



HAL
open science

A traffic approach for profiled Pennington-Worah matrices

Issa Dabo, Camille Male

► **To cite this version:**

Issa Dabo, Camille Male. A traffic approach for profiled Pennington-Worah matrices. 2024. hal-04803396

HAL Id: hal-04803396

<https://hal.science/hal-04803396v1>

Preprint submitted on 25 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A traffic approach for profiled Pennington-Worah matrices

Issa Dabo and Camille Male

Abstract

We study macroscopic observables of large random matrices introduced by Pennington and Worah, of the form $Y(h) = \frac{1}{\sqrt{N_2}} h\left[\left\{\frac{WX}{\sqrt{N_0}}\right\}\right]$, where W and X are random rectangular matrices with independent entries and h is a function evaluated entry-wise. We allow the variance of the entries of the matrices to vary from entry to entry. We complement Pécché perspective from [Electron. Commun. Probab. 24 (2019), no. 66, 1–7] showing a decomposition of $Y(h)$ whose and traffic asymptotic traffic-equivalent for their ingredients, when h belong to the space of odd polynomials. This give a new interpretation of the "linear plus chaos" phenomenon observed for these matrices.

Primary 15B52, 46L54; Keywords: Free Probability, Large Random Matrices

Notations: For an integer $n \geq 1$, we use the notation $[n] := \{1, \dots, n\}$. For a real rectangular matrix A and a function $h : \mathbb{R} \rightarrow \mathbb{R}$, we denote by $h[\{A\}]$ the entry-wise evaluation of h in A , that is the matrix whose entries are the image by h of the corresponding entries of A . We consider matrix sizes N_0, N_1, \dots for rectangular matrices in the classical regime of random matrices, i.e. for each $i \geq 1$, we implicitly assume that N_i is a sequence $N_i(N)$ for a parameter $N \geq 1$ that tends to infinity, and such that the ratio $\frac{N_i}{N}$ converges to a positive limit when $N \rightarrow \infty$.

1 Introduction

In this article, we study random Gram matrices that were introduced by Jeffrey Pennington and Pratik Worah in the context of machine learning [PW19]. We refer the reader to [HMRT22] for the motivation of the authors in this context.

Definition 1.1. *Given two complex random matrices W in $\mathbb{R}^{N_1 \times N_0}$ and X in $\mathbb{R}^{N_0 \times N_2}$, as long as a function $h : \mathbb{R} \rightarrow \mathbb{R}$, we set*

$$Y(h) = \frac{1}{\sqrt{N_2}} h\left[\left\{\frac{WX}{\sqrt{N_0}}\right\}\right] \in \mathbb{R}^{N_1 \times N_2}, \quad (1.1)$$

that we call the Pennington-Worah matrix associated to h, W, X .

The weak convergence of the empirical eigenvalues distribution of YY^t for such a matrix Y is proved in [PW19] for Gaussian entries. Lucas Benigni and Sandrine Péché [BP21] extends their result and study the outliers under the following hypotheses.

Hypothesis 1.2. *The rectangular matrices W and X are independent and have centered real i.i.d. entries with variance one. Moreover there exist constants $\vartheta > 0$, $\alpha > 1$ such that*

$$\mathbb{P}(|W(1, 1)| \geq t \text{ and } |X(1, 1)| \geq t) \leq e^{-\vartheta t^\alpha}.$$

The function $h : \mathbb{R} \rightarrow \mathbb{R}$ is real analytic, and there exist constants $C, c, A_0 > 0$ such that $|h^{(n)}(x)| \leq CA^{cn}$, for all $A \geq A_0$, $n \in \mathbb{N}$ and $x \in [-A, A]$.

Under the above assumptions, the empirical eigenvalues distribution of YY^* converges in probability toward a deterministic limit as N goes to infinity. The limit is described by a consistent system of equations for Stieltjes transforms. Moreover, in [Péc19] Sandrine Péché proposes a presentation of this distribution by exhibiting a simple equivalent model with same limiting distribution. She also use this method with Lucas Benigni in [BP22] to describe the outliers.

To state their results from [Péc19, BP22], we recall that a sequence of sets $S_N \subset \mathbb{R}$ converges in Hausdorff topology toward $S \subset \mathbb{R}$ whenever for $\min\{\varepsilon > 0 \mid S_N + (-\varepsilon, \varepsilon) \subset S, S + (-\varepsilon, \varepsilon) \subset S_N\}$ tends to zero. We use the following terminology.

Definition 1.3. *Let $A = A_N$ and $B = B_N$ be two sequences for $N_1 \times N_2$ rectangular matrices, where $\frac{N_i}{N} \xrightarrow{N \rightarrow \infty} \psi_i > 0$ for $i = 1, 2$. We say that A and B are spectral equivalent if the empirical eigenvalues distributions of AA^t and BB^t converges to the same limit. We say that A and B are strongly spectral equivalent if moreover the spectra of AA^t and BB^t converges to a same set in Hausdorff topology.*

Strongly spectral equivalent matrices have outliers converging to the same positions [CM14, Proposition 2.1]. We denote by $\omega(t) = \frac{e^{-t^2/2}}{\sqrt{2\pi}}$ the density of a real standard Gaussian random variable. A random matrix is say to be a standard Gaussian matrix if it has i.i.d. real standard Gaussian entries.

Theorem 1.4. *A random $Y(h)$ as in Definition 1.1 satisfying Hypothesis 1.2 is spectral equivalent to*

$$Y^{\square \text{lin}}(h) + Y^{\square \text{per}}(h),$$

where $Y^{\square \text{lin}}(h) = \sqrt{\theta_2(h)} \frac{W^{\text{Gau}} X^{\text{Gau}}}{\sqrt{N_0 N_2}}$ and $Y^{\square \text{per}}(h) = \sqrt{\theta_1(h) - \theta_2(h)} \frac{Z^{\text{Gau}}}{\sqrt{N_0}}$. The matrices $W^{\text{Gau}}, X^{\text{Gau}}$ and Z^{Gau} are independent standard Gaussian

matrices and we have set

$$\theta_1(h) = \int_{\mathbb{R}} h^2(t)\omega(t)dt, \quad \theta_2(h) = \left(\int_{\mathbb{R}} h'(t)\omega(t)dt \right)^2.$$

Moreover, if the third moment of the entries of W and X is zero, then $Y(h)$ is strongly spectral equivalent to

$$Y^{\square}(h) = Y^{\square\text{lin}}(h) + Y^{\square\text{per}}(h) + Y^{\square\text{def}}(h),$$

where $Y^{\square\text{def}}(h)$ is an explicit matrix of rank 2.

In words, the first statement says that $Y(h)$ is spectral equivalent to the matrix when h is linear plus an independent i.i.d matrix. We refer this as the *linear plus chaos* phenomenon. Our work is motivated by the following questions:

1. **Stability of the phenomenon.** Do we still have an analogue linear plus chaos phenomenon when the matrices W and X are replaced by more general models with more structure? In this article, we consider profiled matrices, namely matrices with independent entries where the variance of the entries can vary from one variable to another, see Hypothesis 2.1. Working with a different type of equivalent we confirm that the phenomenon holds.
2. **Structure of the noise.** Let h_1 and h_2 be two functions as above, for which both $Y(h_1)$ and $Y(h_2)$ satisfy the linear plus chaos phenomenon. What can be said about the joint distribution of these noises? Can we find functions for which these noises are independent? Or functions for which they are coupled? The roots of this question lie in the very origin of free probability theory which questions the difference between probability spaces generated by different numbers of free random variables [Dyk94]. This article completely characterizes a family of independent matrices that generate the noise arising from profiled Pennington-Worah matrices.
3. **Description of the phenomenon.** What is the intrinsic reason for which Pennington-Worah matrices exhibit such a simple behavior? To progress in this question, we propose in Section [X] a decomposition of a Pennington-Worah as the sum of

$$Y(h) = Y^{\text{lin}} + \sum_{m \geq 2} Y_m^{\text{per}}(h) + Y^{\text{def}}(h) + \epsilon(h),$$

as we state in Section X the joint convergence of each ingredient of the above sum toward an ingredient of a linear plus chaos decomposition. Our interpretation of the emergence of chaos is presented in Remark X.

2 Hypotheses and definition of traffic equivalence

2.1 Matrix model and decomposition

2.1.1 Model of variance profiled matrices

We present in this section our model.

Hypothesis 2.1. *The two random rectangular matrices W and X can be written*

$$W = \Gamma_w \circ W', \quad X = \Gamma_x \circ X'$$

where \circ denotes the entry-wise product of matrices.

1. *The matrices W and X are respectively of size $N_1 \times N_0$ and $N_0 \times N_2$, and setting $N = N_0 + N_1 + N_2$ the sequences defined by $\psi_i = \frac{N_i}{N}$, $i = 0, 1, 2$, converge to positive number.*
2. *The matrices of W' and X' independent and have centered real i.i.d. entries with variance one, the laws of their entries do not depend on N and have finite moments of all order.*
3. *The entries of Γ_w and Γ_x are bounded, and (Γ_w, Γ_x) converges in graphons topology, see Definition 2.6 below. This holds if $\Gamma_w = (\gamma_w(\frac{i}{N}, \frac{j}{N}))_{i,j}$ and $\Gamma_x = (\gamma_x(\frac{i}{N}, \frac{j}{N}))_{i,j}$, where γ_w and γ_x are piecewise continuous maps $[0, 1]^2 \rightarrow \mathbb{R}$. We call Γ_w and Γ_x the variance profiles (or simply profiles) of w and x .*

2.1.2 Hermite polynomials

We recall the special role of Hermite polynomials $(g_n)_{n \geq 0}$ to understand Theorem 1.4. Recall that $\omega(t) = \frac{e^{-t^2/2}}{\sqrt{2\pi}}$ and

$$g_n : t \mapsto \frac{(-1)^n \omega^{(n)}(t)}{\omega(t)}, \quad n \geq 0,$$

where $\omega^{(n)}$ denotes the n -th derivative of ω . The collection $(g_n)_{n \geq 0}$ is an orthonormal basis of the space of $\mathbb{C}[y]$ with respect to the standard Gaussian law, normalized such that $\int_{\mathbb{R}} g_n g_m d\omega = \delta_{n,m} n!$. The symbol $\delta_{n,m}$ stands for the usual Kronecker symbol. For any $n \geq 1$, we have $g'_n = n g_{n-1}$, $\forall n \geq 1$ so in particular, $\int_{\mathbb{R}} g_n^{(m)}(t) \omega(t) dt = \delta_{n,m} n!$ for any $m, n \geq 1$. Hence, for any $h : \mathbb{R} \rightarrow \mathbb{R}$ in $L^2(d\omega)$,

$$\frac{\int h^{(n)}(t) \omega(t) dt}{n!}$$

is the coefficient of to the n -th Hermite polynomial in the basis $(g_n)_{n \geq 0}$.

For any $h : \mathbb{R} \rightarrow \mathbb{R}$, let $\tilde{h} := \frac{h - \sqrt{\theta}g_1}{\sqrt{1-\theta}}$ the orthogonal projection of h on the orthogonal of g_1 with respect to $d\omega$. Since $Y(h)$ is linear in h , it is the sum of the matrices

$$Y(h) = \sqrt{\theta}Y(g_1) + \sqrt{1-\theta}Y(\tilde{h}).$$

From Theorem 1.4, $Y(g_1)$ is spectral equivalent to $\frac{W^{\text{Gau}}X^{\text{Gau}}}{\sqrt{N_0N_2}}$ and $Y(\tilde{h})$ is spectral equivalent to $\frac{Z^{\text{Gau}}}{\sqrt{N_0}}$. The next section introduces a more precise strategy to decompose the matrix, where the Hermite coefficients show up naturally.

2.1.3 A decomposition of Pennington-Worah matrices

Let $Y(h) = \gamma h[\{\gamma_0 W X\}]$ be a Pennington-Worah matrix, where W, X are rectangular matrices and $\gamma, \gamma_0 > 0$. For the polynomial function $h_n : x \mapsto x^n$, the matrix-entry definition gives the expression

$$Y(h_n) = \left(\gamma \gamma_0^n \sum_{\mathbf{d} \in [N_0]^n} \prod_{\ell=1}^n W(i, d_\ell) X(d_\ell, j) \right)_{\substack{i=1, \dots, N_1 \\ j=1, \dots, N_2}}.$$

Our decomposition involves several definitions.

Definition 2.2. 1. A set partition of a set \mathcal{X} , simply called a partition, is a set of non-empty subsets of \mathcal{X} , called its blocks, whose union is \mathcal{X} . We denote by $\mathcal{P}(\mathcal{X})$ the set of partitions of \mathcal{X} . Moreover, for any multi-index $\mathbf{k} = (k_1, \dots, k_n)$, we denote by $\ker \mathbf{k}$ the set partition of $[n] := \{1, \dots, n\}$ such that $p \overset{\ker \mathbf{k}}{\sim} q$ if and only if $k_p = k_q$.

2. An integer partition of an integer $n \geq 1$ is a non-increasing tuple $\lambda^{(n)} = (\lambda_1, \dots, \lambda_\ell)$ of integers, called its parts, which sum up to n . We denote $\lambda^{(n)} \vdash n$ to say that $\lambda^{(n)}$ is an integer partition of n . Moreover, a set partition π of $[n]$ is said to be of type $\lambda^{(n)}$ whenever $\lambda^{(n)}$ is its sequence of blocks size.

3. For any integer partition $\lambda^{(n)}$ of n and any partition π_0 of type $\lambda^{(n)}$, we set

$$Z(\lambda^{(n)}) := \left(\sum_{\substack{\mathbf{d} \text{ s.t.} \\ \ker(\mathbf{d}) = \pi_0}} \prod_{\ell=1}^n W(i, d_\ell) X(d_\ell, j) \right)_{\substack{i=1, \dots, N_1 \\ j=1, \dots, N_2}}.$$

The above expression clearly depends only on the type of π_0 . Therefore we have a canonical decomposition

$$Y(h_n) = \sum_{\lambda^{(n)} \vdash n} c_{\lambda^{(n)}} \times \gamma \gamma_0^n Z(\lambda^{(n)}),$$

where $c_{\lambda^{(n)}}$ is the number of $\pi \in \mathcal{P}(n)$ of type $\lambda^{(n)}$. In particular for $\lambda_2^{(n)} := (2, \dots, 2) \vdash n$, then $c_{\lambda^{(n)}}$ is the number of pair partitions of $[n]$, which is equal to $\int_{\mathbb{R}} h_n(t) d\omega(t)$. Moreover, for any $n \geq 1$, with $\lambda_1^{(n)} := (2, \dots, 2, 1) \vdash n$ and $\lambda_m^{(n)} := (m, 2, \dots, 2) \vdash n$, for $m \geq 2$, we have

$$c_{\lambda_m^{(n)}} = \binom{n}{m} \int_{\mathbb{R}} t^{n-m} d\omega(t) = \frac{1}{m!} \int_{\mathbb{R}} h_n^{(m)}(t) d\omega(t), \quad m \neq 2,$$

where $h_n^{(m)}$ is the m -th derivative of h_n . For any $1 < m \leq n$ we denote $\lambda_1^{(m,n)} = (2, \dots, 2, 1, \dots, 1) \vdash n$ the integer partition with m parts equal to one, for which we also have $c_{\lambda_1^{(m,n)}} = c_{\lambda_m^{(n)}}$.

Definition 2.3. For any $h = \sum_{n \geq 1} a_n h_n$, odd polynomial or an analytic function satisfying Hypothesis 1.2, we denote

$$Y(h) = Y^{\text{lin}}(h) + \sum_{m \geq 1} Y_m^{\text{per}}(h) + Y^{\text{def}}(h) + \epsilon(h),$$

where we have set

$$\begin{aligned} Y^{\text{lin}}(h) &= \sum_{n \geq 1} a_n \left(\int_{\mathbb{R}} h_n'(t) d\omega(t) \right) \times \gamma \gamma_0^n Z(\lambda_1^{(n)}), \\ Y_m^{\text{per}}(h) &= \sum_{n \geq 1} \frac{a_n}{m!} \left(\int_{\mathbb{R}} h_n^{(m)}(t) d\omega(t) \right) \times \gamma \gamma_0^n Z(\lambda_1^{(m,n)}), \quad \forall m \geq 2 \\ Y^{\text{def}}(h) &= \sum_{n \geq 1} \frac{a_n}{6} \left(\int_{\mathbb{R}} h_n'''(t) d\omega(t) \right) \times \gamma \gamma_0^n Z(\lambda_3^{(n)}), \\ \epsilon(h) &= \sum_{n \geq 1} \sum_{\substack{\lambda^{(n)} \vdash n \\ \lambda^{(n)} \neq \lambda_3^{(n)}, \\ \lambda^{(n)} \neq \lambda_1^{(m,n)} \forall m \geq 1}} a_n c_{\lambda^{(n)}} \times \gamma \gamma_0^n Z(\lambda^{(n)}). \end{aligned}$$

Our main result shows that an asymptotic equivalent for these matrices in a sense that is clarified next section, for which the weights can be computed by simple asymptotic rules for the matrices $Z(\lambda^{(n)})$.

2.2 Traffic equivalent

A difficulty arise to consider the decomposition of the previous section from the algebraic aspect. We shall go beyond free probability and use the notion of traffic equivalent to fit the nature of the Pennington-Worah matrix decomposition, by introducing a generalization of non-commutative polynomials. Although this notion is not a rigorously an intermediate notion between spectral and strong spectral equivalence, the reader can skip this

section with this idea in mind, without major consequence for the understand of next section (note that our approach does not prove the convergence of outliers but gives a proposal for the matrix deformation).

A graph is a couple (V, E) where V is non empty set called the vertex set, and E is a multi-ensemble of couples of elements of V , possibly empty, called the edge set. Multi-ensemble means that each element appear with a given multiplicity. The graph are directed: for $e = (v, w) \in E$, we call v the source of e , w its target.

Definition 2.4. Let Ω be a label set and $\mathbf{x} = (x_\omega)_{\omega \in \Omega}$ a collection of formal variables.

1. A test graph labeled by Ω (or in the variables \mathbf{x}) is a triplet $T = (V, E, \gamma)$ where (V, E) is a graph, and $\gamma : E \rightarrow \Omega$ is a map associating the variable $x_{\gamma(e)}$ to the edge $e \in E$.
2. A graph monomial labeled by Ω is the couple $g = (T, \text{in}, \text{out})$ where in and out are two vertices of the test graph T . They are respectively called the input and out of g .

We denote by $\mathcal{G}\langle\Omega\rangle$ the set of connected graph monomial labeled by Ω and by $\mathbb{C}\mathcal{G}\langle\Omega\rangle$ the vector space generated by $\mathcal{G}\langle\Omega\rangle$.

The following definition shows how we can evaluate a graph polynomial in matrices to define a new matrix, generalizing the matrix product.

Definition 2.5. Let $\mathbf{A} = (A_\omega)_{\omega \in \Omega}$ be a collection of matrices $\mathbb{M}_N(\mathbb{R})$ and let $g = (T, \text{in}, \text{out}) \in \mathcal{G}\langle\Omega\rangle$ be a graph monomial, where $T = (V, E, \sigma)$. The evaluation of g in the family \mathbf{A}_N the matrix $g(\mathbf{A}_N)$ with entry $(i, j) \in [N]^2$

$$g(\mathbf{A}_N)(i, j) = \sum_{\substack{\varphi: V \rightarrow [N] \text{ s.t.} \\ \varphi(\text{out})=i, \varphi(\text{in})=j}} \prod_{e=(v,w) \in E} A_{\gamma(e)}(\varphi(w), \varphi(v)). \quad (2.1)$$

Assume that the entries of the matrices have finite moments of all orders. We call traffic distribution of \mathbf{A} the map

$$\Phi_{\mathbf{A}} : g \in \mathcal{G}\langle\Omega\rangle \mapsto \mathbb{E} \left[\frac{1}{N} \text{Tr} [g(\mathbf{A}_N)] \right] \in \mathbb{R}. \quad (2.2)$$

We say that \mathbf{A} converges in traffic distribution whenever $\Phi_{\mathbf{A}}(g)$ converges for all $g \in \mathcal{G}\langle\Omega\rangle$ as the size N goes to infinity, and we say that \mathbf{A} and \mathbf{B} are traffic equivalent if they converge the same limit.

