Eigenvalues of Euclidean Random Matrices

Charles Bordenave [*]

February 2, 2008

Abstract

We study the spectral measure of large Euclidean random matrices. The entries of these matrices are determined by the relative position of n random points in a compact set Ω_n of $\mathbb{R}^d$. Under various assumptions we establish the almost sure convergence of the limiting spectral measure as the number of points goes to infinity. The moments of the limiting distribution are computed, and we prove that the limit of this limiting distribution as the density of points goes to infinity has a nice expression. We apply our results to the adjacency matrix of the geometric graph.

Keywords: random matrix, spectral measure, random geometric graphs, spatial point process, Euclidean distance matrix.

1 Introduction

The main research effort in the theory of random matrices concerns matrices where the coefficients are independent random variables (see Bai [1] for a survey). Few authors have studied the limiting spectral measures of other types of large matrices, in particular, Markov, Hankel and Toeplitz matrices have been studied by Bryc, Dembo and Jiang [3] and Toeplitz matrices by Hammond and Miller [10]. In this paper, we consider another class of random matrices, the Euclidean random matrices (ERM) which have been introduced by Mézard, Parisi and Zee [13]. An ERM is an $n \times n$ matrix, A, whose entries is a function of the positions of n random points in a compact set Ω of $\mathbb{R}^d$. In this paper, Ω will be an hypercube, the n points $\mathcal{X}_n = \{X_1, \cdots, X_n\}$, n uniformly distributed points in Ω and

$$A = (F(X_i - X_j))_{1 \leq i \leq j \leq n}, \tag{1}$$

where F is a measurable mapping from $\mathbb{R}^d$ to $\mathbb{C}$. We will pay attention to the spectral properties of A. In this paper, we will compute some limits of the spectral measure as the number of points n goes to infinity. We will show how the eigenvalues of A are related to the Fourier transform of the mapping F.

Examples of interests in branches of physics are explained in [13] and Offer and Simons [14]. A particularly appealing case is $F(x) = \|x\|$, the Euclidean norm. This subclass of ERM is called the random Euclidean Distance Matrices and some of their spectral properties are derived in Vershik [17], Bogomolny, Bohigas and Schmidt [2]. In [17] the author generalizes the problem considered here and consider an integral operator defined on a metric space. Koltchinskii and Giné [11] have analyzed the convergence of the spectra of the matrix $(n^{-1}h(X_i, X_j))_{1 \leq i,j \leq n}$ to the spectra of the compact integral operator with symmetric kernel h.

[*]Institut de Mathématiques - Université de Toulouse & CNRS (UMR 5219) - 31062 Toulouse - France. Email: bordenave@math.ups-tlse.fr

Another field of application is graph theory. Indeed, if $F(x) = \mathbf{1}(0 \leq \|x\| \leq r)$, then A is the adjacency matrix of the proximity (or geometric) graph (refer to Penrose [15]). More generally if $F(X) = F(-X) \in \{0, 1\}$ then A is the adjacency matrix of a random graph. The spectral properties of the adjacency matrix or related matrices are of prime interest in graph theory. For example the probability of hitting times of random walks on graphs is governed by the spectrum of the transition matrix (for a survey on this subject, see e.g. Section 3 in Lovász [12]). Or, in network epidemics, the time evolution of the infected population is also closely related to the spectral radius and the spectral gap of the adjacency matrix, see Draief, Ganesh and Massoulié [8]. For Erdós-Renyi random graphs, some properties of the spectrum can been computed thanks to the seminal work Wigner of [18] and Füredi and Komlós [9]. For power law graphs and related graphs, see Chung, Lu and Vu [4], [5].

Various generalizations of (1) would be worth to consider. Some extra randomness in the model could be added, and the entry of the matrix i, j could be equal to $F_{ij}(X_i - X_j)$, where $(F_{ij})_{1 \leq i,j \leq n}$ are i.i.d. mappings independent of the point set $\mathcal{X}_n$. Falls into this framework the adjacency matrix of a random graph where there is an edge between two points with a probability which is deterministic function of their distance, such as the small world graphs (see for example Ganesh and Draief [7]).

Another generalization is the original model of Mézard, Parisi and Zee [13] where the entry i, j is equal to

$$F(X_i - X_j) - u\delta_{ij} \sum_k F(X_i - X_k),$$

where δ_{ij} is the Kronecker symbol and $u \in \mathbb{R}$. The case $u = 1$ is of particular interest, the matrix is then a Markov matrix.

In order to obtain the adjacency matrix of more sophisticated geometric graphs, such as the Delaunay triangulation, it would be necessary to consider an entry i, j which depends on the whole point set $\mathcal{X}_n$ and not only on $X_i - X_j$.

