Managing Expectations: Freeness and the Fourier Matrix

Marc Desgroseilliers, Olivier Lévêque, Camille Male

Managing Expectations: Freeness and the Fourier Matrix

Marc Desgroseilliers*, Olivier Lévêque*, Camille Male†
*Swiss Federal Institute of Technology - Lausanne, 1015 Lausanne, Switzerland
Emails: {marc.desgroseilliers, olivier.leveque}@epfl.ch,
†Paris Descartes- Paris, France
Emails: {camille.male}@parisdescartes.fr

***Abstract*—We introduce a new criterion depending on the concept of traffic-freeness that can be used to rule out freeness between an i.i.d. diagonal matrix and another matrix. This sheds new light on a conjecture mentioned in [1], answering the conjecture negatively in many cases.**

A. Motivation

The communication problem prompting our investigation is a wireless channel with frequency and time selective fading ([1]). We hope to reduce the stochasticity involved to make the analysis more robust to real world scenarios.

We begin with a well understood channel. The time-selective coherent channel is modelled as (in vector form):

$$\mathbf{y} = \sqrt{\gamma} H \mathbf{x} + \mathbf{n}$$

where $\mathbf{x}$ is subject to an average power constraint $\mathbb{E}[x_i] \leq P$, $\mathbf{n}$ is a vector whose components are i.i.d. random variables of unit variance, γ is the Signal to Noise Ratio (SNR) and H is a diagonal matrix whose entries come from a fading process known to the receiver, stationary and ergodic.

Assuming the decoder knows the realizations of the fading process, the capacity of this channel (in the limit as the dimension $N \to \infty$) is known:

$$C(\gamma) = \mathbb{E}[\log(1 + \gamma |h|^2)] \tag{1}$$

where h is a random variable distributed according to the stationary distribution of H.

The frequency selective channel is defined analogously

$$\mathbf{y} = \sqrt{\gamma} F G F^* \mathbf{x} + \mathbf{n}$$

where G is a diagonal matrix of fading coefficients, and F is the unitary Fourier matrix, defined as $f_{jk} = \frac{1}{\sqrt{N}} e^{2\pi i (j-1)(k-1)}$ for $j, k = 1, \ldots, N$.

It now seems natural to look at the following channel

$$\mathbf{y} = \gamma H F G F^* \mathbf{x} + \mathbf{n} \tag{2}$$

where we concatenate the effects of the two channels: we have both frequency-selective and time-selective fading. The problem becomes much more difficult to analyze as the interplay between the two types of fading is hard to control.

If we suppose that both fading processes are i.i.d., we have the following theorem, which satisfyingly answers our question concerning capacity.

Theorem 1 ([1]). *Consider the channel model* (2) *with fading unknown at the transmitter, full channel state information at the receiver and H and G having i.i.d. entries. The capacity of this channel is given by*

$$C(\gamma) = \mathbb{E}[\log(1+\alpha\gamma|g|^2)] + \mathbb{E}[\log(1+\nu\gamma|h|^2)] - \log(1+\alpha\nu\gamma) \tag{3}$$

where $0 \leq \alpha \leq \mathbb{E}[|h|^2]$ and $0 \leq \nu \leq \mathbb{E}[|g|^2]$ are coefficients that depend on γ and on the fading distributions, and are defined to be the solution to

$$\mathbb{E}[(1+\alpha\gamma|g|^2)^{-1}] = (1+\alpha\nu\gamma)^{-1} = \mathbb{E}[(1+\nu\gamma|h|^2)^{-1}] \tag{4}$$

The proof of this theorem relies on the fact that the matrices H and FGF^* are asymptotically free, see [2] for the definition of freeness and related material. It was conjectured using numerical simulations that asymptotic freeness between these matrices could be extended to some cases where the matrix G is deterministic. We will provide a criterion that allows us to conclude that in many models, this is not the case.

I. Traffics

Invariance by permutation of one of our matrices is an important ingredient that allows us to make significant progress in proving asymptotic freeness, see the work of [3] on asymptotic liberation. We will now introduce the concept of traffics, which are noncommutative random variables naturally suited to the invariance by permutation.