Definition 2.6. A collection Γ of deterministic matrices converges in graphon distribution whenever for any test graph T labeled by Ω ,

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\prod_{e=(v,w) \in E} \Gamma_{\gamma(e)}(\Phi(w), \Phi(v)) \right] \text{ exists,} \quad (2.3)$$

where $\Phi : V \mapsto [N]$ is a random injective map uniformly distributed.

Example 2.7. Let $Y = h[\{WX\}]$, where W and X are square matrices and $h_n : n \mapsto x^n$. The entry-wise definitions imply that $Y = g_n(W, X)$ where g_n is the graph monomial in two variables w, x is as follow: its vertex set is $\{\text{in}, \text{out}, v_1, \dots, v_n\}$ and for each $i = 1, \dots, n$, there is one edge with source in and target v_i labeled x , and one edge with source v_i and target out labeled w , see Figure 1. We call v_1, \dots, v_n the internal vertices of g_n .

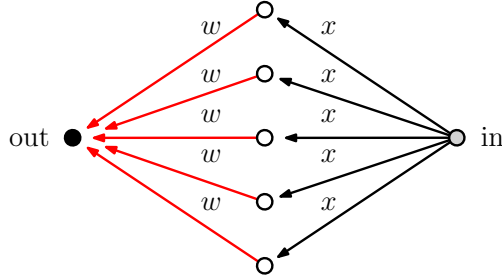


Figure 1: The graph monomial g_n of Pennington-Worah matrices. Note the design symmetry that the graph monomial is unchanged if we permute the internal vertices. Edges colors serves to distinguish edges labels.

The rest of the section defines how we see rectangular matrices as submatrices of large square matrices. Let N_0, N_1, N_2 three matrix size integers as in Section 1 and set $N = N_0 + N_1 + N_2$. A matrix of $M_N(\mathbb{R})$ is seen as a 3×3 rectangular block matrix

$$A = \begin{pmatrix} A_{0,0} & \cdots & A_{2,2} \\ \vdots & & \vdots \\ A_{2,0} & \cdots & A_{2,2} \end{pmatrix},$$

where $A_{i,j} \in M_{N_i, N_j}(\mathbb{R})$ for each $i, j = 0, 1, 2$. If $B \in M_{N_i, N_j}(\mathbb{R})$, we set $\iota_{i,i'}(B) \in M_N(\mathbb{R})$ the matrix A such that $A_{i,j} = B$ and $A_{i',j'} = 0$ for $(i', j') \neq (i, j)$ in the block decomposition. Let Ω be a label set and $\mathbf{s} = (s_\omega, s'_\omega)_{\omega \in \Omega} \in \{0, 1, 2\}^2$ be a collection of indices. We say that $\mathbf{A} = (A_\omega)_{\omega \in \Omega}$ is a collection of \mathbf{s} -rectangular matrices if A_ω is of size $N_{s_\omega} \times N_{s'_\omega}$ for all $\omega \in \Omega$ and we set $\iota_{\mathbf{s}}(\mathbf{A}) = (\iota_{(s_\omega, s'_\omega)}(A_\omega))_{\omega \in \Omega}$. We call traffic distribution of a collection \mathbf{A} of \mathbf{s} -rectangular matrices the traffic distribution of $\iota_{\mathbf{s}}(\mathbf{A})$. Similarly, the definition of graphon convergence extends for rectangular matrices.

3 Presentation of the results

3.1 Statement of our result

We say in short that a random matrix is a standard Gaussian matrix if it has i.i.d. centered real Gaussian entries of variance one. For a matrix A and

an integer $k \geq 1$ with denote $\sqrt[k]{A} := (\sqrt[k]{A(i, j)})_{i, j}$ and $A^{\circ k} = \underbrace{A \circ \dots \circ A}_{k \text{ times}}$.

Theorem 3.1. *Let $\mathbf{Y} = (Y(h))_{h \in \mathbb{C}(\text{odd})[y]}$ be the collection of matrices indexed by odd polynomials*

$$Y(h) = \frac{\sqrt{\psi_0}}{\sqrt{N}} h \left[\left\{ \frac{WX}{\sqrt{N_0}} \right\} \right], \quad \psi_0 = \frac{N_0}{N},$$

where $W = \Gamma_w \circ W'$ and $X = \Gamma_x \circ X'$ satisfy 2.1. We denote by $\mathbf{Y}^{\text{lin}}, \mathbf{Y}^{\text{per}}, \mathbf{Y}^{\text{def}}$ and ϵ the collections of matrices of Definition 2.3 with $\gamma = \frac{\sqrt{\psi_0}}{\sqrt{N}}$ and $\gamma_0 = \frac{1}{\sqrt{N_0}}$.

Then $(\mathbf{Y}^{\text{lin}}, \mathbf{Y}^{\text{per}}, \mathbf{Y}^{\text{def}}, \epsilon)$ is traffic equivalent to $(\mathbf{Y}^{\diamond \text{lin}}, \mathbf{Y}^{\diamond \text{per}}, \mathbf{Y}^{\diamond \text{def}}, \mathbf{0})$ where $\mathbf{Y}^{\diamond \text{lin}}$ and $\mathbf{Y}^{\diamond \text{per}}$ are independent, $\mathbf{Y}^{\diamond \text{def}}$ is deterministic, and are defined as follows. Setting the bounded matrix $M_2 := \sqrt[k]{N_0^{-1} \Gamma_w^{\circ 2} \times \Gamma_x^{\circ 2}}$, we have

$$Y^{\diamond \text{lin}}(h) := \left(\int_{\mathbb{R}} h' \left[\left\{ tM_2 \right\} \right] \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \right) \circ \left(\frac{W^{\text{Gau}}}{\sqrt{N}} \times \frac{X^{\text{Gau}}}{\sqrt{N}} \right).$$

where $W^{\text{Gau}} = \Gamma_w \circ W^{\text{Gau}}$, $X^{\text{Gau}} = \Gamma_x \circ X^{\text{Gau}}$, and $W^{\text{Gau}}, X^{\text{Gau}}$ are independent standard Gaussian matrices. Moreover, for any $m \geq 2$, we have

$$Y_m^{\diamond \text{per}}(h) := \frac{1}{m!} \left(\int_{\mathbb{R}} h^{(m)} \left[\left\{ tM_2 \right\} \right] \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \right) \circ \frac{Z_m^{\text{Gau}}}{\sqrt{N}},$$

where $Z_m^{\text{Gau}}, m \geq 2$, are independent standard Gaussian matrices. Finally, setting $\Lambda_3 = N^{-1} \Gamma_w^{\circ 3} \times \Gamma_x^{\circ 3}$, we have

$$Y^{\diamond \text{def}}(h) := \frac{m_w^{(3)} m_x^{(3)}}{6N} \Lambda_3 \circ \left(\int_{\mathbb{R}} h''' \left[\left\{ tM_2 \right\} \right] \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \right).$$

The expression of $Y^{\diamond \text{lin}}(h)$ is linear in the matrix $\frac{W^{\text{Gau}}}{\sqrt{N}} \times \frac{X^{\text{Gau}}}{\sqrt{N}}$, but now compared to Theorem 1.4 the linear relation means an entry-wise product. Similarly, the noise part $Y^{\diamond \text{def}}(h)$ is now a variance profile Gaussian random matrix. The deformation has finite rank when the profiles are constant, and its entries are $O(N^{-1})$. This deformation is known from [Mal20] to do not change the limit of the limiting singular-values distribution.

Example 3.2. *In the context of Theorem 3.1, assume that one of the random matrices has a constant variance profile, for instance $\Gamma_x(i, j) = 1$ for all i, j . Therefore we see that the profile of the linear part is a rank one matrix: setting $D(h)$ the diagonal matrix whose diagonal entries coincide with those of this profile matrix, we get that $Y(h)$ is traffic equivalent to*

$$\frac{W^{\text{Gau}} X^{\text{Gau}}}{N} D(h) + \Theta(h) \circ \frac{Z^{\text{Gau}}}{\sqrt{N}} + B(h), \quad (3.1)$$

where Θ is a profile and $B(h)$ a deformation that dot not change the limiting empirical singular-values distribution.

3.2 Heuristic and comments

The contribution can be computed from Definition 2.3 with the following approximations.

1. The contribution for the linear part can be approximated jointly in traffic distribution thanks to factorizations rules

$$\begin{aligned}\gamma\gamma_0^n Z(\lambda_1^{(n)}) &\sim \frac{\sqrt{\psi_0}}{\sqrt{N}} \times \frac{1}{\sqrt{N_0}^n} (W^{\text{Gau}} X^{\text{Gau}}) \circ (\Gamma_w^{o2} \times \Gamma_x^{o2})^{\circ \frac{(n-1)}{2}} \\ &= \left(\frac{W^{\text{Gau}}}{\sqrt{N}} \frac{X^{\text{Gau}}}{\sqrt{N}} \right) \circ \left(\frac{1}{N_0} \Gamma_w^{o2} \times \Gamma_x^{o2} \right)^{\circ \frac{(n-1)}{2}}\end{aligned}$$

which indeed gives the expected equivalent

$$Z^{\text{lin}}(h) \sim \sum_{n \geq 1} a_n \left(\int_{\mathbb{R}} h'_n[\{tM_2\}] d\omega(t) \right) \circ \left(\frac{W^{\text{Gau}}}{\sqrt{N}} \frac{X^{\text{Gau}}}{\sqrt{N}} \right) = Y^{\diamond \text{lin}}(h).$$

2. For the deformation term, we have asymptotically

$$Z(\lambda_3^{(n)}) \sim \frac{m_w^{(3)} m_x^{(3)}}{6N} \Lambda_3 \circ M_2^{\circ(n-1)},$$

so again the definition of Y^{def} results in an entry-wise evaluation formula for $Y^{\diamond \text{def}}$.

3. Finally, for each $m \geq 2$,

$$\begin{aligned}\gamma\gamma_0^n Z(\lambda_1^{(m,n)}) &\sim \frac{\sqrt{\psi_0}}{\sqrt{N}} \times \frac{1}{\sqrt{N_0}^n} Z(\lambda_1^{(m,m)}) \circ (\Gamma_w^{o2} \times \Gamma_x^{o2})^{\circ \frac{(n-m)}{2}} \\ &= \sqrt{N_0}^{m-1} \times \frac{\psi_0}{N_0^{2m}} Z(\lambda_1^{(m,m)}) \circ \left(\frac{1}{N_0} \Gamma_w^{o2} \times \Gamma_x^{o2} \right)^{\circ \frac{(n-m)}{2}}\end{aligned}$$

Denoting $Z = \frac{W}{\sqrt{N_0}} \frac{X}{\sqrt{N_0}} = Z_1 \times Z_2$, we have for instance with $m = 2$

$$\frac{1}{N_0^{2m}} Z(\lambda_1^{(2,2)}) = \left(Z \circ Z - Z_1^{\circ 2} \circ Z_2^{\circ 2} \right),$$

which is known to converges to zero in traffic distribution. Hence the perturbative term can be interpreted as fluctuations around this limit. The computations show that $\sqrt{N_0}^{m-1} \times \frac{\psi_0}{N_0^{2m}} Y(\lambda_1^{(m,m)})$ is traffic equivalent to

$$\left(\frac{1}{N_0} \Gamma_w^{o2} \times \Gamma_x^{o2} \right)^{\circ \frac{m}{2}} \circ \left(\int_{\mathbb{R}} h_m d\omega(t) \right) \circ \frac{Z_m^{\text{Gau}}}{\sqrt{N}},$$

which gives the expected equivalent.

In the next section we state the generalization of the above theorem in terms of traffic distribution, for which the analog of Equation (3.1) requires a third additional term.

3.3 Péché perspective and free probability

To emphasize the strength of the equivalent method, recall that free probability gives a **robust and systematic way to compute eigenvalues distributions**. In the context of Theorem 1.4, since the spectral equivalent is a polynomial in independent matrices, the consistent system of equations for the its Stieltjes transform for the distribution of YY^t , which are crucial for numerical applications, can be derived thanks to Dan Virgil Voiculescu’s free probability theory [Voi91]. In particular, a linearization trick developed by [BMS17] allows to reformulate this system as the fundamental subordination property. This approach is quite robust since for more structured spaces the subordination property holds ”with a twist“, formulated thanks to the powerful concept of amalgamation (a non commutative analogue of conditional probability). The method has its limitations, the systems of equations for Stieltjes transforms obtained by free probability may be quite complicated, but it provides explicit algorithms.

Moreover, **the method extends for strong equivalents**. The strong convergence of independent real Gaussian matrices is known by a result of Catherine Donati-Martin and Mireille Capitaine [CDM07] and the latter author show that finite rank deformations can be described in a simple way thanks to the subordination property [Cap13]. Therefore the computation of the outliers for the strong equivalent can be made with the same robust method as for the computation of the eigenvalues distribution.

Beyond free probability, our approach use a specialization of free probability called traffic probability [Mal20], which is motivated by the distributional invariance of Pennington-Worah matrices. Let us denote by $\mathcal{O} = \mathcal{O}_N$ the set of size N orthogonal real matrices and by $\text{Perm} = \text{Perm}_N$ of size N permutation matrices are two compact subgroups $M_N(\mathbb{R})$. Let us describe the two following situations, given a collection of random matrices $\mathbf{A} = (A_j)_{j \in J}$ of size $N_1 \times N_2$, where implicitly N_1 and N_2 goes to infinity such that $\frac{N_i}{N} \xrightarrow{N \rightarrow \infty} \psi_i > 0$ for an auxiliary parameter $N \rightarrow \infty$.

1. The collection of matrices is bi-unitarily invariant in the sense that \mathbf{A} has the same law as $U\mathbf{A}V = (UA_jV)_{j \in J}$ for all orthogonal matrices U, V of corresponding size. **Then we can use Florent Benaych-Georges’s rectangular free probability theory** [BG09], which correspond to the setting of amalgamation over a space of finite dimensional matrices. In the context of Theorem 1.4, we have convergence in joint non-commutative distribution

$$(Y^{\square \text{lin}}(h), Y^{\square \text{per}}(h)) \xrightarrow{N \rightarrow \infty} (y^{\text{lin}}(h), y^{\text{per}}(h))$$

in the sense that for any polynomial P in two non commutative indeterminates consistent with relative matrix sizes, the limit

$$\mathbb{E} \left[\frac{1}{N} P(Y^{\square \text{lin}}(h), Y^{\square \text{per}}(h)) \right] \xrightarrow{N \rightarrow \infty} \Phi(P),$$

exists. The map Φ is a linear form on the space of non commutative polynomials. The framework of non-commutative probability allows to think these indeterminates as random variables by "mimicking" classical probability in a non-commutative way: the limit of matrices are called non-commutative random variables and their are determined by their non-commutative distribution Φ . Non-commutative random variables are understood to be in "generic position" when there are freely independent (or simply free), which means that their distribution satisfies a canonical expression in terms of marginal distribution. The bi-orthogonal invariance of $Y^{\square \text{lin}}(h)$ and $Y^{\square \text{per}}(h)$ explains that their limit are free.

2. The collection of matrices is bi-permutation invariant in the sense that \mathbf{A} has the same law as $U\mathbf{A}V = (UA_jV)_{j \in J}$ for all permutation matrices U, V . **Then we can use Greg Zitelli rectangular traffic probability theory** [Zit24]. Note that the collection of matrices $Y(h)$ itself is bi-permutation invariant. For a collection \mathcal{H} of functions h and under assumptions stated next section, we have convergence in traffic distribution

$$\mathbf{Y} = (Y(h))_{h \in \mathcal{H}} \xrightarrow{N \rightarrow \infty} \mathbf{y} = (y(h))_{h \in \mathcal{H}},$$

in the sense the for any function g in a class of functions defined below and called the split graph-polynomials, the limit

$$\mathbb{E} \left[\frac{1}{N} g(\mathbf{Y}) \right] \xrightarrow{N \rightarrow \infty} \Phi(g),$$

exists. As graph polynomials generalize non commutative polynomials, the traffic distribution extends the non-commutative distribution. The interest is that the traffic distribution gives much more information, capturing certain finite rank deformations and distinguishing easily bi-unitarily invariant matrices. Traffic distributions comes with a notion of traffic-independence which "encompasses" [GC24] three notions of probability: classical probability referred as tensor-probability, the free probability, and the Boolean-probability which is another non-commutative non of independence.

4 Asymptotic traffic distribution

4.1 Definitions and Notations

Our method is the same as in [BP21], using the language of rectangular traffic probability. In this first subsection, we present a slight reformulation of the traffic distribution of matrices relating the square and rectangular cases. We also introduce an important transform, the injective traffic distribution, and the ingredient to link these two notions. We recall that for $i, i' \in \{0, 1, 2\}$, a $N_i \times N_{i'}$ matrix A is canonically associated a square matrix $\iota_{i,i'}(A)$ of size $N = N_0 + N_1 + N_2$ by completing A with zeros.

Let $\mathbf{s} = (s_\omega, s'_\omega)_{\omega \in \Omega} \in \{0, 1, 2\}^2$ be a collection of matrix size indices. A test graph $T = (V, E, \gamma)$ labeled in Ω is split (with respect to \mathbf{s}) if there is a partition $V = V_0 \sqcup V_1 \sqcup V_2$ such it links edges whose label matches the size indices: for any $e = (v, w) \in E$, $v \in V_p$ and $w \in V_q$ implies $(s_{\gamma(e)}, s'_{\gamma(e)}) = (q, p)$. A map $\phi : V \rightarrow [N]$ is split whenever it sends V_i to the range isomorphic to $[N_i]$ in $[N]$.

Definition 4.1. • Let $\mathbf{s} = (s_\omega, s'_\omega)_{\omega \in \Omega} \in \{0, 1, 2\}^2$ and $\mathbf{A} = (A_\omega)_{\omega \in \Omega}$ be a family of \mathbf{s} -rectangular matrices, and denote $\tilde{A}_\omega = \iota_{(s_\omega, s'_\omega)}(A_\omega)$ for all $\omega \in \Omega$. Let $T = (V, E, \gamma)$ be a split test graph labeled by Ω . We call combinatorial trace of T in the family \mathbf{A}_N the complex number

$$\mathrm{Tr}[T(\mathbf{A})] = \sum_{\varphi: V \rightarrow [N]} \prod_{e \in E} \tilde{A}_{\gamma(e)}(\varphi(e)), \quad (4.1)$$

where for $e = (v, w)$ we denote $\varphi(e) = (\varphi(w), \varphi(v))$. The sum can be restricted to split maps since otherwise the summand vanishes. We call injective trace of T in the family \mathbf{A}_N , and denote $\mathrm{Tr}^0[T(\mathbf{A})]$, the quantity defined as above when the sum over φ is restricted to injective split maps. We denote

$$\tau_N[T(\mathbf{A})] := \mathbb{E} \left[\frac{1}{N} \mathrm{Tr}[T(\mathbf{A})] \right], \quad \tau_N^0[T(\mathbf{A})] := \mathbb{E} \left[\frac{1}{N} \mathrm{Tr}^0[T(\mathbf{A})] \right].$$

- For any test graph $T = (V, E, \gamma)$ and any partition π of its vertex set V , we denote by $T^\pi = (V^\pi, E^\pi, \gamma^\pi)$ the test graph obtained by identifying vertices in a same block of the partition π : $V^\pi = \pi$ and each edge $e = (v, w)$ in E induces an edge with source and target the block of π that contains v and w respectively. A partition $\pi \in \mathcal{P}(V)$ of V is split whenever π does not identify vertices of different V_i 's, namely $\mathcal{V}^\pi = V_0^\pi \sqcup V_1^\pi \sqcup V_2^\pi$.