We will consider two models in this note. In the first model, $\Omega = [-1/2, 1/2]^d$ and F is 1-periodic function: if $x, y \in \mathbb{R}^d$ and $x - y \in \mathbb{Z}^d$ then $F(x) = F(y)$. Equivalently, the point set $\mathcal{X}_n = \{X_1, \cdots, X_n\}$ could be on the unit torus $\mathbb{T}^d = \mathbb{R}^d \backslash \mathbb{Z}^d$. We choose a periodic function in order to avoid all boundary effects with the hypercube Ω. The matrix A is defined by (1), where F is a measurable function from $\mathbb{R}^d$ to $\mathbb{C}$. The discrete Fourier transform of F is defined for all $k \in \mathbb{Z}^d$ by $\hat{F}(k) = \int_\Omega F(x)e^{-2i\pi k.x}dx$. Throughout the paper, we assume that a.e. and at 0, the Fourier series of F is equal to F:

$$F(x) = \sum_{k \in \mathbb{Z}^d} \hat{F}(k)e^{2i\pi k.x}.$$

A sufficient condition is $\sum_{k \in \mathbb{Z}^d} |\hat{F}(k)| < \infty$ and F continuous at 0. This Fourier transform plays an important role in the spectrum of A. As an example, consider $U = (U_i)_{1 \leq i \leq n}$ a vector in $\mathbb{C}^d$ and assume F hermitian ($F(-x) = \bar{F}(x)$), then a.s.

$$U^*AU = \sum_{i,j} F(X_i - X_j)U_i\bar{U}_j = \sum_{i,j}\sum_k \hat{F}(k)e^{2i\pi k.(X_i - X_j)}U_i\bar{U}_j = \sum_k \hat{F}(k)\left| \sum_{i=1}^n e^{2i\pi k.X_i}U_i \right|^2.$$

Therefore A is positive if and only if for all $k \in \mathbb{Z}^d$, $\hat{F}(k) \geq 0$.

We will compute explicitly the spectral measure of the matrix $A_n = A/n$ as n tends to ∞,

$$\mu_n = \sum_{i=1}^n \delta_{\lambda_i(n)/n},$$

2

where $\{\lambda_i(n)\}_{1\leq i\leq n}$ is the set of eigenvalues of A. Notice that $\{\lambda_i(n)/n\}_{1\leq i\leq n}$ is the set of eigenvalues of A_n. We define the measure:

$$\mu = \sum_{k\in\mathbb{Z}^d} \delta_{\hat{F}(k)}.$$

Since $\lim_{\|k\|\to\infty} \hat{F}(k) = 0$, μ is a counting measure with an accumulation point at 0.

Theorem 1 *For all Borel sets K with $\mu(\partial K) = 0$ and $0 \notin \bar{K}$, a.s.*

$$\lim_n \mu_n(K) = \mu(K). \tag{2}$$

The convergence of the spectral measure μ_n follows also from Theorem 3.1 in [11]. As an immediate corollary, we obtain the convergence of the spectral radius of A_n, almost surely, $\lim_{n\to\infty} \max_{1\leq i\leq n} \frac{|\lambda_i(n)|}{n} = \max_{k\in\mathbb{Z}^d} |\hat{F}(k)|$. For example if $F(x) = \mathbb{1}(\max_{1\leq i\leq d} |x_i| \leq r)$ then $\hat{F}(k) = r^d \prod_{i=1}^d \text{sinc}(2\pi k_j r)$, where $\text{sinc}(x) = \sin(x)/x$ and $k = (k_1, \cdots, k_d) \in \mathbb{Z}^d$. The spectral radius of A_n converges a.s. to r^d and the second largest eigenvalue to $r^d\text{sinc}(2\pi r)$ if r is small enough, thus the spectral gap is equivalent to $r^{d+2}(2\pi)^2/3!$ as r goes to 0.

Our second model is more challenging. Now, $\mathcal{X}_n = \{X_1, \cdots, X_n\}$ is the set of n independent points uniformly distributed on the hypercube $\delta_n^{-1}\Omega = [-\delta_n^{-1}/2, \delta_n^{-1}/2]^d$ where δ_n goes to 0. In this second model, we scale jointly the number of points and the space. We assume that for some $\gamma > 0$,

$$\lim_n \delta_n^d n = \gamma. \tag{3}$$

γ is the asymptotic density of the point set $\mathcal{X}_n$. Let f be a measurable function from $\mathbb{R}^D$ to $\mathbb{C}$ with support included in Ω, the matrix A is defined by (1) (with F replaced by f).

