Elliptic Approximate Message Passing and an application to theoretical ecology

Mohammed-Younes Gueddari, Walid Hachem, Jamal Najim

ELLIPTIC APPROXIMATE MESSAGE PASSING AND AN APPLICATION TO THEORETICAL ECOLOGY

MOHAMMED-YOUNES GUEDDARI, WALID HACHEM, JAMAL NAJIM

Abstract. Approximate Message Passing (AMP) algorithms have recently gathered significant attention across disciplines such as statistical physics, machine learning, and communication systems. This study aims to extend AMP algorithms to non-symmetric (elliptic) matrices, motivated by analyzing equilibrium properties in ecological systems featuring elliptic interaction matrices.

In this article, we provide the general form of an AMP algorithm associated to a random elliptic matrix, the main change lying in a modification of the corrective (Onsager) term. In order to establish the statistical properties of this algorithm, we use and prove a generalized form of Bolthausen conditioning argument, pivotal to proceed by a Gaussian-based induction.

We finally address the initial motivating question from theoretical ecology. Large foodwebs are often described by Lotka-Volterra systems of coupled differential equations, where the interaction matrix is elliptic random. In this context, we design an AMP algorithm to analyze the statistical properties of the equilibrium point in a high-dimensional regime. We rigorously recover the results established by Bunin [Bun17] and Galla [Gal18] who used techniques from theoretical physics, and extend them with the help of *Propagation of chaos* type arguments.

1. Introduction

Approximate Message Passing (AMP) is a class of versatile and configurable iterative algorithms. The output is a sequence of high-dimensional $\mathbb{R}^n$-valued random vectors $(\boldsymbol{u}^k)_{k\geq 1}$ based on $n\times n$ (usually symmetric) random matrices, see (1). An important feature of AMP is a precise description of the $(\boldsymbol{u}^k)$'s statistical properties as n goes to infinity, mainly via the so-called Density Evolution (DE) equations (3).

Initially used in statistics for solving compressed sensing and sparse signal recovery problems [DMM09, BM11], the AMP algorithms have found numerous applications in the fields of high-dimensional estimation [DAM17, LM19], communication theory [BK17, RGV17], or statistical physics [Mon21], and have undergone extensive developments that have widened their spectrum of applications.

The goal of this article is to extend the AMP procedure to a non-Hermitian setting and to consider large elliptic random matrices. This new setting is mainly motivated by the study of equilibria in large random Lotka-Volterra systems of differential equations, a popular model in theoretical ecology. Interestingly, modifying the matrix nature in the AMP algorithm changes the iteration equation (modification of the *Onsager term*), necessitates extra mathematical developments but does not modify the DE equations.

A primer on Approximate Message Passing. Recall the definition of the Gaussian Orthogonal Ensemble (GOE), a $n\times n$ random matrix with representation $(X+X^\top)/\sqrt{2}$ where $X\in\mathbb{R}^{n\times n}$ has n^2 independent $\mathcal{N}(0,1)$ elements. Consider the following AMP iterative algorithm :

$$\begin{cases} \boldsymbol{u}^1 = A_n\, h_0\left(\boldsymbol{u}^0, \boldsymbol{b}\right) \\ \boldsymbol{u}^{k+1} = A_n\, h_k\left(\boldsymbol{u}^k, \boldsymbol{b}\right) - \left\langle \partial_1 h_k\left(\boldsymbol{u}^k, \boldsymbol{b}\right)\right\rangle_n h_{k-1}\left(\boldsymbol{u}^{k-1}, \boldsymbol{b}\right), \quad \text{for } k\geq 1 \end{cases} \tag{1}$$

where $\boldsymbol{u}^0\in\mathbb{R}^n$ is an initialization vector that can be either random or deterministic, $\boldsymbol{b}=(b_i)_{i\in[n]}$ is a parameter vector and $\boldsymbol{u}^k=(u_i^k)_{i\in[n]}\in\mathbb{R}^n$ is the iterate at step k. Matrix A_n is $n\times n$ random such that $\sqrt{n}A_n$ is drawn from GOE. Functions $h_k:\mathbb{R}^2\to\mathbb{R}$ ($k\geq 0$ with the convention that

Date: February 13, 2024

CNRS, Laboratoire d'informatique Gaspard Monge (LIGM / UMR 8049), Université Gustave Eiffel, France.

$h_{-1} = 0$) are the so-called activation functions and applied componentwise to vectors $\boldsymbol{u}^k = (u_i^k)$ and $\boldsymbol{b} = (b_i)$:

$$h_k(\boldsymbol{u}^k, \boldsymbol{b}) = \big(h_k(u_i^k, b_i)\big) \in \mathbb{R}^n\,.$$

These functions are assumed to be differentiable with respect to the first parameter. We denote the partial derivative $\frac{\partial h}{\partial u}(u,b)$ by $\partial_1 h(u,b)$, and introduce the notation:

$$\big\langle \partial_1 h_k\left(\boldsymbol{u}^k, \boldsymbol{b}\right)\big\rangle_n = \frac{1}{n}\sum_{i\in[n]} \partial_1 h_k\left(u_i^k, b_i\right) \in \mathbb{R}\,.$$

The Onsager term. In the AMP literature, the so-called **Onsager term** defined by

$$-\big\langle \partial_1 h_k\left(\boldsymbol{u}^k, \boldsymbol{b}\right)\big\rangle_n\, h_{k-1}\left(\boldsymbol{u}^{k-1}, \boldsymbol{b}\right)$$

plays a pivotal role to describe the (high dimensional) statistical properties of $(\boldsymbol{u}^0, \cdots, \boldsymbol{u}^k)$. It is designed to asymptotically remove (as n goes to infinity) the non-Gaussian component from $A_n\, h_k\left(\boldsymbol{u}^k, \boldsymbol{b}\right)$.

The joint empirical distribution. The joint empirical distribution of $(\boldsymbol{u}^1, \cdots, \boldsymbol{u}^k)$ is defined as

$$\mu^{\boldsymbol{u}^1,\cdots,\boldsymbol{u}^k} = \frac{1}{n}\sum_{i\in[n]} \delta_{(u_i^1,\cdots,u_i^k)}\,,$$

where $\delta_{(u_1,\cdots u_k)}$ is the Dirac distribution at point $(u_1, \cdots, u_k)$. The techniques developed within the scope of AMP enable to describe the limit of this joint empirical distribution as $n \to \infty$, which turns out to be the distribution of a centered k-dimensional Gaussian vector whose covariance matrix is defined recursively by the **Density Evolution equations**.

Density Evolution (DE) equations. These equations recursively characterize a family of covariance matrices $R^k \in \mathbb{R}^{k\times k}$ with $k \geq 1$ defined as follows.

- Initialization: Let $(\bar{u}, \bar{b})$ be a random vector in $\mathbb{R}^2$. Set

$$R^1 = \mathbb{E}\left[h_0^2\left(\bar{u}, \bar{b}\right)\right] \tag{2}$$

  and let $Z_1 \sim \mathcal{N}(0, R^1)$ be a random variable independent from $(\bar{u}, \bar{b})$.
- Recursion: Suppose that R^{k-1} is given. Let $(Z_1, \cdots, Z_{k-1}) \sim \mathcal{N}_{k-1}(0, R^{k-1})$ be a vector independent from $(\bar{u}, \bar{b})$. Let $R^k = (R_{ij}^k)$ be a $k \times k$ matrix defined by

$$R_{ij}^k = \begin{cases} R^1 & \text{if } i = 1, j = 1 \\ \mathbb{E}\left[h_{i-1}\left(Z_{i-1}, \bar{b}\right) h_{j-1}\left(Z_{j-1}, \bar{b}\right)\right] & \text{if } i \geq 1, j \geq 1 \end{cases} \tag{3}$$

  with the convention that $Z_0 \triangleq \bar{u}$.

Notice that R^{k-1} represents the upper-left corner of matrix R^k. Thus we can also define an operator R as follows which will encode all matrices R^k:

$$\begin{array}{rccccl} R & : & \mathbb{N}^* \times \mathbb{N}^* & \to & \mathbb{R}_+ & \\ & & (i,j) & \mapsto & R_{ij}^k & \text{where } k \geq i, j, \end{array} \tag{4}$$

We will simply denote $R(i,j)$ by R_{ij} for all $(i,j) \in \mathbb{N}^* \times \mathbb{N}^*$.

We now (informally) state the main result of the AMP for a GOE matrix:

Theorem ([BM11, FVRS21]). *Let A_n be a $n \times n$ matrix such that $\sqrt{n}A_n$ is drawn from the GOE. Let $\boldsymbol{u}^0, \boldsymbol{b} \in \mathbb{R}^n$ independent from A_n, and $(\boldsymbol{u}^\ell)_{1\leq \ell \leq k}$ be defined by (1). Suppose that $\mu^{\boldsymbol{u}^0,\boldsymbol{b}} \xrightarrow[n\to\infty]{} \mathcal{L}(\bar{u}, \bar{b})$ and let R^k be defined by the DE equations (2)-(3). Then*

$$\mu^{\boldsymbol{u}^1,\cdots,\boldsymbol{u}^k} \xrightarrow[n\to\infty]{} \mathcal{L}(Z_1, \cdots, Z_k) \qquad \textit{where} \quad (Z_1, \cdots, Z_k) \sim \mathcal{N}_k(\mathbf{0}, R^k)\,.$$

The nature of the convergence of measures $\mu^{\boldsymbol{u}^0,\boldsymbol{b}}$ and $\mu^{\boldsymbol{u}^1,\cdots,\boldsymbol{u}^k}$ will be specified for the main theorem of the article.

About the literature. Numerous studies have extended AMP algorithms to more complex and general scenarios [JM13, BLM15]. For instance, Vector AMP [RSF18] broadens AMP to handle vector observations, considering $\boldsymbol{u}^k$ as a matrix rather than a vector. This adaptation is suited for multi-channel and multi-dimensional signal processing tasks. Fan [Fan21] generalizes AMP algorithms to encompass a wide range of matrices, particularly rotationally invariant ones, Dudeja et al. [DMLS23] explore universality properties of AMP, etc.

Another important generalization known as "Asymmetric AMP" is discussed in [FVRS21, section 2.2]. Given a rectangular matrix A with i.i.d. Gaussian entries, the form of this AMP algorithm can be summarized as the following:

$$\begin{cases} \boldsymbol{v}^k & = A\ g_k\left(\boldsymbol{u}^k, \alpha\right) - \beta_k h_{k-1}\left(\boldsymbol{v}^{k-1}, \boldsymbol{b}\right) \\ \boldsymbol{u}^{k+1} & = A^\top h_k\left(\boldsymbol{v}^k, \boldsymbol{b}\right) - b_k g_k\left(\boldsymbol{u}^k, \alpha\right) \end{cases},$$

which is a two-step algorithm that involves both matrices A and $A^\top$.

None of these extensions cover our model of interest.

Non-Hermitian AMP. In the sequel we will consider an AMP algorithm based on an elliptic matrix instead of a GOE matrix. For the sake of our application, we shall also enable multiple parameter vectors $(\boldsymbol{b}^1, \cdots, \boldsymbol{b}^p)$ instead of a single one (p fixed). These vectors will be stacked into a $n \times p$ matrix

$$B = (\boldsymbol{b}^1, \cdots, \boldsymbol{b}^p) \tag{5}$$

and for $h : \mathbb{R}^{k+1} \to \mathbb{R}$, $\boldsymbol{u} \in \mathbb{R}^n$, $h(\boldsymbol{u}, B)$ will denote the vector

$$h(\boldsymbol{u}, B) = \left(h(u_i, B_{i1}, \cdots, B_{ip}) \right)_{i \in [n]} .$$

Definition 1.1 (Gaussian elliptic matrix model)**.** *A random matrix $M_n = (M_{ij}) \in \mathbb{R}^{n \times n}$ is said to follow the Gaussian elliptic distribution with parameter $\rho \in [-1, 1]$ if the entries' distributions are given by*

$$M_{ii} \sim \mathcal{N}(0, 1+\rho) \quad \textit{for } i \in [n] \qquad \textit{and} \qquad \begin{bmatrix} M_{ij} \\ M_{ji} \end{bmatrix} \sim \mathcal{N}_2\left(\boldsymbol{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right) \quad \textit{for } i < j\,.$$

Moreover, all the elements of the following set are independent:

$$\left\{ M_{ii}\,,\ i \in [n] \right\} \cup \left\{ (M_{ij}, M_{ji})\,,\ i, j \in [n]\,, i < j \right\}.$$

We write $M_n \sim \text{Elliptic}(n, \rho)$ for such matrices. We will also call a *normalized (Gaussian) elliptic matrix* $A_n \in \mathbb{R}^{n \times n}$ a matrix that verifies $\sqrt{n} A_n \sim \text{Elliptic}(n, \rho)$.

Notice that a normalized elliptic matrix A_n is no longer symmetric. Its spectral distribution has been thoroughly studied [Gir86, Nau12, OR14] and it is well-known that it almost surely converges as $n \to \infty$ to the uniform law over the compact set bounded by the ellipse defined for $|\rho| < 1$ by

$$\mathscr{E}_\rho = \left\{ (x, y) \in \mathbb{R}^2 \ :\ \frac{x^2}{(1+\rho)^2} + \frac{y^2}{(1-\rho)^2} \le 1 \right\},$$

hence the name. This model interpolates from antisymmetric matrices ($\rho = -1$) to GOE matrices ($\rho = 1$) with the important special case of matrices with i.i.d. Gaussian entries ($\rho = 0$).

Elliptic AMP. For a normalized elliptic matrix A_n, the AMP algorithm takes the form:

$$\boldsymbol{u}^{k+1} = A_n\ h_k\left(\boldsymbol{u}^k, B\right) - \rho \left\langle \partial_1 h_k\left(\boldsymbol{u}^k, B\right) \right\rangle_n h_{k-1}\left(\boldsymbol{u}^{k-1}, B\right)\,, \tag{6}$$

and we shall prove in Theorem 1 the counterpart of Theorem 1 with the same DE equations, vector $\boldsymbol{b}$ being replaced by matrix B.

Notice however the modified Onsager term, multiplied by the correlation coefficient ρ. If $\rho = 0$, matrix $\sqrt{n} A$ has i.i.d. $\mathcal{N}(0, 1)$ entries and the Onsager term vanishes. For $\rho = 1$ we recover the GOE model.

Bolthausen's conditioning argument. In the framework of the GOE, the challenge for rigorously proving Theorem 1 by induction lies in an essential ingredient often called Bolthausen's conditioning argument [Bol14]. This technique provides a convenient representation of the conditional distribution of the new iterate $\boldsymbol{u}^{k+1}$ given the past $\left(\boldsymbol{u}^1, \cdots, \boldsymbol{u}^k\right)$. In the course of the proof, we establish a generalized version of Bolthausen's technique suited to handle elliptic matrix models.

Application to theoretical ecology. An important challenge in theoretical ecology is to build and analyze mathematical models to describe trophic networks and food-webs [LHM+04] of large dimension. In this regard, Lotka-Volterra (LV) systems of coupled differential equations is a popular model to describe the evolution of the various species' abundances in a food-web. The LV system of equations is written:

$$\frac{d\boldsymbol{x}}{dt}(t) = \boldsymbol{x}(t) \odot \left(\boldsymbol{r} - (I_n - \Sigma_n)\, \boldsymbol{x}(t)\right), \quad \boldsymbol{x}(0) \in (0,\infty)^n, \tag{7}$$

where $\odot$ stands for the Hadamard product, $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the vector of abundances of the n species at time t, $\boldsymbol{r} \in \mathbb{R}^n$ is the vector of intrinsic growth rates of the species and $\Sigma_n = (\Sigma_{ij})$ is the $n \times n$ interaction matrix, Σ_{ij} representing the effect of species j on the growth of species i.

In large dimension, a key feature of LV systems is the use of random matrices[1], see for instance [AT15, ABC+22], to model the interactions between the different species. This choice of a random matrix model is motivated in particular because the estimation of the real interactions is often out of reach.

An elliptic interaction matrix Σ_n, more precisely $\kappa\sqrt{n}\Sigma_n \sim \text{Elliptic}(n,\rho)$ (here $\kappa > 0$ is an extra degree of freedom), covers the case where the reciprocal interactions Σ_{ij} and Σ_{ji} between species are correlated. The elliptic model encompasses the cases of independent and equal reciprocal ecological interactions, and is widely considered in theoretical ecology [AT12, AT15, Bun17, Gal18].

The question we shall address is the description of the statistical properties of an equilibrium $\boldsymbol{x}^\star \in \mathbb{R}^n$ to (7), as n goes to infinity, whenever such an equilibrium exists (sufficient conditions for the existence of a unique and stable equilibrium have been provided in [CEFN22]). More specifically, we will be interested in the number of surviving species at equilibrium, the distribution of the surviving species, the individual distribution of a species, etc.

Based on Theorem 1 and following the strategy developed in [AHMN23] in the context of a symmetric interaction matrix, we will design an AMP algorithm which shall capture the equilibrium's statistical properties. This question has already been addressed by Bunin [Bun17] and Galla [Gal18] who provided a full description of $\boldsymbol{x}^\star$'s statistical properties via a system of non-linear equations at a physical level of rigor. We recover their equations, cf. (13), and provide a rigorous analysis of this system, substantially more demanding than in the symmetric case [AHMN23].

Combining a local AMP result (see Corollary 2) with arguments from the propagation of chaos theory, we also obtain new results on the individual distributions of species with different intrinsic growth rates.

Outline of the article. In Section 2, we present the elliptic AMP algorithm and state the main corresponding results, Theorem 1 (global AMP) and Corollary 2 (blockwise AMP). In Section 3, we present an application of AMP to theoretical ecology and design a specific AMP algorithm to describe the statistical properties of an equilibrium to a large LV system, see Theorem 3. Relying on propagation of chaos arguments, we describe the limiting behaviour of individual species' abundances in Corollary 4 (global exchangeability assumption) and Theorem 5 (blockwise exchangeability assumption). Proofs of AMP results are provided in Section 4. Proofs related to LV equilibria are provided in Section 5.

Technical results of special interest are Lemma 3.4 (description of the key equilibrium parameters via a deterministic system), Propositions 4.7 and 4.8 (extension of Bolthausen conditioning argument to elliptic random matrices) and Proposition 5.2 (chaos propagation for blockwise exchangeable vectors).

[1] In theoretical ecology, the use of random matrices goes back to May [May72].

Main notations and definitions. For a positive integer n, denote $[n] = \{1, \cdots n\}$. For $x \in \mathbb{R}$ let $x_+ = \max(x, 0)$ and $x_- = -\min(x, 0)$ so that $x = x_+ - x_-$. Vectors will be denoted by lowercase bold letters $\boldsymbol{x} = (x_i) \in \mathbb{R}^n$ and matrices by capital letters. For a matrix $A = (A_{ij}) \in \mathbb{R}^{n\times n}$, we denote by $A_{i,*}$ its i-th row and by $A_{*,j}$ its j-th column we also denote by $A_{[i],j}$ the first i elements of the j-th column of A. We denote by $A^\top$ the transpose transpose matrix of A.

For a vector $\boldsymbol{x} \in \mathbb{R}^n$ (respectively a $n \times n$ matrix B), $\|\boldsymbol{x}\|$ (resp. $\|B\|$) denotes its euclidean norm (resp. spectral norm) and $\langle \boldsymbol{x} \rangle_n \triangleq \frac{1}{n}\sum_{i=1}^n x_i$ the arithmetic mean of its coordinates. For two matrices A, B with identical dimensions, denote by $A \odot B = (A_{ij}B_{ij})$ their Hadamard product. The notation applies for two $\mathbb{R}^n$-vectors $\boldsymbol{x} \odot \boldsymbol{y} = (x_i y_i)$.

For $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R}^{p+1} \to \mathbb{R}$ and $\boldsymbol{x}, \boldsymbol{y}^1, \cdots, \boldsymbol{y}^p \in \mathbb{R}^n$, denote by $f(\boldsymbol{x})$ and $g(\boldsymbol{x}, \boldsymbol{y}^1, \cdots, \boldsymbol{y}^p)$ the n-dimensional vectors

$$f(\boldsymbol{x}) = (f(x_i))_{i\in[n]} \qquad \text{and} \qquad g(\boldsymbol{x}, \boldsymbol{y}^1, \cdots, \boldsymbol{y}^p) = (g(x_i, y_i^1, \cdots, y_i^p))_{i\in[n]}\,.$$

In particular, $\boldsymbol{x}_+ = ([x_i]_+)$.

Denote by $\mathcal{L}(X)$ the law of a random variable X. The equality in law between X and Y will be either denoted $X \overset{\mathcal{L}}{=} Y$ or $\mathcal{L}(X) = \mathcal{L}(Y)$. Independence is denoted by $\perp\!\!\!\perp$.

Acknowlegment. We thank all the members of the CNRS project 80-Prime-KARATE where part of this work has been initiated.

2. AMP for random elliptic matrices

We first introduce the notions of complete convergence and Wasserstein spaces. These concepts are crucial for the precise formulation of our main theorem.

2.1. Background.

Complete convergence. Given a sequence of random variables (X_n), we say that X_n converges completely to a constant x if for any other sequence $(Y_n)_n$ such that $X_n \overset{\mathcal{L}}{=} Y_n$ for all n, Y_n converges almost surely to x. We denote this mode of convergence as

$$X_n \xrightarrow[n\to+\infty]{c} x \qquad \text{or} \qquad X_n \xrightarrow[n\to\infty]{} x \quad \text{(completely)}\,.$$

It is worth noticing that the complete convergence of $(X_n)_n$ to x is equivalent to the condition $\sum_{n\in\mathbb{N}} \mathbb{P}(\|X_n - x\| > \varepsilon) < \infty$ for all $\varepsilon > 0$, as per Borel-Cantelli's lemma. One advantage of this convergence mode is that it is transmissible through equality in law, i.e.:

$$\text{If} \quad \begin{cases} X_n \xrightarrow[n\to+\infty]{c} x, \\ Y_n \overset{\mathcal{L}}{=} X_n, \quad \text{for all n} \end{cases} \qquad \text{then} \quad Y_n \xrightarrow[n\to+\infty]{c} x.$$

This property is shared with convergence in probability but not with almost sure convergence.