There is an abuse of notation using the term "trace" in the theorem above since the combinatorial and the injective traces are not defined on matrix spaces. The combinatorial trace is related to the usual trace of

matrices as follows: for any family of matrices \mathbf{A} and any graph monomial $g = (T, \text{in}, \text{out})$ one has $\text{Tr } g(\mathbf{A}_N) = \text{Tr}[T'(A)]$, where T' is obtained from T by identifying in and out. On the other hand, for any test graph T , we have the identity

$$\text{Tr}[T(\mathbf{A})] = \sum_{\pi \in \mathcal{P}(V)} \text{Tr}^0[T^\pi(\mathbf{A})],$$

where $\mathcal{P}(V)$ is the set of partitions of V . The sum can be restricted to the split partitions π of V .

5 Proof of the main theorem

5.1 Traffic method for profiled Pennington-Worah matrices

Let us set $\mathbf{Y} = (Y_h)_{h \in \mathbb{R}[y]}$ the family of matrices

$$Y(h) = \frac{\sqrt{\psi_0}}{\sqrt{N}} h \left[\left\{ \frac{WX}{\sqrt{N_0}} \right\} \right], \quad \psi_0 = \frac{N_0}{N}, \quad (5.1)$$

where W and X are as in Theorem 3.1, and $h[\{ \cdot \}]$ stands for the entry-wise evaluation. Since (5.1) is a linear expression in h , to prove the convergence in traffic distribution of \mathbf{Y} it suffices to prove the convergence of $\mathbb{E}[\frac{1}{N} \text{Tr}[T(\mathbf{Y})]]$ for test graphs indexed by a basis of $\mathbb{R}[y]$.

We call reference graphs the split test graphs T in variables standing for the matrices of \mathbf{Y} labeled by monomials. More precisely, we denote $T = (V, E, \mathbf{n})$ where the index map $\mathbf{n} : E \rightarrow \mathbb{N}$ tells that an edge $e \in E$ is associated to the matrix $Y_{\mathbf{n}(e)} := Y(h_{\mathbf{n}(e)})$ where $h_n : x \mapsto x^n$. We have a partition $V = V_1 \sqcup V_2$ so that each edge of T has its source in V_2 and its target in V_1 . Later at the end of the proof we shall use test graph labeled by Hermite polynomials.

We give an expression of $\mathbb{E}[\frac{1}{N} \text{Tr}[T(\mathbf{Y})]]$ for any reference graph T , we use first the so-called substitution property of graph monomials and test graphs (see [Mal20, Definition 1.7]) to re-write (5.1) in graph language. The proof of the main theorem follows as we can clearly identify the contribution of each components in the Pennington-Worah matrix decomposition.

In the definition below, we use the terminology of [BP21].

Definition 5.1. *Given $T = (V, E, \mathbf{n})$ a reference test graph, we denote by $\mathcal{T}_T = (\mathcal{V}, \mathcal{E}, \delta)$ the (auxiliary) test graph in two variables w and x obtained from T as follows. It is obtained by considering the vertex set of the reference graph (that we then call the reference vertices) and adding for each edge $e \in E$ a collection $\mathcal{N}(e) = (\mathcal{V}_e, \mathcal{E}_e)$ of vertices and edges, that is called the niche of e , defined as follow. There are $\mathbf{n}(e)$ vertices in the niche of e , different from the reference vertices and the vertices of other niches. They are called the internal vertices. Moreover, the niche of e has $2\mathbf{n}(e)$ edges, as each internal vertex is both*

- the source of an edge labeled w whose target is the target of e in T ,
- the target of an edge denoted labeled x whose source is the source of e in T .

Two adjacent edges in \mathcal{T} that share the same internal vertex are say to be companion each other.

When there is no ambiguity about the reference graph T , we write in short $\mathcal{T} := \mathcal{T}_T$. The definitions imply that we have

$$\begin{aligned}\tau_N[T(\mathbf{Y})] &= \left(\frac{\psi_0}{N}\right)^{\frac{|E|}{2}} N_0^{-\sum_{e \in E} \frac{\mathbf{n}(e)}{2}} \tau_N[\mathcal{T}(W, X)] \\ &= \psi_0^{-\sum_{e \in E} \frac{\mathbf{n}(e)-1}{2}} N^{-\frac{|E|}{2} - \sum_{e \in E} \frac{\mathbf{n}(e)}{2}} \sum_{\pi \in \mathcal{P}(\mathcal{V})} \tau_N^0[\mathcal{T}^\pi(W, X)].\end{aligned}\quad (5.2)$$

Recall that Hypothesis 2.1 tells that $W = \Gamma_w \circ W'$ and $X = \Gamma_x \circ X'$, where W', X' are independent i.i.d. matrices and Γ_w, Γ_x are bounded matrices. Next we explain how we can factorize the contribution of profiles and work on the traffic distribution of W' and X' .

Definition 5.2. For any family of \mathbf{s} -rectangular random matrices $\mathbf{A} = (A_\omega)_{\omega \in \Omega}$ and any split test graph $\mathfrak{T} = (\mathfrak{V}, \mathfrak{E}, \gamma)$ labeled in Ω , with the same notations as in Definition 4.1, we set

$$\delta^0[\mathfrak{T}(\mathbf{A})] = \mathbb{E} \left[\prod_{e \in \mathfrak{E}} \tilde{A}_{\gamma(e)}(\Phi(e)) \right],$$

where $\Phi : \mathfrak{V} \rightarrow [N]$ is a injective split map uniformly distributed at random independently of \mathbf{A} .

Let us denote $(m)_n = \frac{m!}{(m-n)!}$ (known as the falling factorial notation), which is the number of injective maps from $[n]$ to $[m]$. From the definition of τ_N^0 , note that we have

$$\begin{aligned}\tau_N^0[\mathfrak{T}(\mathbf{A})] &= \frac{1}{N} \text{Card}\{\pi : \mathfrak{V} \rightarrow [N] \text{ split, injective}\} \delta^0[\mathfrak{T}(\mathbf{A})] \\ &= \frac{1}{N} (N_0)_{|\mathfrak{V}_0|} (N_1)_{|\mathfrak{V}_1|} (N_2)_{|\mathfrak{V}_2|} \delta^0[\mathfrak{T}(\mathbf{A})].\end{aligned}\quad (5.3)$$

On the other hand, for $\mathcal{T} = \mathcal{T}_T$ as in (5.2), we denote by \mathcal{E}_w^π and \mathcal{E}_x^π the set of edges of \mathcal{T} labeled w and x respectively. The independence of W' and X' implies

$$\begin{aligned}\tau_N^0[\mathcal{T}^\pi(W, X)] &= \frac{1}{N} \sum_{\substack{\phi: \mathcal{V}^\pi \rightarrow [N] \\ \text{injective}}} \left(\prod_{e \in \mathcal{E}_w^\pi} \Gamma_w(\phi(e)) \times \prod_{e \in \mathcal{E}_x^\pi} \Gamma_x(\phi(e)) \right) \\ &\quad \times \mathbb{E} \left[\prod_{e \in \mathcal{E}_w^\pi} W'(\phi(e)) \right] \times \mathbb{E} \left[\prod_{e \in \mathcal{E}_x^\pi} X'(\phi(e)) \right].\end{aligned}$$

Since the matrices W' and X' have i.i.d. entries, the values of the expectation does not change if we change the split partition π into another arbitrary split partition Φ . This is still true if Φ is uniformly distributed at random independently of the matrices. Let us define the test graph $\mathcal{T}_w = (\mathcal{V}_w, \mathcal{E}_w)$ in one variable w obtained from \mathcal{T} by removing the edges labeled x and the vertices of V_2 , and let \mathcal{T}_x be defined similarly. Hence we have

$$\begin{aligned} & \tau_N^0[\mathcal{T}^\pi(W, X)] \\ &= \frac{1}{N} \sum_{\substack{\phi: \mathcal{V}^\pi \rightarrow [N] \\ \text{injective}}} \left(\prod_{e \in \mathcal{E}_w} \Gamma_w(\phi(e)) \prod_{e \in \mathcal{E}_x} \Gamma_x(\phi(e)) \right) \delta^0[\mathcal{T}_w^\pi(W')] \delta^0[\mathcal{T}_x^\pi(X')] \\ &= \frac{1}{N} (N_0)_{V_0} (N_1)_{V_1} (N_2)_{V_2} \delta^0[\mathcal{T}^\pi(\Gamma_w, \Gamma_x)] \delta^0[\mathcal{T}_w^\pi(W')] \delta^0[\mathcal{T}_x^\pi(X')], \end{aligned} \tag{5.4}$$

Note that since the entries of Γ_w and Γ_x are bounded the term $\delta^0[\mathcal{T}^\pi(\Gamma_w, \Gamma_x)]$ is bounded. Moreover, $\delta^0[\mathcal{T}_w^\pi(W')] \delta^0[\mathcal{T}_x^\pi(X')] = \delta^0[\mathcal{T}^\pi(W', X')]$ is independent of N . Finally, recalling that $\mathcal{V}^\pi = V_0^\pi \sqcup V_1^\pi \sqcup V_2^\pi$ since π is split, we have

$$(N_0)_{V_0} (N_1)_{V_1} (N_2)_{V_2} = N^{|\mathcal{V}^\pi|} \psi_0^{|V_0^\pi|} \psi_1^{|V_1^\pi|} \psi_2^{|V_2^\pi|} (1 + o(1)).$$

Therefore, by (5.2) and (5.4), setting

$$\begin{aligned} \eta(\pi) &= -1 + |\mathcal{V}^\pi| - \frac{|E|}{2} - \sum_{e \in E} \frac{\mathbf{n}(e)}{2}, \\ \Psi^\pi &= \psi_0^{|V_0^\pi| - \sum_{e \in E} \frac{\mathbf{n}(e)-1}{2}} \psi_2^{|V_2^\pi|} \psi_1^{|V_1^\pi|}, \quad \psi_i = \frac{N_i}{N}, \end{aligned} \tag{5.5}$$

the following expression is valid for any reference graph T

$$\begin{aligned} & \tau_N[T(Y)] \\ &= \sum_{\pi \in \mathcal{P}(\mathcal{V})} N^{\eta(\pi)} \Psi^\pi \delta^0[\mathcal{T}^\pi(\Gamma_w, \Gamma_x)] \delta^0[\mathcal{T}_w^\pi(W')] \delta^0[\mathcal{T}_x^\pi(X')] (1 + o(1)) \end{aligned} \tag{5.6}$$

5.2 Convergence and support of the limit

In a first subsection, we prove that for a reference graph T as in the previous section and for any $\pi \in \mathcal{P}(\mathcal{V})$ such that $\delta^0[\mathcal{T}_w^\pi(W')] \times \delta^0[\mathcal{T}_x^\pi(X')] = \delta^0[\mathcal{T}^\pi(W', X')] \neq 0$, we always have $\eta(\pi) \leq 0$. The validity of this claim implies the convergence in traffic distribution of \mathbf{Y} since therefore in the r.h.s. of Formula (5.6) all terms are bounded. In a second subsection, we state an intermediate result that allows us to identify the partitions such that $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$ and $\eta(\pi) = 0$.

5.2.1 Proof of the convergence of the traffic distribution of \mathbf{Y}

In this subsection, we fix a split reference graph T . Let $\pi \in \mathcal{P}(\mathcal{V})$ be a split partition and denote by $\rho(\pi)$ the restriction of π to the vertex set of T . In the following expression

$$\eta(\pi) = |\mathcal{V}^\pi| - 1 - \frac{|E|}{2} - \sum_{e \in E} \frac{\mathbf{n}(e)}{2}, \quad (5.7)$$

only $|\mathcal{V}^\pi|$ depends on the partition π , so $\eta(\pi)$ is large when the number of vertices of \mathcal{T}^π is large. On the other hand, by independence of the matrices and their entries, we must have edges identifications in order to have $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$. We shall prove that the competition between these two constraints results in terms of the right order.

The strategy consists in translating the condition $\eta(\pi)$ into concrete conditions on two other intermediary graphs. First recall that \mathcal{T}_w is the graph labeled in w obtained from \mathcal{T} by removing the edges labeled x and the vertices of V_2 . It may be a disconnected graph, so we denote by $c_\pi \geq 1$ its number of connected components. Moreover, the number of edges labeled w in \mathcal{T}^π is $|\mathcal{E}_w^\pi| = |\mathcal{E}_w| = \sum_{e \in E} \mathbf{n}(e)$. Finally, since the partition π is split, the number of vertices of this graph is $|\mathcal{V}_w^\pi| = |V_1^\pi| + |V_0^\pi|$. Hence we have the identity

$$\eta_1(\pi) := |\mathcal{V}_w^\pi| - c_\pi - \frac{|\mathcal{E}_w^\pi|}{2} = |V_1^\pi| + |V_0^\pi| - c_\pi - \sum_{e \in E} \frac{\mathbf{n}(e)}{2}. \quad (5.8)$$

We put the emphasis on this formula since η_1 replaces η the computation of the traffic distribution of simpler models, such as independent large Wigner matrices in [Mal20, Chapter 3]. We recall the following definitions.

Definition 5.3. Let $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ be a graph.

1. The graph \mathfrak{G} is a forest whenever the removal of any edge of \mathfrak{G} always increases its number of connected components.
2. The skeleton graph $\bar{\mathfrak{G}} = (\bar{\mathfrak{V}}, \bar{\mathfrak{E}})$ of \mathfrak{G} is obtained by identifying the edges of \mathfrak{G} with the same endpoints, hence forgetting the multiplicity of the edges.

We also recall the following result (for a proof, see [Mal20, Lemma 2.13]).

Lemma 5.4. For any graph $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ with c connected components,

$$|\mathfrak{V}| - c - |\mathfrak{E}| \leq 0,$$

with equality if and only if the graph is a forest.

Therefore we re-rewrite (5.8) as

$$\eta_1(\pi) = \left(|\mathcal{V}_w^\pi| - c_\pi - |\bar{\mathcal{E}}_w^\pi| \right) + \left(|\bar{\mathcal{E}}_w^\pi| - \frac{|\mathcal{E}_w^\pi|}{2} \right),$$

where \mathcal{E}_w^π is the set of edges of the skeleton graph of \mathcal{T}_w^π . The first term in the r.h.s. is non-negative by Lemma 5.4. On the other hand, for any $i \geq 1$ denote by $\mathcal{E}_{w,i}^\pi \subset \mathcal{E}_w^\pi$ the set of edges of multiplicity equal to i in \mathcal{T}_w^π . Then we have

$$\left(|\bar{\mathcal{E}}_w^\pi| - \frac{|\mathcal{E}_w^\pi|}{2} \right) = \frac{|\mathcal{E}_{w,1}^\pi|}{2} - \sum_{i \geq 3} \frac{i-2}{2} |\mathcal{E}_{w,i}^\pi|. \quad (5.9)$$

If $\delta^0[\mathcal{T}_w^\pi(W')] = 0$, then T_w^π has no edge of multiplicity 1 so $(|\bar{\mathcal{E}}_w^\pi| - \frac{|\mathcal{E}_w^\pi|}{2}) \leq 0$ with equality if and only if the edges labeled w are of multiplicity 2 in \mathcal{T}_w^π .

As we cannot find a way to apply the same reasoning η defined in (5.7), we introduce a second graph. We consider $T^{\tilde{\rho}(\pi)}$ the quotient of T by the split partition $\tilde{\rho}(\pi)$ such that

- $\forall i, i' \in V_1, i \sim_{\tilde{\rho}(\pi)} i'$ whenever i and i' belong to the same connected component of the quotient of \mathcal{T}_w by the restriction of π ,
- $\rho(\pi)$ and $\tilde{\rho}(\pi)$ coincide on V_2 ($\forall j, j' \in V_2, j \sim_{\tilde{\rho}(\pi)} j' \Leftrightarrow j \sim_\pi j'$).

Note that the partition $\rho(\pi)$, which is induced by π by restriction, is finer than $\tilde{\rho}(\pi)$ in the sense that each block of $\rho(\pi)$ belong to a block of $\tilde{\rho}(\pi)$. Moreover $T^{\tilde{\rho}(\pi)}$ is a connected graph since T is connected. Its number of edges is the same as for T , namely $|E^{\tilde{\rho}(\pi)}| = |E|$. Its number of vertices is $|V^{\tilde{\rho}(\pi)}| = c_\pi + |V_2^\pi|$. Therefore we have

$$\eta_2(\pi) := |\mathcal{V}^{\tilde{\rho}(\pi)}| - 1 - \frac{|E^{\tilde{\rho}(\pi)}|}{2} = c_\pi + |V_2^\pi| - 1 - \frac{|E|}{2}, \quad (5.10)$$

and Formulae (5.7), (5.8) imply that $\eta(\pi) = \eta_1(\pi) + \eta_2(\pi)$. With $|\bar{E}^{\tilde{\rho}(\pi)}|$ standing for the number of edges of the skeleton of $T^{\tilde{\rho}(\pi)}$, we write as before

$$\eta_2(\pi) = \left(|\bar{\mathcal{E}}_w^\pi| - \frac{|\mathcal{E}_w^\pi|}{2} \right) + \left(|\bar{E}^{\tilde{\rho}(\pi)}| - \frac{|E^{\tilde{\rho}(\pi)}|}{2} \right), \quad (5.11)$$

The first term is non-negative when $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$. By a computation similar to (5.9), we see that last term $|\bar{E}^{\tilde{\rho}(\pi)}| - \frac{|E^{\tilde{\rho}(\pi)}|}{2}$ is bounded by half the number of edges of multiplicity one in $T^{\tilde{\rho}(\pi)}$, which may be positive. We must therefore show that when this quantity is positive, another quantity compensates it. The following definition clarifies phrasing that we use through the proof.

Definition 5.5. Let π be a partition of \mathcal{V} . We say that a group of edges of \mathcal{T}^π are identified by π if their target vertices belong to a same block of π , as long as their sources. An extra-niche (respectively intra-niche) identification is an identification of internal vertices or edges of \mathcal{T}^π from different niches (respectively the same niche). Two edges e and e' of T are niche neighbors (w -niche neighbors or x -niche neighbors) whenever they have in their niche edges forming an extra-niche identification (and these edges are labeled w or x respectively).

The following fact is used several times.

Lemma 5.6. Assume that T and $\pi \in \mathcal{P}(\mathcal{V})$ are split and $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$. If two edges e and e' of T are x -neighbors, then they form a group of edges of multiplicity at least 2 in $T^{\tilde{\rho}(\pi)}$.

Proof. Denote by e_x and e'_x two edges of e and e' respectively forming an extra-niche identification. The targets of e_x and e'_x coincide in \mathcal{T}^π , so the targets of e and e' belong to the the same connected component in \mathcal{T}^π . Hence by definition of $\tilde{\rho}$, e and e' form a group of multiplicity at least 2 in $T^{\tilde{\rho}(\pi)}$. \square

Lemma 5.7. Assume that T and $\pi \in \mathcal{P}(\mathcal{V})$ are split and $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$. Let $e \in E$ be a simple edge in $T^{\tilde{\rho}(\pi)}$. Then e has no x -neighbor, and the internal vertex of its niche form intra-niche identifications.