Considering the change of variable $x \mapsto \delta x$, the matrix A is equal to the matrix B_n defined by

$$B_n = (f_{\delta_n}(X_i - X_j))_{1\leq i\leq j\leq n},$$

where $f_\delta : x \mapsto f(x/\delta)$ and the point set $\mathcal{X}_n = \{X_1, \cdots, X_n\}$ is a set of n independent points uniformly distributed on Ω. The spectrum of B_n is denoted by $(\lambda_1'(n), \cdots, \lambda_n'(n))$, we define the empirical measure of its eigenvalues:

$$\nu_n = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i'(n)},$$

We will prove the following:

Theorem 2 *For all $\gamma > 0$, there exists a measure ν_γ such that for the topology of the weak convergence, a.s.:*

$$\lim_{n\to\infty} \nu_n = \nu_\gamma.$$

Moreover $\gamma \mapsto \nu_\gamma$ is continuous (for the topology of the weak convergence).

The exact computation of ν_γ is a difficult problem, we will compute the value $\nu_\gamma(P_m)$, where P_m is the polynomial $t \mapsto t^m$ (Equation (22)). However, the behavior of ν_γ as γ goes to infinity is simpler. Indeed, we define the Fourier transform of f by, for all $\xi \in \mathbb{R}^d$, $\hat{f}(\xi) = \int_0^\infty e^{-2i\pi\xi.x} f(x)dx$. Since f has a bounded support, $\hat{f}$ is infinitely differentiable. We assume that the following inversion formula holds

$$\text{a.e. and at } 0, \quad f(x) = \int_{\mathbb{R}^d} \hat{f}(\xi)e^{2i\pi\xi.x}d\xi. \tag{4}$$

3

Note that if f is hermitian ($f(-x) = \overline{f}(x)$) then $\hat{f}(\xi) \in \mathbb{R}$ and for $\epsilon > 0$, $\int_{\mathbb{R}^d} \mathbb{1}(|\hat{f}(\xi)| \geq \epsilon)d\xi$ is finite. Hence by the change of variable formula, there exists a function ψ such that for all continuous functions h with $0 \notin supp(h)$:

$$\int_{\mathbb{R}} h(t)\psi(t)dt = \int_{\mathbb{R}^d} h(\hat{f}(x))dx.$$

ψ is the level sets function of $\hat{f}$, if ℓ denotes the d-dimensional Lebesgue measure, for all $t > 0$, $\psi(t) = \lim_{\epsilon \to 0} \ell(\{x : |\hat{f}(x) - t| \leq \epsilon\})/\epsilon$. If $d = 1$ and $\hat{f}$ is a diffeomorphism from $\mathbb{R}$ to K then ψ has support on K and is equal to $\psi(t) = (\hat{f}^{-1})'(t)$.

Theorem 3 *If f is hermitian and (4) holds true, then as γ goes to infinity, for all analytic functions $h(t) = \sum_{m \in \mathbb{N}} h_m t^m$ with $h_0 = 0$ and $\sum_{m \in \mathbb{N}} |h_m| t^m$ finite for all t:*

$$\int_{\mathbb{R}} h(t)\nu_\gamma(dt) \sim \int_{\mathbb{R}} h(t)\gamma^{-2}\psi(\frac{t}{\gamma})dt = \int_{\mathbb{R}} \gamma^{-1}h(\gamma t)\psi(t)dt.$$

This result states that the measure $\nu_\gamma(dt)$ is in a weak sense equivalent (not in the measure theory sense) to the measure $\gamma^{-2}\psi(t/\gamma)dt$ in the high density asymptotic. The measure $\psi(t)dt$ is the continuous analog of the counting measure μ in Theorem 1. As an example, if $d = 1$ and $f(x) = \mathbb{1}(0 \leq |x| \leq r)$ then $\hat{f}(\xi) = r\text{sinc}(2\pi\xi r)$ and $\psi(t)$ is plotted in Figure 1.
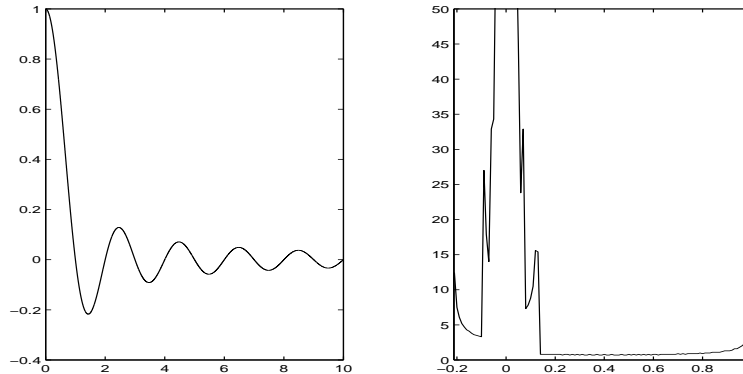

Figure 1: Left: $\hat{f}(\xi)$ for $f(x) = \mathbb{1}(0 \leq |x| \leq r)$. Right: the level set function ψ.