Wasserstein spaces. The Wasserstein space of order $r \geq 2$ denoted by $\mathcal{P}_r\left(\mathbb{R}^d\right)$ is the set of probability distributions μ on $\mathbb{R}^d$ with finite moments of order r:

$$\mathcal{P}_r\left(\mathbb{R}^d\right) = \left\{\mu \in \mathcal{P}(\mathbb{R}^d)\,, \quad \int_{\mathbb{R}^d} \|\boldsymbol{x}\|^r d\mu(\boldsymbol{x}) < \infty\right\}.$$

The Wasserstein distance between $\mu, \nu \in \mathcal{P}_r\left(\mathbb{R}^d\right)$, denoted by $\mathcal{W}(\mu, \nu)$, is defined as:

$$\mathcal{W}(\mu, \nu) = \inf_{\pi\in\Pi(\mu,\nu)} \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} \|\boldsymbol{x} - \boldsymbol{y}\|^r d\pi(\boldsymbol{x}, \boldsymbol{y})\right)^{1/r},$$

where $\Pi(\mu, \nu)$ is the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ having μ and ν as marginals. Given a sequence of probability measures $(\mu_n) \subset \mathcal{P}_r\left(\mathbb{R}^d\right)$ we will say that this sequence converges in the

Wasserstein space to $\mu \in \mathcal{P}_r\left(\mathbb{R}^d\right)$ if $\mathcal{W}(\mu_n, \mu) \xrightarrow[n\to\infty]{} 0$. We write

$$\mu_n \xrightarrow[n\to+\infty]{\mathcal{P}_r\left(\mathbb{R}^d\right)} \mu\,.$$

Convergence of measures in the Wasserstein space can be characterized using *pseudo-Lipschitz* functions. We say that $f : \mathbb{R}^d \to \mathbb{R}$ is a pseudo-Lipschitz function of degree $r \geq 2$ if there exists a constant L such that for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ the following inequality holds:

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq L\|\boldsymbol{x} - \boldsymbol{y}\|\left(1 + \|\boldsymbol{x}\|^{r-1} + \|\boldsymbol{y}\|^{r-1}\right).$$

We denote the set of pseudo-Lipschitz functions by $PL_r\left(\mathbb{R}^d\right)$. The following classical lemma which can be found in [AHMN23, Lemma 1] summarizes the characterizations of the convergence in the Wasserstein space.

Lemma 2.1. *Let $\mu_n, \mu \in \mathcal{P}_r\left(\mathbb{R}^d\right)$ for $r \geq 2$. The following conditions are equivalent:*

(1) $\mu_n \xrightarrow[n\to+\infty]{\mathcal{P}_r\left(\mathbb{R}^d\right)} \mu$,
(2) for all $\varphi \in PL_r\left(\mathbb{R}^d\right)$, $\int \varphi d\mu_n \to \int \varphi d\mu$,
(3) $\mu_n \xrightarrow{w} \mu$ *and* $\int \|\boldsymbol{x}\|^r \mu_n(d\boldsymbol{x}) \to \int \|\boldsymbol{x}\|^r \mu(d\boldsymbol{x})$.

2.2. **Assumptions.** We define an AMP algorithm by a triplet $\left(A_n, \mathcal{H}, \left(\boldsymbol{u}^0, B_n\right)\right)$, where A_n is a random matrix of size $n \times n$, $\mathcal{H} = \{h_k(.,.)\}_{k\in\mathbb{N}}$ is a sequence of functions from $\mathbb{R}^{p+1}$ to $\mathbb{R}$, $\boldsymbol{u}^0 \in \mathbb{R}^n$ is an initialization point and $B_n \in \mathbb{R}^{n\times p}$ is a matrix parameter. For our main theorem, the following assumptions are needed.

(A1) A_n is a normalized elliptic matrix with correlation coefficient $\rho \in [-1, 1]$.
(A2) The random vector $(\boldsymbol{u}^0, B_n) \in \mathbb{R}^{p+1}$ is independent of A_n and there exists a vector $(\bar{u}, \bar{b}_1, \cdots, \bar{b}_p)$ whose distribution belongs to $\mathcal{P}_r(\mathbb{R}^{p+1})$ such that

$$\mu^{\boldsymbol{u}_n^0, B_n} = \mu^{\boldsymbol{u}_n^0, \boldsymbol{b}_n^1, \cdots, \boldsymbol{b}_n^p} \xrightarrow[n\to\infty]{\mathcal{P}_r(\mathbb{R}^{p+1})} \mathcal{L}\left(\left(\bar{u}, \bar{b}_1, \cdots, \bar{b}_p\right)\right) \text{ (completely)}\,.$$

We denote $\bar{\boldsymbol{b}} = (\bar{b}_1, \cdots, \bar{b}_p)$.
(A3) For all $k \geq 0$, the function $h_k : \mathbb{R}^{p+1} \to \mathbb{R}$ is Lipschitz.
(A4) For every $k \geq 0$, $\mathbb{P}\left(\text{The function } x \mapsto h_k\left(x, \bar{\boldsymbol{b}}\right) \text{is constant}\right) < 1$.
(A5) The functions $\partial_1 h_k$ are continuous $\lambda \otimes \mathbb{P}^{\bar{a}}$- almost everywhere, where λ is the Lebesgue measure on $\mathbb{R}$.

Remark 2.2.

- *Assumption (A4) ensures that the covariance matrices R^k defined by the Density Evolution equations in (3) are positive definite, and in particular invertible. This is an important assumption used in the proof of Theorem 1 - see* [FVRS21, Lemma 2.2].
- *Assumption (A5) is a technicality needed to ensure the convergence of $\left\langle \partial_1 h_k\left(\boldsymbol{u}_n^k, B_n\right)\right\rangle_n$ to a deterministic limit.*

2.3. **Main result.** Recall that p is fixed. We first update the DE equations associated to a matrix parameter $B_n \in \mathbb{R}^{n\times p}$. This simply amounts to replace the scalar $\bar{b}$ in (2)-(3) by the vector $\bar{\boldsymbol{b}} = \left(\bar{b}_1, \cdots, \bar{b}_p\right)$. Let $h_i : \mathbb{R}^{p+1} \to \mathbb{R}$, consider vector $(\bar{u}, \bar{\boldsymbol{b}}) \in \mathbb{R}^{p+1}$ then Eq. (2) and (3) write

$$R^1 = \mathbb{E}\left[h_0^2\left(\bar{u}, \bar{\boldsymbol{b}}\right)\right], \qquad R_{ij}^k = \begin{cases} R^1 & \text{if } i = 1, j = 1 \\ \mathbb{E}\left[h_{i-1}\left(Z_{i-1}, \bar{\boldsymbol{b}}\right) h_{j-1}\left(Z_{j-1}, \bar{\boldsymbol{b}}\right)\right] & \text{if } i \geq 1, j \geq 1 \end{cases} \tag{8}$$

We can now state the main result of this section.

Theorem 1. *Let $k \geq 1$, assume (A1)-(A5) and consider a sequence of vectors $\left(\boldsymbol{u}^k\right)_k$ that satisfies the AMP scheme, i.e.*

$$\begin{cases} \boldsymbol{u}^1 = A_n\ h_0\left(\boldsymbol{u}^0, B_n\right), \\ \boldsymbol{u}^{k+1} = A_n\ h_k\left(\boldsymbol{u}^k, B_n\right) - \rho\left\langle \partial_1 h_k\left(\boldsymbol{u}^k, B_n\right)\right\rangle_n h_{k-1}\left(\boldsymbol{u}^{k-1}, B_n\right). \end{cases}$$

Let $(Z_1, \cdots, Z_k)$ be a centered Gaussian vector independent of $(\bar{u}, \bar{\boldsymbol{b}})$ with covariance matrix R^k given by the updated DE equations (8). Then the following convergence of the iterates holds:

$$\mu^{B,\boldsymbol{u}^1,\cdots,\boldsymbol{u}^k} \xrightarrow[n\to+\infty]{\mathcal{P}_r\left(\mathbb{R}^{p+k}\right)} \mathcal{L}\left(\left(\bar{\boldsymbol{b}}, Z_1, \cdots, Z_k\right)\right) \quad \textit{(completely)}. \tag{9}$$

Proof of Theorem 1 is provided in Section 4.

Remark 2.3. *Proof of Theorem 1 crucially relies on the Gaussianity of matrix A_n's entries and an important question would be how to relax this assumption. In Bayati et al. [BLM15], AMP is extended from a GOE model to a general Wigner matrix (symmetric matrix with i.i.d. entries on and above the diagonal) using an alternative strategy based on combinatorial methods. Adaptation of this combinatorial strategy to an elliptic framework will be the subject of a future work.*

Remark 2.4. *Using Lemma 2.1 the convergence result can also be expressed as*

$$\forall \varphi \in PL_r\left(\mathbb{R}^{p+k}\right), \quad \frac{1}{n}\sum_{i=1}^{n} \varphi\left(b_i^1, \cdots, b_i^p, u_i^1, \cdots, u_i^k\right) \xrightarrow[n\to\infty]{} \mathbb{E}\left[\varphi\left(\bar{b}_1, \cdots, \bar{b}_p, Z_1, \cdots, Z_k\right)\right].$$

Notice that the sum is over all integers from 1 to n, and thus each iterate vector $\boldsymbol{u}^\ell$ $(1 \le \ell \le k)$ is flattened. One may want to get a more local information, say the convergence of

$$\frac{1}{|C^{(n)}|}\sum_{i\in C^{(n)}} \varphi\left(b_i^1, \cdots, b_i^p, u_i^1, \cdots, u_i^k\right),$$

where $C^{(n)}$ is a subset of $[n]$.

Corollary 2 generalizes Theorem 1 in this direction. It relies on the following assumption.

(A2′) Let $q \ge 1$ be fixed and consider the following partition of $[n]$:

$$[n] = C_n^{(1)} \cup \cdots \cup C_n^{(q)} \quad \text{where} \quad \frac{|C_n^{(j)}|}{n} \xrightarrow[n\to\infty]{} c_j \in (0,1) \quad \text{for all } j \in [q]. \tag{10}$$

There exist q vectors $\left(\bar{u}_j, \bar{b}_{j,1}, \cdots, \bar{b}_{j,p}\right)$ with $j \in [q]$ such that:

$$\frac{1}{|C_n^{(j)}|}\sum_{i\in C_n^{(j)}} \delta_{\left(u_i^0, b_i^1, \cdots, b_i^p\right)} \xrightarrow[n\to\infty]{\mathcal{P}_r(\mathbb{R}^{p+1})} \mathcal{L}\left(\bar{u}_j, \bar{b}_{j,1}, \cdots, \bar{b}_{j,p}\right) \quad \text{(completely)}.$$

Corollary 2 (blockwise AMP)**.** *Let (A2′) hold and consider the framework of Theorem 1 except for (A2) (replaced by (A2′)). Then for all $j \in [q]$*

$$\frac{1}{|C_n^{(j)}|}\sum_{i\in C_n^{(j)}} \delta_{\left(b_i^1, \cdots, b_i^p, u_i^1, \cdots, u_i^k\right)} \xrightarrow[n\to\infty]{\mathcal{P}_r\left(\mathbb{R}^{p+k}\right)} \mathcal{L}\left(\bar{b}_{j,1}, \cdots, \bar{b}_{j,p}, Z_1, \cdots, Z_k\right) \quad \textit{(completely)}$$

where vector $(Z_1, \cdots, Z_k)$ is defined as in Theorem 1 and is independent from $(\bar{b}_{j,1}, \cdots, \bar{b}_{j,p})$.

Proof of Corollary 2 is postponed to Section 4.5.

3. Application to theoretical ecology: equilibria of large LV systems

3.1. **Large Lotka-Volterra systems.** In an ecological system where there are interactions between n species, the dynamics of these species can be modeled by a set of coupled differential equations called a Lotka-Volterra (LV) system.

Denote by $x_i(t)$ the abundance of species i at time t for $i \in [n]$ and by $\boldsymbol{x}(t) = (x_i(t))_{i\in[n]}$ the vector of abundances of all the species. Denote by $\boldsymbol{r} = (r_i)_{i\in[n]}$ the vector of intrinsic growth rates of all the species, and by Σ_n the $n \times n$ interaction matrix between the species.

The LV system is written

$$\frac{d\boldsymbol{x}}{dt}(t) = \boldsymbol{x}(t) \odot \left(\boldsymbol{r} - (I_n - \Sigma_n)\,\boldsymbol{x}(t)\right), \quad \boldsymbol{x}(0) \in (0,\infty)^n,$$

or equivalently

$$\frac{dx_i}{dt}(t) = x_i(t)\left(r_i - x_i + \sum_{j\in[n]} \Sigma_{ij}x_j(t)\right), \quad x_i(0) > 0 \qquad \text{for all } i \in [n].$$

Here Σ_{ij} represents the effect of species j on the growth of species i. Notice that if $\Sigma_n = 0$ (no interactions), then each species is described by a logistic differential equation. The general properties of a LV system are well-known, see for instance [Tak96, Chapter 3], notice in particular that $\boldsymbol{x}_n(t) > 0$ (componentwise) for all $t \geq 0$ if $\boldsymbol{x}(0) \in (0,\infty)^n$.

We are interested in the regime where n is large and Σ_n is random. Often, the real values of Σ_n are out of reach and an alternative is to choose a random model which statistical properties would reflect a partial knowledge on the ecological interaction network. Among the various matrix models, the Elliptic(n,ρ) model represents a good trade-off between complexity and tractability, see [AT12, AT15, ABC+22]. We will therefore assume that

$$\kappa\sqrt{n}\Sigma_n \sim \text{Elliptic}(n,\rho), \tag{11}$$

Otherwise stated, $\Sigma_n = \frac{A_n}{\kappa}$ where A_n is a normalized elliptic matrix and $\kappa > 0$ is an extra parameter. In this case, $\|\Sigma_n\| = \mathcal{O}(1)$ as $n \to \infty$ and the interaction matrix Σ_n will have a macroscopic effect on the LV system as $n \to \infty$.

3.2. **Existence of a stable and unique equilibrium.** In [CEFN22], sufficient conditions are provided so that system (7) eventually admits a unique and stable equilibrium.

Proposition 3.1 (Prop. 2.3 in [CEFN22])**.** *Consider system (7) where $\kappa\sqrt{n}\Sigma_n \sim \text{Elliptic}(n,\rho)$ and suppose that $\kappa > \sqrt{2(1+\rho)}$. Then almost surely (a.s.) eventually there exists a unique and stable equilibrium $\boldsymbol{x}^{\text{EQ}} = (x_i^{\text{EQ}}(n))_{i\in[n]}$. Otherwise stated, with probability one there exists N such that for all $n \geq N$, there exists a unique $\boldsymbol{x}^{\text{EQ}} \in \mathbb{R}^n$ such that*

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

where $\boldsymbol{x}(t)$ solves (7).

Remark 3.2. *The fact that $\boldsymbol{x}(t) > 0$ for all $t > 0$ only implies that $\boldsymbol{x}^{\text{EQ}} \geq 0$. A vanishing component of $\boldsymbol{x}^{\text{EQ}}$ represents a vanishing species (whose abundance is zero at equilibrium).*

Remark 3.3 (extension of the definition of $\boldsymbol{x}^{\text{EQ}}$)**.** *Notice that in Proposition 3.1, the equilibrium $\boldsymbol{x}^{\text{EQ}}$ is eventually defined. In fact, standard arguments yield*

$$\frac{\|A_n\|}{\kappa} \xrightarrow[n\to\infty]{a.s.} \frac{\sqrt{2(1+\rho)}}{\kappa} < 1$$

by assumption over κ. If the condition $\frac{\|A_n\|}{\kappa} < 1$ is met, which happens a.s. eventually, then the existence of $\boldsymbol{x}^{\text{EQ}}$ is granted by Takeuchi's result [Tak96, Th. 3.2]. We extend the definition of $\boldsymbol{x}^{\text{EQ}}$ by setting

$$\boldsymbol{x}^\star = \begin{cases} \boldsymbol{x}^{\text{EQ}} & \text{if } \frac{\|A_n\|}{\kappa} < 1, \\ 0 & \text{else}. \end{cases} \tag{12}$$

With a slight abuse of notation, one may denote $\boldsymbol{x}^\star = \boldsymbol{x}^{\text{EQ}}\mathbf{1}_{\{\|A_n\|/\kappa\,<\,1\}}$.

3.3. **Statistical properties of the LV equilibrium $\boldsymbol{x}^\star$.** Once the existence of the equilibrium is granted, we shall explore its statistical properties and address questions such as: What is the proportion of surviving species at equilibrium? What is the distribution of surviving species? etc. In this regard, a key device will be the study of the empirical probability measure

$$\mu^{\boldsymbol{x}^\star} = \frac{1}{n}\sum_{i\in[n]} \delta_{x_i^\star(n)},$$

and the design of an appropriate AMP algorithm.

We first introduce a system of three equations whose solutions will play a key role in describing the statistical properties of the equilibrium.

Lemma 3.4. *Let $\rho \in [-1,1]$ and suppose that $\kappa > (1+\rho)/\sqrt{2}$. Consider two independent real random variables $\bar{Z}$ and $\bar{r}$ where $Z \sim \mathcal{N}(0,1)$ and $\bar{r} \geq 0$ with finite second moment and $\mathcal{L}(\bar{r}) \neq \delta_0$. Then the system of equations*

$$\kappa = \delta + \rho\frac{\gamma}{\delta}\,, \tag{13a}$$

$$\sigma^2 = \frac{1}{\delta^2}\mathbb{E}\left[\left(\sigma\bar{Z}+\bar{r}\right)_+^2\right]\,, \tag{13b}$$

$$\gamma = \mathbb{P}\left[\sigma\bar{Z}+\bar{r}>0\right]\,, \tag{13c}$$

admits an unique solution (δ,σ,γ) in $(1/\sqrt{2},\infty)\times(0,\infty)\times(0,1)$.

Proof of Lemma 3.4 is deferred to Appendix A. We follow the lines of the corresponding proof for the symmetric matrix case [ABC+22, Section 3.2] but the case $\rho<0$ requires new arguments.

Remark 3.5. *Notice that condition $\kappa > (1+\rho)/\sqrt{2}$ in Lemma 3.4 is weaker than condition $\kappa > \sqrt{2(1+\rho)}$ provided in Proposition 3.1 unless $\rho = -1$. Otherwise stated if κ satisfies the condition*

$$\frac{1+\rho}{\sqrt{2}} < \kappa \leq \sqrt{2(1+\rho)}\,, \quad (\rho > -1)$$

the system may admit a unique solution but the existence of a stable equilibrium is not granted.

We can now state the main result of this section.

Theorem 3. *Let A_n be a normalized elliptic matrix, $\boldsymbol{r} \in \mathbb{R}^n$ a random vector independent from A_n satisfying:*

$$\mu^{\boldsymbol{r}} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}(\bar{r}) \quad \textit{(completely)}\,,$$

where $\bar{r} \geq 0$ is a real valued random variable with finite second moment and $\mathcal{L}(\bar{r}) \neq \delta_0$. Let $\bar{Z}$ be a $\mathcal{N}(0,1)$ random variable independent from $\bar{r}$.

Let $\kappa > \sqrt{2(1+\rho)}$ and consider the LV system (7) where $\Sigma_n = \frac{A_n}{\kappa}$. Let $\boldsymbol{x}^\star$ be defined by (12) and $(\delta,\sigma,\gamma) \in (1/\sqrt{2},\infty)\times(0,\infty)\times(0,1)$ be the solution of (13) in Lemma 3.4, then

$$\mu^{\boldsymbol{x}^\star} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \pi := \mathcal{L}\left(\left(1+\rho\gamma/\delta^2\right)\left(\sigma\bar{Z}+\bar{r}\right)_+\right) \quad \textit{(completely)}\,. \tag{14}$$

Proof of Theorem 3 is outlined in Section 5.1. It closely follows the strategy developed in [AHMN23] in the context of a symmetric interaction matrix (see in particular the outline of the proof in [AHMN23, Section 3.1]). This strategy is adapted to the elliptic case with the help of Theorem 1 and the existence of a unique solution to (13).

Remark 3.6 (Proportion of surviving species)**.** *Strictly speaking, the proportion of surviving species at equilibrium is given by:*

$$\mu^{\boldsymbol{x}^\star}(0,\infty) = \frac{1}{n}\sum_{i\in[n]} \mathbf{1}_{(0,\infty)}(x_i^\star)\,.$$

As a consequence, convergence (14) in Theorem 3 does not apply for $x \mapsto \mathbf{1}_{(0,\infty)}(x)$ is not continuous at zero, a discontinuity point of the limiting cumulative function. However for any continuous function f_ε satisfying

$$f_\varepsilon(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 1 & \text{for } x \geq \varepsilon \end{cases}$$

for a small $\varepsilon > 0$, one has $\frac{1}{n}\sum_i f_\varepsilon(x_i^\star) \xrightarrow[n\to\infty]{} \mathbb{E}f_\varepsilon\left[\left(1+\rho\gamma/\delta^2\right)\left(\sigma\bar{Z}+\bar{r}\right)_+\right]$ and

$$\mathbb{E}f_\varepsilon\left[\left(1+\rho\gamma/\delta^2\right)\left(\sigma\bar{Z}+\bar{r}\right)_+\right] \xrightarrow[\varepsilon\to 0]{} \gamma = \mathbb{P}(\sigma\bar{Z}+\bar{r}>0)\,.$$

Hence γ appears as a good approximation of the proportion of surviving species. This is confirmed by simulations, see Figure 1a.