Traffic spaces are spaces of noncommutative random variables equipped with extra structure. There exists an associated notion of traffic freeness and it also allows one to compute joint moments of noncommutative random variables from the individual moments, much like freeness. It is the interplay between freeness and traffic freeness that will interest us. We will mostly be concerned with large random matrices, but a more general setting can be elaborated and we encourage the algebraically-inclined reader to read [4] where this is developed.

Hypothesis 2. *Throughout this section, we let A_N be a diagonal matrix with i.i.d. entries, with distribution having bounded support, and $B_N = F_N \Delta_N F_N^*$ of the form a diagonal deterministic matrix $\Delta_N = \operatorname{diag}(\delta_1, \ldots, \delta_N)$ conjugated by the unitary Fourier matrix F_N. We suppose that the δ_i are bounded and converge in distribution.*

The idea behind traffics is to evaluate matrices not only on polynomials, but more generally on graphs, which have a richer structure. We will now define what we mean by evaluating a matrix on a graph. For the remainder of this document, we will consider T a finite connected oriented graph. Also, for V the vertex set of T, let $\Pi_{|V|}$ be the set of partitions of the vertices of T. For $\pi \in \Pi_{|V|}$, we let T^π be the graph obtained by identifying the vertices of T according to the partition π.

Definition 3 ((injective) trace)**.** *Let T be a (finite, connected, oriented) graph with edge set E and vertex set V, and $\mathbf{X}_N = (X_i)_{i\in I}$ a family of matrices of dimension N. Let there be given two functions $\epsilon : E \to \{1, *\}$ and $\gamma : E \to I$. Here, γ chooses a matrix for each edge, and for this matrix, ϵ picks between the matrix and its adjoint.*

We define $\tau_N(T(\mathbf{X}_N))$ as

$$\tau_N(T(\mathbf{X}_N)) := \frac{1}{N}\mathbb{E}\mathrm{Tr}(T(\mathbf{X}_N)) = \\ = \frac{1}{N}\mathbb{E}\sum_{\phi:V\to[N]}\prod_{e=(v,w)\in E} X_{\gamma(e)}^{\epsilon(e)}(\phi(v),\phi(w))$$

where $[N] = \{1,2,\dots,N\}$. If the summation in the above definition is over all functions ϕ which are injective, we call this object the injective trace and denote it $\tau_N^0(T(\mathbf{X}_N))$.

This object $\tau_N(T(X_N))$ can be interpreted as follows: to each edge is attached a matrix from the family $\mathbf{X}_N$. The summation is over all possible assignments of indices for the vertices, which for every edge picks an entry from the attached matrix.

Remark 4. *We have the following relationship between the trace and the injective trace.*

$$\tau_N(T(\mathbf{X}_N)) = \sum_{\pi\in\Pi_{|V|}} \tau_N^0(T^\pi(\mathbf{X}_N))$$

We can express τ^0 in terms of τ by using the Möbius inversion formula on the lattice $\Pi_{|V|}$.

$$\tau_N^0(T(\mathbf{X}_N)) = \sum_{\pi\in\Pi_{|V|}} \mu(\pi)\tau_N(T^\pi(\mathbf{X}_N)) \tag{5}$$

where μ is the so-called Möbius function that depends on the structure of the partially ordered set. In our case, this is the lattice $\Pi_{|V|}$ of all partitions of the set V, ordered by the relationship of inclusion, i.e. a finer partition is below a coarser one. This means the bottom element is the partition of singletons, and the top element is the partition consisting of one block. See [5] for many applications, as well as explicit computations of the Möbius function.

Definition 5. *Similarly the moments of a distribution, we define the map $T \to \tau_N(T(\mathbf{X}_N))$ as the traffic distribution in moments. By convergence in traffic distribution, we mean the pointwise convergence of this map.*

Remark 6. *In the above definition, we only ask that for every fixed graph T, the sequence $\tau_N(T(X_N))$ has a finite limit as $N\to\infty$. This will be sufficient for our purposes. The appropriate concept of spaces of traffics where this convergence takes place has been introduced in [4], and we refer the reader to this article for the concepts surrounding spaces of traffics.*

Observe that by taking the graph to be a cycle, we can recover any given $*$-monomial. This means that convergence in traffic distribution implies convergence in *-distribution [1]. Given Hypothesis 2 on A_N and B_N, we have the following result concerning the joint distribution of (A_N, B_N).