Proof. A edge $e \in E$ that is simple in $T^{\tilde{\rho}(\pi)}$ has no x -neighbor since otherwise it would contradict Lemma 5.6. Moreover, by the centering and the independent of the matrices W', X' and their entries, the condition $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$ implies that the edges labeled x in the niche of e must be identified somewhere, so necessary they form intra-niche identifications. \square

We can now prove that $\eta(\pi) \leq 0$. Let us denote by l^π the number of simple edges of $T^{\tilde{\rho}(\pi)}$ and by l'^π the number of edge of multiplicity at least 3 in \mathcal{T}_w^π . Since \mathcal{T}_w^π has no simple edge when $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$, then (5.9) shows that $|\bar{\mathcal{E}}_w^\pi| - \frac{|\mathcal{E}_w^\pi|}{2} \leq -\frac{l'^\pi}{2}$. On the other hand, since each niche contains an odd number of internal vertices, Lemma 5.7 implies that the niches corresponding to simple edges of $T^{\tilde{\rho}(\pi)}$ must have an edge of multiplicity at least 3, and so we have $l^\pi \leq l'^\pi$. Moreover, from the definitions of l^π we get $|\bar{E}^{\tilde{\rho}(\pi)}| - \frac{|E^{\tilde{\rho}(\pi)}|}{2} \leq \frac{l^\pi}{2}$. Formula (5.11) and the above arguments imply that $\eta(\pi) \leq \frac{l^\pi - l'^\pi}{2} \leq 0$ whenever $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$ and so $\eta(\pi) \leq 0$. This proves the convergence $C_N(\mathcal{T}^\pi)$, and so the convergence in traffic distribution of \mathbf{Y} as explain in the presentation of Section 5.2.

5.2.2 Support of the traffic distribution

The condition $\eta(\pi) = 0$ is hence equivalent to the following four conditions for the intermediary graphs \mathcal{T}_w^π and $T^{\tilde{\rho}(\pi)}$ introduced before:

$$-1 + c_\pi + |V_2^\pi| - |\bar{E}^{\tilde{\rho}(\pi)}| = 0, \quad (5.12)$$

$$|\bar{E}^{\tilde{\rho}(\pi)}| - \frac{|E^{\tilde{\rho}(\pi)}|}{2} = \frac{l^\pi}{2}, \quad (5.13)$$

$$|\bar{\mathcal{E}}_w^\pi| - \frac{|\mathcal{E}_w^\pi|}{2} = -\frac{l^\pi}{2} \quad (5.14)$$

$$|\mathcal{V}_w^\pi| - c_\pi - |\bar{\mathcal{E}}_w^\pi| = 0, \quad (5.15)$$

with l^π the number of simple edges of $T^{\tilde{\rho}(\pi)}$.

- Definition 5.8.**
1. A simple cycle in a graph is a sequence of pairwise-distinct vertices v_1, \dots, v_k , such that v_i and v_{i+1} are adjacent, with indices modulo k (there is no restriction on the directions of the edges).
 2. A cut edge of a graph is an edge whose removal increases the number of connected components. The set of cut edges of a graph G is denoted $\mathcal{C}_1(G)$.
 3. A cactus (respectively a pseudo-cactus) is a graph such that each edge belongs to exactly (respectively at most) one simple cycle.
 4. A strong component of a pseudo-cactus G is whether a simple cycle or a cut edge of G , and we denote by $\mathcal{SC}(G)$ the set of strong components of G , see an example Figure 2. We call cut vertex a vertex of a graph that belong to several strong components.

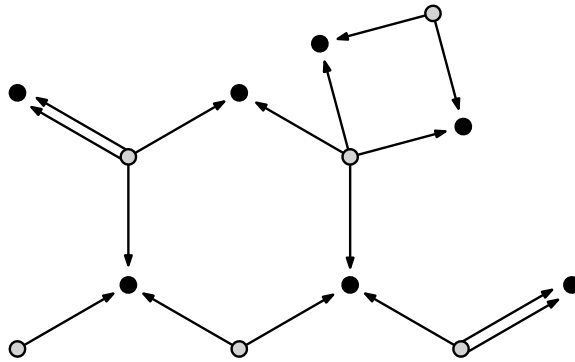


Figure 2: A pseudo-cactus with two cut edges, two length 2 simple cycles, one length 4 and one length 6 cycles.

Proposition 5.9. *In the above setting, let $\pi \in \mathcal{P}(\mathcal{V})$ such that $\eta(\pi) = 0$ and $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$. Then necessarily $T^{\rho(\pi)}$ is a pseudo-cactus, and the partition π do no identify edges or internal vertices of \mathcal{T}^π from different strong components of $T^{\rho(\pi)}$.*

The motivations of this statement are presented next section in Proposition 5.13. The aim of this section is to prove the Proposition 5.9 and provide elements for the computation of the contribution on each strong component. The first lemma indicates the multiplicity of the edges labeled w in the graphs that contribute in the large N limit.

Lemma 5.10. *Assume that T and $\pi \in \mathcal{P}(\mathcal{V})$ are split, that $\delta^0[\mathcal{T}^\pi(W', X')]$ is nonzero and $\eta(\pi) = 0$. Then there are exactly one group of edges of multiplicity 3 labeled w in \mathcal{T}^π within each niche corresponding to a simple edge of $T^{\tilde{\rho}(\pi)}$, and all other groups of edges labeled w have multiplicity 2.*

Proof. Recall that $\mathcal{E}_{w,3}^\pi \subset \mathcal{E}_w^\pi$ stands for the set of edges of multiplicity equal to 3 in \mathcal{T}_w^π and denote by $\mathcal{E}_{w,>3}^\pi$ those of multiplicity greater than 3. Since \mathcal{T}_w^π has no edge of multiplicity 1 (otherwise $\delta^0[\mathcal{T}_w^\pi(W')] = 0$), then (5.14) reformulates as

$$\frac{l^\pi - |\mathcal{E}_{w,3}^\pi|}{2} + (|\bar{\mathcal{E}}_{w,>3}^\pi| - \frac{|\mathcal{E}_{w,>3}^\pi|}{2}) = 0.$$

The second term $(|\bar{\mathcal{E}}_{w,>3}^\pi| - \frac{|\mathcal{E}_{w,>3}^\pi|}{2})$ is non positive and vanishes whenever $\mathcal{E}_{w,>3}^\pi = \emptyset$. Moreover, since the number of edges in each niche is odd, and since by Lemma 5.7 simple edges of $T^{\tilde{\rho}(\pi)}$ have no x -neighbor, each of these niches contain at least a group of edges labeled x of multiplicity at least three. But if some edges forms an intra-niche pairing, so do their compagnons (recall from Definition 5.1 that two edges of \mathcal{T} are compagnon whenever they share the same internal vertex). Hence each niche associated to a simple edge of $T^{\tilde{\rho}(\pi)}$ contains at least one group of edges of multiplicity ≥ 3 labeled w . Hence we get, $l^\pi \leq |\bar{\mathcal{E}}_{w,\geq 3}^\pi|$, so that $(l^\pi - |\mathcal{E}_{w,3}^\pi|)$ is non positive and vanishes whenever $|\mathcal{E}_{w,3}^\pi| = l^\pi$. All together, this proves the lemma. \square

First we use the above lemma to relate the edges of multiplicity 1 in $T^{\tilde{\rho}(\pi)}$ with the cut edges of $T^{\rho(\pi)}$.

Lemma 5.11. *Assume that T and $\pi \in \mathcal{P}(\mathcal{V})$ are split, that $\delta^0[\mathcal{T}^\pi(W', X')]$ is nonzero and $\eta(\pi) = 0$. Let e be an edge of E of multiplicity 1 in $T^{\tilde{\rho}(\pi)}$. Then it is a cut edge of $T^{\rho(\pi)}$ and its niche have no extra-niche identification.*

Proof. Let e be an edge of E of multiplicity 1 in $T^{\tilde{\rho}(\pi)}$ and denote in short by \mathcal{N} its niche. We first prove the second part of the lemma. It is known from Lemma 5.7 for the edges labeled x in \mathcal{N} have no extra-niche identification, let us prove it for the edges labeled w . Lemma 5.7 also tells that all internal

vertices in \mathcal{N} have always intra-identifications. Hence there is one group of internal vertices of multiplicity at least 3 in $\mathcal{T}^{\rho(\pi)}$, and the other groups may have multiplicity greater than or equal to 2. But Lemma 5.10 tells that the edges labeled w in this niche form one group of multiplicity 3 and other groups of multiplicity at least 2. All together, this prove that there are no additional extra-niche identifications.

Let us now prove that there are no extra-niche identification of internal vertices in \mathcal{N} . The following argument, referred as the argument of separation, is use several times in the sequel. Denote by \check{V}_0 the set of internal vertices of \mathcal{N} that are identified with a vertex outside its niche and let us prove that $\check{V}_0 = \emptyset$. Denote by $\tilde{\pi} \in \mathcal{P}(V)$ the modification of π obtained by isolating the vertices of \check{V}_0 in the following sense

$$\tilde{\pi} = \{B \setminus \check{V}_0, B \in \pi\} \sqcup \{B \cap \check{V}_0, B \in \pi\}. \quad (5.16)$$

This modification does not change the δ^0 weight associated to the partition, that is $\delta^0[\mathcal{T}^\pi(W', X')] = \delta^0[\mathcal{T}^{\tilde{\pi}}(W', X')] \neq 0$. Hence by the previous section, we have $\eta(\tilde{\pi}) \leq 0$. But by definition $\eta(\tilde{\pi}) = \eta(\pi) + |\check{V}_0|$ and by hypothesis $\eta(\pi) = 0$, so $|\check{V}_0| = 0$. We have prove that there are no extra-niche identifications in the niche of e .

It remains to prove that e is necessarily a cut edge of $T^{\rho(\pi)}$. For clarity of the presentation, assume first that e is not a cut edge of T (in which case it cannot be a cut edge of $T^{\rho(\pi)}$) and let us find a contradiction. We use a similar argument as in the previous paragraph, considering a modification \check{T} of T obtained by separating a vertex v of e as follow: we add a vertex v' and decide that e is adjacent to v' instead of v . The resulting graph is connected. Moreover, separating $\check{V}_0 := \{v\}$ in π as in (5.16) gives a partition $\tilde{\pi}$ of $\mathcal{T}_{\check{T}}$ such that $\delta^0[\mathcal{T}_{\check{T}}^{\tilde{\pi}}(W', X')] = \delta^0[\mathcal{T}^\pi(W', X')] \neq 0$ and $\eta(\tilde{\pi}) = \eta(\pi) + 1$. Hence a contradiction.

Finally, we consider the case where e is a cut edge of T but not of $T^{\rho(\pi)}$. Let S be the disconnected graph obtained from T by separating a vertex v of e . Denote by S_1 the connected component of e in S and S_2 its complementary. As before, separating v in π provides a partition $\tilde{\pi}$ of the vertices of S with same δ^0 weight. A priori we cannot use the previous reasoning since S is disconnected, but actually we get the same contradiction by considering another graph. Let \check{T} be obtained from S by identifying a vertex of S_1 and a vertex of S_2 that are identified by π (it always exists since e is not a cut edge of $T^{\rho(\pi)}$). The partition $\tilde{\pi}$ induces a partition $\tilde{\pi}'$ of the vertices of \check{T} with same δ^0 weight and such that $\eta(\tilde{\pi}') = \eta(\pi) + 1$ (we have an additional vertex without changing the number of edges), obtaining the same contradiction. \square

We now related the edges of $T^{\rho(\pi)}$ that are of multiplicity 2 in $T^{\tilde{\rho}(\pi)}$ with the simple cycles of $T^{\rho(\pi)}$.

Lemma 5.12. *Assume that T and $\pi \in \mathcal{P}(\mathcal{V})$ are split, that $\delta^0[\mathcal{T}^\pi(W', X')]$ is nonzero and $\eta(\pi) = 0$. Let $e \in E$ be an edge of multiplicity 2 in $T^{\tilde{\rho}(\pi)}$. Then e belongs to a unique simple cycle of $T^{\rho(\pi)}$. Moreover, none of the edges or internal vertices in the niches of the edges of this cycle is identified out the union of these niches.*

Proof. Let e_0 be an edge of E of multiplicity 2 in $T^{\tilde{\rho}(\pi)}$ and denote in short by \mathcal{N}_0 its niche. We want to prove that e_0 has a unique x -neighbor e_1 and has a unique w -neighbor e_{-1} . In other words, there exists a couple of compagnon vertices $e_w^{(0)}, e_x^{(0)}$ in \mathcal{N}_0 with extra-niche identifications. The latter couple is not necessarily unique, but our analysis below shows that when it is not unique then e_0 belong to a cycle of length two.

The existence of a w -neighbor e_{-1} of e_0 follows from Lemma 5.10, which indicates that the multiplicity of the edges labeled w of $\mathcal{T}^{\tilde{\rho}(\pi)}$ is two. Indeed, since the number of edges of each label in a niche is odd, a parity argument ensures that there exists an edge $e_w^{(0)} \in \mathcal{N}_0$ and an edge $e_w^{(-1)}$ in the niche \mathcal{N}_{-1} of another edge e_{-1} of T that are identified.

To prove the existence a x -neighbor e_1 of e_0 , let us prove that the compagnon $e_x^{(0)}$ of $e_w^{(0)}$ has an extra-niche identification. Assume momentarily that $e_x^{(0)}$ has no extra-niche identification and let us find a contradiction. Since $\delta^0[\mathcal{T}^\pi(W', X')]$ is nonzero, the multiplicity of $e_x^{(0)}$ is greater than 1, so it forms an identification which is necessarily an intra-niche identifications if the latter assumption is valid. But when two edges form an intra-niche identification, so do their compagnons. This implies that $e_w^{(0)}$ has both intra and extra-niche identification, so its multiplicity in \mathcal{T}^π is at least 3. This is in contradiction with Lemma 5.10 which says that this multiplicity is 2. Hence $e_x^{(0)}$ has an extra-niche identification with an edge $e_x^{(1)}$. We denote by \mathcal{N}_1 the niche of $e_x^{(1)}$ and by e_1 the associated x -neighbor of e_0 .

We now show the unicity of the x -neighbor e_1 . Let e'_1 be any x -neighbor of e_0 , and denote by $e'_x{}^{(1)}$ and $e'_x{}^{(0)}$ two edges in the respective niches that are identified. This identification implies that the sources of e_0 and e'_1 are equal in $\mathcal{T}^{\tilde{\rho}(\pi)}$. Also, the sources of the compagnons of $e'_x{}^{(1)}$ and $e'_x{}^{(0)}$ are also equal, so their targets belong to the same connected component. Hence e_0 and e'_1 form an edge of multiplicity at least two in $T^{\tilde{\rho}(\pi)}$. On the other hand, $\eta(\pi) = 0$ implies that $T^{\tilde{\rho}(\pi)}$ has edges of multiplicity two. This proves that $e'_1 = e_1$ since otherwise this will exhibit an edge of multiplicity greater than 2.

We now prove the uniqueness of the w -neighbor e_{-1} of e_0 . Recall that $e_w^{(-1)}$ and $e_w^{(0)}$ denote two edges in the respective niches forming an extra-niche identification. Assume that there is another edge $e'_w{}^{(0)} \in \mathcal{N}_0$ forming an extra-niche identification. We then consider the compagnon $e'_x{}^{(0)}$ of this new edge $e'_w{}^{(0)}$. By the above, the x -neighbor is unique, so $e'_x{}^{(0)}$ has also an extra-niche identification with an edge $e'_x{}^{(1)}$ in the niche \mathcal{N}_1 of e_1 . Note that

$e'_x^{(1)} \neq e_x^{(1)}$ otherwise $e_w^{(0)}, e'_w^{(0)}$ and $e_w^{(-1)}$ might form an edge of multiplicity at least 3 in \mathcal{T}_w^π . Now consider the graph consisting in $e_w^{(0)}, e'_w^{(0)}$, and the companions of $e_x^{(1)}$ and $e'_x^{(1)}$. It consists in a cycle of length 4, which is a simple cycle if the vertices are pair-wise distinct. By construction, one sees directly that the edges are pairwise distinct and so are their source vertices. On the other hand, the condition $\eta(\pi) = 0$ implies that the skeleton of \mathcal{T}_w^π is a forest. Hence the graph is not a simple cycle, which means that the targets of the edges forming this graph are identified. Hence $e_w^{(0)}$ and $e_w^{(1)}$ are identified and so $e_{-1} = e_1$. This prove the uniqueness of e_{-1} , and that $e_{-1} = e_1$ when e_0 and e_{-1} have more than one group of edges identified to form extra-niche identification (this fact will be relevant later on).

We can now prove the lemma. Starting with e_0 we construct a sequence $e_i, i = 0, \dots, r$ of edges in T such that, if i is even, e_i and e_{i+1} are x -neighbors, and if i is odd they are w -neighbors, until we come back to an edge we have already visited ($r = \min\{s | e_s \in \{e_0, \dots, e_{s-1}\}\}$). By uniqueness of extra-niche neighbors, we necessarily have $e_r = e_0$ (and so r is even).

Let us denote by $\mathcal{C} = \{e_0, \dots, e_{r-1}\}$ the set of edges forming this cycle and by $\mathcal{N}(\mathcal{C})$ the union of the niches of elements of \mathcal{C} . By construction, the edges of $\mathcal{N}(\mathcal{C})$ are not identified outside of $\mathcal{N}(\mathcal{C})$. We then deduce that tis property holds for the internal vertices of $\mathcal{N}(\mathcal{C})$ with the usual argument of separation. We modify π into $\tilde{\pi}$ to separate the set \check{V}_0 of vertices of $\mathcal{N}(\mathcal{C})$ that are identified outside of $\mathcal{N}(\mathcal{C})$. This does not change the δ^0 weight since it does not modify the multiplicity of the edges. Hence it produces a partition such that $\eta(\tilde{\pi}) \geq 0$ and $\eta(\tilde{\pi}) = \eta(\pi) + |\check{V}_0|$, the assumption $\eta(\pi)$ implying that $\check{V}_0 = \emptyset$. This proves the second part of the lemma.

Finally we can prove that e_0 belongs to a unique cycle of $T^{\rho(\pi)}$. Let S be the graph obtained by removing to $T^{\rho(\pi)}$ the edges of \mathcal{C} , and with a small abuse use \mathcal{C} to refer to the subgraph of $T^{\rho(\pi)}$ formed by these edges. Note that e_0 belongs to a unique cycle if and only if the connected components of S have exactly one vertex in \mathcal{C} . On the other hand, the presence of a connected component C with at least two vertices in \mathcal{C} would yield a contradiction thanks to the separation argument: one modifies $T^{\rho(\pi)}$ by separating a vertex common to \mathcal{C} and C , which does not disconnect the graph, producing a quotient of higher but bounded contribution. Hence e_0 belong to a unique cycle, which concludes the proof of the lemma. \square

Since Condition (5.14) says that the edges of $T^{\tilde{\rho}(\pi)}$ are of cardinality 1 or 2, the two above lemmas show Proposition 5.9.

5.3 Asymptotic expression of the traffic distribution

Given a partition $\rho_0 \in \mathcal{P}(V)$ such that T^{ρ_0} is a pseudo-cactus and a partition of the auxiliary graph $\pi \in \mathcal{P}(\mathcal{V})$, we write $\rho(\pi) = \rho_0$ is a shortcut to say

that π is a split partition such that $\delta^0[\mathcal{T}^\pi(W', X')] \neq 0$, $\eta(\pi) = 0$ and the restriction of π on V is $\rho(\pi) = \rho_0$.