Remark 3.7 (Distribution of surviving species)**.** *Denote by $\boldsymbol{s}(\boldsymbol{x}^\star)$ the subvector of $\boldsymbol{x}^\star$ with the positive components of $\boldsymbol{x}^\star$. Its dimension $|\boldsymbol{s}(\boldsymbol{x}^\star)|$ is random and the distribution of the surviving species is given by:*

$$\mu^{\boldsymbol{s}(\boldsymbol{x}^\star)} = \frac{1}{|\boldsymbol{s}(\boldsymbol{x}^\star)|} \sum_{i\in[|\boldsymbol{s}(\boldsymbol{x}^\star)|]} \delta_{[\boldsymbol{s}(\boldsymbol{x}^\star)]_i} \ .$$

A formal convergence of $\mu^{\boldsymbol{s}(\boldsymbol{x}_n^\star)}$ is out of reach (see the arguments in Remark 3.6) but a good proxy should be

$$\mathcal{L}\left(\left(1+\rho\gamma/\delta^2\right)\left(\sigma\bar{Z}+\bar{r}\right)_+ \,\middle|\, \sigma\bar{Z}+\bar{r}>0 \right) \ , \tag{15}$$

the density of which is explicit. Let

$$f_{\sigma\bar{Z}+\bar{r}}(y) = \int_{\mathbb{R}} \frac{e^{-\frac{(y-r)^2}{2\sigma^2}}}{\sqrt{2\pi}\,\sigma} \mathbb{P}_{\bar{r}}(dr) \quad \textit{and} \quad 1+\rho\frac{\gamma}{\delta^2} = \frac{\kappa}{\delta} \ ,$$

then the density of (15) denoted by f_{surv} is written

$$f_{surv}(y) \ = \ \frac{\delta}{\kappa} f_{\sigma\bar{Z}+\bar{r}}\left(\frac{\delta\, y}{\kappa}\right) \frac{\mathbf{1}_{(y>0)}}{\gamma} \ . \tag{16}$$

One can now easily express the relation between π as defined in (14) and f_{surv}:

$$\pi(dy) = \gamma f_{surv}(y)\, dy + (1-\gamma)\delta_0(dy) \ . \tag{17}$$

Notiece that if the r.v. $\bar{r}$ is constant then f_{surv} is the density of a truncated Gaussian distribution.

Simulations show a very good fit between this distribution and the histogram associated to $\mu^{\boldsymbol{s}(\boldsymbol{x}^\star)}$ for large n, see Fig. 1b.

3.4. **Propagation of chaos.** Combining Theorem 3 with propagation of chaos type results [Szn91], we are able, with extra exchangeability assumptions on the vector $\boldsymbol{r} \in \mathbb{R}^n$, to describe the limiting behaviour of individual abundances.

We obtain two kinds of results. If $\boldsymbol{r}$ is exchangeable, then the distribution of every individual abundance converges toward the same limit given by π in (14), see Corollary 4. If $\boldsymbol{r}$ is blockwise exchangeable (to be defined), then within each block each abundance can have a specific limit, see Theorem 5, which may differ from π.

Corollary 4. *Consider the framework of Theorem 3 and assume moreover that vector $\boldsymbol{r}$ is exchangeable. Let $\boldsymbol{x}^\star = (x_i^\star(n))$ be defined by (12) and recall the definition of the distribution π:*

$$\pi = \mathcal{L}\left(\left(1+\rho\gamma/\delta^2\right)\left(\sigma\bar{Z}+\bar{r}\right)_+ \right) \ .$$

Then $(x_1^\star(n), \cdots, x_n^\star(n))$ is an exchangeable sequence and for any fixed $K \geq 1$,

$$(x_1^\star(n), \cdots, x_K^\star(n)) \xrightarrow[n\to\infty]{\mathcal{L}} \pi^{\otimes K} \ .$$

Proof of Corollary 4 is postponed to Section 5.2.

Remark 3.8. *This result should be compared to Geman and Hwang* [GH82, Theorem 3].

For the next result, we need some extra definitions. Let $q \geq 1$ be a fixed integer. Consider q sequences $n_1(n), \cdots, n_q(n)$ satisfying

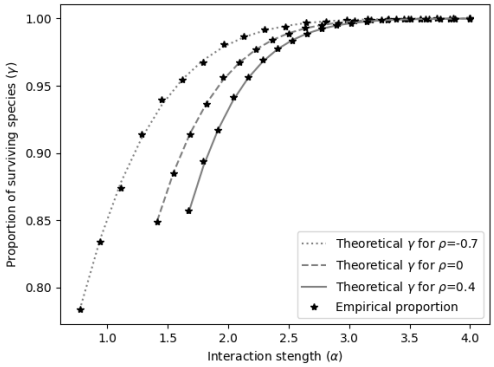
$$n_1 + \cdots + n_q = n \qquad \text{and} \qquad \frac{n_j}{n} \xrightarrow[n\to\infty]{} c_j \in (0,1) \ , \quad j \in [q] \ . \tag{18}$$

Consider the following partition of $[n]$:

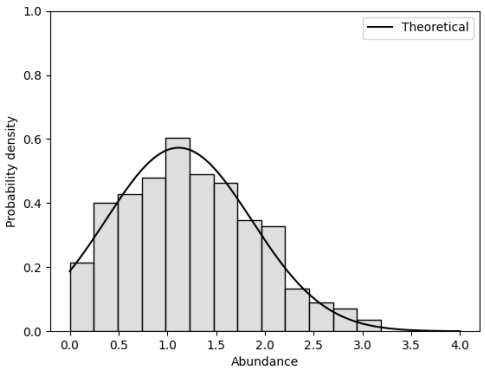
$$\begin{cases} C_n^{(1)} &= \{1, \cdots, n_1\} \ , \\ C_n^{(j)} &= \{n_1 + \cdots + n_{j-1} + 1, \ \cdots, n_1 + \cdots + n_j\} \ , \quad 1 < j \leq q \end{cases} \tag{19}$$

so that

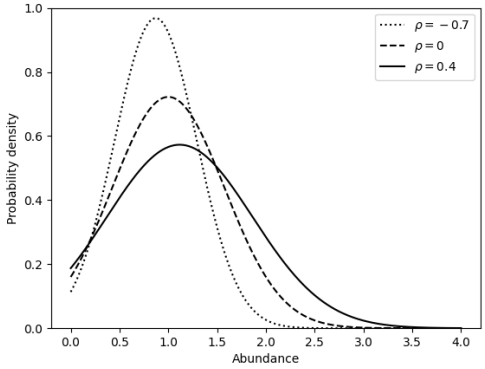
$$[n] = \bigcup_{i=1}^{q} C_n^{(j)} \qquad \text{and} \qquad |C_n^{(j)}| = n_j \ , \quad j \in [q] \ .$$



(A) Experimental proportions of surviving species vs theoretical values γ for three correlation coefficients $\rho = -0.7, 0, 0.4$ w.r.t. the interaction strength (κ).



(B) Histogram of positive abundances vs the density function f_{surv} described in (16) for $\rho = 0.4$ with the interaction strength fixed to $\kappa = 2$.



(C) Plot of the density function f_{surv} for $\rho = -0.7, 0, 0.4$.

FIGURE 1. Comparison between the theoretical solution of the fixed point equations (13) and their empirical Monte Carlo counterpart obtained by computing equilibria $\boldsymbol{x}^\star$ for various realizations of matrix A. Every $\boldsymbol{x}^\star$ is the solution of a Linear Complementarity Problem (see (44)) and is thus computed by Lemke algorithm. For Figure 1a and Figure 1b we chose a matrix of size 200 and we fixed the number of Monte Carlo experiments to 100 and 500 respectively.

Any $\mathbb{R}^n$-valued vector $\boldsymbol{v}$ can be decomposed into q $\mathbb{R}^{n_j}$-valued subvectors $\boldsymbol{v}^{(j)}$:

$$\boldsymbol{v} = (\boldsymbol{v}^{(j)},\ j \in [q]) \qquad \text{where} \qquad \boldsymbol{v}^{(j)} = (v_k)_{k \in C_n^{(j)}}\,.$$

Given a permutation $\sigma_j \in \mathcal{S}_{n_j}$, we denote by $\boldsymbol{v}^{(j,\sigma_j)}$ the vector $\boldsymbol{v}^{(j)}$ where each component has been permuted according to σ_j.

Consider now $\sigma = (\sigma_1, \cdots, \sigma_q)$ where $\sigma_j \in \mathcal{S}_{n_j}$. Given a vector $\boldsymbol{v} \in \mathbb{R}^n$ we denote by $\boldsymbol{v}^\sigma$ the vector

$$\boldsymbol{v}^\sigma = \left(\boldsymbol{v}^{(j,\sigma_j)};\ j \in [q]\right)\,.$$

Example 3.9. *Consider $n = 6$ and a number of blocks $q = 2$ such that $n_1 = 4$ and $n_2 = 2$. Let $\boldsymbol{v} = (v_1, v_2, v_3, v_4, v_5, v_6)^\top$ and consider two permutations*

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \in \mathcal{S}_4 \quad \textit{and} \quad \sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in \mathcal{S}_2\,.$$

Then $\boldsymbol{v}^{(1,\sigma_1)} = (v_3, v_4, v_1, v_2)^\top$ *and* $\boldsymbol{v}^{(2,\sigma_2)} = (v_6, v_5)^\top$. *Let* $\sigma = (\sigma_1, \sigma_2)$, *then* $\boldsymbol{v}^\sigma = (v_3, v_4, v_1, v_2, v_6, v_5)^\top$.

Definition 3.10. *A random* $\mathbb{R}^n$*-valued vector* $\boldsymbol{v}$ *is blockwise exchangeable with respect to the partition defined in* (19) *if for any* $\sigma = (\sigma_1, \cdots, \sigma_q) \in \mathcal{S}_{n_1} \times \cdots \times \mathcal{S}_{n_q}$,

$$\boldsymbol{v}^\sigma \overset{\mathcal{L}}{=} \boldsymbol{v}\,.$$

Otherwise stated, for any bounded continuous test functions $\varphi_j : \mathbb{R}^{n_j} \to \mathbb{R}$ *with* $j \in [q]$,

$$\mathbb{E}\left(\varphi_1\left(\boldsymbol{v}^{(1)}\right) \times \cdots \times \varphi_q\left(\boldsymbol{v}^{(q)}\right)\right) = \mathbb{E}\left(\varphi_1\left(\boldsymbol{v}^{(1,\sigma_1)}\right) \times \cdots \times \varphi_q\left(\boldsymbol{v}^{(q,\sigma_q)}\right)\right).$$

If there is no confusion, we simply say that $\boldsymbol{v}$ *is blockwise exchangeable.*

We are now in position to state our final result.

Theorem 5. *Consider the framework of Theorem 3 and let* $q \geq 1$ *be a fixed integer. Let* $(n_j, j \in [q])$ *and* $(C_n^{(j)}, j \in [q])$ *be given by* (18)-(19). *Let* $\boldsymbol{x}^\star$ *be defined in* (12). *Assume that* $\boldsymbol{r}$ *is blockwise exchangeable and that for all* $j \in [q]$

$$\mu^{\boldsymbol{r}^{(j)}} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \bar{r}_j \quad \textit{(completely)}$$

where $\bar{r}_j \geq 0$ *is a random variable with finite second moment.*

Then for any sequence $\psi_n \in C_n^{(j)}$ *where* $j \in [q]$ *is fixed,*

$$x^\star_{\psi_n} \xrightarrow[n\to\infty]{\mathcal{L}} \pi_j := \mathcal{L}\left(\left(1 + \rho\gamma/\delta^2\right)\left(\sigma\bar{Z} + \bar{r}_j\right)_+\right).$$

Moreover, let $k_1, \cdots, k_q \geq 1$ *be fixed integers and consider subsets*

$$\mathcal{K}_n^{(j)} \subset C_n^{(j)} \quad \textit{with} \quad |\mathcal{K}_n^{(j)}| = k_j \quad \textit{and} \quad k = k_1 + \cdots + k_q\,,$$

then the $\mathbb{R}^k$*-valued vector*

$$\boldsymbol{x}^\star_{[k_1,\cdots,k_q]} := \left(x^\star_\ell\,,\ \ell \in \mathcal{K}_n^{(1)} \cup \cdots \cup \mathcal{K}_n^{(q)}\right)$$

satisfies

$$\tilde{\boldsymbol{x}}^\star_{[k_1,\cdots,k_q]} \xrightarrow[n\to\infty]{\mathcal{L}} \prod_{j=1}^q \pi_j^{\otimes k_j}\,.$$

Proof of Theorem 5 is postponed to Section 5.3.

Remark 3.11 (Global versus local distributions)**.** *Unlike the case where the intrinsic growth rates vector* $\boldsymbol{r}$ *is exchangeable (see Corollary 4), notice now that each block of the equilibrium vector* $\boldsymbol{x}^\star$ *converges to a different law which locally depends on the structure of* $\boldsymbol{r}$. *In particular the* j*-th block of* $\boldsymbol{x}^\star$ *converges to*

$$\pi_j = \mathcal{L}\left(\left(1 + \rho\gamma/\delta^2\right)\left(\sigma\bar{Z} + \bar{r}_j\right)_+\right),$$

which is different than the overall asymptotic behaviour of $\boldsymbol{x}^\star$,

$$\pi = \mathcal{L}\left(\left(1 + \rho\gamma/\delta^2\right)\left(\sigma\bar{Z} + \bar{r}\right)_+\right).$$

Simulations based on a three-block piece-wise constant vector $\boldsymbol{r} = (\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(2)}, \boldsymbol{r}^{(3)})$ *are provided in Figure 2.*

Remark 3.12 (Distribution π is a mixture of the π_j's)**.** *Recall the definitions of the* q *random variables* $(\bar{r}_j)_{j\in[q]}$ *in Theorem 5 and the definition of* $\bar{r}$ *in Theorem 3, we can notice that the law of* $\bar{r}$ *is the mixture the laws of* $(\bar{r}_j)_{j\in[q]}$ *with coefficients* $(c_j)_{j\in[q]}$, *i.e.*

$$\mathcal{L}(\bar{r}) = \sum_{j=1}^q c_j \mathcal{L}(\bar{r}_j).$$

This also means that the limiting distribution of the whole equilibrium vector $\boldsymbol{x}^\star$ is a mixture of laws, i.e.

$$\mu^{\boldsymbol{x}^\star} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \sum_{j=1}^{q} c_j \pi_j \quad \textit{where} \quad \pi_j = \mathcal{L}\left(\left(1+\rho\gamma/\delta^2\right)\left(\sigma\bar{Z}+\bar{r}_j\right)_+\right).$$

Remark 3.13 (surviving species within block j)**.** *From an empirical point of view, simulations easily provide the number of surviving species within a block j of size n_j, that is*

$$\frac{\#\{x^*_\ell > 0, \ell \in C_n^{(j)}\}}{n_j},$$

and the value of their positive abundance. Following Remarks 3.6 and 3.7, Theorem 5 provides their analytical counterparts. The quantity

$$\gamma_j = \mathbb{P}(\sigma\bar{Z}+\bar{r}_j > 0)$$

is a good approximation for the proportion of surviving species within block j and the density

$$f^j_{surv}(y) \;=\; \frac{\delta}{\kappa} f_{\sigma\bar{Z}+\bar{r}_j}\left(\frac{\delta\, y}{\kappa}\right)\frac{\mathbf{1}_{(y>0)}}{\gamma_j} \quad \textit{where} \quad f_{\sigma\bar{Z}+\bar{r}_j}(y) = \int_{\mathbb{R}} \frac{e^{-\frac{(y-r)^2}{2\sigma^2}}}{\sqrt{2\pi}\,\sigma}\mathbb{P}_{\bar{r}_j}(dr),$$

for the distribution of the surviving species in block j.

One can notice that f_{surv} is a mixture of the f^j_{surv}'s:

$$f_{surv}(y) = \sum_{j=1}^{q} \frac{c_j\gamma_j}{\gamma} f^j_{surv}(y) \quad \textit{with} \quad \sum_{j=1}^{q} \frac{c_j\gamma_j}{\gamma} = 1\,.$$

Based on a three-block piece-wise constant vector $\boldsymbol{r} = (\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(2)}, \boldsymbol{r}^{(3)})$, the densities f^1_{surv}, f^2_{surv} and f^3_{surv} are compared to the corresponding simulation based histograms in Figure 2.

4. Proofs of Theorem 1 and Corollary 2

In Section 4.1 we provide various results related to elliptic random matrices. Sections 4.2–4.4 are devoted to the proof of Theorem 1. After introducing new notations in Section 4.2, we provide an adaptation of Bolthausen conditioning argument to elliptic random matrices in Section 4.3, see Propositions 4.7 and 4.8. This represents the crux of the proof of Theorem 1 and our main contribution to this section. Section 4.4 is devoted to the end of proof of Theorem 1 and closely follows [FVRS21]. Proof of Corollary 2 is established in Section 4.5.

4.1. Preliminary results on elliptic matrices.

Let X be a $n \times n$ matrix with independent $\mathcal{N}(0,1)$ entries. If the $n \times n$ matrices G and $\widetilde{G}$ satisfy

$$G \overset{\mathcal{L}}{=} \frac{X+X^\top}{\sqrt{2}} \qquad \text{and} \qquad \widetilde{G} \overset{\mathcal{L}}{=} \frac{X-X^\top}{\sqrt{2}},$$

then we say that G is a GOE matrix and $\widetilde{G}$ an antisymmetric GOE matrix.

From the definition 1.1 of an elliptic matrix, it is easy to check that a matrix $M \sim \text{Elliptic}(n,\rho)$ for $\rho \in [-1,1]$ can be characterized as

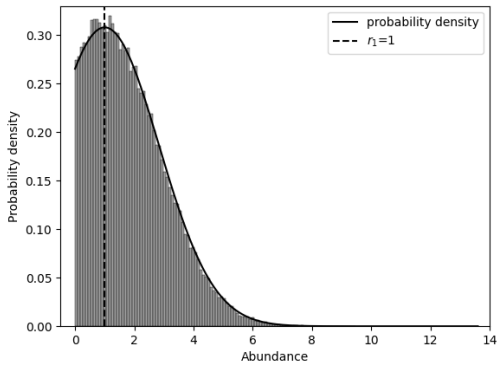
$$M \overset{\mathcal{L}}{=} \sqrt{\frac{1+\rho}{2}}G + \sqrt{\frac{1-\rho}{2}}\widetilde{G}, \tag{20}$$

where G is a GOE matrix, $\widetilde{G}$ is an antisymmetric GOE matrix, and $G \perp\!\!\!\perp \widetilde{G}$.

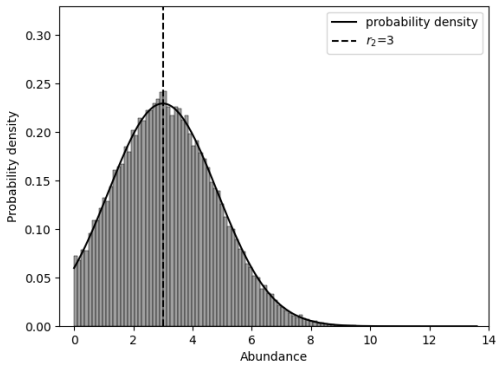
We begin by two elementary results on GOE matrices:

Lemma 4.1. *Let G and $\widetilde{G}$ be respectively a symmetric and an antisymmetric $n \times n$ GOE matrix. Consider two deterministic vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, then:*

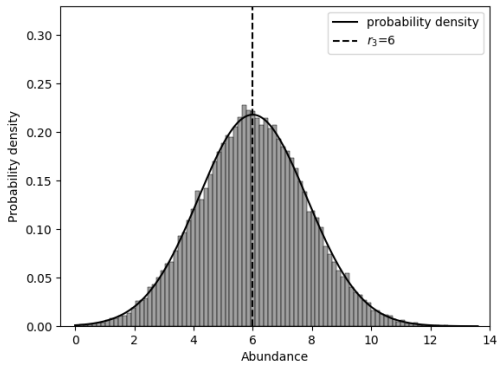
$$(i) \quad \mathbb{E}\, G\boldsymbol{u}\boldsymbol{v}^\top G = (\boldsymbol{v}^\top\boldsymbol{u})I_n + \boldsymbol{v}\boldsymbol{u}^\top \quad \textit{and} \quad \mathbb{E}\,\widetilde{G}\boldsymbol{u}\boldsymbol{v}^\top\widetilde{G} = -(\boldsymbol{v}^\top\boldsymbol{u})I_n + \boldsymbol{v}\boldsymbol{u}^\top,$$
$$(ii) \quad G\boldsymbol{u} \sim \mathcal{N}\left(0, I_n + \boldsymbol{u}\boldsymbol{u}^\top\right) \quad \textit{and} \quad \widetilde{G}\boldsymbol{u} \sim \mathcal{N}\left(0, I_n - \boldsymbol{u}\boldsymbol{u}^\top\right).$$



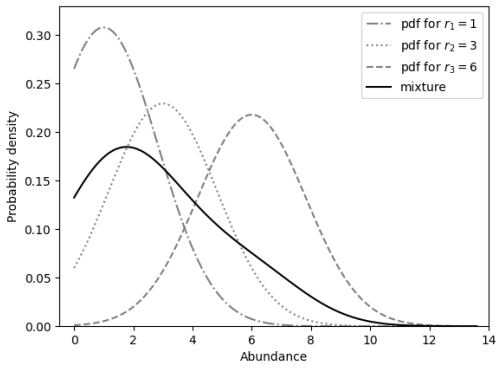
(A) The density f^1_{surv} compared to the histogram of surviving species in block 1.



(B) The density f^2_{surv} compared to the histogram of surviving species in block 2.



(C) The density f^3_{surv} compared to the histogram of surviving species in block 3.



(D) The Distribution of $\boldsymbol{x}^\star$ as a mixture of the three blocks' distributions.

FIGURE 2. The vector $\boldsymbol{r}$ is decomposed into three consecutive blocks, for each block we fix a constant value $r_1, r_2, r_3 = 1, 3, 6$ respectively, we also choose different sizes of the blocks to be $n/2, 3n/10, n/5$. We then solve the Linear Complementarity Problem problem (see (44)) with the help of Lemke algorithm for $\rho = 0$ and $\kappa = 2$.

Proof. We prove the two statements for G, the corresponding results for $\widetilde{G}$ can be shown similarly. For (i), we start by writing $\mathbb{E}[G\boldsymbol{u}\boldsymbol{v}^\top G]_{kk} = \sum_{ij} \mathbb{E}G_{ki}G_{kj}u_iv_j = \sum_i \mathbb{E}G^2_{ki}u_iv_i = u_kv_k + \boldsymbol{v}^\top\boldsymbol{u}$, and $\mathbb{E}[G\boldsymbol{u}\boldsymbol{v}^\top G]_{kl} = \sum_{ij} \mathbb{E}G_{ki}G_{jl}u_iv_j = u_lv_k$ for $k \neq l$. Thus, $\mathbb{E}G\boldsymbol{u}\boldsymbol{v}^\top G = (v^\top u)I_n + \boldsymbol{v}\boldsymbol{u}^\top$. We now prove (ii). As in the proof of [FVRS21, Lemma 6.14], let us complete the vector $\boldsymbol{u}$ in a deterministic orthogonal matrix $U = \left[\boldsymbol{u}\ \widetilde{U}\right]$. By the orthogonal invariance of GOE matrices, we have

$$G\boldsymbol{u} \overset{\mathcal{L}}{=} UGU^\top\boldsymbol{u} = UGe_1 \sim \mathcal{N}\left(0, U(e_1e_1^\top + I_n)U^\top\right) = \mathcal{N}\left(0, I_n + \boldsymbol{u}\boldsymbol{u}^\top\right),$$

hence the desired result. □

Using these results, we now have the two following propositions of elliptic matrices.