Theorem 7 ([4], Theorem 3.4)**.** *Suppose that B_N is deterministic and that for all graphs T, $\tau_N(T(B_N))$ converges as $N\to\infty$. Then (A_N, B_N) converges in traffic distribution, and hence in $*$-distribution.*

The matrix $B_N = F_N\Delta F_N^*$ is of particular interest to us. Let ω be a primitive N^{th} root of unity, $\mathbb{1}$ denote the indicator function and (N) the operation of reducing modulo N. Using the fact that $\sum_{j=0}^{N-1}\omega^{jk} = N\mathbb{1}_{k=0(N)}$, we have the following lemma.

[1] i.e. the noncommutative moments of the random variables

Lemma 8. *The value of $\tau_N[T(B_N)]$ is*

$$\frac{1}{N^{1+|E|-|V|}}\sum_{j_1,\dots j_{|E|}\in[N]} \delta_{j_1}\dots\delta_{j_{|E|}}\prod_{v\in V}\mathbb{1}_{I(v)=0(N)}.$$

Where, for a vertex v, we let $I(v) := \sum_{e \text{ entering } v} j_e - \sum_{e \text{ exiting } v} j_e$ be the difference between the sum of the values associated to the edges entering v and those exiting v.

Proof. First, observe that the fact that $B_N = F_N\Delta_N F_N^*$ immediately yields that $(b_n)_{jk} = \frac{1}{N}\sum_{l=1}^N \delta_l\omega^{(j-k)l}$. By Definition 3,

$$\tau_N[T(B_N)] = \frac{1}{N}\sum_{\substack{i_1,\dots i_{|V|}\\ \in[N]}}\prod_{e\in E} B_N(e_{source}, e_{destination})$$

$$= \frac{1}{N}\sum_{\substack{i_1,\dots i_{|V|}\\ \in[N]}}\sum_{\substack{j_1,\dots j_{|E|}\\ \in[N]}}\prod_{e\in E}\frac{\omega^{(i_{source}-1)(j_e-1)}}{\sqrt{N}}\delta_{j_e}\frac{\omega^{-(i_{dest}-1)(j_e-1)}}{\sqrt{N}}$$

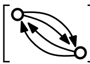
$$= \frac{1}{N^{1+|E|}}\sum_{\substack{j_1,\dots j_{|E|}\\ \in[N]}}\delta_{j_1}\dots\delta_{j_{|E|}}\sum_{\substack{i_1,\dots i_{|V|}\\ \in[N]}}\prod_{v\in V}\omega^{(i_v-1)I(v)}$$

$$= \frac{1}{N^{1+|E|-|V|}}\sum_{\substack{j_1,\dots j_{|E|}\\ \in[N]}}\delta_{j_1}\dots\delta_{j_{|E|}}\prod_{v\in V}\mathbb{1}_{I(v)=0(N)}$$

□

We give an example which will be of use later: we evaluate the matrix B_N on a particular graph with 2 vertices and 4 edges.

Corollary 9. *Evaluated on the matrix B_N, we have*



$$\tau_N\left[\ \cdot\ \right] = \frac{1}{N^3}\sum_{j_1\dots j_4}\delta_{j_1}\dots\delta_{j_4}\mathbb{1}_{j_1+j_2-j_3-j_4=0(N)}$$

We now have the necessary background to introduce the concept of traffic freeness.

Definition 10. *Let $\mathbf{X}_1,\dots,\mathbf{X}_p$ be families of different random matrices and let T be a graph with the associated ϵ and γ functions. Denote by $T_1,\dots,T_k$ the connected components of T that are labelled by variables in the same family. Consider $\bar{T}$ the graph whose vertices are given by $T_1,\dots,T_K$, as well as the vertices v_i of T that are common to at least 2 components T_j. There is an edge between v_i and T_j in $\bar{T}$ if $v_i\in T_j$ in T. Finally, we say that T is a* free product *in $\mathbf{X}_1,\dots,\mathbf{X}_p$ whenever $\bar{T}$ is a tree.*