The previous section proves that for any reference graph T we have the asymptotic expression

$$\begin{aligned} \tau_N[T(\mathbf{Y})] &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \text{ s.t.} \\ T^{\rho_0} \text{ pseudo cactus}}} \sum_{\substack{\pi \in \mathcal{P}(V) \\ \rho(\pi) = \rho_0}} \Psi^\pi \delta^0[\mathcal{T}^\pi(\Gamma_w, \Gamma_x)] \delta^0[\mathcal{T}^\pi(W', X')] + o(1), \end{aligned} \quad (5.17)$$

where we recall (5.5) and Definition 5.2:

$$\Psi^\pi = \psi_0^{|V_0^\pi| - \sum_{e \in E} \frac{n(e)-1}{2}} \psi_2^{|V_2^\pi|} \psi_1^{|V_1^\pi|}, \quad \delta^0[\mathfrak{T}(\mathbf{A})] = \mathbb{E} \left[\prod_{e \in \mathfrak{E}} \tilde{A}_{\gamma(e)}(\Phi(e)) \right],$$

$\psi_i = \frac{N_i}{N}$ for $i = 0, 1, 2$ and $\Phi : \mathfrak{W} \rightarrow [N]$ is split, injective and uniformly distributed independently of \mathbf{A} .

It may be useful to compare Formula 5.17 with the following consequence of [Mal20, Chapter 6] and [GC24, Part II]. In the proposition below, we call well-oriented (w.o.) pseudo-cactus a pseudo-cactus for which the edges of each simple cycle follow a same orientation along their cycle.

Proposition 5.13. *Let $\mathbf{A} = (A_j)_{j \in J}$ be a family of N square random matrices that are unitarily invariant in law. Assume that \mathbf{A} converges in non commutative distribution and satisfies the asymptotic factorization property. Denote by \mathbb{J}_N the matrix whose all entries are $\frac{1}{N}$. Then for all $\beta \in \mathbf{C}$ the collection $\mathbf{A} + \beta \mathbb{J}_N := (A_j + \beta \mathbb{J}_N)_{j \in J}$ converges in traffic distribution. Moreover, for any test graph T in the variables $\mathbf{a} = (a_j)_{j \in J}$ and any partition ρ_0 of the vertex set of T , we have*

$$\tau_N[T(\mathbf{A} + \beta \mathbb{J}_N)] \xrightarrow{N \rightarrow \infty} \sum_{\substack{\rho_0 \in \mathcal{P}(V) \text{ s.t.} \\ T^{\rho_0} \text{ w.o. pseudo cactus}}} \prod_{C \in \mathcal{SC}(T_0^\rho)} \tau^0(C),$$

where for any $C \in \mathcal{SC}(T^\rho)$, if C is a cut-edge then $\tau^0(C) = \beta$, and if C is a simple cycle then $\tau^0(C) = \kappa_n(a_1, \dots, a_n)$ is the n -th free cumulant function applied to the limit of the matrices A_1, \dots, A_n along the cycle of C .

The rectangular analogue can be deduced from the computations of [Zit24] (explaining our scaling factor) and the real analogue (with orthogonal invariance) from [Au18]. The method consists hence in re-writing (5.17) by summing over the partitions of the reference graph T rather than $\mathcal{T} = \mathcal{T}_T$ and exhibiting some factorization structure with respect to the strong components.

The latter is the motivation for the second part of the statement in Proposition 5.9: it shows that these partitions $\pi \in \mathcal{P}(\mathcal{V})$ that contribute

can be factorized with respect to the niches of the strong components S of T^{ρ_0} in the following way. For any strong component $S \in SC(T_0^{\rho_0})$, and any π such that $\rho(\pi) = \rho_0$, denote by π_S the restriction of π to the vertices in S and its niche vertex. Then since π does not identify extra-strong components internal vertices, π is the finest of all partitions in $\mathcal{P}(\mathcal{V})$ such that the internal blocks of π_S are contained in blocks of π , for any $S \in SC(T^{\rho_0})$.

In the three following subsections, we consider a strong component $S = S(\rho_0)$ of \mathcal{T}^{ρ_0} of a given type, i.e. a cut-edge, a length 2 cycle or a higher length cycle. In each case, we describe the partition π_S and the subgraph of \mathcal{T}^π induced by the niches of the edges of S , and gives an illustration in Figure 3. We denote from now by $n_1(S), \dots, n_{L_S}(S)$ an enumeration of the edge labels in S , with the shortcuts $n(S) := n_1(S)$ if $L_S \leq 2$ and $n'(S) := n_2(S)$ if $L_S = 2$. We write " $n = 2k + 1$ " with all variations of indices.

5.3.1 Focus on cut edges

We denote by $\binom{p}{q} = \frac{p!}{q!(p-q)!}$ the usual binomial coefficient counting the number of choice of q element among p , and by $\mathcal{P}_2(n)$ the set of pairings of n elements. Recall that if ξ denotes a standard real Gaussian random variable, then $\mathbb{E}[\xi^n] = |\mathcal{P}_2(n)|$. For each $n \geq 1$, we recall that $h_n : x \mapsto x^n$ denotes the n -th power function. Then the Gaussian integration formula reads $n\mathbb{E}[h_{n-1}(\xi)] = \mathbb{E}[h'_n(\xi)]$.

Assume that $S \in SC(T^\pi)$ consists in a cut edge e . Lemma (5.10) says that π_S identifies 3 internal vertices to form a first block and pairs the other internal vertices. Hence $n(S) \geq 3$ otherwise we cannot form a group of 3 internal edges of same label. We have a total number of

$$\binom{n(S)}{3} |\mathcal{P}_2(n(S) - 3)| = \binom{n(S)}{3} \mathbb{E}[\xi^{n(S)-3}] = \mathbb{E}[h'''_{n(S)}(\xi)]/6$$

partitions as above, any of them having $k(S)$ internal vertex blocks.

5.3.2 Focus on length 2 simple cycles

Assume that S consists in a cycle of length 2. Lemma 5.10 implies that the internal vertices are paired. It must have at least one block formed by an internal vertex of the niche of each edge, but since there is an odd number of vertices in each niche, each pairing satisfies this property. We have a total number of

$$|\mathcal{P}_2(n(S) + n'(S))| = \mathbb{E}[\xi^{n(S)+n'(S)}] = \mathbb{E}[h_{n(S)}(\xi)h_{n'(S)}(\xi)]$$

partitions π_S as above, any of them having with a total of $k(S) + k'(S) + 1$ internal vertex blocks.

5.3.3 Focus on higher lengths

Assume that S consists in a simple cycle of length $L \geq 3$ of extra-niche successive neighbors. Constructing the cycles in the proof of Lemma 5.12, we have shown that in the niche of each edge of S there is an edge labeled x (the one realizing the neighboring) such that π identifies the targets of all these edges. This forms a first block of π_S , that we refer as the central block (note that it contains at least one internal vertex from the niche of each edge of S). This central block actually cannot contain more than one vertex from each niche: otherwise one sees easily that this will produce an edge labeled w with multiplicity greater than 2 with the usual compaign argument. Moreover, the proof of Lemma 5.12 shows that each niche has a single edge labeled w forming an extra-niche identification unless the cycle is of length 2. This implies the same property for edges labeled x (since the compaign of edge labeled x forming extra-niche identification also form an extra-niche identification by the multiplicity 2 constraint). The conclusion is that π consists in the central block together with pairings of the remaining vertices in order to pair the edges labeled w .

To chose π_S we can first chose its central block by choosing one internal vertex in each niche, and then we chose intra-niche pairing of the remaining vertices, which gives a total number of

$$\prod_{\ell=1}^{L_S} n_\ell(S) |\mathcal{P}_2(n_\ell(S) - 1)| = \prod_{\ell=1}^{L_S} n_\ell(S) \mathbb{E}[\xi^{n_\ell(S)-1}] = \prod_{\ell=1}^{L_S} \mathbb{E}[h'_{n_\ell(S)}(\xi)]$$

possibilities for partition π_S satisfying the above condition, any of these partitions having with a total of $1 + \sum_{\ell=1}^{L_S} k_\ell(S)$ internal vertex blocks.

5.3.4 Conclusion

Recall that two test graphs are isomorphic whenever there exists a bijection between their vertex set that preserves adjacency and labels.

For a partition ρ_0 of V such that T^{ρ_0} is a pseudo-cactus and the previous section shows that for any π such that $\rho(\pi) = \rho_0$, the isomorphic class of \mathcal{T}^π depends only on ρ_0 , not on π . On the other hand, the summand in the sum over π in (5.17) is a function of the isomorphism class of \mathcal{T}^π , so it is a function of ρ_0 . In this section we write explicitly the dependance in ρ_0 except for the contribution of the profiles $\delta^0[\mathcal{T}^\pi(\Gamma_w, \Gamma_x)]$ which is considered later.

To write $\delta^0[\mathcal{T}^\pi(W', X')]$ in terms of ρ_0 recall that for each cut edge of $T^{\rho(\pi)}$, there is in \mathcal{T}^π a group of edges of multiplicity 3 in each variables and the other groups are all of multiplicity 2. Denote by

- $\mathcal{C}_1(\rho_0)$ the set of cut edges of T^{ρ_0} , $c_1(\rho_0) = |\mathcal{C}_1(\rho_0)|$,
- $m_w^{(3)} = \mathbb{E}[W'(1, 1)^3]$ and $m_x^{(3)} = \mathbb{E}[X'(1, 1)^3]$ the third moments of the matrix entries.

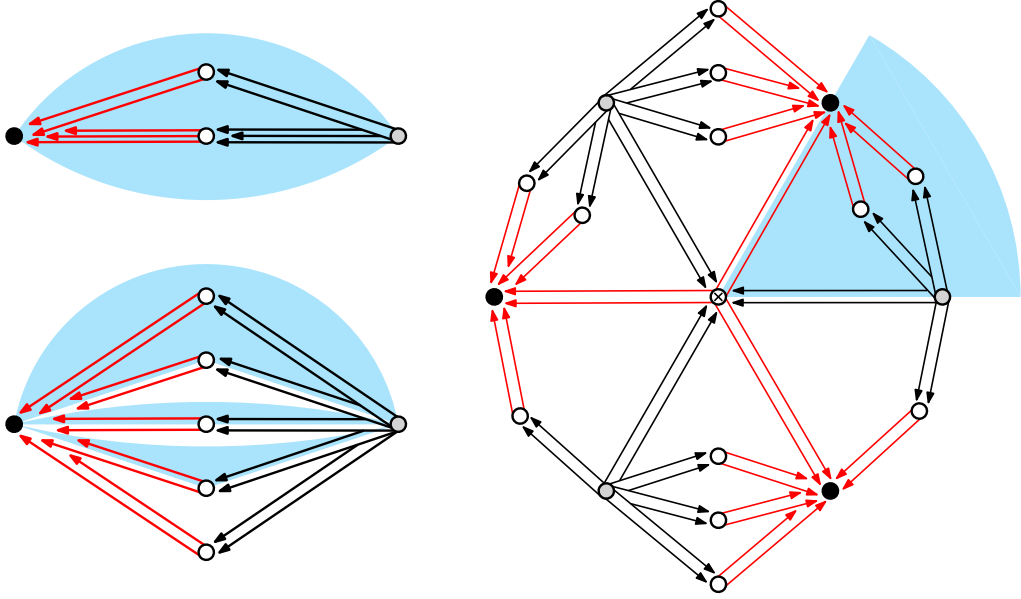


Figure 3: Three types of strong components that contribute in the limiting traffics distribution: the upper leftmost figure represents the strong component associated to a cut edge of \mathcal{T}^{ρ_0} , the bottom leftmost stands for a length two cycle, and the rightmost one for a cycle of length 6. The light blue region represents the niche of an edge labeled by $x \mapsto x^5$.

Recall that the variables are normalized ($\mathbb{E}[W'(1, 1)^2] = \mathbb{E}[X'(1, 1)^2] = 1$). We hence deduce that $\delta^0[\mathcal{T}^\pi(W', X')] = 0$ if T^{ρ_0} has a cut edge S with label $n(S) = 1$, and otherwise

$$\delta^0[\mathcal{T}^\pi(W', X')] := \mathbb{E} \left[\prod_{e \in \mathcal{E}_w^\pi} \tilde{W}'(\varphi(e)) \times \prod_{e \in \mathcal{E}_x^\pi} \tilde{X}'(\varphi(e)) \right] = (m_w^{(3)} m_x^{(3)})^{c_1(\rho_0)}$$

which is independent of the edge labels Y_n of T with index $n \geq 3$.

In order to write Ψ^π as a function of ρ_0 note first that $|V_1^\pi| = |V_1^{\rho_0}|$ and $|V_2^\pi| = |V_2^{\rho_0}|$ by definition. Moreover, the number of internal vertices does not depend on π : indeed, denote

- $C_2(\rho_0)$ the set of simple cycles of length 2 of T^{ρ_0} , $c_2(\rho_0) = |C_2(\rho_0)|$,
- $C_3(\rho_0)$ the set of higher length simple cycles, and $c_3(\rho_0) = |C_3(\rho_0)|$.

Sections 5.3.1, 5.3.2 and 5.3.3 yield

$$\begin{aligned} |V_0^\pi| &= \sum_{S \in \mathcal{C}_1(\rho_0)} k(S) + \sum_{S \in \mathcal{C}_2(\rho_0)} (k(S) + k'(S) + 1) \\ &\quad + \sum_{S \in \mathcal{C}_3(\rho_0)} \left(1 + \sum_{\ell=1}^{L_S} k_\ell(S) \right) = c_2 + c_3 + \sum_{e \in E} \frac{\mathbf{n}(e) - 1}{2}. \end{aligned}$$

So the definition of Ψ^π gives

$$\Psi^\pi := \psi_0^{|V_0^\pi| - \sum_{e \in E} \frac{\mathbf{n}(e) - 1}{2}} \psi_2^{|V_2^\pi|} \psi_1^{|V_1^\pi|} = \psi_0^{c_2 + c_3} \psi_1^{|V_1^{\rho_0}|} \psi_2^{|V_2^{\rho_0}|},$$

where $V_i^{\rho_0}$ is the set of vertices of T^{ρ_0} in V_i . As announced, this expression is a function of ρ_0 . It also does not depend on the labeling.

The number of partitions π such that $\rho(\pi) = \rho_0$ is given by the count of Subsections 5.3.1 to 5.3.3. We recall $h_n : x \mapsto x^n$ and ξ a standard real Gaussian random variable. With the above computations and (5.17), we have finally obtained the following asymptotic formula.

Lemma 5.14. *Let \mathbf{Y} be the family of profiled Pennington-Worah matrices defined in (5.1). Then under our assumption, for any reference test graph $T_{\mathbf{n}} = (V, E, \mathbf{n})$, with $\mathcal{T}_{\mathbf{n}} = \mathcal{T}_{T_{\mathbf{n}}}$ the associated auxiliary test graph and $\pi_0 \in \mathcal{P}(\mathcal{V})$ such that $\rho(\pi_0) = \rho_0$, we have*

$$\begin{aligned} \tau_N[T_{\mathbf{n}}(\mathbf{Y})] &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \\ T_{\mathbf{n}}^{\rho_0} \text{ pseudo cactus}}} \psi_2^{|V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \delta^0[\mathcal{T}_{\mathbf{n}}^{\pi_0}(\Gamma_w, \Gamma_x)] \\ &\quad \times \left(\frac{m_w^{(3)} m_x^{(3)}}{6} \right)^{c_1(\rho_0)} \prod_{S \in \mathcal{C}_1(\rho_0)} \mathbb{E}[h_{n(S)}'''(\xi)] \\ &\quad \times \psi_0^{c_2(\rho_0)} \prod_{S \in \mathcal{C}_2(\rho_0)} \mathbb{E}[h_{n(S)}(\xi) h_{n'(S)}(\xi)] \\ &\quad \times \psi_0^{c_3(\rho_0)} \prod_{S \in \mathcal{C}_3(\rho_0)} \prod_{i=1}^{L_S} \mathbb{E}[h'_{n_i(S)}(\xi)] + o(1). \end{aligned} \quad (5.18)$$

6 Construction of the asymptotic equivalent

In this section, we analysis the expression (5.18), using Proposition 5.13 and technics from traffic probability [Mal20] in order to construct three explicit families of matrices \mathbf{Y}^{lin} , \mathbf{Y}^{per} and \mathbf{B} indexed by $\mathbb{C}[y]$ such that, for the collections restricted to odd polynomials, \mathbf{Y} has the same limiting traffic distribution as

$$\mathbf{Y}^{\text{lin}} + \mathbf{Y}^{\text{per}} + \mathbf{B} := (Y^{\text{lin}}(h) + Y^{\text{per}}(h) + B(h))_{h \in \mathbb{C}[y]}.$$

Moreover, each matrix is a linear function of its argument $h \in \mathbb{C}[y]$. The construction of each collection comes from the analysis the contributions of each type of strong component in (5.18).

The limiting traffic distribution of \mathbf{B} has the same expression as (5.18) if we set to zero the contributions that are not associated to the cut-edge set $\mathcal{C}_1(\rho_0)$, see Lemma 6.2. The matrices of \mathbf{B} are deterministic and their entries are of order N^{-1} (they are called of Boolean type in [Mal20]).

The construction of the collection \mathbf{Y}^{lin} is motivated by the contribution from $\mathcal{C}_3(\rho_0)$ in (5.18). The matrices of \mathbf{Y}^{lin} are obtained by applying profiles to the matrix $Y_1 = \frac{WX}{\sqrt{N}\sqrt{N}}$. Comparing with Proposition 5.13, it may be useful to recall that the product of two independent Ginibre matrices converges toward a non commutative random variables whose free cumulants are constant (they do not depend on the order of the cumulant), called a free Poisson variable. By Shlyakhtenko [Shl96], since we can expect that a good notion of free Poisson variable over the diagonal holds to describe the asymptotic of $\frac{WX}{\sqrt{N}\sqrt{N}}$ in canonical terms.

Therefore the linear matrix \mathbf{Y}^{lin} comes also with contributions for length 2 cycle that we must subtract from the length 2 cycle contribution (5.18) in order to surmise the perturbation family \mathbf{Y}^{per} . Recall that for a collection of non commutative random variables, all free cumulants of order greater than 2 vanish if and only if the collection is circular or semi-circular, which are the limit GOE and Ginibre matrix ensembles. Shlyakhtenko proves in [Shl96] the analogue for variance profiled matrices. The consequence is that to construct \mathbf{Y}^{per} it suffices to understand a covariance structure. This is in particular the moment where we switching from reference test graphs to test graphs labeled by the Hermite polynomials.

Each case relies a same lemma stated in the following subsection.

6.1 A property of the function δ^0

In Section 5.1, after the definition of the auxiliary graph \mathcal{T}_T , we use the substitution property while replacing the edge of a test graph by graph operations. While this property is obvious for the evaluation of the combinatorial traces, it is not longer true for the injective trace. The lemma below shows that the substitution property can be applied for the map δ^0 evaluated in bounded matrices. We restrict our statement to the situation we meet later on.