Proposition 4.2. *Let $M \sim$ Elliptic(n, ρ). Let $\boldsymbol{q} \in \mathbb{R}^n$ and $U \in \mathbb{R}^{n\times k}$ be deterministic with $\boldsymbol{q}^\top U = 0$. Let $P \in \mathbb{R}^{n\times n}$ be a deterministic matrix that satisfies $PU = 0$. Then*

$$(M - \rho PM^\top)\boldsymbol{q} \quad \perp\!\!\!\perp \quad \begin{bmatrix}(M - \rho M^\top)U \\ U^\top M^\top U\end{bmatrix}.$$

Proof. Since the considered quantities form a Gaussian vector, it is enough to show that $(M - \rho PM^\top)\boldsymbol{q}$ is decorrelated from all columns of the two matrices $(M - \rho M^\top)U$ and $U^\top M^\top U$. To this end, we use the characterization (20) and write

$$\begin{bmatrix} M \\ M - \rho M^\top \end{bmatrix} \quad \stackrel{\mathcal{L}}{=} \quad \begin{bmatrix} \sqrt{\frac{1+\rho}{2}} G + \sqrt{\frac{1-\rho}{2}} \widetilde{G}, \\ (1-\rho)\sqrt{\frac{1+\rho}{2}} G + (1+\rho)\sqrt{\frac{1-\rho}{2}} \widetilde{G} \end{bmatrix}.$$

Let $\boldsymbol{u}$ be any column of U. We first show that $(M - \rho PM^\top)\boldsymbol{q}$ and $(M - \rho M^\top)\boldsymbol{u}$ are decorrelated. To compute $\mathbb{E}(M - \rho PM^\top)\boldsymbol{q}\boldsymbol{u}^\top(M - \rho M^\top)^\top$, we use Lemma 4.1-(i), noticing that the cross terms involving G and $\widetilde{G}$ in the previous characterization are zero, and that $\boldsymbol{q}^\top \boldsymbol{u} = 0$. This leads to

$$\begin{aligned} \mathbb{E}(M - \rho PM^\top)\boldsymbol{q}\boldsymbol{u}^\top(M - \rho M^\top)^\top &= \left(\frac{1-\rho^2}{2}(I - \rho P) - \frac{1-\rho^2}{2}(I + \rho P) \right) \boldsymbol{u}\boldsymbol{q}^\top \\ &= -\rho(1-\rho^2) P \boldsymbol{u}\boldsymbol{q}^\top \;=\; 0\,, \end{aligned}$$

by noticing that $P\boldsymbol{u}\boldsymbol{q}^\top = 0$.

To obtain that $(M - \rho PM^\top)\boldsymbol{q}$ and $U^\top M^\top \boldsymbol{u}$ are decorrelated for each column $\boldsymbol{u}$ of U, we easily notice that $\mathbb{E}\left(M - \rho PM^\top\right)\boldsymbol{q}\boldsymbol{u}^\top M = \text{scalar} \times \boldsymbol{u}\boldsymbol{q}^\top$ from the structures of $M - \rho M^\top$ and M provided above and from Lemma 4.1-(i). Thus, $\mathbb{E}\left(M - \rho PM^\top\right)\boldsymbol{q}\boldsymbol{u}^\top MU = 0$ since $\boldsymbol{q}^\top U = 0$. □

Proposition 4.3. *Let $M \sim \text{Elliptic}(n, \rho)$. Let $\boldsymbol{q} \in \mathbb{R}^n$ be a unit-norm deterministic vector, and let P be a deterministic orthogonal projection matrix on a subspace of $\mathbb{R}^n$ such that $P\boldsymbol{q} = 0$. Then,*

$$(M - \rho PM^\top)\boldsymbol{q} \sim \mathcal{N}\left(0, I - \rho^2 P + \rho \boldsymbol{q}\boldsymbol{q}^\top\right).$$

Proof. Using the same principle as in the previous proof, we write

$$M - \rho PM^\top \stackrel{\mathcal{L}}{=} \sqrt{\frac{1+\rho}{2}}(I - \rho P)G + \sqrt{\frac{1-\rho}{2}}(I + \rho P)\widetilde{G}.$$

By Lemma 4.1-(ii), we then have that

$$(I - \rho P)G\boldsymbol{q} \sim \mathcal{N}\left(0, (I - \rho P)^2 + \boldsymbol{q}\boldsymbol{q}^\top\right) \quad \text{and} \quad (I + \rho P)\widetilde{G}\boldsymbol{q} \sim \mathcal{N}\left(0, (I + \rho P)^2 - \boldsymbol{q}\boldsymbol{q}^\top\right).$$

Finally, we get $(M - \rho PM^\top)\boldsymbol{q} \sim \mathcal{N}(0, \Sigma)$ with

$$\Sigma \;=\; \frac{1+\rho}{2}\left((I - \rho P)^2 + \boldsymbol{q}\boldsymbol{q}^\top\right) + \frac{1-\rho}{2}\left((I + \rho P)^2 - \boldsymbol{q}\boldsymbol{q}^\top\right) \;=\; I - \rho^2 P + \rho \boldsymbol{q}\boldsymbol{q}^\top,$$

which yields the desired result. □

4.2. **Proof of Theorem 1: notations and some preparation.** We introduce hereafter notations used throughout the proof.

Conditioning. The conditional equality in distribution of two random variables X and Y given a σ-field $\mathcal{F}$ will be denoted as $X \stackrel{\mathcal{L}}{=}_{|\mathcal{F}} Y$. Formally

$$X \stackrel{\mathcal{L}}{=}_{|\mathcal{F}} Y \qquad \text{iff} \qquad \mathbb{E}[\varphi(X)\,|\,\mathcal{F}] \;=\; \mathbb{E}[\varphi(Y)\,|\,\mathcal{F}] \quad (a.s.)$$

for every non-negative measurable function φ.

The conditionnal independence of X and Y givent $\mathcal{F}$ will be denoted by $X \perp\!\!\!\perp_{|\mathcal{F}} Y$. Formally,

$$X \perp\!\!\!\perp_{|\mathcal{F}} Y \qquad \text{iff} \qquad \mathbb{E}\left[\varphi(X)\psi(Y)|\,\mathcal{F}\right] \;=\; \mathbb{E}\left[\varphi(X)|\,\mathcal{F}\right]\mathbb{E}\left[\psi(Y)|\,\mathcal{F}\right] \;(a.s.)$$

for all non-negative measurable functions φ and ψ.

The following lemma will be of use later.

Lemma 4.4. *Let $\mathcal{F}$, $\mathcal{G}$ be two σ-fields and $Y, X, \overline{X}$ be random variables. Suppose that (i) Y is $\mathcal{F}$-measurable, (ii) $X \stackrel{\mathcal{L}}{=} \overline{X}$ and (iii) $\overline{X} \perp\!\!\!\perp \mathcal{F}$. Suppose moreover that for some measurable function φ:*

$$(iv)\ \varphi\left(X, Y\right) \stackrel{\mathcal{L}}{=}_{|\mathcal{F}} \varphi\left(\overline{X}, Y\right) \quad \textit{and} \quad (v)\ \varphi(X, Y) \perp\!\!\!\perp_{|\mathcal{F}} \mathcal{G}\,.$$

Let $\mathcal{H} = \sigma(\mathcal{F} \cup \mathcal{G})$, then there exists a random variable $\widetilde{X}$ such that $\widetilde{X} \overset{\mathcal{L}}{=} X$, $\widetilde{X} \perp\!\!\!\perp \mathcal{H}$ and

$$\varphi(X,Y) \overset{\mathcal{L}}{=}_{|\mathcal{H}} \varphi(\widetilde{X},Y).$$

Proof. We skip all the integrability issues. Let Z be an $\mathcal{H}$-measurable random variable, and ψ any measurable function. Let $\Phi = \psi \circ \varphi$, we have

$$\begin{aligned}
\mathbb{E}\left[Z\Phi\left(X,Y\right)\right] \;=\; \mathbb{E}\left[\mathbb{E}\left[Z\Phi\left(X,Y\right) \mid \mathcal{F}\right]\right] &\overset{(a)}{=} \mathbb{E}\left[\mathbb{E}\left[Z \mid \mathcal{F}\right]\mathbb{E}\left[\Phi\left(X,Y\right) \mid \mathcal{F}\right]\right], \\
&\overset{(b)}{=} \mathbb{E}\left[\mathbb{E}\left[Z \mid \mathcal{F}\right]\mathbb{E}\left[\Phi\left(\overline{X},Y\right) \mid \mathcal{F}\right]\right], \\
&\overset{(c)}{=} \mathbb{E}\left[\mathbb{E}\left[Z \mid \mathcal{F}\right]\mathbb{E}_{\overline{X}}\left[\Phi\left(\overline{X},Y\right)\right]\right],
\end{aligned}$$

where (a) follows from the fact that $\varphi(X,Y) \perp\!\!\!\perp_{\mathcal{F}} \mathcal{G}$ implies $\varphi(X,Y) \perp\!\!\!\perp_{\mathcal{F}} \mathcal{H}$, hence $\varphi(X,Y) \perp\!\!\!\perp_{\mathcal{F}} Z$ (see for instance [Kal02, Corollary 6.7]), (b) follows from assumption (ii) and (c) from (iii).

Consider now a r.v. $\widetilde{X}$ such that $\widetilde{X} \overset{\mathcal{L}}{=} X$ and $\widetilde{X} \perp\!\!\!\perp \mathcal{H}$ then $\mathbb{E}_{\overline{X}}\Phi(\overline{X},Y) = \mathbb{E}_{\widetilde{X}}\Phi(\widetilde{X},Y)$ and

$$\mathbb{E}\left[Z\Phi\left(X,Y\right)\right] \;=\; \mathbb{E}\left[\mathbb{E}\left(Z \mid \mathcal{F}\right)\mathbb{E}_{\widetilde{X}}\Phi\left(\widetilde{X},Y\right)\right] \;=\; \mathbb{E}\left[Z\,\mathbb{E}_{\widetilde{X}}\Phi\left(\tilde{X},Y\right)\right] \overset{(d)}{=} \mathbb{E}\left[Z\Phi\left(\widetilde{X},Y\right)\right],$$

where (d) follows from Fubini's theorem. This completes the proof. □

Notational shortcuts. The following notations, related to the AMP iterations, will be of constant use in the sequel.

$$\begin{aligned}
&\boldsymbol{q}^k \triangleq h_k\left(\boldsymbol{u}_n^k, B_n\right) \;\in \mathbb{R}^n && \text{for } k \geq 0 \quad \text{and } \boldsymbol{q}^{-1} = \mathbf{0}_n\,, \\
&Q_k \triangleq \left[\boldsymbol{q}^0, \cdots, \boldsymbol{q}^{k-1}\right] \;\in \mathbb{R}^{n\times k} && \text{for } k \geq 1\,, \\
&U_k \triangleq \left[\boldsymbol{u}^1, \cdots, \boldsymbol{u}^k\right] \;\in \mathbb{R}^{n\times k} && \text{for } k \geq 1\,, \\
&d_k \triangleq \left\langle \partial_1 h_k\left(\boldsymbol{u}_n^k, B_n\right)\right\rangle_n \in \mathbb{R} && \text{for } k \geq 1, \quad \text{and } d_0 = 0\,.
\end{aligned}$$

Using these notations, Eq. (6) can be written in the following compact form

$$\boldsymbol{u}^{k+1} = A\boldsymbol{q}^k - \rho d_k \boldsymbol{q}^{k-1}, \quad k \geq 0. \tag{21}$$

Projections. Denote by $\Pi_{\text{span}\,Q}$ the orthogonal projection matrix on the column span of the matrix Q, and as $\Pi_{\text{span}\,Q^\perp}$ the orthogonal projection matrix on the orthogonal of the latter subspace. It is well-known that

$$\Pi_{\text{span}\,Q_k} = Q_k\left(Q_k^\top Q_k\right)^\dagger Q_k^\top\,,$$

where $A^\dagger$ represents a pseudo-inverse of matrix A.

For $k \geq 1$, we introduce the notations

$$\begin{cases} P_k \;\triangleq \Pi_{\text{span}\,Q_k} \;= Q_k\left(Q_k^\top Q_k\right)^\dagger Q_k^\top \\ P_k^\perp \triangleq \Pi_{\text{span}\,Q_k{}^\perp} = I_n - P_k \end{cases} \qquad \text{and} \qquad \boldsymbol{\alpha}^k \;\triangleq\; \left(Q_k^\top Q_k\right)^\dagger Q_k^\top \boldsymbol{q}^k = \begin{bmatrix} \alpha_1^k \\ \vdots \\ \alpha_k^k \end{bmatrix} \in \mathbb{R}^k.$$

By convention, $P_0 = 0_{n\times n}$ and $P_0^\perp = I_n$.

Filtration. We define the filtration $(\mathcal{F}_k)_{k\geq 1}$ by

$$\mathcal{F}_k \triangleq \sigma\left(B, \boldsymbol{u}^0, \boldsymbol{u}^1, \cdots, \boldsymbol{u}^k\right)$$

Preparing the induction. Recall the definitions of $\boldsymbol{q}^k$, Q_k, U_k, P_k and $\boldsymbol{\alpha}^k$ introduced above. The first step of the proof is to establish the following structural result.

Proposition 4.5. *Let $k, \ell \geq 1$ and define the vectors*

$$\boldsymbol{v}^{k,\ell} \triangleq U_k^\top \boldsymbol{q}^\ell - d_\ell Q_k^\top \boldsymbol{q}^{\ell-1} \quad \text{and} \quad \mathcal{I}_k(A) \triangleq (A - \rho P_k A^\top) P_k^\perp \boldsymbol{q}^k \,.$$

Then $\boldsymbol{u}^1 = A\boldsymbol{q}^0 = \mathcal{I}_0(A)$ and

$$\boldsymbol{u}^{k+1} \quad = \quad \sum_{\ell=1}^k \alpha_\ell^k \boldsymbol{u}^\ell + \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \left(\boldsymbol{v}^{k,k} - \sum_{\ell=1}^k \alpha_\ell^k \boldsymbol{v}^{k,\ell-1} \right) + \mathcal{I}_k(A) \,, \quad (k \geq 1) \,.$$

Proof of Proposition 4.5 is postponed to Appendix C.

Remark 4.6 (mesurability issues)**.** *Consider the decomposition of $\boldsymbol{u}^{k+1}$ in Proposition 4.5 then*

$$\sum_{\ell=1}^k \alpha_\ell^k \boldsymbol{u}^\ell + \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \left(\boldsymbol{v}^{k,k} - \sum_{\ell=1}^k \alpha_\ell^k \boldsymbol{v}^{k,\ell-1} \right),$$

is $\mathcal{F}_k$–measurable while in general the term $\mathcal{I}_k(A)$ is not. The strategy developed by Bolthausen amounts to replace matrix A in $\mathcal{I}_k(A)$ by some matrix $\widetilde{A} \perp\!\!\!\perp_{\mathcal{F}_k} A$ before proceeding to the induction. This is the goal of next section.

4.3. Proof of Theorem 1: adaptation of Bolthausen conditioning argument.

Proposition 4.7. *For $k \geq 1$, there exists a $n \times n$ matrix $\widetilde{A}$ such that $\widetilde{A} \overset{\mathcal{L}}{=} A$, $\widetilde{A} \perp\!\!\!\perp \mathcal{F}_k$ and*

$$(A - \rho P_k A^\top) P_k^\perp \overset{\mathcal{L}}{=}_{|\mathcal{F}_k} (\widetilde{A} - \rho P_k \widetilde{A}^\top) P_k^\perp \,.$$

In particular,

$$\mathcal{I}_k(A) \overset{\mathcal{L}}{=}_{|\mathcal{F}_k} \mathcal{I}_k(\widetilde{A}) \,.$$

Proposition 4.7, the proof of which is postponed to the end of Section 4.3, is a consequence of a more general result stated in Proposition 4.8.

Recall that $Q_k = [\boldsymbol{q}^0, \cdots, \boldsymbol{q}^{k-1}]$ and that $P_k = \Pi_{\operatorname{span} Q_k} = Q_k \left(Q_k^\top Q_k\right)^\dagger Q_k^\top$. Denote by

$$\overset{+}{\boldsymbol{q}}{}^k \triangleq P_k^\perp \boldsymbol{q}^k \,. \tag{22}$$

Notice that $\overset{+}{\boldsymbol{q}}{}^k$ is $\mathcal{F}_k$-measurable. Let r_k be the rank of the matrix Q_k - notice that r_k is $\mathcal{F}_{k-1}$–measurable. Let O_k a $\mathcal{F}_{k-1}$–measurable $n \times (n - r_k)$ matrix which columns form an orthonormal basis of $\operatorname{span} Q_k^\perp$. Such a matrix exists: for instance, consider $\boldsymbol{q}^0, \cdots, \boldsymbol{q}^{k-1}$ and the deterministic canonical base $(\boldsymbol{e}^\ell)_{\ell \in [n]}$ of $\mathbb{R}^n$ and construct by Gram-Schmidt procedure an orthonormal basis of $\mathbb{R}^n$ whose first r_k vectors span Q_k. Build O_k out of the remaining $n - r_k$ vectors. In particular,

$$O_k O_k^\top = P_k^\perp \qquad \text{and} \qquad O_k^\top O_k = I_{n-r_k} \,.$$

Proposition 4.8. *For every $k \geq 1$, it holds that*

$$\begin{bmatrix} (A - \rho A^\top) O_k \\ O_k^\top A^\top O_k \end{bmatrix} \overset{\mathcal{L}}{=}_{|\mathcal{F}_k} \begin{bmatrix} (\widetilde{A} - \rho \widetilde{A}^\top) O_k \\ O_k^\top \widetilde{A}^\top O_k \end{bmatrix},$$

where $\widetilde{A} \overset{\mathcal{L}}{=} A$ and $\widetilde{A}$ is independent of $\mathcal{F}_k$.

Proof. For $A \in \mathbb{R}^{n \times n}$ and $O \in \mathbb{R}^{n \times n'}$ $(n' \geq 1)$, let

$$\mathcal{K}(A, O) = \begin{bmatrix} (A - \rho A^\top) O \\ O^\top A^\top O \end{bmatrix} .$$

We prove the statement by induction on $k \geq 1$ and begin by proving it for $k = 1$. Notice that $Q_1 = [\boldsymbol{q}^0]$, O_1 has dimension $n \times (n-1)$ and $(\boldsymbol{q}^0)^\top O_1 = 0$. Recall that $\mathcal{F}_1 = \sigma\left(B, \boldsymbol{u}^0, \boldsymbol{u}^1\right)$ and $\boldsymbol{u}^1 = A\boldsymbol{q}^0$. Taking into account the fact that $A \perp\!\!\!\perp \mathcal{F}_0$ and applying Proposition 4.2 with P (in the proposition) equal to zero, we have:

$$\mathcal{K}(A, O_1) \quad \perp\!\!\!\perp_{|\mathcal{F}_0} \quad \boldsymbol{u}^1 \,.$$

Now consider $\overline{A}$ independent from all the considered quantities, then

$$\mathcal{K}(A, O) \overset{\mathcal{L}}{=}_{|\mathcal{F}_0} \mathcal{K}(\overline{A}, O)$$

We can now apply Lemma 4.4 to prove the existence of $\widetilde{A}$ independent of $\mathcal{F}_1 = \sigma(\mathcal{F}_0, \{\boldsymbol{u}_1\})$ satisfying

$$\mathcal{K}(A, O_1) \overset{\mathcal{L}}{=}_{|\mathcal{F}_1} \mathcal{K}(\widetilde{A}, O_1).$$

The statement is proved for $k = 1$. Suppose now that

$$\mathcal{K}(A, O_{k-1}) \overset{\mathcal{L}}{=}_{|\mathcal{F}_{k-1}} \mathcal{K}(\widetilde{A}, O_{k-1}) \tag{23}$$

where $\widetilde{A}$ is independent of $\mathcal{F}_{k-1}$, and let us prove that this equality holds for k. Notice that one can assume that $\widetilde{A}$ is independent of $\mathcal{F}_k$, since this does not change the conditional distribution of $\mathcal{K}(\widetilde{A}, O_{k-1})$ in (23).