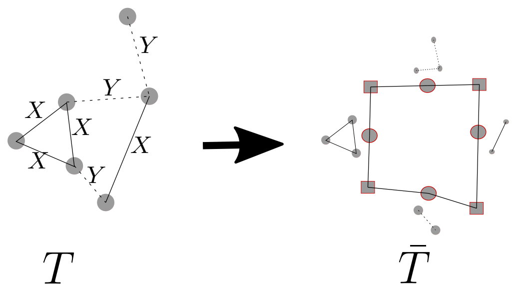


Fig. 1. Example of the construction of $\bar{T}$. On the righthand side, the round nodes represent connected components of T and the square nodes represent shared vertices.

Definition 11. *Two families of random matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ are said to be asymptotically traffic free if their joint distribution converges to some limit and for any graph T, we have that*

$$\tau^0[T(\mathbf{X}_1, \mathbf{X}_2)] = \begin{cases} \prod_{j=1}^{k} \tau^0[T_j(\mathbf{X}_{i_j})] & \text{if } T \text{ is a free product in } \mathbf{X}_i \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Concretely this means that if two families of matrices are traffic free, we can compute the injective trace (and therefore also the trace) on any graph from the traffic distribution of the marginals, as it is a product according to equation (6).

Theorem 12 ([4] Theorem 3.4)**.** *Let $E_N^{(1)}$ and $E_N^{(2)}$ be two independent families of matrices such that:*

1) *$E_N^{(1)}$ is invariant by permutation.*
2) *$E_N^{(1)}$ and $E_N^{(2)}$ converge in traffic distribution.*
3) *For $j = 1, 2$ and for any family of finite graphs T_i in the variable x_j, we have that*

$$\mathbb{E}\left[\prod_{i=1}^{n} \frac{1}{N}\mathrm{Tr}[T_i(E_N^{(j)})]\right] \underset{N\to\infty}{\to} \prod_{i=1}^{n} \tau[T_i(x_j)]$$

Then the matrices $E_N^{(1)}$ and $E_N^{(2)}$ are asymptotically traffic free.

Corollary 13. *With Hypothesis 2, A_N and B_N are asymptotically traffic free. Indeed, the only condition to check in Theorem 12 is point 3, which is true both for the i.i.d. diagonal matrices A_N and for the deterministic matrix B_N.*

A. Criteria ensuring the lack of asymptotic freeness

In this section, we compute explicit criteria that prevents matrices from being asymptotically free. Let A^0 be the recentred matrix $A_N - \frac{1}{N}\mathbb{E}\mathrm{Tr}(A_N)I$ and let δ_i^0 be the diagonal entries of the recentred matrix Δ_N^0. The strategy is the following: we use the rules of traffic freeness to compute the quantity $\frac{1}{N}\mathbb{E}\mathrm{Tr}[(A_N^0 B_N^0)^k]$. If the two matrices are free, this quantity should converge to 0. If this is not the case, we know that our matrices cannot be free.

1) Criterion of order 2: This approach has already been used in [4] to deduce the following theorem. Here, since the entries the matrix conjugated by the Fourier matrix could be random, we use a different notation, $M_N = F_N \Xi_N F_N^*$ with $\Xi_N = \mathrm{diag}(\xi_1, \dots, \xi_N)$.

Theorem 14 ([4] Corollary 3.5)**.** *Let M_N be a random matrix asymptotically traffic free from a diagonal matrix A_N, and let P and Q be *-polynomials. If the limiting empirical eigenvalue distribution of neither M_N nor A_N is a Dirac mass and*

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}\mathrm{Tr}[P(M_N) \circ Q(M_N)] \neq \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\mathrm{Tr}[P(M_N)] \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\mathrm{Tr}[Q(M_N)] \quad (7)$$

*then the matrices A_N and M_N are not asymptotically *-free. Here, $\circ$ denotes the Hadamard product, i.e. $(M \circ N)_{jk} = m_{jk} n_{jk}$.*

By choosing $P(X) = X$ and $Q(X) = X$, we obtain a simple, easy to apply criterion. Observe that the diagonal entries of M_N are all equal to $\frac{1}{N}\sum_{i=1}^{N} \xi_i$. We can reformulate the criterion (7) as

