):4333–4336, Apr. 1989.
[14] M. Dougoud, L. Vinckenbosch, R. P. Rohr, L.-F. Bersier, and C. Mazza. The feasibility of equilibria in large ecosystems: A primary but neglected concept in the complexity-stability debate. *PLoS computational biology*, 14(2):e1005988, 2018.
[15] T. Galla. Dynamically evolved community size and stability of random Lotka-Volterra ecosystems. *EPL (Europhysics Letters)*, 123(4):48004, Sept. 2018. arXiv: 1808.06660.
[16] M. R. Gardner and W. R. Ashby. Connectance of large dynamic (cybernetic) systems: critical values for stability. *Nature*, 228(5273):784, 1970.
[17] T. Gibbs, J. Grilli, T. Rogers, and S. Allesina. Effect of population abundances on the stability of large random ecosystems. *Physical Review E*, 98(2):022410, 2018.
[18] V. Girko. Elliptic law. *Theory of Probability & Its Applications*, 30(4):677–690, 1986.
[19] V. Girko. The elliptic law: ten years later i. 1995.
[20] K. Gopalsamy. Global asymptotic stability in volterra's population systems. *JOURNAL of Mathematical Biology*, 19(2):157–168, 1984.
[21] R. Hering. Oscillations in lotka-volterra systems of chemical reactions. *Journal of mathematical chemistry*, 5(2):197–202, 1990.
[22] J. Hofbauer and K. Sigmund. *Evolutionary games and population dynamics*. Cambridge university press, 1998.
[23] S. Janson. *Gaussian hilbert spaces*. Number 129. Cambridge university press, 1997.
[24] K. Kiss and S. Kovács. Qualitative behavior of n-dimensional ratio-dependent predator–prey systems. *Applied Mathematics and Computation*, 199(2):535–546, 2008.
[25] R. Latała. Some estimates of norms of random matrices. *Proceedings of the American Mathematical Society*, 133(5):1273–1282, 2005.
[26] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer Science & Business Media, 2012.
[27] R. May. Will a large complex system be stable? *Nature*, 238(5364):413, 1972.
[28] A. Naumov. Elliptic law for real random matrices. *arXiv preprint arXiv:1201.1639*, 2012.
[29] H. H. Nguyen and S. O'Rourke. The elliptic law. *International Mathematics Research Notices*, 2015(17):7620–7689, 2015.
[30] M. Opper and S. Diederich. Phase transition and 1/f noise in a game dynamical model. *Physical Review Letters*, 69(10):1616–1619, Sept. 1992. Publisher: American Physical Society.
[31] S. O'Rourke and D. Renfrew. Low rank perturbations of large elliptic random matrices. *Electronic Journal of Probability*, 19(none):1 – 65, 2014.
[32] L. Stone. The feasibility and stability of large complex biological networks: a random matrix approach. *Scientific reports*, 8(1):8246, 2018.
[33] Y. Takeuchi. *Global dynamical properties of Lotka-Volterra systems*. World Scientific, 1996.

Maxime Clenet,
Laboratoire d'Informatique Gaspard Monge, UMR 8049
CNRS & Université Gustave Eiffel
5, Boulevard Descartes,
Champs sur Marne, 77454 Marne-la-Vallée Cedex 2, France
e-mail: `maxime.clenet@univ-eiffel.fr`

Hafedh El Ferchichi,
Ecole Normale Supérieure Paris Saclay
e-mail: `hafedh.el_ferchichi@ens-paris-saclay.fr`

Jamal Najim,
Laboratoire d'Informatique Gaspard Monge, UMR 8049
CNRS & Université Gustave Eiffel
5, Boulevard Descartes,

Champs sur Marne, 77454 Marne-la-Vallée Cedex 2, France
e-mail: najim@univ-mlv.fr

Appendix A. Proof of Theorem 2.1: adaptations to the case of a covariance profile

In this section, we provide the arguments to prove Theorem 2.1 in the case where matrix B follows the model (4), i.e.

$$B = \frac{\tilde{A}_n}{\alpha_n \sqrt{n}} + \frac{\mu}{n}\mathbf{1}_n\mathbf{1}_n^{\mathsf{T}}\,,$$

where $\tilde{A}_n$'s entries are i.i.d. $\mathcal{N}(0,1)$ on and above the diagonal ($i \leqslant j$), and $(\tilde{A}_{ij}, \tilde{A}_{ji})$ is a standard bivariate Gaussian vector ($i<j$) with covariance $\operatorname{cov}(\tilde{A}_{ij}, \tilde{A}_{ji}) = \rho_{ij}$, and independent from the remaining random variables.