Lemma 6.1. $\mathbf{A} = (A_\omega)_{\omega \in \Omega}$ a family of \mathbf{s} -rectangular random matrices with bounded entries and let $\mathfrak{T} = (\mathfrak{V}, \mathfrak{E}, \gamma)$ be a test graph. Assume that \mathfrak{T} has an edge $e_0 \in V_2 \times V_1$ associated to a matrix of the form $A_{\gamma(e_0)} = g_\sigma(\mathbf{A})$, where $g = (\mathfrak{S}, \text{in}, \text{out})$. Assume that in is the only vertex of \mathfrak{S} in V_2 , and out is the only vertex of \mathfrak{S} in V_1 . Let us denote \mathcal{T} the test graph obtained from \mathfrak{T} by replacing the edge e_0 by the graph \mathfrak{S} , identifying the input or g with the

source of e_0 , as well as the output of g with the source of e_0 . Then, denoting $R = N^{-v_0} g_{\mathfrak{S}}(\mathbf{A})$ where v_0 is the number of internal vertices of \mathfrak{S} , we have as N goes to infinity

$$\delta^0[\mathfrak{I}(\mathbf{A})] = \delta^0[\mathfrak{I}'(\mathbf{A}, R)] + o(1).$$

Proof. Recall that for Φ a uniform split injection $\mathfrak{Y} \rightarrow [N]$ distributed independently of \mathbf{A} , denoting $\Phi(e) := (\Phi(w), \Phi(v))$ when $e = (v, w)$, we have

$$\delta^0[\mathfrak{I}(\mathbf{A})] := \mathbb{E}\left[\prod_{e \in \mathfrak{E}} A_{\gamma(e)}(\Phi(e))\right].$$

Setting $\mathfrak{Y} = \mathfrak{Y}_0 \sqcup \mathfrak{Y}_1 \sqcup \mathfrak{Y}_2$, the boundedness of the matrix entries implies

$$\begin{aligned} \delta^0[\mathfrak{I}(\mathbf{A})] &= \frac{1}{N_1^{|\mathfrak{Y}_1|} N_2^{|\mathfrak{Y}_2|}} \sum_{\substack{\phi_1: \mathfrak{Y}_1 \rightarrow [N_1] \\ \text{injective}}} \sum_{\substack{\phi_2: \mathfrak{Y}_2 \rightarrow [N_2] \\ \text{injective}}} \\ &\quad \left(\frac{1}{N_0^{|\mathfrak{Y}_0|}} \sum_{\phi_0: \mathfrak{Y}_0 \rightarrow [N_0]} \prod_{e \in \mathfrak{E}} \iota_{(1,2)}(A_{\gamma(e)})(\phi_{0,1,2}(e)) \right) + o(1) \end{aligned} \quad (6.1)$$

where in the last formula $\phi_{0,1,2}$ coincide with ϕ_i on $V_i^{\rho_0}$ for each $i = 0, 1, 2$. Recall that $\iota_{(1,2)}$ denotes the canonical injection of $M_{N_1, N_2}(\mathbb{R}) \rightarrow M_{N, N}(\mathbb{R})$. By assumption, denoting $\mathfrak{S} = (V_{\mathfrak{S}}, E_{\mathfrak{S}}, \gamma_{\mathfrak{S}})$, we have

$$\iota_{(1,2)}(A_{\gamma(e)})(\phi_{0,1,2}(e_0)) = \frac{1}{N^{v_0}} \sum_{\phi'_0: \mathfrak{Y}_{\mathfrak{S}} \rightarrow [N_0]} \prod_{e \in \mathfrak{E}_{\mathfrak{S}}} \iota_{(1,2)}(A_{\gamma_{\mathfrak{S}}(e)})(\phi'_{0,1,2}(e)),$$

where $\phi'_{0,1,2}$ is defined as $\phi_{0,1,2}$ with ϕ' instead of ϕ . We therefore can write our expression in terms of \mathcal{T} . Denoting $\mathcal{V} = \mathfrak{Y}_1 \sqcup \mathfrak{Y}_2 \sqcup \mathcal{V}_0$ its vertex set, \mathcal{E} its edge set, and keeping the notation γ for the label map,

$$\begin{aligned} &\frac{1}{N_0^{|\mathfrak{Y}_0|}} \sum_{\phi_0: \mathfrak{Y}_0 \rightarrow [N_0]} \prod_{e \in \mathfrak{E}} \iota_{(1,2)}(A_{\gamma(e)})(\phi_{0,1,2}(e)) \\ &= \frac{1}{N_0^{|\mathcal{V}_0|}} \sum_{\phi_0: \mathcal{V}_0 \rightarrow [N_0]} \prod_{e \in \mathfrak{E}} \iota_{(1,2)}(A_{\gamma(e)})(\phi_{0,1,2}(e)) \end{aligned}$$

By the same reasoning as for showing (6.1) in the reverse sense, we get that the latter expression equals $\delta^0[\mathcal{T}(\mathbf{A}, R)]$ up to a negligible error resulting from the replacement of the map on \mathcal{V}_0 by an injective map. \square

Our strategy is to apply the above lemma in each edge of T^{ρ_0} in Lemma 5.14 reducing the complicated structure of profiled PW matrices to simpler matrix ensembles.

6.2 The Boolean type deterministic deformation

In the context of Lemma 5.14, let $e \in T^{\rho_0}$ be a cut-edge with label $\mathbf{n}(e) \geq 3$. Section 5.3.1 describes the subgraph $\mathfrak{S} = \mathfrak{S}(e)$ of \mathcal{T}^{π_0} induced by the niche of e . Denote $g_{\mathfrak{S}}$ the graph monomial whose test graph is \mathfrak{S} and whose input and output are the vertices of this graph in V_2 and V_1 respectively. The graph \mathfrak{S} has $v_0 = 1 + \frac{\mathbf{n}(e)-3}{2}$ internal vertices, corresponding to the $\frac{\mathbf{n}(e)-3}{2}$ internal vertex pairing plus one group of 3 vertices. The corresponding identification for the edges implies that the following expression holds

$$\begin{aligned} & N^{-v_0} g_{\mathfrak{S}}(\Gamma_w, \Gamma_x) \\ &= \left[\left(\frac{1}{N} \sum_{\ell=1}^{N_0} \Gamma_w(i, \ell)^3 \Gamma_x(\ell, j)^3 \right) \times \left(\frac{1}{N} \sum_{\ell=1}^{N_0} \Gamma_w(i, \ell)^2 \Gamma_x(\ell, j)^2 \right)^{\frac{\mathbf{n}(e)-3}{2}} \right]_{i,j} \\ &= \Lambda_3 \circ \left(\sqrt[\circ]{\Lambda_2}^{\circ(\mathbf{n}(e)-3)} \right). \end{aligned} \quad (6.2)$$

where we have set

$$\Lambda_\ell := N^{-1} \Gamma_w^{\circ \ell} \times \Gamma_x^{\circ \ell}, \quad \ell = 1, 2, \quad (6.3)$$

and used the concise notations $A^{\circ n} = (A_{p,q}^n)_{p,q}$ and $\sqrt[\circ]{A} = (\sqrt{A_{p,q}})_{p,q}$. Note that with $h_n : x \mapsto x^n$, for any matrix A we have $h_n'''[\{xA\}] = h_n'''(x)A^{\circ(n-3)}$. We hence propose to introduce the following collection of matrices.

Lemma 6.2. *Let $\mathbf{B} = (B(h))_{h \in \mathbb{C}[y]}$ be the collection of $N_1 \times N_2$ deterministic matrices defined as follows: for ξ a standard real Gaussian random variable,*

$$B(h) := \frac{m_w^{(3)} m_x^{(3)}}{6N} \Lambda_3 \circ \mathbb{E} \left[h''' \left[\left\{ \xi \sqrt[\circ]{\Lambda_2} \right\} \right] \right]$$

Then \mathbf{B} converges in traffic distribution and for any reference graph $T_{\mathbf{n}} = (V, E, \mathbf{n})$, with same notations as (5.18), we have

$$\begin{aligned} \tau_N [T_{\mathbf{n}}(\mathbf{B})] &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \\ T_{\mathbf{n}}^{\rho_0} \text{ tree}}} \psi_2^{|V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \delta^0 [T_{\mathbf{n}}^{\rho_0}(\Gamma_w, \Gamma_x)] \\ &\quad \times \left(\frac{m_w^{(3)} m_x^{(3)}}{6} \right)^{|E|} \prod_{e \in E} \mathbb{E} [h_{\mathbf{n}(e)}'''(\xi)]. \end{aligned}$$

Note that since all the edges of a tree are cut edges, then $C_1(\rho_0) = E$ and the above expression coincides with (5.18) when T^{ρ_0} is a tree.

Proof. For any edge $e \in E$ and any split injective map $\phi : V \rightarrow [N]$,

$$\begin{aligned} & \iota_{(1,2)}(B(h_{\mathbf{n}(e)}))(\phi(e)) \\ &= \frac{m_w^{(3)} m_x^{(3)}}{6N} \mathbb{E} [h_{\mathbf{n}(e)}'''(\xi)] \iota_{(1,2)} \left(\Lambda_3 \circ \sqrt[\circ]{\Lambda_2}^{\circ(\mathbf{n}(e)-3)} \right) (\phi(e)). \end{aligned}$$

Therefore, writing τ_N in terms of the injective distribution and using the definitions, we have

$$\begin{aligned}
\tau_N[T_{\mathbf{n}}(\mathbf{B})] &= \frac{1}{N} \sum_{\rho_0 \in \mathcal{P}(V)} \sum_{\substack{\phi: V \rightarrow [N] \\ \text{injective}}} \prod_{e \in E} \iota_{(1,2)}(B(h_{\mathbf{n}(e)}))(\phi(e)) \\
&= \frac{1}{N} \left(\frac{m_w^{(3)} m_x^{(3)}}{6N} \right)^{|E|} \prod_{e \in E} \mathbb{E}[h_{\mathbf{n}(e)}'''(\xi)] \\
&\quad \times \sum_{\rho_0 \in \mathcal{P}(V)} \sum_{\substack{\phi: V \rightarrow [N] \\ \text{injective}}} \prod_{e \in E} \iota_{(1,2)}\left(\Lambda_3 \circ \sqrt[\circ]{\Lambda_2}^{\circ(\mathbf{n}(e)-3)}\right)(\phi(e))
\end{aligned}$$

We can hence substitute the subgraphs $\mathfrak{S}(e)$ defined in the beginning of this section and apply Lemma 6.1 for each vertex e of T^{ρ_0} , getting

$$\begin{aligned}
&\sum_{\substack{\phi: V \rightarrow [N] \\ \text{injective}}} \prod_{e \in E} \iota_{(1,2)}\left(\Lambda_3 \circ \sqrt[\circ]{\Lambda_2}^{\circ(\mathbf{n}(e)-3)}\right)(\phi(e)) \\
&= (N_1)_{|V_1^{\rho_0}|} (N_2)_{|V_2^{\rho_0}|} \mathbb{E} \left[\prod_{e \in E} \iota_{(1,2)}\left(N^{-v_0} g_{\mathfrak{S}(e)}(\Gamma_w, \Gamma_x)\right)(\Phi(e)) \right] \\
&= N^{|V_1^{\rho_0}| + |V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \psi_2^{|V_2^{\rho_0}|} (1 + o(1)) \delta^0[\mathcal{T}_{\mathbf{n}}^{\pi_0}(\Gamma_w, \Gamma_x)],
\end{aligned}$$

We hence have

$$\begin{aligned}
\tau_N[T_{\mathbf{n}}(\mathbf{B})] &= \left(\frac{m_w^{(3)} m_x^{(3)}}{6} \right)^{|E|} \prod_{e \in E} \mathbb{E}[h_{\mathbf{n}(e)}'''(\xi)] \\
&\quad \times \sum_{\rho_0 \in \mathcal{P}(V)} N^{\eta_B(\rho_0)} \psi_1^{|V_1^{\rho_0}|} \psi_2^{|V_2^{\rho_0}|} \delta^0[\mathcal{T}_{\mathbf{n}}^{\pi_0}(\Gamma_w, \Gamma_x)] (1 + o(1)).
\end{aligned}$$

where $\eta_B(\rho_0) = -1 - |E| + |V^{\rho_0}|$. By Lemma 5.4, $\eta_B(\pi) \leq 0$ with equality if T^{ρ_0} is a tree. We hence get the expected asymptotic expression. \square

6.3 The linear collection \mathbf{Y}^{lin}

We follow the same strategy as in the previous section, with the difference that we want to compare our expression with the expression of the linear model. In the context of Lemma 5.14, let e be an edge of T^{ρ_0} contained in a simple cycle of length greater than 2. Section 5.3.1 describes the subgraph of \mathcal{T}^{π_0} induced by the cycle, so in particular the subgraph induced by the niche of e . If $\mathbf{n}(e) = 1$, the niche of e consists in two edges forming the extra-niche identification. Assume that $\mathbf{n}(e)$ is greater than one. We denote

by $\mathfrak{S}(e)$ the subgraph generated by the other edges, that form intra-niche identifications.

The graph $\mathfrak{S}(e)$ has one vertex out in V_1 , on vertex in in V_2 , $v_0 = \frac{\mathbf{n}(e)-1}{2}$ internal vertices, an edge of multiplicity two labeled w between each internal vertex and out, and an edge of multiplicity two labeled x between the input and in. Comparing with the context of Lemma 6.1, note that the graph monomial $g_{\mathbf{n}(e)} = (\mathfrak{S}(e), \text{in}, \text{out})$ satisfies

$$N^{-v_0} g_{\mathbf{n}(e)}(\Gamma_w, \Gamma_x) = \left(\sqrt[\circ]{\Lambda_2} \right)^{\circ(\mathbf{n}(e)-1)}, \quad (6.4)$$

where we recall that $\Lambda_2 := N^{-1}\Gamma_w^{o2} \times \Gamma_x^{o2}$. Moreover removing from \mathcal{T}^π the edges and internal vertices of $\mathfrak{S}(e)$ gives the subgraph we obtain assuming $\mathbf{n}(e) = 1$. In consequence, our operation is equivalent to replace $Y_{\mathbf{n}(e)}$ by the entry-wise product of the above matrix with Y_1 .

Similarly, let e, e' be two edges of T^{ρ_0} that form a simple cycle of length 2. Assume that $\mathbf{n}(e) + \mathbf{n}(e')$ is greater than 2, and denote by $\mathfrak{S}(e, e')$ a subgraph generated by all internal vertices in the niche of e but one pairing, and all edges attached to it. This graph has one vertex out in V_1 , on vertex in in V_2 , $v_0(e, e') := \frac{\mathbf{n}(e)+\mathbf{n}(e')-2}{2}$ internal vertices, and the same configuration of double edges as in $\mathfrak{S}(e)$ of the previous paragraph. The important fact is that the graph operation $g_{\mathbf{n}(e), \mathbf{n}(e')} = (\mathfrak{S}(e, e'), \text{in}, \text{out})$ factorizes

$$\begin{aligned} N^{-v_0(e, e')} g_{\mathbf{n}(e), \mathbf{n}(e')}(\Gamma_w, \Gamma_x) \\ = \left(N^{-\frac{\mathbf{n}(e)-1}{2}} g_{\mathbf{n}(e)} \right)(\Gamma_w, \Gamma_x) \circ \left(N^{-\frac{\mathbf{n}(e')-1}{2}} g_{\mathbf{n}(e')} \right)(\Gamma_w, \Gamma_x), \end{aligned}$$

where \circ is the entry-wise product. While removing from \mathcal{T}^π the edges and internal vertices of $\mathfrak{S}(e, e')$ also gives the subgraph we obtain assuming $\mathbf{n}(e) = 1$, we can distribute the induced contribution as profiles applied to e and to e' separately.

This motivates the introduction of the following collection of matrices.

Lemma 6.3. *Let $W = \Gamma_w \circ W'$ and $X = \Gamma_x \circ X'$ be as in our main theorem. Let $\mathbf{Y}^{\text{lin}} = (Y^{\text{lin}}(h))_{h \in \mathbb{C}[y]}$ be the collection of $N_1 \times N_2$ profiled matrices such that for any polynomial h ,*

$$Y^{\text{lin}}(h) = \mathbb{E}[h'[\{\xi \sqrt[\circ]{\Lambda_2}\}]] \circ \left(\frac{W}{\sqrt{N}} \times \frac{X}{\sqrt{N}} \right),$$

where $\Lambda_2 := N^{-1}\Gamma_w^{o2} \times \Gamma_x^{o2}$ and $\sqrt[\circ]{\Lambda_2} := (\sqrt{\Lambda_2(i, j)})_{i, j}$. Then \mathbf{Y}^{lin} converges in traffic distribution and for any reference graph $T_{\mathbf{n}} = (V, E, \mathbf{n})$, with notations as in (5.18) we have

$$\begin{aligned} \tau_N[T_{\mathbf{n}}(\mathbf{Y}^{\text{lin}})] &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \\ T_{\mathbf{n}}^{\rho_0} \text{ cactus}}} \psi_2^{|V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \delta^0 [T_{\mathbf{n}}^{\pi_0}(\Gamma_w, \Gamma_x)] \\ &\times \psi_0^{c(\rho_0)} \prod_{e \in E} \mathbb{E}[h'_{\mathbf{n}(e)}(\xi)] + o(1). \end{aligned}$$

Note that the expression coincides with (5.18) on cacti whose strong components are simple cycles of length greater than 2. Moreover, the map $(h, W, X) \mapsto \mathbf{Y}^{\text{lin}}$ is 3-linear.

Proof. Let $T_{\mathbf{n}} = (V, E, \mathbf{n})$ be a reference graph and denote by $T_{\mathbf{1}} = (V, E, \mathbf{1})$ the test graph obtained from T by assuming all edges are labeled 1. We denote $\mathcal{T}_{\mathbf{1}} = (\mathcal{V}_{\mathbf{1}}, \mathcal{E}_{\mathbf{1}}, \gamma_{\mathbf{1}}) := \mathcal{T}_{T_{\mathbf{1}}}$. Let $\mathfrak{T} = (\mathfrak{V}, \mathfrak{E}, \gamma)$ be the test graph in the variables w, x , and a collection of variables $\mathbf{r} = (r_n)_{n \geq 1}$ obtained from $\mathcal{T}_{\mathbf{1}}$ as follows: for each edge e of T , we add a so-called corrective edge in $\mathcal{T}_{\mathbf{1}}$ between the endpoints of e labeled $r_{\mathbf{n}(e)}$. Note that \mathfrak{T} and $\mathcal{T}_{\mathbf{1}}$ have same vertex set. Then similarly to (5.6) for computing $\tau_N[T_{\mathbf{1}}(Y_1)]$, with $\Psi^\pi = \psi_0^{|V_0^\pi|} \psi_1^{|V_1^\pi|} \psi_2^{|V_2^\pi|}$ we have

$$\begin{aligned} & \tau_N[T_{\mathbf{n}}(\mathbf{Y}^{\text{lin}})] \\ &= \sum_{\pi \in \mathcal{P}(\mathcal{V}_{\mathbf{1}})} N^{\eta(\pi)} \Psi^\pi \delta^0[\mathfrak{T}(\Gamma_w, \Gamma_x, \mathbf{R})] \delta^0[\mathcal{T}_{\mathbf{1}}^\pi(W', X')](1 + o(1)). \end{aligned}$$

where η can simply be written

$$\eta(\pi) = -1 - \frac{\mathcal{E}}{2} + |\mathcal{V}^\pi| = \left(-1 - |\bar{\mathcal{E}}^\pi| + |\mathcal{V}^\pi|\right) + \left(|\bar{\mathcal{E}}^\pi| - \frac{\mathcal{E}}{2}\right)$$

Most of the arguments in the sequel can be deduced from the previous section, but it may be of interest to have an independent sketch of proof. Lemma 5.4 and $\delta^0[\mathcal{T}_{\mathbf{1}}^\pi(W', X')] \neq 0$ imply that \mathcal{T}^π is a cactus whose cycles are of length two (that we call double tree later on). Moreover a pair of double edge have same label, either w or x . For each edge e of T^{ρ_0} , we call w -neighbor of e the edge e' such that the edges labeled w in their niche are identified, and similarly we define the x neighbors. Two edges are called niche neighbor if they are either w or x -neighbors. The definitions imply that the edges of T in a same equivalent class of equivalence for the niche-neighboring relation form a simple cycle in $T^{\rho(\pi)}$. For simple cycles of length greater than 2, their niches induce the usual star-shape subgraph of Section 5.3.3 (where the edges labels $\mathbf{n}(e)$ are equal to one), see an example Figure 4.