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{E}\left[\left[\frac{1}{N}\sum_{i=1}^{N} \xi_i\right]^2\right] - \left[\frac{1}{N}\mathbb{E}\sum_{i=1}^{N} \xi_i\right]^2 \\ &= \lim_{N\to\infty} \mathrm{Var}\left[\frac{1}{N}\sum_{i=1}^{N} \xi_i\right] \neq 0 \end{aligned} \quad (8)$$

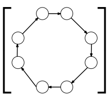
In Section II, we will use this criterion, which relies on the moment of order 2, to deduce that some matrices are not asymptotically free. However, if the matrix M_N in (7) is deterministic, this criterion will always evaluate to 0. This prompts us to compute a criterion to check asymptotic freeness which relies on the moment of order 4.

Remark 15. *A_N and M_N are asymptotically traffic free, as long as M_N is independent of A_N and that it converges in traffic-distribution. This can be deduced using the fact that A_N is diagonal from the proof of [4, Theorem 3.4].*

2) Criterion of order 4: We now compute the criterion of order 4 for deterministic matrices B_N. Let us remark that if the deterministic matrix B_N is replaced by the random matrix M_N from the previous section, a slightly more complex criterion can be obtained, but the essential features are the same.

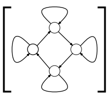
We have

$$\frac{1}{N}\mathbb{E}\mathrm{Tr}((A_N^0 B_N^0)^4) = \tau_N \left[\quad\right]$$



where the edges are alternating in A_N^0 and B_N^0

$$= \tau_N \left[\quad\right] \quad (9)$$



since A^0 is diagonal

$$= \sum_{\pi \in \Pi_{|V|}} \tau_N^0[T^\pi(A_N^0, B_N^0)]$$

Let us call C_k the cyclic graph with loops, C_4 appearing in Equation (9). Recall that T^π is the graph induced from T by quotienting by the relation induced by the partition π. We now use the fact that A_N and B_N are asymptotically traffic free. The following is a reformulation for this particular case of Definition 11.

For a graph T, denote by T_b be the graph obtained by keeping only the edges with labels B_N. We then have that

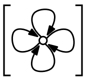
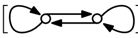
$$\begin{aligned} &\lim_{N\to\infty} \tau_N^0[T^\pi(A_N, B_N)] \\ &= \lim_{N\to\infty} \tau^0[T_b^\pi(B_N)] \prod_{\beta\in\pi} \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\mathrm{Tr}(A_N^{|\beta|}) \end{aligned} \quad (10)$$

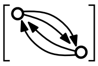
where for $\beta \in \pi$ a block of the partition π, we denote $|\beta|$ the size of this block. Observe that since A_N is diagonal, the graphs T_b^π are always connected.

In principle, this gives us a formula to compute any moment of the product $A_N B_N$. However, the complexity grows very quickly when the order of the moment increases and this is why we use the moment of order 4.

Moreover, Equation (10) involves τ^0 and Lemma 8 gives a formula for τ. We will therefore use the formula $\tau = \sum_{\pi\in\Pi_{|V|}} \tau^0$ to relate τ and τ^0. Since the partitions are small, we can do it by hand, but, in general, we would use Möbius inversion to express τ^0 in terms of τ.

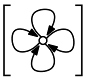
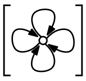
We must therefore enumerate all possible partitions of the 4 vertices and look at their contribution in Equation (10). Observe that since the entries are centred, if a partition has a singleton, the contribution is 0. Moreover, since A_N^0 and B_N^0 are traffic free, the induced graph $\overline{T^\pi}$ should be a tree for the contribution to be non-zero. This last condition is always satisfied. Only 4 partitions are left, 2 of them yielding the same graph. Writing out Equation (10), $\frac{1}{N}\mathbb{E}\mathrm{Tr}[(A_N, B_N)^4]$ equals

τ_N^0 [] $\times \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}A_N^4\right] + 2\tau_N^0$ [] $\times \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}A_N^2\right]^2$

$+\tau_N^0$ [] $\times \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}A_N^2\right]^2 + o(1)$ (11)