There are essentially 3 issues to resolve, to fully adapt the proof developed in Section 3 to the covariance profile case:

(1) The decomposition (16) yields $\frac{\tilde{A}_n}{\sqrt{n}} = W_n^1 + \boldsymbol{i} W_n^2$, where W_n^1, W_n^2 are Hermitian matrices with
$$\operatorname{var}\left([W_n^1]_{ij}\right) = \frac{1+\rho_{ij}}{2n} \qquad \text{and} \qquad \operatorname{var}\left([W_n^2]_{ij}\right) = \frac{1-\rho_{ij}}{2n}\,.$$
Since W_n^1, W_n^2 are no longer Wigner matrices, but rather matrices with a variance profile, an extra argument is needed to obtain an almost-sure upper bound for $\limsup_n \|W_n^1\| + \limsup_n \|W_n^2\|$.

(2) The Lipschitz property for $\widetilde{R}_{k,n}(\tilde{A}_n)$. Essentially, we need the counterpart of Lemma 3.6 to the context of a covariance profile.

(3) The control of the term $\mathbb{E}\,\widetilde{R}_{k,n}(\tilde{A})$.

A.1. **Proof of issue 1: Control of the spectral norm of a Hermitian matrix with a variance profile.** Applying Latała's theorem [25], we easily show that

$$\mathbb{E}\|W_n^1\| + \mathbb{E}\|W_n^2\| \leqslant C$$

where C is a constant independent from n.

Now write

$$W_n^1 = \frac{\Upsilon_n \circ X_n}{\sqrt{n}} \qquad \text{where} \qquad \Upsilon_n = (\Upsilon_{ij})\ ,\quad \Upsilon_{ij} = \sqrt{\frac{1+\rho_{ij}}{2}},$$

matrix $X_n = (X_{ij})$ is a Wigner matrix with i.i.d. $\mathcal{N}(0,1)$ entries on and above the diagonal, and $\circ$ stands for the Hadamard product, i.e. $\Upsilon_n \circ X_n = (\Upsilon_{ij} X_{ij})$. Notice that $\sqrt{n} W_n^1$ is 1-Lipschitz with respect to the Frobenius norm

$$\|X_n\|_{\text{Frob}} = \sqrt{\sum_{ij} |X_{ij}|^2}\ .$$

Hence by Gaussian concentration, we have

$$\mathbb{P}\left\{\left|\sqrt{n}\|W_n^1\| - \sqrt{n}\mathbb{E}\|W_n^1\|\right| > \delta\right\} \leqslant 2e^{-\frac{\delta^2}{2}}\ .$$

Taking $\delta = \varepsilon\sqrt{n}$, we obtain

$$\mathbb{P}\left\{\left|\|W_n^1\| - \mathbb{E}\|W_n^1\|\right| > \varepsilon\right\} \leqslant 2e^{-\frac{n\varepsilon^2}{2}}\ .$$

The same holds for W_n^2, hence the upper control:

$$\limsup_n \left(\|W_n^1\| + \|W_n^2\|\right) \quad \leqslant \quad \limsup_n \left(\mathbb{E}\|W_n^1\| + \mathbb{E}\|W_n^2\|\right) \quad \leqslant \quad C$$

almost surely. It remains to replace the truncation function φ in (22) by the smooth function

$$\psi(x) = \begin{cases} 1 & \text{if } x \leqslant C + \eta, \\ 0 & \text{else.} \end{cases}$$

to proceed.