This implies that $T^{\rho(\pi)}$ is a cactus whose number of strong components is $c(\rho) = |V_0|^\pi$, and there is a single π corresponding to ρ_0 . Using $\delta^0[\mathcal{T}_{\mathbf{1}}^\pi(W', X')] = 1$, $\Psi^\pi = \psi_2^{|V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \psi_0^{c(\rho_0)}$ in the following computation

$$\begin{aligned} & \tau_N[T_{\mathbf{n}}(\mathbf{Y}^{\text{lin}})] \\ &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \text{ s.t.} \\ T_{\mathbf{n}}^{\rho_0} \text{ cactus}}} \sum_{\substack{\pi \in \mathcal{P}(\mathcal{V}_{\mathbf{1}}) \text{ s.t.} \\ \rho(\pi) = \rho_0}} \Psi^\pi \delta^0[\mathfrak{T}^\pi(\Gamma_w, \Gamma_x, \mathbf{R})] \delta^0[\mathcal{T}_{\mathbf{1}}^\pi(W', X')] + o(1) \\ &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \text{ s.t.} \\ T_{\mathbf{n}}^{\rho_0} \text{ cactus}}} \psi_2^{|V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \delta^0[\mathfrak{T}^{\pi_0}(\Gamma_w, \Gamma_x, \mathbf{R})] \times \psi_0^{c(\rho_0)} + o(1), \end{aligned}$$

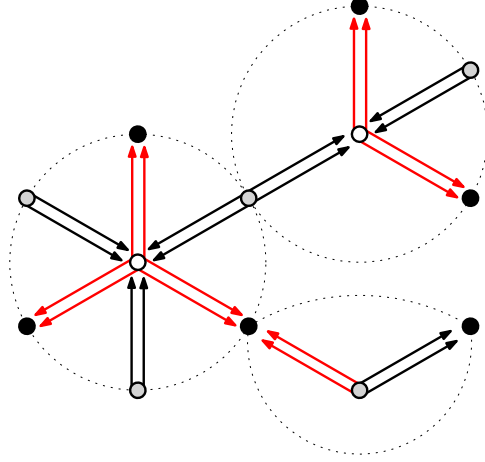


Figure 4: The test graph that contribute to the injective traffic distribution of $\frac{W'X'}{N}$, where we have encircled the equivalent class for the niche neighboring relation.

where π_0 is any partition of \mathcal{V}_1

On the other hand, recalling that $\iota_{(1,2)}$ is the canonical injection of $M_{N_1, N_2}(\mathbb{R}) \rightarrow M_{N, N}(\mathbb{R})$, for any edge $e \in E$ and any split injective map $\phi : V^{\pi_0} \rightarrow [N]$,

$$\begin{aligned} \iota_{(1,2)}(R(h_{\mathbf{n}(e)}))(\phi(e)) &= \mathbb{E}[h'_{\mathbf{n}(e)}(\xi)] (\iota_{(1,2)}(\Lambda_2))^{\circ \frac{\mathbf{n}(e)-1}{2}} (\phi(e)) \\ &= \mathbb{E}[h'_{\mathbf{n}(e)}(\xi)] \times \iota_{(1,2)}(N^{-v_0} g_{\mathbf{n}(e)}(\Gamma_w, \Gamma_x))(\phi(e)), \end{aligned}$$

where $g_{\mathbf{n}(e)}$ is given in (6.4). Therefore using Lemma 6.1 on \mathcal{T}_n for each subgraph $\mathfrak{S}(e)$ and $\mathfrak{S}(e, e')$ defined in the introduction of this section, we have

$$\delta^0[\mathfrak{T}^{\pi_0}(\Gamma_w, \Gamma_x, \mathbf{R})] = \prod_{e \in E} \mathbb{E}[h'_{\mathbf{n}(e)}(\xi)] \times \delta^0[\mathcal{T}_n(\Gamma_w, \Gamma_x)].$$

Altogether, this proves the asymptotic expression stated in the lemma. Hence the convergence holds for all reference test graph, so \mathbf{Y}^{lin} converges in traffic distribution. \square

6.4 Identification of an additive circular noise

6.4.1 Constant profiles

In this section we assume that all entries of Γ_w and Γ_x are equal to 1. Since we look for a collection \mathbf{Y}^{per} such that $\mathbf{Y}^{\text{lin}} + \mathbf{Y}^{\text{per}} + \mathbf{B}$ as the same limiting traffic distribution as \mathbf{Y} , we shall consider the bilinear map

$$f(h_1, h_2) = \mathbb{E}[h_1(\xi)h_2(\xi)] - \mathbb{E}[h'_1(\xi)] \times \mathbb{E}[h'_2(\xi)],$$

obtained from the length 2 cycle contribution of (5.18) minus the expression valid only for higher length cycles. The family $(g_n)_{n \geq 0}$ of Hermite polynomials is a basis of $\mathbb{C}[y]^{\mathbb{N}}$ by

$$g_n : y \mapsto (-1)^n e^{\frac{y^2}{2}} \frac{d^n}{dy^n} e^{-\frac{y^2}{2}}.$$

It is an orthogonal sequence for the standard Gaussian law

$$\mathbb{E}[g_n(\xi)g_m(\xi)] = \delta_{n,m}n!, \quad \forall n, m \geq 0.$$

The Leibniz formula implies the formula

$$\frac{d^{n+1}}{dy^{n+1}} e^{-\frac{y^2}{2}} = -y \frac{d^n}{dy^n} e^{-\frac{y^2}{2}} - 2n \frac{d^{n-1}}{dy^{n-1}} e^{-\frac{y^2}{2}}$$

which yields the identity $g'_n(y) = ng_{n-1}(y)$, $\forall n \geq 1$. Moreover the orthogonality relation implies $\mathbb{E}[g_n(\xi)] = \mathbb{E}[g_n(\xi)g_0(\xi)] = \delta_{n,0}$ for any $n \geq 1$ and so $\mathbb{E}[g'_n(\xi)] = n\mathbb{E}[g_{n-1}(\xi)] = \delta_{n,1}$. We therefore have for all $n, m \geq 1$

$$f(g_n, g_m) = n!\delta_{n,m} - \delta_{n,1}\delta_{m,1}.$$

We can hence propose the following collection of matrices. We call double-tree a cactus whose simple cycles are all of length two.

Lemma 6.4. *Let $\mathbf{Y}^{per} = (Y^{per}(h))_{h \in \mathbb{C}[y]}$ be a collection of $N_1 \times N_2$ random matrices such that the map $h \mapsto Y^{per}(h)$ is linear and*

- $Y^{per}(g_0) = Y^{per}(g_1) = 0$,
- the matrices $\sqrt{\psi_0 n! N^{-1}} Y^{per}(g_n)$, $n \geq 2$ labeled by Hermite polynomials are i.i.d. with i.i.d. real standard Gaussian variables.

Then \mathbf{Y}^{per} converges in traffic distribution and for any reference test graph $T_{\mathbf{n}}$ whose edges are labeled by integer greater than 1,

$$\tau_N [T_{\mathbf{n}}(\mathbf{Y}^{per})] \tag{6.5}$$

$$= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \\ T_{\mathbf{n}}^{\rho_0} \text{ double tree}}} \psi_2^{|V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \psi_0^{c(\rho_0)} \\ \times \prod_{C \in \mathcal{C}(\rho_0)} \left(\mathbb{E}[h_{n(C)}(\xi)h_{n'(C)}(\xi)] - \mathbb{E}[h'_{n(C)}(\xi)]\mathbb{E}[h'_{n'(C)}(\xi)] \right) \tag{6.6}$$

where notations are as in Lemma 5.14.

Example 6.5. For instance, one can compute the first odd Hermite polynomials $g_1(y) = y$, $g_3(y) = y^3 - 3y$ and $g_5(y) = y^5 - 10y^3 + 14y$. Hence we have for the power functions $h_3 = g_3 + 3g_1$ and $h_5 = g_5 + 10g_3 + 16g_1$. So we can write $Y^{per}(h_3) = \sqrt{6}Z_1$, $Y^{per}(h_5) = 10\sqrt{6}Z_1 + 2\sqrt{30}Z_2$, where Z_1 and Z_2 are independent matrices with i.i.d. real centered Gaussian variable of variance ψ_0 .

Remark 6.6. Let consider \mathbf{Y}^{lin} defined in Lemma 6.3. If the profile matrices Γ_w, Γ_x are constant equal to one, then so is the matrix $\sqrt[\circ]{\Lambda_2}$. Hence the property $\mathbb{E}[g'_n(\xi)] = \delta_{n,1}$ implies that $Y^{\text{lin}}(g_n) = 0$ if g_n is the Hermite polynomial of order $n \geq 2$. The same reasoning shows that, for \mathbf{B} defined in Lemma 6.2, then $B(g_n) = 0$ unless $n = 3$. This is not longer true when $\sqrt[\circ]{\Lambda_2}$ is not constant equal to one.

Proof. Let T be a test graph whose edges are labeled by Hermite polynomials of order greater than 1. The associated rectangular matrices are independent with i.i.d. entries, for which the computation of the traffic distribution is similar to [Mal20, Chapter 3]

$$\begin{aligned} \tau_N[T_{\mathbf{g}}(\mathbf{Y}^{per})] &= \sum_{\rho_0 \in \mathcal{P}(V)} \frac{1}{N} \sum_{\substack{\phi: V^{\rho_0} \rightarrow [N] \\ \text{injective}}} \mathbb{E} \left[\prod_{e \in E} \iota_{1,2} \left(Y^{per}(g_e) \right) (\phi(e)) \right] \\ &= \sum_{\rho_0 \in \mathcal{P}(V)} \sum_{\substack{\phi: V^{\rho_0} \rightarrow [N] \\ \text{injective}}} N^{\eta(\rho_0)} (1 + o(1)) \psi_1^{|V_1^{\rho_0}|} \psi_2^{|V_2^{\rho_0}|} \delta^0 [T_{\mathbf{g}}(\sqrt{N} \mathbf{Y}^{per})] \end{aligned}$$

where $\eta(\rho_0) = -1 - \frac{|E|}{2} + |V^{\rho_0}|$. By Lemma 5.4 we get

$$\begin{aligned} \tau_N[T_{\mathbf{g}}(\mathbf{Y}^{per})] &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \text{ s.t.} \\ T^{\rho_0} \text{ double tree}}} \sum_{\phi: V^{\rho_0} \rightarrow [N] \text{ injective}} \psi_1^{|V_1^{\rho_0}|} \psi_2^{|V_2^{\rho_0}|} \\ &\quad \times \prod_{\{e, e'\} \in \bar{E}} \mathbb{E} [y_{p,q}^{(g_e)} y_{p,q}^{(g_{e'})}] + o(1) \end{aligned}$$

In the above sum, \bar{E} denotes the skeleton of E and the product is over all elements of \bar{E} that are denoted $\{e, e'\}$ where e and e' are two edges of $T_{\mathbf{g}}$ that are identified by ρ . By definition of the matrices Y^{per} and by orthogonality of the Hermite polynomials, we have $\mathbb{E} [y_{p,q}^{(g_e)} y_{p,q}^{(g_{e'})}] = \delta_{g_e, g_{e'}} \times \psi_0 n(\mathbf{e})! = \psi_0 \mathbb{E} [g_e(\xi) g_{e'}(\xi)]$. We hence obtain the asymptotic formula

$$\begin{aligned} \tau_N[T_{\mathbf{g}}(\mathbf{Y}^{per})] &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \\ T_n^{\rho_0} \text{ double tree}}} \psi_2^{|V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \psi_0^{c(\rho_0)} \prod_{C \in \mathcal{C}(\rho_0)} \mathbb{E} [g_{n(C)}(\xi) g_{n'(C)}(\xi)]. \end{aligned}$$

Since $\mathbb{E}[g'_n(\xi)] = 0$ for all $n > 1$, the asymptotic formula (6.5) is valid for all test graphs labeled by Hermite polynomials of positive order. Indeed, if there is an edge labeled g_1 then, since $\mathbb{E}[g_1(\xi)g_n(\xi)] - \mathbb{E}[g_1(\xi)]\mathbb{E}[g_n(\xi)] = \mathbb{E}[\xi g_n(\xi)] - \mathbb{E}[g'_n(\xi)] = 0$ then (6.5) vanishes, and otherwise the expressions coincides. Since the Hermite polynomials form a basis of $\mathbb{C}[y]$ and the map $h \mapsto Y^{per}(h)$ is linear, this proves the convergence in traffic distribution of \mathbf{Y}^{per} . Finally, since (6.5) is a multilinear function of the labels, the formula is also valid for all test graphs labeled in $\mathbb{C}[y]$. \square

In the rest of this section, we emphasis a property that we use later. We say that a family \mathbf{A} that converges in traffic distribution satisfies the asymptotic factorization property whenever for any test graph T_1, \dots, T_L , we have

$$\mathbb{E}\left[\prod_{\ell=1}^L \frac{1}{N} \text{Tr}[T(\mathbf{A})]\right] = \prod_{\ell=1}^L \tau_N[T(\mathbf{A})] + o(1).$$

Lemma 6.7. *The couple $(W'/\sqrt{N}, X'/\sqrt{N})$ and the collection $\mathbf{Y}^{per'}$ defined in Lemma 6.3 and Lemma 6.8 satisfy the asymptotic factorization property.*

Proof. The proof follows from minor modification of the convergence $\tau_N[T(\mathbf{Y}^{\text{lin}})]$ by considering a unconnected graph T and normalizing the trace by N^{-c} where c is the number of connected components. The independence of the matrix entries shows that we can factorize the contributions of each connected component. \square

6.4.2 General profiles

We consider a collection of matrices as in Lemma 6.4 that we denote $\mathbf{Y}^{per'} = (Y^{per'}(h))_{h \in \mathbb{C}[y]}$. We look for a collection of matrices $\mathbf{Y}^{per} = (Y^{per}(h))_{h \in \mathbb{C}[y]}$ of the form

$$Y^{per}(h) = \tilde{R}(h) \circ \mathbf{Y}^{per'}(h) \tag{6.7}$$

for $\tilde{R}(h)$ chosen in order to match the remaining terms.

Section 6.3 gives an expression of the traffic distribution of a couple of $N_1 \times N_2$ profiled Gaussian matrix $Z_1 = R_1 \circ Z'$, $Z_2 = R_2 \circ Z'$, where $\sqrt{N}Z'$ has i.i.d. standard real Gaussian entries. Their injective traffic distribution is supported on double trees, and the contribution of a double edge with labels z_1 and z_2 is as follows: for the graph monomial g with two vertices out $\in V_1$, in $\in V_2$, and one double edge labeled z_1 and z_2 from in to out, we have

$$g(R_1, R_2) = R_1 \circ R_2.$$

Of independence interest, note that $Ng(R_1, R_2)$ is the matrix of the co-called \mathcal{R} -transform over the diagonal computed by Shlyakhtenko [Shl96].

On the other hand, let us S be a double edge formed by two edges e, e' of T^{ρ_0} as in Section 5.3.2. Its vertices are in $\in V_2, \text{out} \in V_1$ and a set $\mathcal{V}_0(S)$ of $\frac{\mathbf{n}(e)+\mathbf{n}(e')}{2}$ internal vertices, with double edges between internal and non internal vertices as usual. The associated graph monomial $g_{\mathbf{n}(e)+\mathbf{n}(e')} = (S, \text{in}, \text{out})$ satisfies

$$\begin{aligned} & \frac{1}{N} g_{\mathbf{n}(e)+\mathbf{n}(e')}(\Gamma_w, \Gamma_x) \\ &= \left[\left(\frac{1}{N} \sum_{k=1}^{N_0} \Gamma_w^{\circ 2}(i, k) \Gamma_w^{\circ 2}(k, j) \right)^{\frac{\mathbf{n}(e)+\mathbf{n}(e')}{2}} \right]_{i,j} = \sqrt[\circ]{\Lambda_2}^{\text{on}(e)} \circ \sqrt[\circ]{\Lambda_2}^{\text{on}(e')}. \end{aligned}$$

We can hence propose the following collection of matrices, using the simple relation $y^n h'_n(x) = y h'_n(xy)$ for $h_n : x \mapsto x^n$.

Lemma 6.8. *Let $\mathbf{Y}^{per} = (Y^{per}(h))_{h \in \mathbb{C}[y]}$ be the collection of matrices*

$$Y^{per}(h) = h[\{\sqrt[\circ]{\Lambda_2}\}] \circ Y^{per'}(h),$$

where the collection $\mathbf{Y}^{per'}$ is as in Lemma 6.4. Then \mathbf{Y}^{per} converges in traffic distribution, and for any reference test graph $T_{\mathbf{n}}$, we have

$$\begin{aligned} & \tau_N [T_{\mathbf{n}}(\mathbf{Y}^{per})] \\ &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \\ T_{\mathbf{n}}^{\rho_0} \text{ double tree}}} \psi_2^{|V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \delta^0 [T_{\mathbf{n}}^{\pi_0}(\Gamma_w, \Gamma_x)] \psi_0^{c(\rho_0)} \prod_{C \in \mathcal{C}(\rho_0)} f(h_{\mathbf{n}(e)}, h_{\mathbf{n}(e')}), \end{aligned}$$

where we recall that $f(h_1, h_2) = \mathbb{E}[h_1(\xi)h_2(\xi)] - \mathbb{E}[h_1'(\xi)] \times \mathbb{E}[h_2'(\xi)]$.