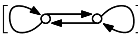
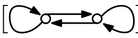
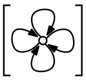
and we have to compute the value of the 3 different graphs in B_N. This is what follows:

3) 1st Graph:

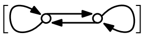
τ_N^0 [] $= \tau_N$ [] $= \frac{1}{N^4}\sum_{j_1,\dots,j_4}\delta_{j_1}^0\dots\delta_{j_4}^0$

$$= \left[\frac{1}{N}\sum_j \delta_j^0\right]^4 = 0$$

4) 2nd Graph: Observe that

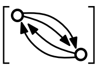
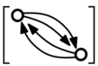
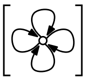
τ_N^0 [] $= \tau_N$ [] $- \tau_N^0$ []

and

τ_N [] $= \frac{1}{N^3}\sum_{j_1,\dots,j_4}\mathbb{1}_{j_2-j_3=0(N)}\delta_{j_1}^0\dots\delta_{j_4}^0$

$$= \frac{1}{N^3}\sum_{j_1,j_2j_4}\delta_{j_1}^0(\delta_{j_2}^0)^2\delta_{j_4}^0 = \left[\frac{1}{N}\mathrm{Tr}(\Delta^0)\right]^2\left[\frac{1}{N}\mathrm{Tr}(\Delta^0)^2\right] = 0$$

5) 3rd Graph:

τ_N^0 [] $= \tau_N$ [] $- \tau_N^0$ []

$$= \frac{1}{N^3}\sum_{j_1,\dots,j_4}\mathbb{1}_{j_1+j_2-j_3-j_4=0(N)}\delta_{j_1}^0\dots\delta_{j_4}^0$$

Putting the last 3 computations together into Equation (11), we conclude that

$$\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\mathrm{Tr}((A_N^0B_N^0)^4) = \mathbb{E}\left[\lim_{N\to\infty}\frac{1}{N}Tr((A_N^0)^2)\right]^2 \times \frac{1}{N^3}\sum_{j_1,\dots,j_4}\delta_{j_1}^0\dots\delta_{j_4}^0\mathbb{1}_{j_1+j_2=j_3+j_4(N)}$$

The conclusion that we can draw from this is that if the quantity

$$S_N(\Delta_N^0) := \frac{1}{N^3}\sum_{j_1,\dots,j_4}\delta_{j_1}^0\dots\delta_{j_4}^0\mathbb{1}_{j_1+j_2-j_3-j_4=0(N)} \quad (12)$$

is not asymptotically 0, the matrices A_N and B_N cannot be asymptotically free.

II. Examples of lack of freeness

We return to our problem concerning wireless communication. In order to generalize the capacity theorem from [1] to a wider family of channels, the two channel matrices modelling the time and frequency fading must be freely independent: these are the matrices A_N and B_N (or M_N). We will therefore apply our criteria to these matrices with particular models for B_N (or M_N) to identify cases where the matrices cannot possibly be asymptotically free.

We begin with a deterministic example. More examples and computations can be found in the doctoral thesis [6].

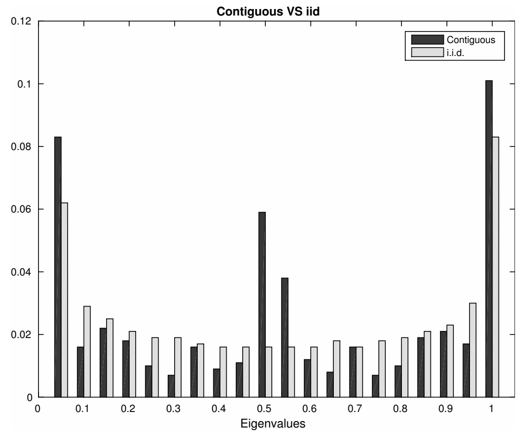


Fig. 2. We see very different behaviour for the eigenvalues. Here $N = 1000$ and we plot the eigenvalues of $AF\Delta F^*$ where A is i.i.d. Bernoulli(0.5) and Δ is either i.i.d. or contiguous (we subtract a weight of 0.5 at the origin for better visibility).