Recall that $\overset{+}{\boldsymbol{q}}{}^k = P_k^{\perp} \boldsymbol{q}^k$, and observe that the event

$$E_k \triangleq \{r_k = r_{k-1} + 1\} \in \mathcal{F}_{k-1}$$

coincides with the event $\{\overset{+}{\boldsymbol{q}}{}^{k-1} \neq 0\}$. Define the matrix W_k such that $W_k \triangleq \left[\overset{+}{\boldsymbol{q}}{}^{k-1} \,|\, O_k\right]$ on E_k and $W_k \triangleq O_k$ on E_k^c. The random matrix W_k is $\mathcal{F}_{k-1}$-measurable, so is $O_{k-1}^{\top} W_k$. Moreover, $P_{k-1}^{\perp} W_k = W_k$. Write

$$\begin{aligned}\begin{bmatrix}(A - \rho A^{\top}) W_k \\ W_k^{\top} A^{\top} W_k\end{bmatrix} &= \begin{bmatrix}(A - \rho A^{\top}) P_{k-1}^{\perp} W_k \\ (P_{k-1}^{\perp} W_k)^{\top} A^{\top} P_{k-1}^{\perp} W_k\end{bmatrix} \\ &\overset{(a)}{=} \begin{bmatrix}(A - \rho A^{\top}) O_{k-1} \left(O_{k-1}^{\top} W_k\right) \\ \left(O_{k-1}^{\top} W_k\right)^{\top} O_{k-1}^{\top} A^{\top} O_{k-1} \left(O_{k-1}^{\top} W_k\right)\end{bmatrix},\end{aligned}$$

where equality (a) holds because $P_{k-1}^{\perp} W_k = W_k$, then using the induction hypothesis (23), we get

$$\begin{bmatrix}(A - \rho A^{\top}) W_k \\ W_k^{\top} A^{\top} W_k\end{bmatrix} \overset{\mathcal{L}}{=}_{|\mathcal{F}_{k-1}} \begin{bmatrix}(\widetilde{A} - \rho \widetilde{A}^{\top}) W_k \\ W_k^{\top} \widetilde{A}^{\top} W_k\end{bmatrix},$$

Substituting by the expression of W_k, we have proved that:

$$\begin{bmatrix}(A - \rho A^{\top}) \overset{+}{\boldsymbol{q}}{}^{k-1} & (A - \rho A^{\top}) O_k \\ (\overset{+}{\boldsymbol{q}}{}^{k-1})^{\top} A^{\top} \overset{+}{\boldsymbol{q}}{}^{k-1} & (\overset{+}{\boldsymbol{q}}{}^{k-1})^{\top} A^{\top} O_k \\ O_k^{\top} A^{\top} \overset{+}{\boldsymbol{q}}{}^{k-1} & O_k^{\top} A^{\top} O_k\end{bmatrix} \mathbb{1}_{E_k} \overset{\mathcal{L}}{=}_{|\mathcal{F}_{k-1}} \begin{bmatrix}(\widetilde{A} - \rho \widetilde{A}^{\top}) \overset{+}{\boldsymbol{q}}{}^{k-1} & (\widetilde{A} - \rho \widetilde{A}^{\top}) O_k \\ (\overset{+}{\boldsymbol{q}}{}^{k-1})^{\top} \widetilde{A}^{\top} \overset{+}{\boldsymbol{q}}{}^{k-1} & (\overset{+}{\boldsymbol{q}}{}^{k-1})^{\top} \widetilde{A}^{\top} O_k \\ O_k^{\top} \widetilde{A}^{\top} \overset{+}{\boldsymbol{q}}{}^{k-1} & O_k^{\top} \widetilde{A}^{\top} O_k\end{bmatrix} \mathbb{1}_{E_k}, \tag{24}$$

and

$$\begin{bmatrix}(A - \rho A^{\top}) O_k \\ O_k^{\top} A^{\top} O_k\end{bmatrix} \mathbb{1}_{E_k^c} \overset{\mathcal{L}}{=}_{|\mathcal{F}_{k-1}} \begin{bmatrix}(\widetilde{A} - \rho \widetilde{A}^{\top}) O_k \\ O_k^{\top} \widetilde{A}^{\top} O_k\end{bmatrix} \mathbb{1}_{E_k^c}. \tag{25}$$

Recall that $\mathcal{F}_k = \sigma\left(\mathcal{F}_{k-1}, \{\boldsymbol{u}^k\}\right)$, let us study the quantity $\boldsymbol{u}^k$.

$$\begin{aligned}\boldsymbol{u}^k &= A\boldsymbol{q}^{k-1} - d_{k-1}\boldsymbol{q}^{k-2} \\ &= AP_{k-1}\boldsymbol{q}^{k-1} + AP_{k-1}^{\perp}\boldsymbol{q}^{k-1} - d_{k-1}\boldsymbol{q}^{k-2} \\ &= AP_{k-1}\boldsymbol{q}^{k-1} + \rho\left(AP_{k-1}\right)^{\top} \overset{+}{\boldsymbol{q}}{}^{k-1} - d_{k-1}\boldsymbol{q}^{k-2} + \left(A - \rho P_{k-1} A^{\top}\right) \overset{+}{\boldsymbol{q}}{}^{k-1}.\end{aligned}$$

We can re-write this expression as

$$\boldsymbol{u}^k = \boldsymbol{m}^{k-1} + \boldsymbol{z}^k,$$

with

$$\boldsymbol{z}^k \triangleq \left(A - \rho P_{k-1} A^{\top}\right) \overset{+}{\boldsymbol{q}}{}^{k-1} \quad \text{and} \quad \boldsymbol{m}^{k-1} \text{ is } \mathcal{F}_{k-1}\text{-measurable.}$$

We now want to prove that

$$\boldsymbol{u}^k \perp\!\!\!\perp_{|\mathcal{F}_{k-1}} \mathcal{K}(A, O_k), \tag{26}$$

which is equivalent to

$$\boldsymbol{z}^k \perp\!\!\!\perp_{|\mathcal{F}_{k-1}} \mathcal{K}(A, O_k), \tag{27}$$

which can also be reduced into two smaller problems (see, e.g., [FVRS21, Lemma 7.9.a]),

$$\boldsymbol{z}^k \mathbb{1}_{E_k} \perp\!\!\!\perp_{|\mathcal{F}_{k-1}} \mathcal{K}(A, O_k)\mathbb{1}_{E_k}, \tag{28}$$

$$\boldsymbol{z}^k \mathbb{1}_{E_k^c} \perp\!\!\!\perp_{|\mathcal{F}_{k-1}} \mathcal{K}(A, O_k)\mathbb{1}_{E_k^c}. \tag{29}$$

Let us begin by showing (28). By Equation (24) we have the following equality of joint laws,

$$\left(\boldsymbol{z}^k \mathbb{1}_{E_k},\ \mathcal{K}(A, O_k)\mathbb{1}_{E_k}\right) \overset{\mathcal{L}}{=}_{|\mathcal{F}_{k-1}} \left(\boldsymbol{z}^k \mathbb{1}_{E_k},\ \mathcal{K}(\widetilde{A}, O_k)\mathbb{1}_{E_k}\right),$$

so in order to show (28), it suffices to show that

$$\boldsymbol{z}^k \mathbb{1}_{E_k} \perp\!\!\!\perp_{|\mathcal{F}_{k-1}} \mathcal{K}(\widetilde{A}, O_k)\mathbb{1}_{E_k}.$$

But since $\widetilde{A}$ is independent of $\mathcal{F}_{k-1}$, this is a direct consequence of Proposition 4.2. Similarly, by noticing that $\boldsymbol{z}^k \mathbb{1}_{E_k^c} = 0$ and using Equation (25) and the fact that 0 is independent of any gaussian vector we can prove (29). Now using (26) and the induction hypothesis (23) we can finally apply Lemma 4.4 to complete the proof. □

We are now in position to prove Proposition 4.7.

Proof of Proposition 4.7. To prove Proposition 4.7 using Proposition 4.8, we write $(A-\rho P_k A^\top)P_k^\perp = (A - \rho A^\top)O_k O_k^\top + \rho O_k O_k^\top A^\top O_k O_k^\top$, and we use Proposition 4.8 along with the following well-known result (see, e.g., [FVRS21, Lemma 7.6.c]): If X, X', and Y are random vectors on a probability space, and $\mathcal{F}$ is a σ–field on this space such that $X \overset{\mathcal{L}}{=}_{|\mathcal{F}} X'$ and that Y is $\mathcal{F}$–measurable, then, for each measurable function φ, it holds that $\varphi(X, Y) \overset{\mathcal{L}}{=}_{|\mathcal{F}} \varphi(X', Y)$. □

Taking advantage of Proposition 4.7, we can improve Proposition 4.5 by replacing $\mathcal{I}_k(A)$ by $\mathcal{I}_k(\widetilde{A})$. We also replace some random quantities by their deterministic equivalents.

Recall the definition of matrices R^k given by the Density Evolution equations (3), we define three related quantities $\sigma_k^2 \in \mathbb{R}^+$, $\mathring{\sigma}_k \in \mathbb{R}^+$ and $\bar{\boldsymbol{\alpha}}^k \in \mathbb{R}^k$, such as

$$\sigma_k^2 = R_{k,k}\,, \tag{30}$$

$$\mathring{\sigma}_{k+1}^2 = \sigma_{k+1}^2 - \left(R^{k+1}_{[k],k+1}\right)^\top \left(R^k\right)^{-1} \left(R^{k+1}_{[k],k+1}\right), \tag{31}$$

$$\bar{\boldsymbol{\alpha}}^k = \left(R^k\right)^{-1} R^{k+1}_{[k],k+1}\,. \tag{32}$$

Remark 4.9. *Notice that $\mathring{\sigma}_{k+1}^2$ is the Schur complement of R^k in the matrix R^{k+1}. One should think of $\mathring{\sigma}_k$ and $\bar{\boldsymbol{\alpha}}^k$ as the deterministic equivalents of $\frac{\|\mathring{\boldsymbol{q}}^k\|}{\sqrt{n}}$ and $\boldsymbol{\alpha}^k$ respectively when n is large.*

Proposition 4.10. *Using the previous notations we have the following decomposition of $\boldsymbol{u}^{k+1}$.*

$$\begin{aligned}\boldsymbol{u}^{k+1} &\overset{\mathcal{L}}{=}_{|\mathcal{F}_k} \sum_{\ell=1}^k \alpha_\ell^k \boldsymbol{u}^\ell + \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \left(\boldsymbol{v}^{k,k} - \sum_{\ell=1}^k \alpha_\ell^k \boldsymbol{v}^{k,\ell-1}\right) + (\tilde{A} - \rho P_k \tilde{A}^\top)\mathring{\boldsymbol{q}}^k \\ &\overset{\mathcal{L}}{=}_{|\mathcal{F}_k} \bar{\boldsymbol{u}}^{k+1} + \Delta^{k+1},\end{aligned}$$

where

$$\bar{\boldsymbol{u}}^{k+1} \triangleq \sum_{\ell=1}^k \bar{\alpha}_\ell^k \boldsymbol{u}^\ell + \mathring{\sigma}_k \boldsymbol{\xi}^{k+1}\,, \tag{33}$$

and Δ^{k+1} is defined on the events $E_k \triangleq \{\mathring{\boldsymbol{q}}^k \neq \boldsymbol{0}\}$ and $E_k^c = \{\mathring{\boldsymbol{q}}^k = \boldsymbol{0}\}$ by

$$\begin{aligned}\Delta^{k+1} \triangleq\ & \sum_{\ell=1}^k \left(\alpha_\ell^k - \bar{\alpha}_\ell^k\right) \boldsymbol{u}^\ell + \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \left(\boldsymbol{v}^{k,k} - \sum_{\ell=1}^k \alpha_\ell^k \boldsymbol{v}^{k,\ell-1}\right) \\ &+ \left(\frac{\|\mathring{\boldsymbol{q}}^k\|}{\sqrt{n}} - \mathring{\sigma}_k\right) \boldsymbol{\xi}^{k+1} + \left(\sqrt{1-\rho^2} - 1\right) \frac{\|\mathring{\boldsymbol{q}}^k\|}{\sqrt{n}} P_k \boldsymbol{\xi}^{k+1} + \left(\sqrt{1+\rho} - 1\right) \frac{\mathring{\boldsymbol{q}}^k (\mathring{\boldsymbol{q}}^k)^\top}{\sqrt{n}\|\mathring{\boldsymbol{q}}^k\|} \boldsymbol{\xi}^{k+1}\end{aligned}$$

and

$$\Delta^{k+1} \triangleq \sum_{\ell=1}^{k} \left(\alpha_\ell^k - \bar{\alpha}_\ell^k\right) \boldsymbol{u}^\ell + \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \left(\boldsymbol{v}^{k,k} - \sum_{\ell=1}^{k} \alpha_\ell^k \boldsymbol{v}^{k,\ell-1}\right),$$

respectively.

A very similar result is obtained in [FVRS21, Section 6.4].

Remark 4.11. *The aim of this proposition is to approximate the asymptotic behavior of the distribution of the iterates $\boldsymbol{u}^{k+1}$ with the distribution of $\bar{\boldsymbol{u}}^{k+1}$ which is easier to handle provided that $\bar{\alpha}_\ell^k$ and $\mathring{\sigma}_k$ are deterministic quantities. This is achieved by proving that the difference Δ^{k+1} is asymptotically negligible.*

Proof. We only address the case where $\mathring{\boldsymbol{q}}^k \neq 0$. In the other case, the term $(\tilde{A} - \rho P_k \tilde{A}^\top)\mathring{\boldsymbol{q}}^k = 0$ does not need to be handled. The starting point is the decomposition of $\boldsymbol{u}^{k+1}$ in Proposition 4.5. By Proposition 4.3, the conditional distribution of $(\widetilde{A} - \rho P_k \widetilde{A}^\top)P_k^\perp \boldsymbol{q}^k = (\widetilde{A} - \rho P_k \widetilde{A}^\top)\mathring{\boldsymbol{q}}^k$ given $\mathcal{F}_k$ is

$$\mathcal{L}\left((\widetilde{A} - \rho P_k \widetilde{A}^\top)\mathring{\boldsymbol{q}}^k \,|\, \mathcal{F}_k\right) = \mathcal{N}\left(0, \frac{1}{n}\|\mathring{\boldsymbol{q}}^k\|^2 \left(I - \rho^2 P_k + \rho \mathring{\boldsymbol{q}}^k (\mathring{\boldsymbol{q}}^k)^\top / \|\mathring{\boldsymbol{q}}^k\|^2\right)\right).$$

Therefore, letting $\boldsymbol{\xi}^{k+1} \sim \mathcal{N}(0, I_n)$ be independent of $\mathcal{F}_k$, it holds by Proposition 4.7 that

$$(A - \rho P_k A^\top)P_k^\perp \boldsymbol{q}^k \overset{\mathcal{L}}{=}_{|\mathcal{F}_k} \frac{1}{\sqrt{n}}\|\mathring{\boldsymbol{q}}^k\| \left(I - \rho^2 P_k + \rho \mathring{\boldsymbol{q}}^k (\mathring{\boldsymbol{q}}^k)^\top / \|\mathring{\boldsymbol{q}}^k\|^2\right)^{1/2} \boldsymbol{\xi}^{k+1}.$$

It is clear that we can take

$$\left(I - \rho^2 P_k + \rho \mathring{\boldsymbol{q}}^k (\mathring{\boldsymbol{q}}^k)^\top / \|\mathring{\boldsymbol{q}}^k\|^2\right)^{1/2} = I + \left(\sqrt{1-\rho^2} - 1\right) P_k + \left(\sqrt{1+\rho} - 1\right) \mathring{\boldsymbol{q}}^k (\mathring{\boldsymbol{q}}^k)^\top / \|\mathring{\boldsymbol{q}}^k\|^2.$$

The proof of the proposition is completed. □

4.4. **Proof of Theorem 1: end of proof.** The remainder of the proof of Theorem 1 follows almost verbatim the proof provided by [FVRS21, Section 6]. For completeness, we provide here the important steps of the proof without rigorously justifying all technical details.

We will proceed by induction on $k \geq 1$. Suppose that for any pseudo-Lipschitz function φ,

$$(\mathfrak{H}_k) \qquad \frac{1}{n}\sum_{i=1}^{n} \varphi(B_{i,*}, u_i^1, \cdots, u_i^k) \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[\varphi(\bar{\boldsymbol{b}}, Z_1, \cdots, Z_k)\right],$$

and we want to prove this same convergence for $k+1$.
To do this, fix any pseudo-Lipschitz function ψ and consider the following random variable

$$S = \frac{1}{n}\sum_{i=1}^{n} \psi(B_{i,*}, u_i^1, \cdots, u_i^k, u_i^{k+1}).$$

Recall the definition (33) of $\bar{\boldsymbol{u}}^{k+1}$ and write S as follows:

$$\begin{aligned} S &\overset{\mathcal{L}}{=}_{|\mathcal{F}_k} \frac{1}{n}\sum_{i=1}^{n} \psi(B_{i,*}, u_i^1, \cdots, u_i^k, \bar{u}_i^{k+1}) \\ &\quad + \frac{1}{n}\left(\sum_{i=1}^{n} \psi(B_{i,*}, u_i^1, \cdots, u_i^k, u_i^{k+1}) - \sum_{i=1}^{n} \psi(B_{i,*}, u_i^1, \cdots, u_i^k, \bar{u}_i^{k+1})\right), \end{aligned}$$

where $\boldsymbol{u}^{k+1} \overset{\mathcal{L}}{=}_{|\mathcal{F}_k} \bar{\boldsymbol{u}}^{k+1} + \Delta^{k+1}$. The idea will be then to prove that

$$(34) \quad S_1 \triangleq \frac{1}{n}\sum_{i=1}^{n} \psi(B_{i,*}, u_i^1, \cdots, u_i^k, \bar{u}_i^{k+1}) \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[\psi(\bar{\boldsymbol{b}}, Z_1, \cdots, Z_k)\right],$$

$$(35) \quad S_2 \triangleq \frac{1}{n}\left(\sum_{i=1}^{n} \psi(B_{i,*}, u_i^1, \cdots, u_i^k, u_i^{k+1}) - \sum_{i=1}^{n} \psi(B_{i,*}, u_i^1, \cdots, u_i^k, \bar{u}_i^{k+1})\right) \xrightarrow[n\to+\infty]{c} 0.$$

Let us begin by proving the convergence in (34). This proof can be decomposed into two steps, the first step is to prove that the conditional expectation $\mathbb{E}[S_1 \mid \mathcal{F}_k]$ converges to the desired limit,

and the second step (which is omitted) is to show that S_1 concentrates around its conditional expectation.

In the sequel, we shall rely on the following notation. Let X a random variable and $X \perp\!\!\!\perp Y$. We denote by $\mathbb{E}_X$ the expectation with respect to the distribution of X. In particular,

$$\mathbb{E}_X f(X,Y) = \int f(x,Y)\mathbb{P}_X(dx) = \mathbb{E}(f(X,Y) \mid Y)\,.$$

Let us compute the conditional expectation of S_1 given $\mathcal{F}_k$.

$$\begin{aligned}
\mathbb{E}\left[S_1 \mid \mathcal{F}_k\right] &= \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n \psi\left(B_{i,*}, u_i^1, \cdots, u_i^k, \sum_{l=1}^k \bar{\alpha}_l^k u_i^l + \mathring{\sigma}_{k+1}\xi_i^{k+1}\right)\middle|\ \mathcal{F}_k\right] \\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{E}_{\xi_i^{k+1}}\left[\psi\left(B_{i,*}, u_i^1, \cdots, u_i^k, \sum_{l=1}^k \bar{\alpha}_l^k u_i^l + \mathring{\sigma}_{k+1}\xi_i^{k+1}\right)\right] \\
&\triangleq \frac{1}{n}\sum_{i=1}^n \Psi\left(B_{i,*}, u_i^1, \cdots, u_i^k\right).
\end{aligned}$$

By [FVRS21, Lemma 7.23], Ψ is also a pseudo-lipschitz function, thus using the induction hypothesis ($\mathfrak{H}_k$) we can write

$$\frac{1}{n}\sum_{i=1}^n \Psi\left(B_{i,*}, u_i^1, \cdots, u_i^k\right) \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[\Psi\left(Z_1, \cdots, Z_k\right)\right].$$

Now, given a random variable $\tilde{Z} \sim \mathcal{N}(0,1)$ independent of $\mathcal{F}_k$ we can write

$$\mathbb{E}\left[\Psi\left(Z_1, \cdots, Z_k\right)\right] = \mathbb{E}\left[\mathbb{E}_{\tilde{Z}}\left[\psi\left(Z_1, \cdots, Z_k, \sum_{\ell=1}^k \bar{\alpha}_\ell^k Z_\ell + \mathring{\sigma}_{k+1}\tilde{Z}\right)\right]\right],$$

Put $Z_{k+1} \triangleq \sum_{\ell=1}^k \bar{\alpha}_\ell^k Z_\ell + \mathring{\sigma}_{k+1}\tilde{Z}$, and observe that $(Z_1, \cdots, Z_{k+1}) \sim \mathcal{N}\left(0, R^{k+1}\right)$, then

$$\mathbb{E}\left[\Psi\left(Z_1, \cdots, Z_k\right)\right] = \mathbb{E}\left[\psi\left(Z_1, \cdots, Z_{k+1}\right)\right].$$

Now let us give some proof elements for the convergence in (35). The idea is to simply use the pseudo-Lipschitz property and bound the term S_2 by the distance $\|\boldsymbol{u}^k - \bar{\boldsymbol{u}}^k\|$. Thus the main ingredient of this proof is to show that

$$\frac{1}{\sqrt{n}}\|\Delta^{k+1}\| \xrightarrow[n\to+\infty]{c} 0. \tag{36}$$

Recall the expression of Δ^{k+1} in Proposition 4.10 which can be written as the sum of five terms:

$$\Delta^{k+1} = \Delta^{(1)} + \Delta^{(2)} + \Delta^{(3)} + \Delta^{(4)} + \Delta^{(5)}, \tag{37}$$

where

$$\begin{aligned}
\Delta^{(1)} &= \sum_{\ell=1}^k \left(\alpha_\ell^k - \bar{\alpha}_\ell^k\right)\boldsymbol{u}^\ell \\
\Delta^{(2)} &= \rho Q_k\left(Q_k^\top Q_k\right)^\dagger\left(\boldsymbol{v}^{k,k} - \sum_{\ell=1}^k \boldsymbol{\alpha}_\ell^k \boldsymbol{v}^{k,\ell-1}\right) \\
\Delta^{(3)} &= \left(\frac{1}{\sqrt{n}}\|\mathring{\boldsymbol{q}}^k\| - \mathring{\sigma}_k\right)\xi^{k+1} \\
\Delta^{(4)} &= \left(\sqrt{1-\rho^2} - 1\right)\frac{1}{\sqrt{n}}\|\mathring{\boldsymbol{q}}^k\| P_k \xi^{k+1} \\
\Delta^{(5)} &= \left(\sqrt{1+\rho} - 1\right)\frac{\mathring{\boldsymbol{q}}^k(\mathring{\boldsymbol{q}}^k)^\top}{\sqrt{n}\|\mathring{\boldsymbol{q}}^k\|}\boldsymbol{\xi}^{k+1}.
\end{aligned} \tag{38}$$

In order to prove (36), it suffices to show that the normalized norm of each of $\{\Delta^{(j)},\ j \in [5]\}$ converges to 0. The key arguments of this proof can be summarized in the following lemma which are direct consequences of the induction hypothesis ($\mathfrak{H}_k$).