Proof. Let $T = (V, E, \mathbf{n})$ be a reference test graph and denote $\tilde{\mathbf{R}} = (\Lambda_2^{\circ \frac{n}{2}})_{n \geq 1}$. The collection of matrices $\mathbf{Y}^{per'}$ is invariant in law by left and right multiplication by permutation matrices, so we can factorize the profiles under the injective trace

$$\begin{aligned} & \tau_N [T_{\mathbf{n}}(\mathbf{Y}^{per})] \\ &= \sum_{\rho_0 \in \mathcal{P}(V)} \delta^0 [T_{\mathbf{n}}^{\rho_0}(\tilde{\mathbf{R}})] \times \tau_N^0 [T_{\mathbf{n}}^{\rho_0}(\mathbf{Y}^{per'})]. \end{aligned}$$

Therefore we can use the expression $\tau_N^0 [T_{\mathbf{n}}(\mathbf{Y}^{per'})]$ from the previous case.

$$\begin{aligned} & \tau_N [T_{\mathbf{n}}(\mathbf{Y}^{per})] \\ &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \\ T_{\mathbf{n}}^{\rho_0} \text{ double tree}}} \psi_2^{|V_2^{\rho_0}|} \psi_1^{|V_1^{\rho_0}|} \psi_0^{c(\rho_0)} \delta^0 [T_{\mathbf{n}}^{\rho_0}(\tilde{\mathbf{R}})] \\ & \quad \times \prod_{C \in \mathcal{C}(\rho_0)} \left(\mathbb{E}[h_{n(C)}(\xi)h_{n'(C)}(\xi)] - \mathbb{E}[h'_{n(C)}(\xi)] \mathbb{E}[h'_{n'(C)}(\xi)] \right). \end{aligned}$$

On the other hand, recalling that $\iota_{(1,2)}$ is the canonical injection of $M_{N_1, N_2}(\mathbb{R}) \rightarrow M_{N, N}(\mathbb{R})$, for any double edge $\bar{e} = \{e, e'\}$ of $T_{\mathbf{n}}^{\rho_0}$ and any split injective map $\phi : V^{\rho_0} \rightarrow [N]$, the definition of $\tilde{\mathbf{R}}$ implies

$$\begin{aligned} & \iota_{(1,2)}(\tilde{R}_{n(e)})(\phi(e)) \times \iota_{(1,2)}(\tilde{R}_{n(e')})(\phi(e')) \\ &= \iota_{(1,2)}(\Lambda_2^{\text{on}(e)})(\phi(e)) \times \iota_{(1,2)}(\Lambda_2^{\text{on}(e')})(\phi(e')) \\ &= \iota_{(1,2)}\left(N^{-\frac{\mathbf{n}(e)+\mathbf{n}(e')}{2}} g_{\mathbf{n}(e)+\mathbf{n}(e')}(\Gamma_w, \Gamma_x)\right)(\phi(e)), \end{aligned}$$

with $g_{\mathbf{n}(e)+\mathbf{n}(e')}$ as in the beginning of the section. Therefore using Lemma 6.1 for each double edge, we have

$$\delta^0\left[T_{\mathbf{n}}^{\rho_0}(\sqrt{\Lambda_2} \circ \mathbf{R}(h))\right] = \delta^0[\mathcal{T}_{\mathbf{n}}^{\pi_0}(\Gamma_w, \Gamma_x)].$$

Altogether this proves the expected asymptotic formula. \square

6.5 Conclusion

We shall now use the asymptotic traffic independence principle for the collections $\mathbf{Y}^{\text{lin}}, \mathbf{Y}^{\text{per}}$ and \mathbf{B} . The drawback of our presentation is that since we have variance profiles we cannot use existing theorems to conclude. Although it is easy to use this theorem we are going to repeat the arguments of the three last section.

Let $\mathbf{Z}_1, \mathbf{Z}_2$ and \mathbf{Z}_3 be three independent families of rectangular matrices indexed by some set J , that converges in traffic distribution and satisfy the asymptotic factorization property. Assume furthermore the families are bi-permutation invariant, that is \mathbf{Z}_j has the same law as the collection $(U\mathbf{Z}_jV)_{j \in J}$ for any permutation matrices U and V of appropriate size, for $j = 1, 2, 3$.

Then the asymptotic traffic independence theorem for rectangular matrices [Zit24] proves that $(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)$ converges in traffic distribution. Let us recall the sketch of the proof. Let \mathcal{T} be a testing graph in three families of variables $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$ with vertex set $\mathcal{V} = V_0 \sqcup V_1 \sqcup V_2$. Let π be a split partition of its vertex set and denote for any $i = 1, 2, 3$ by \mathcal{T}_i is the graph obtained from \mathcal{T} by removing edge that are not in \mathbf{z}_i . Its vertex set is denoted $\mathcal{V}_i = V_{0,i} \sqcup V_{1,i} \sqcup V_{2,i}$. Setting $(N)_{\mathcal{T}^\pi} = \prod_{j=0,1,2} (N_j)_{|\mathcal{V}_j^\pi|}$, we have as before

$$\tau_N^0[\mathcal{T}^\pi(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)] = N^{-1}(N)_{\mathcal{T}^\pi} \times \delta^0[\mathcal{T}^\pi(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)]$$

and reciprocally

$$\begin{aligned} \delta^0[\mathcal{T}^\pi(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)] &= \delta^0[\mathcal{T}^\pi(\mathbf{Z}_1)] \delta^0[\mathcal{T}^\pi(\mathbf{Z}_2)] \delta^0[\mathcal{T}^\pi(\mathbf{Z}_3)] \\ &= (N)_{\mathcal{T}_1^\pi}^{-1} (N)_{\mathcal{T}_2^\pi}^{-1} (N)_{\mathcal{T}_3^\pi}^{-1} \times \mathbb{E}\left[\prod_{S \in \text{CC}(\mathcal{T}^\pi)} \text{Tr}^0 S(\mathbf{Z}_{i_S})\right], \end{aligned}$$

where the product is over the union all connected components S of \mathcal{T}_1^π , \mathcal{T}_2^π and \mathcal{T}_3^π , and i_S denotes the edge labels in 1, 2, 3 of matrices associated to S .

We therefore get, setting $\Psi^\pi = \prod_{j=0,1,2} \psi_j^{|\mathcal{V}_j^\pi| - \sum_{i=1,2,3} |\mathcal{V}_{j,i}^\pi|}$ we get

$$\tau_N^0[\mathcal{T}^\rho(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)] = N^{\eta(\pi)}(1 + o(1))\Psi^\pi \prod_{S \in \mathcal{CC}(\mathcal{T}^\pi)} \tau_N^0[S(\mathbf{Z}_{i_S})]$$

for some $\eta(\pi)$ whose expression can be made explicite from the above computation. The important properties are that

1. $\eta(\pi) \leq 0$ with equality if and only if a certain graph, called the graph of colored component of \mathcal{T}^π , is a tree [Mal20, Zit24]
2. if the elements of $\mathcal{CC}(\mathcal{T}^\pi)$ are pseudo-cacti, then the graph of colored component of \mathcal{T}^π is a tree if and only if \mathcal{T}^π is a pseudo-cactus whose strong components have edges labels associate to a single family among $\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3$ [GC24].

Assuming that for any test graph S that is not a pseudo-cactus, $\lim_{N \rightarrow \infty} \tau_N^0[S(\mathbf{Z}_i)] = 0$ for $i = 1, 2, 3$, we get

$$\tau_N[\mathcal{T}(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)] = \sum_{\substack{\pi \in \mathcal{P}(\mathcal{V}) \text{ s.t.} \\ \mathcal{T}^\pi \text{ w.c. p.-cactus}}} \Psi^\pi \prod_{S \in \mathcal{CC}(\mathcal{T}^\pi)} \tau_N^0[S(\mathbf{Z}_{i_S})] + o(1)$$

where "w.c. p – cactus" is a shortcut for well-colored pseudo-cactus, meaning that all edges of each simple cycle of the cactus \mathcal{T}^π are labeled either by $\mathbf{Z}_1, \mathbf{Z}_2$ or \mathbf{Z}_3 .

The collections of matrices \mathbf{Y}^{lin} , \mathbf{Y}^{per} and \mathbf{B} are not bi-permutation invariant when the profiles are not constant. But they are defined by applying profiles to bi-permutation collections of independent matrices and we can use this property.

More precisely, with $\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3$ and the test graph \mathcal{T} given as above, assume we are also given a collection of matrices with bounded entries $\mathbf{\Gamma}$ and a test graph \mathcal{T}_+ obtained from \mathcal{T} by adding a set \mathcal{E}_+ of edges labeled for matrices in $\mathbf{\Gamma}$, but without adding vertices. Let \mathcal{T}_- be the union of test graphs obtained from \mathcal{T}_+ by removing the edges that are not standing for a matrix in $\mathbf{\Gamma}$. Since \mathcal{T} and \mathcal{T}_+ have same vertex set we have

$$\begin{aligned} & \tau_N[\mathcal{T}_+^\rho(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3, \mathbf{\Gamma})] \\ &= \sum_{\rho \in \mathcal{P}(\mathcal{V})} N^{-1} (N_1)_{|\mathcal{V}_1^\rho|} (N_2)_{|\mathcal{V}_2^\rho|} \delta^0[\mathcal{T}^\pi(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)] \delta^0[\mathcal{T}_-^\pi(\mathbf{\Gamma})] \\ &= \sum_{\substack{\pi \in \mathcal{P}(\mathcal{V}) \text{ s.t.} \\ \mathcal{T}^\pi \text{ w.c. p.-cactus}}} \Psi^\pi \delta^0[\mathcal{T}_-^\pi(\mathbf{\Gamma})] \prod_{S \in \mathcal{CC}(\mathcal{T}^\pi)} \tau_N^0[S(\mathbf{Z}_{i_S})] + o(1) \end{aligned}$$

Informally, we can factorize the profile contribution under the injective trace. We recall the definitions and set the following notations

$$\begin{aligned}
Y^{\text{lin}}(h) &:= \mathbb{E}[h'[\{\xi \overset{\circ}{\sqrt{\Lambda_2}}\}]] \circ \left(\frac{W}{\sqrt{N}} \times \frac{X}{\sqrt{N}} \right) \\
&=: \Gamma_1(h) \circ \left((\Gamma_w \circ \frac{W'}{\sqrt{N}}) \times (\Gamma_x \circ \frac{X'}{\sqrt{N}}) \right) \\
Y^{\text{per}}(h) &:= h[\{\overset{\circ}{\sqrt{\Lambda_2}}\}] \circ Y^{\text{per}'}(h) =: \Gamma_2(h) \circ Y^{\text{per}'}(h), \\
B(h) &= \frac{m_w^{(3)} m_x^{(3)}}{6N} \Lambda_3 \circ \mathbb{E} \left[h''' \left[\{\xi \overset{\circ}{\sqrt{\Lambda_2}}\} \right] \right] =: \Gamma_3(j) \circ \mathbb{J}_N
\end{aligned}$$

where \mathbb{J}_N is the matrix whose entries are $\frac{1}{N}$. We set $\mathbf{Z}_1 = (W'/\sqrt{N}, X'/\sqrt{N})$, $\mathbf{Z}_2 = \mathbf{Y}^{\text{per}'}$, $\mathbf{Z}_3 = (\mathbb{J}_N)$ and the collection $\mathbf{\Gamma}$ consisting in Γ_w, Γ_x and the matrices $\Gamma_1(h), \Gamma_2(h), \Gamma_3(h)$ for all $h \in \mathbb{C}[y]$. For any test graph T in three collections of variables

$$\tau_N [T(\mathbf{Y}^{\text{lin}}, \mathbf{Y}^{\text{per}}, \mathbf{B})] = \tau_N [\mathcal{T}_+(\mathbf{\Gamma}, \mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)]$$

where \mathcal{T}_+ is obtained from T

- adding for each edge labeled of T associated to $Y^{\text{per}}(h)$ and edge with label $\Gamma_2(h)$ and same endpoints,
- adding for each edge labeled of T associated to $B(h)$ and edge with label $\Gamma_3(h)$ and same endpoints,
- adding for each edge labeled of T associated to $Y^{\text{lin}}(h)$ and edge with label $\Gamma_2(h)$ and same endpoints,
- replacing each edge of T associated to $Y^{\text{lin}}(h)$ by a niche with one internal vertex, one edge labeled w' and one edge labeled x'
- adding for each edge in labeled w' and edge with label Γ_w and same endpoints, and for each edge in labeled x' and edge with label Γ_x similarly.

So we can apply the above observation: with same notations as above

$$\begin{aligned}
&\tau_N [\mathcal{T}_+(\mathbf{\Gamma}, \mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)] \\
&= \sum_{\substack{\pi \in \mathcal{P}(\mathcal{V}) \text{ s.t.} \\ \mathcal{T}^\pi \text{ w.c. p-cactus}}} \Psi^\pi \delta^0 [\mathcal{T}^\pi(\mathbf{\Gamma})] \prod_{S \in \mathcal{CC}(\mathcal{T}^\rho)} \tau_N^0 [S(\mathbf{Z}_{i_S})] + o(1).
\end{aligned}$$

We have for any test graph S

$$\begin{aligned}\tau_N^0[S(\mathbf{Z}_3)] &\xrightarrow{N \rightarrow \infty} \mathbb{1}(S \text{ is a tree}), \\ \tau_N^0[S(\mathbf{Z}_2)] &\xrightarrow{N \rightarrow \infty} \mathbb{1}(S \text{ is a double tree}) \prod_{C \in \mathcal{SC}(S)} f(h_n(S), h_{n'}(S)) \\ \tau_N^0[S(\mathbf{Z}_1)] &\xrightarrow{N \rightarrow \infty} \mathbb{1}(S \text{ is a w.c. - double tree}),\end{aligned}$$

where in the second formula $h_n(S), h_{n'}(S)$ are the edges labels in the double edge S and in the third one "w.c. - double tree" means that for each doubles edge, both edges labels are w or are x .

Let as before $\rho(\pi)$ be the restriction of π to the vertices of \mathcal{V} in $V_1 \sqcup V_2$. The previous niche-neighbor argument shows that T^ρ is a cactus such that for the simple cycles of T^ρ labeled by the edges associated to \mathbf{Y}^{lin} , the edges of their niche forms the star-shape test graph of section 5.3.3 when all edges labels are 1. We therefore have, when T is a reference test graph

$$\begin{aligned}\tau_N[T(\mathbf{Y}^{\text{lin}}, \mathbf{Y}^{\text{per}}, \mathbf{B})] &= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \text{ s.t.} \\ T^{\rho_0} \text{ pseudo-cactus}}} \Psi^\rho \prod_{S \in \mathcal{CC}_1(T^{\rho_0})} \mathbb{1}(\text{the label is in } \mathbf{z}_3) \\ &\times \prod_{S \in \mathcal{CC}_2(T^{\rho_0})} \left(\mathbb{1}(\text{the labels are in } \mathbf{z}_2) \psi_0 \prod_{C \in \mathcal{SC}(S)} f(h_n(S), h_{n'}(S)) \right) \\ &\times \prod_{S \in \mathcal{CC}_3(T^{\rho_0})} \mathbb{1}(\text{the labels are in } \mathbf{z}_1) \psi_0 \times \delta^0[\mathcal{T}^{\pi_0}(\Gamma)] + o(1).\end{aligned}$$

where π_0 is the only partition of \mathcal{V} such that $\rho(\pi) = \rho_0$ and \mathcal{T}^π is a pseudo-cactus with simple cycles of length two. Going back to the definition as the

profile matrices as in the previous sections shows

$$\begin{aligned}
& \tau_N [T(\mathbf{Y}^{\text{lin}}, \mathbf{Y}^{\text{per}}, \mathbf{B})] \\
&= \sum_{\substack{\rho_0 \in \mathcal{P}(V) \text{ s.t.} \\ T^{\rho_0} \text{ pseudo-cactus}}} \delta^0 [\mathcal{T}_n^{\pi_0}(\mathbf{\Gamma})] \Psi^\rho \\
&\times \prod_{\substack{S \in \mathcal{CC}_1(T^{\rho_0}) \\ \text{with label in } \mathbf{z}_3}} \left(\frac{m_w^{(3)} m_x^{(3)}}{6} \right)^{c_1(\rho_0)} \prod_{S \in \mathcal{C}_1(\rho_0)} \mathbb{E}[h_{n(S)}'''(\xi)] \\
&\times \prod_{\substack{S \in \mathcal{CC}_2(T^{\rho_0}) \\ \text{with label in } \mathbf{z}_2}} \psi_0 \prod_{C \in \mathcal{SC}(S)} f(h_{n(S)}, h_{n'(S)}) \\
&\times \prod_{\substack{S \in \mathcal{CC}_3(T^{\rho_0}) \\ \text{with label in } \mathbf{z}_1}} \psi_0 \prod_{\ell=1}^{L_S} \mathbb{E}[h'_{n_\ell(S)}(\xi)] + o(1).
\end{aligned}$$

Hence, if $T = (V, E, \gamma)$ is a reference test graph in a single collection \mathbf{z} , denoting for any $\theta : E \rightarrow \{1, 2, 3\}$ by T_θ the test graph obtained from T by changing for each edge e its label $\gamma(e) = z(h)$ into $z_{\theta(e)}(h)$, we have

$$\tau_N^0 [T^{\rho_0}(\mathbf{Y}^{\text{lin}} + \mathbf{Y}^{\text{per}} + \mathbf{B})] = \sum_{\theta: E \rightarrow \{1, 2, 3\}} \tau_N^0 [T_\theta^{\rho_0}(\mathbf{Y}^{\text{lin}}, \mathbf{Y}^{\text{per}}, \mathbf{B})].$$

If S is a double edge of T^{ρ_0} , either θ attributes labels in \mathbf{z}_1 for both edges, or labels in \mathbf{z}_2 . All other contributions factorizing, these two termes adds up to give the expected formula. This proves that any trace of $\mathbf{Y}^{\text{lin}} + \mathbf{Y}^{\text{per}} + \mathbf{B}$ in a reference test graph labeled by odd polynomials satisfies the same asymptotic formula than the collection of Pennington-Worah matrices. Hence they have the same limiting traffic distribution, which conclude the proof of the main theorem.

References

- [Au18] Benson Au. Traffic distributions of random band matrices. Electron. J. Probab., 23:Paper No. 77, 48, 2018.
- [BG09] Florent Benaych-Georges. Rectangular random matrices, related convolution. Probab. Theory Related Fields, 144(3-4):471–515, 2009.
- [BMS17] Serban T. Belinschi, Tobias Mai, and Roland Speicher. Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem. J. Reine Angew. Math., 732:21–53, 2017.

- [BP21] Lucas Benigni and Sandrine Péché. Eigenvalue distribution of some nonlinear models of random matrices. Electron. J. Probab., 26:Paper No. 150, 37, 2021.
- [BP22] Lucas Benigni and Sandrine Péché. Largest Eigenvalues of the Conjugate Kernel of Single-Layered Neural Networks. arXiv e-prints, page arXiv:2201.04753, January 2022.
- [Cap13] M. Capitaine. Additive/multiplicative free subordination property and limiting eigenvectors of spiked additive deformations of Wigner matrices and spiked sample covariance matrices. J. Theoret. Probab., 26(3):595–648, 2013.
- [CDM07] M. Capitaine and C. Donati-Martin. Strong asymptotic freeness for Wigner and Wishart matrices. Indiana Univ. Math. J., 56(2):767–803, 2007.
- [CM14] Benoît Collins and Camille Male. The strong asymptotic freeness of Haar and deterministic matrices. Ann. Sci. Éc. Norm. Supér. (4), 47(1):147–163, 2014.
- [Dyk94] Ken Dykema. Interpolated free group factors. Pacific J. Math., 163(1):123–135, 1994.
- [GC24] Camille Male Guillaume Cébron, Antoine Dahlqvist. Traffic distributions and independence ii: Universal constructions for traffic spaces. Documenta Mathematica, 29(1):39–114, 2024.
- [HMRT22] Trevor Hastie, Andrea Montanari, Saharon Rosset, and Ryan J. Tibshirani. Surprises in high-dimensional ridgeless least squares interpolation. Ann. Statist., 50(2):949–986, 2022.
- [Mal20] Camille Male. Traffic distributions and independence: permutation invariant random matrices and the three notions of independence. Mem. Amer. Math. Soc., 267(1300), 2020.
- [Péc19] Sandrine Péché. A note on the Pennington-Worah distribution. Electron. Commun. Probab., pages Paper No. 66, 7, 2019.
- [PW19] Jeffrey Pennington and Pratik Worah. Nonlinear random matrix theory for deep learning. J. Stat. Mech. Theory Exp., pages 124005, 14, 2019.
- [Shl96] Dimitri Shlyakhtenko. Random Gaussian band matrices and freeness with amalgamation. Internat. Math. Res. Notices, (20):1013–1025, 1996.
- [Voi91] Dan-Virgil Voiculescu. Limit laws for random matrices and free products. Invent. Math., 104(1):201–220, 1991.

- [Zit24] Gregory Zitelli. Traffic probability for rectangular random matrices. Random Matrices Theory Appl., 13(3):Paper No. 2450013, 55, 2024.