A. Contiguous values

Fix $0 < \eta < 1$ and consider the diagonal matrix

$$\delta_i = \begin{cases} 1 & \text{if } i < \eta N \\ 0 & \text{otherwise} \end{cases}$$

This matrix has two blocks of values, and represents on off fading where the fading blocks are very long. The associated Δ_N^0 matrix is

$$\delta_i^0 = \begin{cases} 1-\eta & \text{if } i < \eta N \\ -\eta & \text{otherwise} \end{cases}$$

We can make the explicit computation of $S_N(\Delta_N^0)$ in this case. For example, we take $\eta = 1/2$ and N even, $N' = N/2$. Similar results can be obtained for $0 < \eta < 1, \eta \neq 1/2$.

We add the contributions with an even number of positive terms and subtract those with an odd number and obtain

$$\frac{8}{N^3} * \left[N'^2 + \frac{N'(N'-1)(N'-2)}{3}\right] * \eta^4 = \frac{1}{16}\left(\frac{1}{3} + \frac{8}{3N^2}\right)$$

This will not converge to 0 as $N \to \infty$, showing the lack of asymptotic freeness in this case.

B. Markov Chains

Let us consider the case where the matrix M_N is not deterministic and the entries ξ_i are distributed according to a stationary Markov chain. It is sufficient to use the criterion of order 2, Equation (8).

For concreteness, we examine a Markov chain ξ_i with two states, but the discussion carries to any Markov chain whose stationary distribution has zero expectation.

Let the chain have two states s_1 and s_2 with transition probabilities α and β. The chain remains in its current state with probabilities $1-\alpha$ and $1-\beta$ respectively. The stationary distribution is given by

$$\pi = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$$

and so we require that $\beta s_1 + \alpha s_2 = 0$. We set ξ_0 the initial state to have distribution π.

The transition probabilities $p_{ij}(n)$ can easily be computed by diagonalizing the transition matrix.

$$p_{ij}(n) = \begin{pmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n \\ \frac{\beta}{\alpha+\beta} - \frac{\beta}{\alpha+\beta}(1-\alpha-\beta)^n & \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta}(1-\alpha-\beta)^n \end{pmatrix}$$

Lemma 16. *On the above Markov chain, we have that the expression in the criterion of order 2, Equation* (8)*, evaluates to the equation at the bottom of the page.*

Consequence. *Equation* (16) *will be zero if we take* α *and* β *such that* $\frac{1}{\alpha}, \frac{1}{\beta} \in o(N)$*. Note that in the case where* α *and* β *are constants, the asymptotic freeness of* A_N *and* M_N *has already been proven in [1]. Equation* (16) *will however be nonzero if we take* $\alpha = k/N$, $\beta = k'/N$ *for* k, k' *constants. When looking at the Markov chain* ξ_i*, this would typically result in blocks whose length is of order* N*. In our communication scenario, this would translate to a bursty channel. See Figure 3 for the empirical eigenvalue distribution of the product* $A_N M_N$ *associated to these Markov chains.*

Proof of Lemma 16. According to Equation (8), we must compute $\mathbb{E}[(\frac{1}{N}\sum_i^N \xi_i)^2] = \frac{1}{N^2}\mathbb{E}[\sum_{ij}^N \xi_i\xi_j]$ since the diagonal entries of $F\Xi F^*$ are the same across the diagonal. It suffices to compute the second moment and by stationarity we can suppose $i = 0$. First, $\mathbb{E}(\xi_i^2) = \frac{\beta}{\alpha+\beta}s_1^2 + \frac{\alpha}{\alpha+\beta}s_2^2$.Now, for $k > 0$,

$$\begin{aligned}
\mathbb{E}(\xi_0\xi_k) &= \mathbb{E}(\xi_0\xi_k|\xi_0 = s_1)\mathbb{P}(\xi_0 = s_1) + \mathbb{E}(\xi_0\xi_k|\xi_0 = s_2) \\
&\quad \times \mathbb{P}(\xi_0 = s_2) \\
&= [\mathbb{P}(\xi_k = s_1|\xi_0 = s_1)s_1^2 + \mathbb{P}(\xi_k = s_2|\xi_0 = s_1)s_1s_2] \\
&\quad \times \frac{\beta}{\alpha+\beta} \\
&+ [\mathbb{P}(\xi_k = s_1|\xi_0 = s_2)s_1s_2 + \mathbb{P}(\xi_k = s_2|\xi_0 = s_2)s_2^2] \\
&\quad \times \frac{\alpha}{\alpha+\beta} \\
&= \frac{\alpha\beta}{(\alpha+\beta)^2}(s_1 - s_2)^2(1-\alpha-\beta)^k
\end{aligned}$$