Proposition 4.12. *Let $k \geq 1$ be a fixed integer and suppose that the induction hypothesis ($\mathfrak{H}_k$) is satisfied for rank k, i.e.*

$$\forall \varphi \in PL_2\left(\mathbb{R}^{p+k}\right)\ \frac{1}{n}\sum_{i=1}^{n}\varphi\left(B_{i,*}, u_i^1, \cdots, u_i^k\right) \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[\varphi\left(\bar{\boldsymbol{b}}, Z_1, \ldots, Z_k\right)\right]$$

where $(Z_1, \cdots, Z_k)$ is a centered gaussian vector of covariance matrix $R^k \in \mathbb{R}^{k\times k}$ which is defined recursively using the Density Evolution equations (3). Then we have the following consequences

(c-1) For all $j \leq k$, $\frac{1}{n}\|\boldsymbol{u}^j\|^2 \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[Z_j^2\right]$.

(c-2) For all $j \leq k$, $\frac{1}{n}\|\boldsymbol{q}^j\|^2 \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[h_j\left(Z_j, \bar{\boldsymbol{b}}\right)^2\right]$.

(c-3) For all $i, j \leq k$, $\frac{1}{n}\langle \boldsymbol{q}^{i-1}, \boldsymbol{q}^{j-1}\rangle \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[h_{i-1}\left(Z_{i-1}, \bar{\boldsymbol{b}}\right) h_{j-1}\left(Z_{j-1}, \bar{\boldsymbol{b}}\right)\right] = R^k_{i,j}$.

(c-4) $\frac{1}{n}Q_k^\top Q_k = \left(\frac{1}{n}\langle \boldsymbol{q}^{i-1}, \boldsymbol{q}^{j-1}\rangle\right)_{1\leq i,j\leq k} \xrightarrow[n\to+\infty]{c} R^k$.

(c-5) $\boldsymbol{\alpha}^k = \left(Q_k^\top Q_k\right)^\dagger Q_k^\top q^k \xrightarrow[n\to+\infty]{c} \left(R^k\right)^{-1} R^{k+1}_{[1,k],k+1} = \bar{\boldsymbol{\alpha}}^k$.

(c-6) $\frac{1}{n}\|\overset{+}{\boldsymbol{q}}{}^k\|^2 \xrightarrow[n\to+\infty]{c} \overset{+}{\sigma}{}^2_{k+1}$.

(c-7) For $j \leq k$, $d_j = \frac{1}{n}\sum_{i=1}^n h_j'\left(u_i^j, B_{i,*}\right) \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[h_j'\left(Z_j, \bar{\boldsymbol{b}}\right)\right] \triangleq \bar{d}_j$.

Using this proposition, we can already see that from (c-1) and (c-5) we have $\frac{1}{\sqrt{n}}\|\Delta^{(1)}\| \xrightarrow[n\to+\infty]{c} 0$ and from (c-2) we have $\frac{1}{\sqrt{n}}\|\Delta^{(3)}\| \xrightarrow[n\to+\infty]{c} 0$. The quantities $\Delta^{(4)}$ and $\Delta^{(5)}$ are small rank projections of some gaussian vectors and thus their normalized norms converge to 0. It remains to show that the term

$$\Delta^{(2)} = Q_k\left(\frac{1}{n}Q_k^\top Q_k\right)^\dagger\left(\frac{1}{n}\boldsymbol{v}^{k,k} - \frac{1}{n}\sum_{\ell=1}^{k}\alpha_\ell^k \boldsymbol{v}^{k,\ell-1}\right), \tag{39}$$

has a normalized norm that converges to 0. This can be achieved by showing that

$$\frac{1}{n}\boldsymbol{v}^{k,\ell} \xrightarrow[n\to+\infty]{c} 0.$$

The j-th row of $\frac{1}{n}\boldsymbol{v}^{k,\ell}$ can be written as:

$$\frac{1}{n}v_j^{k,\ell} = \frac{1}{n}\langle \boldsymbol{u}^j, \boldsymbol{q}^\ell\rangle - \frac{1}{n}d_\ell\langle \boldsymbol{q}^{j-1}, \boldsymbol{q}^{\ell-1}\rangle,$$

where $\langle u, v\rangle = u^\top v$. By Proposition 4.12 we have:

- $d_\ell \xrightarrow[n\to+\infty]{c} \bar{d}_\ell$,
- $\frac{1}{n}\langle \boldsymbol{q}^{j-1}, \boldsymbol{q}^{\ell-1}\rangle \xrightarrow[n\to+\infty]{c} R^k_{j,\ell}$,
- $\frac{1}{n}\langle \boldsymbol{u}^j, \boldsymbol{q}^\ell\rangle = \frac{1}{n}\sum_{i=1}^n u_i^j h_\ell(u_i^\ell) \xrightarrow[n\to+\infty]{c} \mathbb{E}\left(Z_j h_\ell(Z_\ell)\right)$.

Using Stein's integration by parts formula and the density evolution equations we get:

$$\begin{aligned}\mathbb{E}\left(Z_j h_\ell(Z_\ell)\right) &= \mathbb{E}\left(Z_j Z_\ell\right)\mathbb{E}\left(h_\ell'(Z_\ell)\right)\\ &= \mathbb{E}\left(h_{j-1}(Z_{j-1})h_{\ell-1}(Z_{\ell-1})\right)\bar{d}_\ell\\ &= \bar{d}_\ell R^k_{j,\ell}.\end{aligned}$$

Which leads us to the desired result $\frac{1}{n}\boldsymbol{v}^{k,\ell} \xrightarrow[n\to+\infty]{c} 0$.
This completes the proof of Theorem 1.

4.5. **Proof of Corollary 2: blockwise convergence of AMP.** Assume (A1), (A2′) and (A3) - (A5). To prove Corollary 2 for a parameter matrix B of size $n \times p$ we will have to use Theorem 1 for an augmented parameter matrix $B' = [B \mid \boldsymbol{s}]$ of size $n \times (p+1)$, where $\boldsymbol{s}$ represents a "selector vector". This explains the utility of presenting our main Theorem 1 using multiple parameter vector $(\boldsymbol{b}^1, \cdots, \boldsymbol{b}^p)$ in (5) instead of a single one as what we usually see in the literature.
Recall the partition defined in (10), and let $\boldsymbol{s} \in \mathbb{R}^n$ be a blockwise constant vector with q different values $\tilde{s}_1, \cdots, \tilde{s}_q$ such that

$$s_i = \tilde{s}_j \quad \text{if and only if} \quad i \in C_n^{(j)} \quad \text{for } i \in [n] \text{ and } j \in [q]. \tag{40}$$

Put $B' = [\boldsymbol{s} \mid B]$ and re-write the AMP iteration defined in (6) as

$$\boldsymbol{u}^{k+1} = A h_k\left(\boldsymbol{u}^k, B'\right) - \rho \left\langle \partial_1 h_k\left(\boldsymbol{u}^k, B'\right)\right\rangle_n h_{k-1}\left(\boldsymbol{u}^{k-1}, B'\right),$$

where we abuse the notation for h_k which will depend only on the first $p+1$ coordinates. In this setting, the Density Evolution equations (8) remain unchanged. To use Theorem 1 we should verify that the parameter matrix B' satisfies Assumption (A2).

Lemma 4.13. *Assume that $(\boldsymbol{u}^0, B)$ satisfies Assumption (A2′) and let $\boldsymbol{s}$ be defined by (40). Then $(\boldsymbol{u}^0, B')$ satisfies Assumption (A2), i.e. there exists a vector $(\bar{u}, \bar{b}_1, \cdots, \bar{b}_p, \bar{s})$ whose distribution belongs to $\mathcal{P}_2\left(\mathbb{R}^{p+2}\right)$ such that*

$$\mu^{\boldsymbol{u}^0, B'} = \mu^{\boldsymbol{u}^0, \boldsymbol{b}^1, \cdots, \boldsymbol{b}^p, \boldsymbol{s}} \xrightarrow[n\to\infty]{\mathcal{P}_r\left(\mathbb{R}^{p+2}\right)} \mathcal{L}\left((\bar{u}, \bar{b}_1, \cdots, \bar{b}_p, \bar{s})\right).$$

In this case $\mathcal{L}\left(\bar{u}, \bar{b}_1, \cdots, \bar{b}_p, \bar{s}\right) = \sum_{j=1}^q c_j \mathcal{L}\left(\bar{u}_j, \bar{b}_{j,1}, \cdots, \bar{b}_{j,p}\right) \otimes \delta_{\tilde{s}_j}$, and in particular

$$\bar{s} \sim \sum_{j=1}^q c_j \delta_{\tilde{s}_j} \quad \bar{u} \sim \sum_{j=1}^q c_j \mathcal{L}(\bar{u}_j), \quad \text{and} \quad \bar{b}_\ell \sim \sum_{j=1}^q c_j \mathcal{L}(\bar{b}_{j,\ell}) \text{ for all } \ell \in [p].$$

Proof. Let Υ be a pseudo-Lipschitz test function, we want to show the existence of $\left(\bar{u}, \bar{s}, \bar{b}_1, \cdots, \bar{b}_p\right)$ such that

$$\frac{1}{n}\sum_{i=1}^n \Upsilon\left(u_i^0, b_i^1, \cdots, b_i^p, s_i\right) \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[\Upsilon\left(\bar{u}, \bar{b}_1, \cdots, \bar{b}_p, \bar{s}\right)\right].$$

The previous sum can be expressed with respect to the partition (10) as

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n \Upsilon\left(u_i^0, b_i^1, \cdots, b_i^p, s_i\right) &= \sum_{j=1}^q \frac{n_j}{n}\frac{1}{n_j}\sum_{i\in C_n^{(j)}} \Upsilon\left(u_i^0, b_i^1, \cdots, b_i^p, s_i\right) \\ &= \sum_{j=1}^q \frac{n_j}{n}\frac{1}{n_j}\sum_{i\in C_n^{(j)}} \Upsilon\left(u_i^0, b_i^1, \cdots, b_i^p, \tilde{s}_j\right).\end{aligned}$$

Thus using Assumption (A2′) with the test function $\Upsilon\left(\cdots, \tilde{s}_j\right)$ yields

$$\frac{1}{n}\sum_{i=1}^n \Upsilon\left(u_i^0, b_i^1, \cdots, b_i^p, s_i\right) \xrightarrow[n\to+\infty]{c} \sum_{j=1}^q c_j \mathbb{E}\left[\Upsilon\left(\bar{u}_j, \bar{b}_{j,1}, \cdots, \bar{b}_{j,p}, \tilde{s}_j\right)\right].$$

Let $\left(\bar{u}, \bar{b}_1, \cdots, \bar{b}_p, \bar{s}\right) \sim \sum_{j=1}^q c_j \mathcal{L}\left(\bar{u}_j, \bar{b}_{j,1}, \cdots, \bar{b}_{j,p}\right) \otimes \delta_{\tilde{s}_j}$, we have

$$\sum_{j=1}^q c_j \mathbb{E}\left[\Upsilon\left(\bar{u}_j, \bar{b}_{j,1}, \cdots, \bar{b}_{j,p}, \tilde{s}_j\right)\right] = \mathbb{E}\left[\Upsilon\left(\bar{u}, \bar{b}_1, \cdots, \bar{b}_p, \bar{s}\right)\right],$$

hence the result. □

We can now apply Theorem 1 which gives the following convergence result

$$\mu^{B', \boldsymbol{u}^1, \cdots, \boldsymbol{u}^k} \xrightarrow[n\to\infty]{\mathcal{P}_r\left(\mathbb{R}^{p+k+1}\right)} \mathcal{L}\left((B', Z_1, \cdots, Z_k)\right) \quad \text{(completely)} \tag{41}$$

Fix j to be an integer in $[q]$, let φ be any pseudo-Lipschitz function on $\mathbb{R}^{p+k}$ and let ψ be a continuous bounded function defined on $\mathbb{R}$ such that:

$$\psi(\tilde{s}_\ell) = \begin{cases} 1 & \text{if } \ell = j, \\ 0 & \text{otherwise.} \end{cases} \tag{42}$$

Finally, consider the test function $\phi \in PL_r\left(\mathbb{R}^{p+k+1}\right)$ defined as

$$\phi\left(x, \boldsymbol{y}\right) = \psi(x)\varphi(\boldsymbol{y}) \quad \forall (x, \boldsymbol{y}) \in \mathbb{R} \times \mathbb{R}^{p+k},$$

and apply the result in (41) to get

$$\frac{1}{n}\sum_{i=1}^{n} \phi\left(\tilde{s}_i, b_i^1, \cdots, b_i^p, u_i^1, \cdots, u_i^k\right) \xrightarrow[n\to+\infty]{c} \mathbb{E}\left[\phi\left(\bar{s}, \bar{b}_1, \cdots, \bar{b}_p, Z_1, \cdots, Z_k\right)\right], \tag{43}$$

where $\left(\bar{s}, \bar{b}_1, \cdots, \bar{b}_p\right)$ is defined as in Lemma 4.13 and $(Z_1, \cdots, Z_k)$ is an independent gaussian vector that satisfies the Density evolution equations (8) that depend only on $\left(\bar{u}, \bar{b}_1, \cdots, \bar{b}_p\right)$. The structure of ψ in (42) implies that the left hand side of (43) can be expressed as

$$\frac{1}{n}\sum_{i=1}^{n} \phi\left(\tilde{s}_i, b_i^1, \cdots, b_i^p, u_i^1, \cdots, u_i^k\right) = \frac{n_j}{n}\frac{1}{n_j}\sum_{i\in C_n^{(j)}} \varphi\left(b_i^1, \cdots, b_i^p, u_i^1, \cdots, u_i^k\right)$$

thus

$$\frac{1}{n_j}\sum_{i\in C_n^{(j)}} \varphi\left(b_i^1, \cdots, b_i^p, u_i^1, \cdots, u_i^k\right) \xrightarrow[n\to+\infty]{c} \frac{1}{c_j}\mathbb{E}\left[\phi\left(\bar{s}, \bar{b}_1, \cdots, \bar{b}_p, Z_1, \cdots, Z_k\right)\right],$$

Now recall the law of the vector $(\bar{s}, \bar{\boldsymbol{b}}) = \left(\bar{s}, \bar{b}_1, \cdots, \bar{b}_p\right)$ which is independent of $\boldsymbol{Z} = (Z_1, \cdots, Z_k)$, we have

$$\begin{aligned}
\mathbb{E}\left[\phi\left(\bar{s}, \bar{b}_1, \cdots, \bar{b}_p, Z_1, \cdots, Z_k\right)\right] &= \mathbb{E}_{(\bar{s},\bar{\boldsymbol{b}})}\left[\mathbb{E}_{\boldsymbol{Z}}\left[\phi\left(\bar{s}, \bar{b}_1, \cdots, \bar{b}_p, Z_1, \cdots, Z_k\right)\right]\right] \\
&= \sum_{\ell=1}^{q} c_\ell \mathbb{E}_{(\bar{\boldsymbol{b}},\boldsymbol{Z})}\left[\phi\left(\tilde{s}_\ell, \bar{b}_{\ell,1}, \cdots, \bar{b}_{\ell,p}, Z_1, \cdots, Z_k\right)\right] \\
&= c_j \mathbb{E}_{(\bar{\boldsymbol{b}},\boldsymbol{Z})}\left[\varphi\left(\bar{b}_{j,1}, \cdots, \bar{b}_{j,p}, Z_1, \cdots, Z_k\right)\right].
\end{aligned}$$

Finally, we get

$$\frac{1}{n_j}\sum_{i\in C_n^{(j)}} \varphi\left(b_i^1, \cdots, b_i^p, u_i^1, \cdots, u_i^k\right) \xrightarrow[n\to+\infty]{c} \mathbb{E}_{(\bar{\boldsymbol{b}},\boldsymbol{Z})}\left[\varphi\left(\bar{b}_{j,1}, \cdots, \bar{b}_{j,p}, Z_1, \cdots, Z_k\right)\right],$$

which ends the proof.

5. Remaining proofs of Section 3

5.1. Proof of Theorem 3: AMP algorithm to describe the LV equilibrium's statistics. Notice that the existence of an equilibrium $\boldsymbol{x}_n^\star$ is granted by Proposition 3.1 under the condition $\kappa \geq \sqrt{2(1+\rho)}$. The rest of the proof follows very closely [AHMN23, Section 3.3] with Theorems 1 and 3 to handle the elliptic case. We shall often drop subscript n to lighten the notations.

Related Linear Complementarity Problem. For a LV system, it is well-known that the equilibrium satisfies a non-linear optimization problem called Linear Complementarity Problem (LCP), see [Tak96]. The LCP problem $\mathrm{LCP}\left(I - \Sigma, -\boldsymbol{r}\right)$ with parameters matrices $I, \Sigma \in \mathbb{R}^{n\times n}$ and vector $\boldsymbol{r} \in \mathbb{R}^n$ consists in finding a vector $\boldsymbol{x}^\star$ satisfying

$$\begin{cases} x_i^\star & \geq & 0\,, \\ x_i^\star\left(r_i - [(I-\Sigma)\boldsymbol{x}^\star]_i\right) & = & 0\,, \\ r_i - [(I-\Sigma)\boldsymbol{x}^\star]_i & \leq & 0\,, \end{cases} \qquad \text{for } i \in [n]\,. \tag{44}$$

If such a vector exists, we write $\boldsymbol{x}^\star \in \mathrm{LCP}\left(I - \Sigma, -\boldsymbol{r}\right)$. The first condition follows from the fact that $\boldsymbol{x}_n(t)$ is always (component-wise) positive for a LV system, the second condition simply express the nullity of the derivative at equilibrium. The last condition is a necessary condition (see for instance [Tak96, Th. 3.2.5]) for Lyapunov stability and has also an ecological interpretation

of non-invasibility (see [ABC+22, Section 3(a)]). In [AHMN23, Prop. 4], it is proved that the solution of a LCP $(I-\Sigma,-\boldsymbol{r})$ equivalently satisfies a fixed point equation in the sense that:

$$\boldsymbol{z}=\Sigma\boldsymbol{z}_{+}+\boldsymbol{r}_{n} \quad \Leftrightarrow \quad \boldsymbol{z}_{+}\in \mathrm{LCP}\left(I-\Sigma,-\boldsymbol{r}\right). \tag{45}$$

Otherwise stated, in case of uniqueness, $\boldsymbol{z}_{+}=\boldsymbol{x}^{\star}$.

An AMP algorithm. Let (δ,σ,γ) be the solution of System (13). Define the activation function h_k by:

$$h_k(u,a)=\frac{(u+a)_{+}}{\delta} \quad \text{for} \quad k\geq 0 \qquad \text{with} \qquad \partial_1 h_k(u,a)=\frac{\mathbf{1}_{(u+a>0)}}{\delta},$$

and consider the following AMP algorithm

$$\boldsymbol{u}^{k+1}=\frac{A_n}{\delta}\left(\boldsymbol{u}^{k}+\boldsymbol{a}\right)_{+}-\rho\frac{\langle \mathbf{1}_{(\boldsymbol{u}^k+\boldsymbol{a}>0)}\rangle_n\left(\boldsymbol{u}^{k-1}+\boldsymbol{a}\right)_{+}}{\delta^2}, \tag{46}$$

where

$$\begin{cases}\boldsymbol{u}^{0} &= \mathbf{1}_n\\ \boldsymbol{a} &= \left(1+\rho\frac{\gamma}{\delta^2}\right)\boldsymbol{r}=\frac{\kappa}{\delta}\boldsymbol{r}\end{cases}.$$

Notice that by Assumption $\mu^{\boldsymbol{a}} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \bar{a}$ where $\bar{a}=\left(1+\rho\frac{\gamma}{\delta^2}\right)\bar{r}$. We can easily check that Assumptions(A1)-(A5) are satisfied and hence can apply Theorem 1. If one is only interested in the limiting law of $\mu^{\boldsymbol{u}^k}$, the DE equations write

$$\begin{cases}\theta_1^2 &= \frac{1}{\delta^2}\mathbb{E}(1+\bar{a})_{+}^2\\ \theta_{k+1}^2 &= \frac{1}{\delta^2}\mathbb{E}(\theta_k\bar{Z}+\bar{a})_{+}^2\end{cases}, \tag{47}$$

where $\bar{Z}\sim\mathcal{N}(0,1)$ is independent from $\bar{a}$. Theorem 1 yields

$$\mu^{\boldsymbol{u}^k} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} Z_k\sim\mathcal{N}(0,\theta_k^2).$$

Departing from (46), we establish a perturbed LCP with respect to (45). Denote

$$\boldsymbol{\xi}^k=\boldsymbol{u}^k+\boldsymbol{a} \qquad \text{and} \qquad \gamma^k=\langle \mathbf{1}_{(\boldsymbol{u}^k+\boldsymbol{a}>0)}\rangle_n.$$

Taking advantage of the definition of $\boldsymbol{a}$ and the relations between δ,σ and γ from (13), easy (but lengthy) computations yield

$$\boldsymbol{\xi}_{+}^k-\frac{\boldsymbol{\xi}_{-}^k}{1+\rho\frac{\gamma}{\delta^2}}=\Sigma\boldsymbol{\xi}_{+}^k+\boldsymbol{r}+\frac{\boldsymbol{\varepsilon}^k}{1+\rho\frac{\gamma}{\delta^2}},$$

where

$$\boldsymbol{\varepsilon}^k=\frac{\rho\gamma}{\delta^2}\left(\boldsymbol{\xi}_{+}^k-\boldsymbol{\xi}_{+}^{k-1}\right)+\frac{\rho}{\delta^2}(\gamma-\gamma^k)\boldsymbol{\xi}_{+}^{k-1}+\left(\boldsymbol{\xi}^k-\boldsymbol{\xi}^{k+1}\right).$$

Defining $\boldsymbol{z}^k=\boldsymbol{\xi}_{+}^k-\frac{\boldsymbol{\xi}_{-}^k}{1+\rho\frac{\gamma}{\delta^2}}$ and $\tilde{\boldsymbol{\varepsilon}}^k=\frac{\boldsymbol{\varepsilon}^k}{1+\rho\frac{\gamma}{\delta^2}}$, we remark that $\boldsymbol{z}_{+}^k=\boldsymbol{\xi}_{+}^k$ and end up with the fixed-point equation $\boldsymbol{z}^k=\Sigma_n\boldsymbol{z}_{+}^k+\boldsymbol{r}+\tilde{\boldsymbol{\varepsilon}}^k$. Otherwise stated

$$\boldsymbol{z}_{+}^k\in \mathrm{LCP}(I-\Sigma,-\boldsymbol{r}-\tilde{\boldsymbol{\varepsilon}}^k). \tag{48}$$

We first focus on the asymptotic distribution of $\mu^{\boldsymbol{z}_{+}^k}$. Setting

$$\sigma_k=\frac{\delta}{\kappa}\theta_k$$

and noticing that function $(u+a)_{+}$ is Lipschitz, we obtain by Theorem 1 that

$$\mu^{\boldsymbol{z}_{+}^k} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}\left(\left(1+\frac{\rho\gamma}{\delta^2}\right)(\sigma_k\bar{Z}+\bar{r})_{+}\right).$$

Replacing θ_k by σ_k in the DE equations yields the equation

$$\sigma_{k+1}^2=\frac{1}{\delta^2}\mathbb{E}(\sigma_k\bar{Z}+\bar{r})_{+}^2$$

which by [AHMN23, Lemma 2] yields that $\sigma_k \xrightarrow[k\to\infty]{} \sigma$, the solution of (13b). Hence

$$\mathcal{L}\left(\left(1+\frac{\rho\gamma}{\delta^2}\right)(\sigma_k\bar{Z}+\bar{r})_+\right) \xrightarrow[k\to\infty]{\mathcal{P}_2(\mathbb{R})} \mathcal{L}\left(\left(1+\frac{\rho\gamma}{\delta^2}\right)(\sigma\bar{Z}+\bar{r})_+\right). \tag{49}$$

The arguments to establish convergence (14) in Theorem 3 from (45), (48) and (49) follow exactly those in [AHMN23, Section 3.4] and are thus omitted.