A.2. **Proof of issue 2: $\widetilde{R}_k(\tilde{A})$ is a Lipschitz function of Gaussian i.i.d. random variables.** It suffices to replace function Γ in Lemma 3.6 by

$$\widetilde{\Gamma} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$$

where

$$\widetilde{\Gamma}_{ii}(X) = X_{ii} \qquad \text{and} \qquad \begin{cases} \widetilde{\Gamma}_{ij}(X) = \sqrt{\frac{1+\rho_{ij}}{2}} X_{ij} + \sqrt{\frac{1-\rho_{ij}}{2}} X_{ji} & (i<j)\ , \\ \widetilde{\Gamma}_{ji}(X) = \sqrt{\frac{1+\rho_{ij}}{2}} X_{ij} - \sqrt{\frac{1-\rho_{ij}}{2}} X_{ji} & (i<j)\ . \end{cases}$$

and to modify accordingly the Lipschitz constant by $\widetilde{K} = 2\sqrt{2} \geqslant 2\sqrt{1 + \max_{ij}|\rho_{ij}|}$.

A.3. **Proof of issue 3: Magnitude of $\mathbb{E}\,\widetilde{R}_{k,n}(\tilde{A}_n)$.** To address this issue, we provide a quick argument which relies on Isserlis' theorem also called Wick's formula (see [23, Th. 1.28]), highly dependent on the Gaussiannity of the entries.

Theorem A.1 (Isserlis Theorem). *if $(X_1, \cdots, X_n)$ is a centered normal vector, then*

$$\mathbb{E}(X_1 X_2 \cdots X_n) = \sum_{\Pi} \prod_{\{i,j\}\in\Pi} \mathbb{E}(X_i X_j) \tag{29}$$

where the sum is over all the partitions Π of $[n]$ into pairs $\{i,j\}$, and the product over all the pairs contained in Π.

Recall that:

$$\mathbb{E}\widetilde{R}_k(\tilde{A}_n) = \sum_{\ell\geqslant 2} \mathbb{E}\left[\frac{\boldsymbol{e}_k^{\mathsf{T}}}{\alpha^{\ell-2}}\left(\frac{\tilde{A}}{\sqrt{n}}\right)^{\ell} \mathbf{1}\right] = \sum_{\ell\geqslant 2} \frac{\mathbb{E}\boldsymbol{e}_k^{\mathsf{T}}\tilde{A}^{\ell}\mathbf{1}}{\alpha^{\ell-2} n^{\frac{\ell}{2}}} =: \sum_{\ell\geqslant 2} \frac{C_\ell}{\alpha^{\ell-2} n^{\frac{\ell}{2}}}\ .$$

Consider a matrix $\overline{A}_n$ where the pairwise covariance $\mathrm{cov}(\overline{A}_{ij}, \overline{A}_{ji}) = 1$. Denote by $\overline{C}_\ell = \mathbb{E}\boldsymbol{e}_k^{\mathsf{T}}\overline{A}^{\ell}\mathbf{1}$. We will show that each quantity $|C_\ell|$ is bounded by $\overline{C}_\ell$. Notice that:

$$C_\ell = \sum_{i_1,\dots,i_{\ell+1}} \mathbb{E}(\tilde{A}_{i_1 i_2}\tilde{A}_{i_2 i_3}\dots\tilde{A}_{i_\ell, i_{\ell+1}})\ . \tag{30}$$

by Isserlis' theorem, we have:

$$\begin{aligned} \left|\mathbb{E}(\tilde{A}_{i_1,i_2}\tilde{A}_{i_2,i_3}\dots\tilde{A}_{i_\ell,i_{\ell+1}})\right| \quad &\leqslant \quad \left|\sum_{\Pi}\prod_{\{j,k\}\in\Pi}\mathbb{E}(\overline{A}_{i_j i_{j+1}}\overline{A}_{i_k i_{k+1}})\right| \\ &\leqslant \quad \sum_{\Pi}\prod_{\{j,k\}\in\Pi}\mathbb{E}(\overline{A}_{i_j i_{j+1}}\overline{A}_{i_k i_{k+1}}) \quad = \quad \mathbb{E}\left(\overline{A}_{i_1 i_2}\cdots\overline{A}_{i_\ell i_{\ell+1}}\right)\ . \end{aligned}$$

From this, we deduce that $|C_\ell| \leqslant \overline{C}_\ell$, hence $|\mathbb{E}\widetilde{R}_k(\tilde{A})| \leqslant \mathbb{E}\widetilde{R}_k(\overline{A})$. This gives the desired bound since $\mathbb{E}\widetilde{R}_k(\overline{A}) = \mathcal{O}(1)$.