With this expression in hand, we can compute the second moment

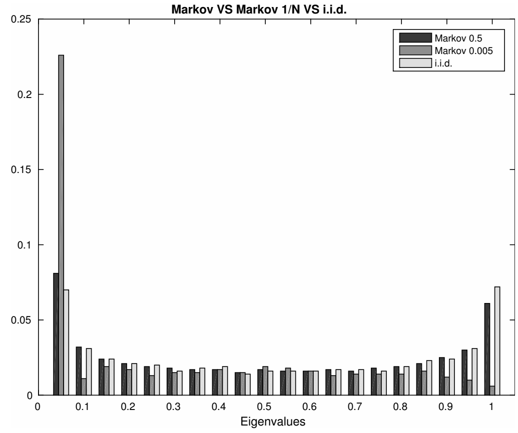


Fig. 3. We see here that if the transition probability between states is very low ($\alpha = 0.005$), the eigenvalue distribution is significantly different (we subtract a weight of 0.5 at the origin for better visibility).

$$\begin{aligned}
2\sum_{i>j}\mathbb{E}(\xi_i\xi_j) &= 2\frac{\alpha\beta}{(\alpha+\beta)^2}(s_1-s_2)^2\sum_{i>j}(1-\alpha-\beta)^{i-j} \\
&= 2\frac{\alpha\beta}{(\alpha+\beta)^2}(s_1-s_2)^2\left[\sum_{j=2}^N \frac{1-(1-\alpha-\beta)^j}{\alpha+\beta} - 1\right] \\
&= 2\frac{\alpha\beta}{(\alpha+\beta)^2}(s_1-s_2)^2 \\
&\times \left[(N-1)\frac{1-\alpha-\beta}{\alpha+\beta} + \frac{(1-\alpha-\beta)^2}{(\alpha+\beta)^2} - \frac{(1-\alpha-\beta)^{N+1}}{(\alpha+\beta)^2}\right]
\end{aligned}$$

Putting all this together we get the equation from the statement of the lemma.

□

III. Conclusion and perspectives

In this chapter, we have introduced the concepts of freeness and traffic-freeness in noncommutative probability spaces and seen how these allow us to describe the eigenvalue distribution of sums and products of large random matrices. We have discussed two criteria preventing asymptotic freeness between particular families of matrices, hence providing an answer to questions concerning freeness raised in [1]. This sheds new light and gives a better understanding of channels experiencing both time and frequency domain fading.

References

[1] A. M. Tulino, G. Caire, S. Shamai, and S. Verdu, "Capacity of channels with frequency-selective and time-selective fading," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1187–1215, 2010.

[2] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*. Cambridge University Press, 2006.

[3] B. Farrell and G. Anderson, "Asymptotically liberating sequences of random unitary matrices," *Advances in Mathematics*, vol. 255, pp. 381–413, 2014.

[4] C. Male, "The distribution of traffics and their free product," *Arxiv*, 2011.

[5] R. Stanley, *Enumerative Combinatorics Volume 1*. Cambridge University Press, 1997.

[6] M. Desgroseilliers, "Reducing randomness in matrix models for wireless communications," Ph.D. dissertation, EPFL, 2015.

$$\mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^N \xi_i\right) = \frac{2}{N^2}\frac{\alpha\beta}{(\alpha+\beta)^2}(s_1-s_2)^2\left[(N-1)\frac{1-\alpha-\beta}{\alpha+\beta} + \frac{(1-\alpha-\beta)^2}{(\alpha+\beta)^2} - \frac{(1-\alpha-\beta)^{N+1}}{(\alpha+\beta)^2}\right] + \frac{1}{N}\left[\frac{\beta}{\alpha+\beta}s_1^2 + \frac{\alpha}{\alpha+\beta}s_2^2\right]$$