5.2. **Proof of Corollary 4.** We rely on the following result, see Sznitman [Szn91]:

Proposition 5.1. *Chaos propagation,* [Szn91] *Let $(X_1,\cdots,X_n)$ be a random vector of law P_n and let μ_n be its empirical measure $\mu_n=\frac{1}{n}\sum_{i=1}^n\delta_{X_i}\in\mathcal{P}(\mathbb{R})$. Assume the following*

(i) There exists an probability measure $\mu\in\mathcal{P}(\mathbb{R})$, such that the random probability measure μ_n converges to μ in law.

(ii) The vector $(X_1,\cdots,X_n)$ is exchangeable, that is for each permutation $\sigma\in\mathcal{S}_n$:

$$\left(X_{\sigma(1)},\cdots,X_{\sigma(n)}\right)\overset{\mathcal{L}}{=}(X_1,\cdots,X_n)\,.$$

Under these assumptions, the probability distribution P_n is μ-chaotic, that is for each fixed integer K we have

$$(X_1,\cdots,X_K)\xrightarrow[n\to\infty]{\mathcal{L}}\mu^{\otimes K}.$$

Corollary 4 is a direct consequence of Proposition 5.1, we only need to verify that $\boldsymbol{x}^\star=(x_1^\star,\cdots,x_n^\star)$ satisfies the two assumptions.

Proof of Assumption (i). By Theorem 3 we have the following convergence

$$\mu^{\boldsymbol{x}^\star}\xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})}\pi:=\mathcal{L}\left(\left(1+\rho\gamma/\delta^2\right)\left(\sigma\bar{Z}+\bar{r}\right)_+\right)\quad(\textit{completely})\,,$$

which implies the convergence in probability of $\mu^{\boldsymbol{x}^\star}$ and thus the convergence in law.

Proof of Assumption (ii). We now prove that $\boldsymbol{x}^\star$ is exchangeable. Let the permutation $\sigma\in\mathcal{S}_n$ be fixed and $P_\sigma\in\mathbb{R}^{n\times n}$ its associated permutation matrix. We introduce the set

$$\mathcal{E}(A)=\left\{\frac{\|A\|}{\kappa}<1\right\}.$$

Suppose now that $\frac{\|A\|}{\kappa}<1$ then $\boldsymbol{x}^\star=\boldsymbol{x}^\star=\boldsymbol{z}_+$ where $\boldsymbol{z}=\Sigma\boldsymbol{z}_++\boldsymbol{r}$. The function

$$\boldsymbol{y}\mapsto\Sigma\boldsymbol{y}_++\boldsymbol{r}$$

is Lipschitz with Lipschitz parameter $\|\Sigma\|<1$ hence $\boldsymbol{z}=\lim_p\boldsymbol{z}(p)$ where $\boldsymbol{z}(p)$ is defined by:

$$\begin{cases}\boldsymbol{z}(0)=0\,,\\ \boldsymbol{z}(p+1)=\Sigma\boldsymbol{z}_+(p)+\boldsymbol{r}\,.\end{cases}$$

Consider the following notations:

$$A^\sigma=P_\sigma^{-1}AP_\sigma\,,\quad \boldsymbol{y}^\sigma=P_\sigma\boldsymbol{y}\quad\text{for any }\boldsymbol{y}\in\mathbb{R}^n\,.$$

At first, we consider the $\boldsymbol{z}(p)$'s regardless of the condition $\frac{\|A\|}{\kappa}<1$ and prove by induction that

$$\forall\, p\geq 0\,,\quad \mathcal{L}(\boldsymbol{z}(p),A,\boldsymbol{r})=\mathcal{L}(\boldsymbol{z}^\sigma(p),A^\sigma,\boldsymbol{r}^\sigma)\,. \tag{50}$$

Since P_σ is orthogonal, the invariance property of elliptic matrices implies that $A^\sigma\overset{\mathcal{L}}{=}A$. For $p=1$, $\boldsymbol{z}(1)=\boldsymbol{r}$ and $\mathcal{L}(\boldsymbol{r}^\sigma,A^\sigma)=\mathcal{L}(\boldsymbol{r},A)$ (recall that $\boldsymbol{r}\perp\!\!\!\perp A$) hence the induction property. Now

$$\mathcal{L}(\boldsymbol{z}(p+1),A,\boldsymbol{r})\;=\;\mathcal{L}(A\boldsymbol{z}_+(p)+\boldsymbol{r},A,\boldsymbol{r})\overset{(a)}{=}\mathcal{L}(A^\sigma\boldsymbol{z}_+^\sigma(p)+\boldsymbol{r},A,\boldsymbol{r}^\sigma)\;=\;\mathcal{L}(\boldsymbol{z}^\sigma(p+1),A^\sigma,\boldsymbol{r}^\sigma)\,,$$

where (a) follows from the induction hypothesis. Eq.(50) is proved.

We can now transfer the exchangeability to $\boldsymbol{z}$ conditionnally on $\mathcal{E}(A)$. Notice that $\mathcal{E}(A)=\mathcal{E}(A^\sigma)$ and take any bounded continuous test function Φ, then

$$\mathbb{E}\Phi(\boldsymbol{z}(p))\mathbf{1}_{\mathcal{E}(A)}=\mathbb{E}\Phi(\boldsymbol{z}^\sigma(p))\mathbf{1}_{\mathcal{E}(A^\sigma)}\,.$$

Letting $p \to \infty$ yields

$$\mathcal{L}(\boldsymbol{z} \mid \mathcal{E}(A)) = \mathcal{L}(\boldsymbol{z}^{\sigma} \mid \mathcal{E}(A)) \tag{51}$$

Now if $\frac{\|A\|}{\kappa} > 1$ then $\boldsymbol{x}^\star = 0$ and $P_\sigma \boldsymbol{x}^\star = 0$. Combining this remark with (51) finally yields that $\mathcal{L}(\boldsymbol{x}^\star) = \mathcal{L}(P_\sigma \boldsymbol{x}^\star)$. The exchangeability of $\boldsymbol{x}^\star$ is proved.

5.3. **Proof of Theorem 5.** We follow the same strategy as in the proof of Corollary 4 except that we need the blockwise form of the AMP theorem (see Corollary 2) and a the generalized version of Proposition 5.1 stated hereafter.

Proposition 5.2. *Consider the partition* $\left\{C_n^{(j)}\right\}_{j\in[q]}$ *defined by (18)-(19). Let* $X = (X_1, \cdots, X_n)$ *be a random vector of law* P_n *and let* $\mu_n^{(1)}, \cdots, \mu_n^{(q)} \in \mathcal{P}(\mathbb{R})$ *be the empirical measures of the* q *blocks of* X *respectively, i.e.*

$$\mu_n^{(j)} \triangleq \frac{1}{n_j} \sum_{i \in C_n^{(j)}} \delta_{X_i} \quad \text{for all } j \in [q].$$

Assume the following

(i) There exist q *probability measures* $\mu_1, \cdots, \mu_q \in \mathcal{P}(\mathbb{R})$ *such that the random vector*

$$\left(\mu_n^{(1)}, \cdots, \mu_n^{(q)}\right)$$

converges in law to the vector $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_q)$ *in the product space* $\mathcal{P}(\mathbb{R})^q$.

(ii) The vector $X = (X_1, \cdots, X_n)$ *is blockwise exchangeable (see Definition 3.10), that is for each permutations* $\boldsymbol{\sigma} = (\sigma_1, \cdots, \sigma_q) \in \mathcal{S}_{n_1} \times \cdots \times \mathcal{S}_{n_q}$ *we have the following* $X^{\boldsymbol{\sigma}} \overset{\mathcal{L}}{=} X$.

Under these assumptions, the probability distribution P_n *is* $\boldsymbol{\mu}$*-chaotic, that is for each fixed* q*-uplet of integers* $(k_1, \cdots, k_q) \in [n_1] \times \cdots, \times [n_q]$ *we have*

$$X_{[k_1,\cdots,k_q]} \xrightarrow[n\to\infty]{\mathcal{L}} \prod_{j=1}^{q} \mu_j^{\otimes k_j},$$

where $X_{[k_1,\cdots,k_q]}$ *is the* $k_1 + \cdots + k_q$*-dimensional vector obtained by a concatenation the vectors* $(X_i)_{i\in\mathcal{K}_n^{(1)}}, \cdots, (X_i)_{i\in\mathcal{K}_n^{(q)}}$ *such that* $\mathcal{K}_n^{(j)}$ *is the subset of the* k_j *first elements of* $C_n^{(j)}$.

Proof of Proposition 5.2 is postoned to Appendix B.

Theorem 5 is a direct consequence of Proposition 5.2 and we only need to check that vector $\boldsymbol{x}^\star = (x_1^\star, \cdots, x_n^\star)$ satisfies assumptions (i) and (ii).

Proof of Assumption (i). Consider the empirical measure of the coordinates of the j-th block of $\boldsymbol{x}^\star$

$$\mu_n^{(j)} = \frac{1}{n_j} \sum_{i \in C_n^{(j)}} \delta_{x_i^\star}.$$

By blockwise AMP we already have

$$\mu_n^{(j)} \xrightarrow[n\to\infty]{\mathcal{P}_2(\mathbb{R})} \pi_j \quad (\textit{completely}).$$

If we endow the space of probability measures with the following distance

$$\tilde{d}(\mu, \nu) := \sup_{\psi} \left| \int_{\mathbb{R}} \psi d\mu - \int_{\mathbb{R}} \psi d\nu \right|,$$

then we also have convergence in probability of the sequence of measures $\left(\mu_n^{(j)}\right)_n$ (considered as random variables living in the space of probability measures) to π_j, the underlying distance is $\tilde{d}$,

$$\mu_n^{(j)} \xrightarrow[n\to+\infty]{\mathbb{P},\tilde{d}} \pi_j,$$

the convergence in probability of the components $\left(\mu^{(j)}\right)_{j\in[q]}$ implies the joint convergence in probability

$$\left(\mu_n^{(1)},\cdots,\mu_n^{(q)}\right) \xrightarrow[n\to+\infty]{\mathbb{P},d} (\mu_1,\cdots,\mu_q)$$

where the underlying distance d is a distance on the product of q probability measure spaces. Finally, and in particular we have convergence in law.

Proof of Assumption (ii). The proof that $\boldsymbol{x}^\star = (x_1^\star,\cdots,x_n^\star)$ is blockwise exchangeable closely follows the lines of the proof of exchangeability presented in Section 5.2 and is thus omitted.

References

[ABC+22] I. Akjouj, M. Barbier, M. Clenet, W. Hachem, M. Maïda, F. Massol, J. Najim, and V-C. Tran. Complex systems in ecology: A guided tour with large lotka-volterra models and random matrices. arXiv preprint arXiv:2212.06136, 2022.

[AHMN23] I. Akjouj, W. Hachem, M. Maïda, and J. Najim. Equilibria of large random lotka-volterra systems with vanishing species: a mathematical approach, 2023.

[AT12] Stefano Allesina and Si Tang. Stability criteria for complex ecosystems. Nature, 483(7388):205–208, 2012.

[AT15] S. Allesina and S. Tang. The stability–complexity relationship at age 40: a random matrix perspective. Population Ecology, 57(1):63–75, 2015.

[BK17] Jean Barbier and Florent Krzakala. Approximate Message-Passing decoder and capacity achieving sparse superposition codes. IEEE Transactions on Information Theory, 63(8):4894–4927, 2017.

[BLM15] M. Bayati, M. Lelarge, and M. Montanari. Universality in polytope phase transitions and message passing algorithms. The Annals of Applied Probability, 25(2), 4 2015.

[BM11] M. Bayati and A. Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. IEEE Transactions on Information Theory, 57(2):764–785, 2011.

[Bol14] E. Bolthausen. An iterative construction of solutions of the tap equations for the sherrington–kirkpatrick model. Communications in Mathematical Physics, 325(1):333–366, 2014.

[Bun17] Guy Bunin. Ecological communities with lotka-volterra dynamics. Phys. Rev. E, 95:042414, 2017.

[CEFN22] M. Clenet, H. El Ferchichi, and J. Najim. Equilibrium in a large lotka–volterra system with pairwise correlated interactions. Stochastic Processes and their Applications, 153:423–444, 2022.

[DAM17] Yash Deshpande, Emmanuel Abbe, and Andrea Montanari. Asymptotic mutual information for the balanced binary stochastic block model. Information and Inference: A Journal of the IMA, 6(2):125–170, 2017.

[DMLS23] R. Dudeja, Y. M. Lu, and S. Sen. Universality of approximate message passing with semirandom matrices. The Annals of Probability, 51(5):1616–1683, 2023.

[DMM09] David L. Donoho, Arian Maleki, and Andrea Montanari. Message-passing algorithms for compressed sensing. Proceedings of the National Academy of Sciences, 106(45):18914–18919, 2009.

[Fan21] Zhou Fan. Approximate message passing algorithms for rotationally invariant matrices, 2021.

[FVRS21] Oliver Y. Feng, Ramji Venkataramanan, Cynthia Rush, and Richard J. Samworth. A unifying tutorial on approximate message passing, 2021.

[Gal18] Tobias Galla. Dynamically evolved community size and stability of random lotka-volterra ecosystems(a). Europhysics Letters, 123(4):48004, 2018.

[GH82] S. Geman and C-R. Hwang. A chaos hypothesis for some large systems of random equations. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 60(3):291–314, 1982.

[Gir86] V. Girko. Elliptic law. Theory of Probability & Its Applications, 30(4):677–690, 1986.

[JM13] A. Javanmard and A. Montanari. State evolution for general approximate message passing algorithms, with applications to spatial coupling. Information and Inference: A Journal of the IMA, 2(2):115–144, 2013.

[Kal02] O. Kallenberg. Foundations of modern probability. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[LHM+04] M. Leibold, M. Holyoak, N. Mouquet, P. Amarasekare, J. Chase, M. Hoopes, R. Holt, J. Shurin, R. Law, D. Tilman, M. loreau, and A. Gonzalez. The metacommunity concept: a framework for multi-scale community ecology. Ecology letters, 7(7):601–613, 2004.

[LM19] Marc Lelarge and Léo Miolane. Fundamental limits of symmetric low-rank matrix estimation. Probability Theory and Related Fields, 173:859–929, 2019.

[May72] R. M. May. Will a large complex system be stable? Nature, 238(5364):413–414, 1972.

[Mon21] Andrea Montanari. Optimization of the Sherrington–Kirkpatrick hamiltonian. SIAM Journal on Computing, 0(0):FOCS19–1–FOCS19–38, 2021.

[Nau12] Alexey Naumov. Elliptic law for real random matrices, 2012.

[OR14] Sean O'Rourke and David Renfrew. Low rank perturbations of large elliptic random matrices. Electronic Journal of Probability, 19(none):1 – 65, 2014.

[RGV17] Cynthia Rush, Adam Greig, and Ramji Venkataramanan. Capacity-achieving sparse superposition codes via Approximate Message Passing decoding. IEEE Transactions on Information Theory, 63(3):1476–1500, 2017.
[RSF18] Sundeep Rangan, Philip Schniter, and Alyson K. Fletcher. Vector approximate message passing, 2018.
[Szn91] A-S. Sznitman. Topics in propagation of chaos. In Paul-Louis Hennequin, editor, Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989, pages 165–251, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.
[Tak96] Y. Takeuchi. Global dynamical properties of Lotka-Volterra systems. World Scientific, 1996.

Appendix A. Proof of Lemma 3.4

Recall the assumptions of Lemma 3.4 and system (13a)-(13c).

Notice that [AHMN23, Section 3.2] give the proof of existence and uniqueness of a solution (δ, σ, γ) in the case where $\rho = 1$ and a careful reading indicates that the proof remains true for $\rho \geq 0$. We thus assume $\rho \in [-1, 0)$ in the sequel.

We remind a few facts from [AHMN23, Lemma 2] which concern Eq. (13b)-(13c) and remain true for $\rho < 0$:

- For every $\delta > 1/\sqrt{2}$, the solution $\sigma(\delta)$ of (13b) exists,
- For every $\delta > 1/\sqrt{2}$, inequality $\gamma(\delta) < \delta^2$ holds true,
- Function $\delta \mapsto \sigma(\delta)$ is a decreasing function such that $\lim_{\delta \to (1/\sqrt{2})^+} \sigma(\delta) = +\infty$.

Given δ we obtain $\sigma(\delta)$ by solving (13b) and $\gamma(\delta)$ by (13c). In order to find a solution δ which satisfies (13a), we study function

$$h(\delta) = \delta + \rho \frac{\gamma(\delta)}{\delta}$$

and prove that there exists a unique $\delta^\star \in (1/\sqrt{2}, \infty)$ such as $h(\delta^\star) = \kappa$. The existence of $\delta^\star$ is easy to establish. In fact, function $\delta \mapsto h(\delta)$ is continuous,

$$\lim_{\delta \to \infty} h(\delta) = +\infty > \kappa \qquad \text{and} \qquad \lim_{\delta \to (1/\sqrt{2})^+} h(\delta) = \frac{1}{\sqrt{2}} + \rho \frac{1/2}{1/\sqrt{2}} = \frac{1+\rho}{\sqrt{2}} < \kappa \,,$$

which implies the existence of some $\delta^\star$ satisfying $h(\delta^\star) = \kappa$. Now in order to prove the uniqueness of $\delta^\star$, it suffices to prove that h is strictly increasing. For $\rho < 0$ and $\delta > 1/\sqrt{2}$ we have

$$h'(\delta) = 1 + \rho \frac{\gamma'(\delta)}{\delta} - \rho \frac{\gamma(\delta)}{\delta^2} \geq 1 + \rho \frac{\gamma'(\delta)}{\delta} \geq 1 - \frac{\gamma'(\delta)}{\delta} \,. \tag{52}$$

In order to prove that h is increasing, it suffices to establish the following inequality:

Lemma A.1. *Let $\delta \in (1/\sqrt{2}, +\infty)$ then $\gamma'(\delta) < \delta$.*

Before proceeding to the proof of Lemma A.1, we introduce auxiliary functions that will simplify the forthcoming computations. Let

$$Q(x) = \int_x^{+\infty} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \,, \qquad f(x) = \left(1 + x^2\right) Q(x) - x \times \frac{e^{-x^2/2}}{\sqrt{2\pi}} \,,$$

the derivative of which are given by:

$$Q'(x) = -\frac{e^{-x^2/2}}{\sqrt{2\pi}} \,, \qquad f'(x) = 2\left(xQ(x) - \frac{e^{-x^2/2}}{\sqrt{2\pi}}\right) .$$

We consider the change of variable

$$x(\delta) = -\frac{1}{\sigma(\delta)} < 0$$

and rewrite the system of equations (13a)-(13c) using Q, f and $x(\delta)$:

$$\kappa = \delta + \rho \frac{\gamma(\delta)}{\delta} \,, \tag{53}$$

$$\delta^2 = \mathbb{E}\left[f(\bar{r}x(\delta))\right] \,, \tag{54}$$

$$\gamma(\delta) = \mathbb{E}\left[Q(\bar{r}x(\delta))\right] \,, \tag{55}$$

with $(\delta, x, \gamma) \in (1/\sqrt{2}, \infty) \times (-\infty, 0) \times (0, 1)$. In order to obtain Eq.(54), we start from (13c) which is rewritten as

$$\begin{aligned}
\delta^2 &= \frac{1}{\sigma^2}\mathbb{E}\int_{-\frac{\bar{r}}{\sigma}}^{\infty}(\sigma^2 t^2 + 2\sigma t\bar{r} + \bar{r}^2)\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,dt \quad = \quad \mathbb{E}\int_{x\bar{r}}^{\infty}(t^2 - 2t\bar{r}x + (x\bar{r})^2)\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,dt\,, \\
&= \mathbb{E}(x\bar{r})^2 Q(x\bar{r}) + \mathbb{E}\int_{x\bar{r}}^{\infty}(t^2 - 2xt\bar{r})\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,dt\,, \\
&= \mathbb{E}(1 + (x\bar{r})^2)Q(x\bar{r}) - \mathbb{E}x\bar{r}\frac{e^{-\frac{(x\bar{r})^2}{2}}}{\sqrt{2\pi}}\,, \qquad (56)
\end{aligned}$$

where the last equality follows from an integration by parts.

With notations Q, f, $x(\delta)$ and equation (55) at hand, notice that γ depends on δ via x. The next lemma expresses the derivatives of x and γ.

Lemma A.2. *Let $\delta > 1/\sqrt{2}$ then x and γ's derivatives with respect to δ write:*

$$x'(\delta) = \frac{\delta x(\delta)}{\delta^2 - \mathbb{E}\left[Q\left(\bar{r}x(\delta)\right)\right]}\,, \qquad \gamma'(\delta) = -\frac{\delta x(\delta)}{\delta^2 - \mathbb{E}\left[Q\left(\bar{r}x(\delta)\right)\right]} \times \frac{\mathbb{E}\left[\bar{r}e^{-(\bar{r}x(\delta))^2/2}\right]}{\sqrt{2\pi}}\,.$$

Momentarily assuming Lemma A.2, we are now in position to prove Lemma A.1.

Proof of Lemma A.1. For $\delta > \frac{1}{\sqrt{2}}$, using Lemma A.2 and Eq. (56) we get

$$\frac{\gamma'(\delta)}{\delta} = \frac{1 - \mathbb{E}\left[\left(1 + \left(\bar{r}x(\delta)\right)^2\right)\frac{Q(\bar{r}x(\delta))}{\delta^2}\right]}{1 - \mathbb{E}\left[\frac{Q(\bar{r}x(\delta))}{\delta^2}\right]}.$$

The inequality $\gamma'(\delta) < \delta$ is equivalent to

$$1 - \mathbb{E}\left[\left(1 + \left(\bar{r}x(\delta)\right)^2\right)\frac{Q\left(\bar{r}x(\delta)\right)}{\delta^2}\right] < 1 - \mathbb{E}\left[\frac{Q\left(\bar{r}x(\delta)\right)}{\delta^2}\right] \quad \Leftrightarrow \quad 0 < \mathbb{E}\left[\left(\bar{r}x(\delta)\right)^2\frac{Q\left(\bar{r}x(\delta)\right)}{\delta^2}\right],$$

the last inequality being true because $\mathcal{L}(\bar{r}) \neq \delta_0$. □

Proof of Lemma A.2. By differentiating (54) with respect to δ we get:

$$2\delta = x'(\delta)\,\mathbb{E}\left[\bar{r}f'\left(\bar{r}x(\delta)\right)\right],$$

from which we extract $x'(\delta)$. Using the explicit form of f' yields:

$$x'(\delta) = \frac{\delta}{\mathbb{E}\left[\bar{r}^2 x(\delta)\,Q\left(\bar{r}x(\delta)\right)\right] - \mathbb{E}\left[\left(\bar{r}/\sqrt{2\pi}\right)e^{-(\bar{r}x(\delta))^2/2}\right]}.$$

From (54) and (56), we get:

$$\mathbb{E}\left[\frac{\bar{r}e^{-(\bar{r}x(\delta))^2/2}}{\sqrt{2\pi}}\right] = \mathbb{E}\left[\left(\frac{1}{x(\delta)} + \bar{r}^2 x(\delta)\right)Q\left(\bar{r}x(\delta)\right)\right] - \frac{\delta^2}{x(\delta)},$$

thus

$$x'(\delta) = \frac{\delta}{\mathbb{E}\left[\bar{r}^2 x(\delta)\,Q\left(\bar{r}x(\delta)\right)\right] - \mathbb{E}\left[\left(\frac{1}{x(\delta)} + \bar{r}^2 x(\delta)\right)Q\left(\bar{r}x(\delta)\right)\right] + \frac{\delta^2}{x(\delta)}} = \frac{\delta x(\delta)}{\delta^2 - \mathbb{E}\left[Q\left(\bar{r}x(\delta)\right)\right]}.$$

□

Proof of Lemma 3.4 is completed.

Appendix B. Generalized propagation of chaos

In this section we prove the generalized version of Sznitman's propagation of chaos result presented in section 5.3 (see Proposition 5.2) to cover the case of blockwise structured random vectors.

We provide the following proof of Proposition 5.2 that follows the same ideas of [Szn91].

Proof of Proposition 5.2. We want to prove that for any test functions $\varphi^{(1)},\cdots,\varphi^{(q)}$ where $\varphi^{(j)}$ is the tensor product of k_j test functions $\varphi_1^{(j)},\cdots,\varphi_{k_j}^{(j)}$ we have the following limit

$$\mathbb{E}\left[\varphi^{(1)}\left(X^{(1)}\right)\times\cdots\times\varphi^{(q)}\left(X^{(q)}\right)\right]\xrightarrow[n\to\infty]{}\prod_{i=1}^{k_1}\mathbb{E}\left[\varphi_i^{(1)}\left(\tilde{X}_1\right)\right]\times\cdots\times\prod_{i=1}^{k_q}\mathbb{E}\left[\varphi_i^{(q)}\left(\tilde{X}_q\right)\right], \tag{57}$$

where $(\tilde{X}_1,\cdots,\tilde{X}_q)\in\mathbb{R}^q$ is a random vector having the law $\boldsymbol{\mu}$. The term $\varphi^{(j)}\left(X^{(j)}\right)$ is equal to the following product

$$\varphi^{(j)}\left(X^{(j)}\right)=\varphi_1^{(j)}\left(X_1^{(j)}\right)\times\cdots\times\varphi_{k_j}^{(j)}\left(X_{k_j}^{(j)}\right).$$

In order to prove (57), we will consider the following intermediate term

$$B_n=\mathbb{E}\left[\prod_{k=1}^{k_1}\left(\frac{1}{n_1}\sum_{i=1}^{n_1}\varphi_k^{(1)}\left(X_i^{(1)}\right)\right)\times\cdots\times\prod_{k=1}^{k_q}\left(\frac{1}{n_q}\sum_{i=1}^{n_q}\varphi_k^{(q)}\left(X_i^{(q)}\right)\right)\right],$$

we will also denote the left hand side and the right hand side of (57) by A_n and C_n respectively. So it is sufficient to prove that $A_n-B_n\xrightarrow[n\to\infty]{}0$ and $B_n-C_n\xrightarrow[n\to\infty]{}0$.

Lemma B.1. *We have* $B_n-C_n\xrightarrow[n\to\infty]{}0$.

Proof. This immediate by Assumption (i). In fact, consider the continuous bounded test function $F:\mathcal{P}(\mathbb{R})^q\to\mathbb{R}$ defined for any q-uplet of probability measures $(\nu_1,\cdots,\nu_q)$ by

$$F(\nu_1,\cdots,\nu_q)=\prod_{i_1=1}^{k_1}\left(\int_{\mathbb{R}}\varphi_{i_1}^{(1)}d\nu_1\right)\times\cdots\times\prod_{i_q=1}^{k_q}\left(\int_{\mathbb{R}}\varphi_{i_q}^{(q)}d\nu_q\right).$$

Using Assumption (i) yields to the desired result. □

Lemma B.2. *We have* $A_n-B_n\xrightarrow[n\to\infty]{}0$

Proof. We have

$$A_n=\mathbb{E}\left[\prod_{i_1=1}^{k_1}\varphi_{i_1}^{(1)}\left(X_{i_1}^{(1)}\right)\times\cdots\times\prod_{i_q=1}^{k_q}\varphi_{i_q}^{(q)}\left(X_{i_q}^{(q)}\right)\right].$$

By Assumption (ii) we can re-write A_n as follows

$$A_n=\frac{1}{\prod_{j=1}^q(n_j!)}\sum_{\sigma_1\in\mathfrak{S}_1,\cdots,\sigma_q\in\mathfrak{S}_q}\mathbb{E}\left[\prod_{i_1=1}^{k_1}\varphi_{i_1}^{(1)}\left(X_{\sigma_1(i_1)}^{(1)}\right)\times\cdots\times\prod_{i_q=1}^{k_q}\varphi_{i_q}^{(q)}\left(X_{\sigma_q(i_q)}^{(q)}\right)\right],$$

Now observe that in the j-th factor in the formula above, the product is taken only over the subset $[k_j]$ of $[n_j]$, thus we consider the equivalence relation $\sim_j$ defined on $\mathfrak{S}_j$ by

$$\forall\sigma,\nu\in\mathfrak{S}_{n_j}\quad[\sigma\sim_j\nu]\ \Leftrightarrow\ [\sigma(i)=\nu(i),\ \forall i\in[k_j]],$$

in oder words, we will identify permutations that agree on the set $[k_j]$, this means that for all $j\in[q]$, if $\sigma_j\sim_j\nu_j$ then

$$\prod_{i_1=1}^{k_1}\varphi_{i_1}^{(1)}\left(X_{\sigma_1(i_1)}^{(1)}\right)\times\cdots\times\prod_{i_q=1}^{k_q}\varphi_{i_q}^{(q)}\left(X_{\sigma_q(i_q)}^{(q)}\right)=\prod_{i_1=1}^{k_1}\varphi_{i_1}^{(1)}\left(X_{\nu_1(i_1)}^{(1)}\right)\times\cdots\times\prod_{i_q=1}^{k_q}\varphi_{i_q}^{(q)}\left(X_{\nu_q(i_q)}^{(q)}\right).$$

Now consider the quotient group of $\mathfrak{S}_{n_j}$ with respect to $\sim_j$ denoted as $\tilde{\mathfrak{S}}_j$, then we have the following:

$$A_n=\left(\prod_{j=1}^q\frac{(n_j-k_j)!}{n_j!}\right)\sum_{\sigma_1\in\tilde{\mathfrak{S}}_1,\cdots,\sigma_q\in\tilde{\mathfrak{S}}_q}\mathbb{E}\left[\prod_{i_1=1}^{k_1}\varphi_{i_1}^{(1)}\left(X_{\sigma_1(i_1)}^{(1)}\right)\times\cdots\times\prod_{i_q=1}^{k_q}\varphi_{i_q}^{(q)}\left(X_{\sigma_q(i_q)}^{(q)}\right)\right].$$

Now let us analyse the term B_n. If we develop the products inside the $\mathbb{E}$ symbol, we get

$$B_n = \mathbb{E}\left[\frac{1}{n_1^{k_1}} \sum_{i_1^{(1)},\cdots,i_{k_1}^{(1)}=1}^{n_1} \prod_{\ell=1}^{k_1} \varphi_\ell^{(1)}\left(X_{i_\ell^{(1)}}^{(1)}\right) \times \cdots \times \frac{1}{n_q^{k_q}} \sum_{i_1^{(q)},\cdots,i_{k_q}^{(q)}=1}^{n_q} \prod_{\ell=1}^{k_q} \varphi_\ell^{(q)}\left(X_{i_\ell^{(q)}}^{(q)}\right)\right]$$

$$= \frac{1}{\prod_{j=1}^q n_j^{k_j}} \sum_{\substack{i_1^{(1)},\cdots,i_{k_1}^{(1)}\in[n_1] \\ \vdots \\ i_1^{(q)},\cdots,i_{k_q}^{(q)}\in[n_q]}} \mathbb{E}\left[\prod_{\ell=1}^{k_1} \varphi_\ell^{(1)}\left(X_{i_\ell^{(1)}}^{(1)}\right) \times \cdots \times \prod_{\ell=1}^{k_q} \varphi_\ell^{(q)}\left(X_{i_\ell^{(q)}}^{(q)}\right)\right]$$

For all $j \in [q]$, let S_j be the subset of $[n_j]^{k_j}$ of k_j-uplets with different elements. The sum in B_n is over the grid $[n_1]^{k_1} \times \cdots \times [n_q]^{k_q}$, we will decompose it into a sum over $S_1 \times \cdots \times S_q$ and a sum over the complementary set $[n_1]^{k_1} \times \cdots \times [n_q]^{k_q} \setminus S_1 \times \cdots \times S_q$, let us denote these two sums by B'_n and B''_n respectively. We will show that

$$A_n - B'_n \xrightarrow[n\to\infty]{} 0 \quad \text{and} \quad B''_n \xrightarrow[n\to\infty]{} 0.$$

Let us first prove that $A_n - B'_n \xrightarrow[n\to\infty]{} 0$. We have

$$B'_n = \frac{1}{\prod_{j=1}^q n_j^{k_j}} \sum_{\substack{(i_1^{(1)},\cdots,i_{k_1}^{(1)})\in S_1 \\ \vdots \\ (i_1^{(q)},\cdots,i_{k_q}^{(q)})\in S_q}} \mathbb{E}\left[\prod_{\ell=1}^{k_1} \varphi_\ell^{(1)}\left(X_{i_\ell^{(1)}}^{(1)}\right) \times \cdots \times \prod_{\ell=1}^{k_q} \varphi_\ell^{(q)}\left(X_{i_\ell^{(q)}}^{(q)}\right)\right].$$

Observe now that we can identify each k_j-uplet $(i_1^{(j)},\cdots,i_{k_j}^{(j)})$ of S_j with an element $\sigma_j \in \tilde{\mathfrak{S}}_j$, and this is because by definition of the set S_j the indices $\{i_1^{(j)},\cdots,i_{k_j}^{(j)}\}$ are different, i.e. one can construct $\sigma_j \in \tilde{\mathfrak{S}}_j$ such that

$$\sigma_j(\ell) = i_\ell^{(j)} \quad \forall \ell \in [k_j],$$

this essentially means that we can index the sum in B'_n using permutations,

$$B'_n = \frac{1}{\prod_{j=1}^q n_j^{k_j}} \sum_{\sigma_1\in\tilde{\mathfrak{S}}_1,\cdots,\sigma_q\in\tilde{\mathfrak{S}}_q} \mathbb{E}\left[\prod_{\ell=1}^{k_1} \varphi_\ell^{(1)}\left(X_{\sigma_1(\ell)}^{(1)}\right) \times \cdots \times \prod_{\ell=1}^{k_q} \varphi_\ell^{(q)}\left(X_{\sigma_q(\ell)}^{(q)}\right)\right].$$

Except the multiplicative factor, the terms A_n and B'_n are exactly the same, so let us consider the difference

$$A_n - B'_n = \left(\prod_{j=1}^q \frac{(n_j-k_j)!}{n_j!} - \frac{1}{\prod_{j=1}^q n_j^{k_j}}\right) \times \sum_{\sigma_1\in\tilde{\mathfrak{S}}_1,\cdots,\sigma_q\in\tilde{\mathfrak{S}}_q} \mathbb{E}\left[\prod_{\ell=1}^{k_1} \varphi_\ell^{(1)}\left(X_{\sigma_1(\ell)}^{(1)}\right) \times \cdots \times \prod_{\ell=1}^{k_q} \varphi_\ell^{(q)}\left(X_{\sigma_q(\ell)}^{(q)}\right)\right],$$

denote M an upper bound of the test functions $\varphi_1^{(1)},\cdots,\varphi_{k_q}^{(q)}$ and notice that the set $\tilde{\mathfrak{S}}_1 \times \cdots \times \tilde{\mathfrak{S}}_q$ is of cardinal $\prod_{j=1}^q n_j! / \prod_{j=1}^q (n_j-k_j)!$, thus we get

$$|A_n - B'_n| \leq \left(\prod_{j=1}^q \frac{(n_j-k_j)!}{n_j!} - \frac{1}{\prod_{j=1}^q n_j^{k_j}}\right)\left(\prod_{j=1}^q \frac{n_j!}{(n_j-k_j)!}\right) M^{k_1+\cdots+k_q}$$

$$= \left(1 - \prod_{j=1}^q \frac{n_j!}{n_j^{k_j}(n_j-k_j)!}\right) M^{k_1+\cdots+k_q}.$$

Recall that when $n \to \infty$ we also have $n_j(n) \to \infty$ because $c_j > 0$. As $k_j \leq n_j$ is a constant integer we have

$$\frac{n_j!}{n_j^{k_j}(n_j - k_j)!} \xrightarrow[n\to\infty]{} 1 \quad \forall j \in [q],$$

finally we get $A_n - B'_n \xrightarrow[n\to\infty]{} 0$. It remains to show that $B''_n \xrightarrow[n\to\infty]{} 0$. Now, B''_n is a sum over the complementary part of $S_1 \times \cdots \times S_q$ in $[n_1]^{k_1} \times \cdots \times [n_q]^{k_q}$, this complementary subset is of size

$$n_1^{k_1} \times \cdots \times n_q^{k_q} - \frac{n_1! \times \cdots \times n_q!}{(n_1 - k_1)! \times \cdots \times (n_q - k_q)!},$$

if we upper bound the test functions by M again, we obtain the following inequality

$$|B''_n| \leq \left(1 - \prod_{j=1}^{q} \frac{n_j!}{n_j^{k_j}(n_j - k_j)!}\right) M^{k_1 + \cdots + k_q},$$

thus $B''_n \xrightarrow[n\to\infty]{} 0$. □

In summary, we have proved that each term of the following expression:

$$A_n - C_n = (A_n - B'_n) + B''_n + (B_n - C_n)$$

converges to 0, which ends the proof of Proposition 5.2.

Appendix C. Proof of Proposition 4.5

Define the matrix

$$D_k \triangleq \begin{bmatrix} d_0 & & & \\ & d_1 & & \\ & & \ddots & \\ & & & d_{k-1} \end{bmatrix} \text{ for } k \geq 1,$$

Recall the compact form of the AMP iteration (21)

$$\boldsymbol{u}^{k+1} = A\boldsymbol{q}^k - \rho d_k \boldsymbol{q}^{k-1}, \quad k \geq 0.$$

Considering the iterates $\boldsymbol{u}^1, \ldots, \boldsymbol{u}^k$ provided by the previous equation and using the notations we just introduced, it is easy to see that

$$AQ_k = U_k + \rho \begin{bmatrix} 0 & Q_{k-1} \end{bmatrix} D_k, \quad k \geq 1,$$

where Q_0 is by convention the empty matrix, (i.e. $\begin{bmatrix} 0 & Q_0 \end{bmatrix} = [0] \in \mathbb{R}^{1\times 1}$). For $k \geq 1$, we now write

$$\begin{aligned} \boldsymbol{u}^{k+1} &= A\boldsymbol{q}^k - \rho d_k \boldsymbol{q}^{k-1} = A(P_k + P_k^\perp)\boldsymbol{q}^k - \rho d_k \boldsymbol{q}^{k-1} \\ &= AP_k\boldsymbol{q}^k + AP_k^\perp \boldsymbol{q}^k + \rho(AP_k)^\top P_k^\perp \boldsymbol{q}^k - \rho(AP_k)^\top P_k^\perp \boldsymbol{q}^k - \rho d_k \boldsymbol{q}^{k-1} \\ &= AP_k\boldsymbol{q}^k + \rho(AP_k)^\top P_k^\perp \boldsymbol{q}^k - \rho d_k \boldsymbol{q}^{k-1} + \mathcal{I}_k(A). \end{aligned}$$

The first term at the right hand side can be rewritten as

$$AP_k\boldsymbol{q}^k = AQ_k \left(Q_k^\top Q_k\right)^\dagger Q_k^\top \boldsymbol{q}^k = \left(U_k + \rho \begin{bmatrix} 0 & Q_{k-1} \end{bmatrix} D_k\right) \boldsymbol{\alpha}^k,$$

and the next term is rewritten as

$$\begin{aligned} \rho(AP_k)^\top P_k^\perp \boldsymbol{q}^k &= \rho \left(AQ_k \left(Q_k^\top Q_k\right)^\dagger Q_k^\top\right)^\top P_k^\perp \boldsymbol{q}^k \\ &= \rho \left(\left(U_k + \rho \begin{bmatrix} 0 & Q_{k-1} \end{bmatrix} D_k\right) \left(Q_k^\top Q_k\right)^\dagger Q_k^\top\right)^\top P_k^\perp \boldsymbol{q}^k \\ &= \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \left(U_k^\top P_k^\perp \boldsymbol{q}^k + \rho D_k^\top \begin{bmatrix} 0 & Q_{k-1} \end{bmatrix}^\top P_k^\perp \boldsymbol{q}^k\right) \\ &= \rho Q_k \left(Q_k^\top Q_k\right)^\dagger U_k^\top P_k^\perp \boldsymbol{q}^k, \end{aligned}$$

since $P_k^{\perp} Q_{k-1} = 0$. This gives

$$\boldsymbol{u}^{k+1} = \sum_{l=1}^{k} \alpha_l^k \boldsymbol{u}^l + \rho \begin{bmatrix} 0 & Q_{k-1} \end{bmatrix} D_k \boldsymbol{\alpha}^k + \rho Q_k \left(Q_k^\top Q_k\right)^\dagger U_k^\top P_k^{\perp} \boldsymbol{q}^k - \rho d_k P_k \boldsymbol{q}^{k-1} + \mathcal{I}_k(A), \tag{58}$$

since $\boldsymbol{q}^{k-1} \in \operatorname{span} P_k$. Developing in turn the three middle terms at the right hand side of this equation, and observing that $P_k \boldsymbol{q}^l = \boldsymbol{q}^l$ for each $l \leq k-1$, we obtain

$$\rho \begin{bmatrix} 0 & Q_{k-1} \end{bmatrix} D_k \boldsymbol{\alpha}^k = \rho \sum_{l=1}^{k} \alpha_l^k d_{l-1} \boldsymbol{q}^{l-2} = \rho P_k \sum_{l=1}^{k} \alpha_l^k d_{l-1} \boldsymbol{q}^{l-2} = \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \sum_{l=1}^{k} \alpha_l^k d_{l-1} Q_k^\top \boldsymbol{q}^{l-2}$$

(with $\boldsymbol{q}^{-1} = 0$),

$$\begin{aligned} \rho Q_k \left(Q_k^\top Q_k\right)^\dagger U_k^\top P_k^{\perp} \boldsymbol{q}^k &= Q_k \left(Q_k^\top Q_k\right)^\dagger \left(U_k^\top \boldsymbol{q}^k - U_k^\top P_k \boldsymbol{q}^k\right) \\ &= \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \left(U_k^\top \boldsymbol{q}^k - U_k^\top Q_k \boldsymbol{\alpha}^k\right) \\ &= \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \left(U_k^\top \boldsymbol{q}^k - \sum_{l=1}^{k} \alpha_l^k U_k^\top \boldsymbol{q}^{l-1}\right), \end{aligned}$$

and

$$-\rho d_k P_k \boldsymbol{q}^{k-1} = -\rho d_k Q_k \left(Q_k^\top Q_k\right)^\dagger Q_k^\top \boldsymbol{q}^{k-1}.$$

Injecting these equations into (58), we obtain

$$\boldsymbol{u}^{k+1} = \sum_{l=1}^{k} \alpha_l^k \boldsymbol{u}^l + \rho Q_k \left(Q_k^\top Q_k\right)^\dagger \left(U_k^\top \boldsymbol{q}^k - d_k Q_k^\top \boldsymbol{q}^{k-1} - \sum_{l=1}^{k} \alpha_l^k \left(U_k^\top \boldsymbol{q}^{l-1} - d_{l-1} Q_k^\top \boldsymbol{q}^{l-2}\right)\right) + \mathcal{I}_k(A),$$

which is the equation provided in the statement of the proposition.