Remark. Let $\gamma_n = n\delta_n^d$, if γ_n tends to infinity and δ_n goes to 0, with the material of this note, we may also prove the convergence of $\delta_n^d \sum_{i=1}^n \delta_{\lambda_i'(n)/\gamma_n}$ to the measure $\psi(t)dt$ on all continuous function h with compact support and $0 \notin supp(h)$.

The spectral radius of the matrix B_n is not computed explicitly in this paper. However, the following upper bound is available:

Proposition 4 *If $d \geq 2$ and $Po(\gamma)$ denotes a random variable with Poisson distribution of intensity γ, then with a probability tending to 1 as n goes to infinity,*

$$\max_{1 \leq i \leq n} |\lambda_i'(n)| \leq j(n) \sup_{x \in \Omega} |f(x)|,$$

where $j(n)$ is solution of: $n\mathbb{P}(Po(\gamma) \geq j(n)+1) \leq 1 < n\mathbb{P}(Po(\gamma) \geq j(n))$.

For n large enough, using the inequality $\mathbb{P}(Po(\gamma) \geq k) \leq \exp(-\frac{k}{2}\ln(\frac{k}{\gamma}))$, for $k \geq e^2\gamma$, we deduce that $j(n) \leq 3\ln n/\ln\ln n$ for n large enough.

The remainder of this paper is organized as follows. In Section 2, we prove Theorem 1, In Section 3, we prove Theorems 2, 3 and Proposition 4. Finally, in Section 4, we state some simple results on the eigenvectors of A and on the correlation of the eigenvalues.

By convention C will denote a constant which does not depend on n. Its exact value may change throughout the paper. Also we define: $\|F\|_\infty = \sup_{x \in \mathbb{R}^d} |F(x)|$ and $B(x,r)$ will denote the open ball of radius r and center x on the torus $\mathbb{T}^d$.

2 Proof of Theorem 1

The proof of Theorem 1 relies on the classical Wigner's method [18] to compute the empirical mean distribution measure of eigenvalues. We will compute for all $m \in \mathbb{N}$:

$$\mathbb{E}\mathrm{tr}A_n^m = \frac{1}{n^m}\mathbb{E}\mathrm{tr}A^m = \frac{1}{n^m}\sum_{i=1}^n \lambda_i^m = \mu_n(P_m).$$

We will then use a Talagrand's concentration inequality to prove that $\mathrm{tr}A_n^m$ is not far from its mean and conclude. About the rate of convergence of μ_n to μ, we will state (in the forthcoming Lemma 6) that, if $P_m(t) = t^m$, $m \geq 1$,

$$\lim_n n\left(\mathbb{E}\mu_n(P_m) - \mu(P_m)\right) = \sum_{q=1}^{m-1} q\mu(P_q)\mu(P_{m-q}) - \frac{m(m-1)}{2}\mu(P_m). \tag{5}$$

We begin with a technical lemma.

Lemma 5 *For $0 \leq p \leq m$, let $\Sigma_{m,p}$ be the set of surjective mappings from $\{1,\cdots,m\}$ to $\{1,\cdots,p\}$. We have:*

$$\mathbb{E}trA^m = \sum_{p=1}^m \binom{n}{p} \sum_{\phi \in \Sigma_{m,p}} \int_{\Omega^p} \prod_{j=1}^m F(x_{\phi(j)} - x_{\phi(j+1)})dx_1\cdots dx_p, \tag{6}$$

with $\phi(m+1) = \phi(1)$ and with the convention that $\binom{n}{p} = 0$ for $p > n$.

Proof. By definition:

$$\mathrm{tr}A^m = \sum_{i_1,\cdots i_m} \prod_{j=1}^m F(X_{i_j} - X_{i_{j+1}}), \tag{7}$$

with $i_{m+1} = i_1$ and the sum is over all n-tuples of integers $\mathbf{i} = (i_1,\cdots,i_m)$ in $\{1,n\}^m$. Let $p(\mathbf{i})$ be the set of distinct indices in $\mathbf{i}$. We can define a surjective mapping $\phi_\mathbf{i}$ in $\Sigma_{m,p(\mathbf{i})}$ such that $i_j = i_{\phi_\mathbf{i}(j)}$. Taking the expectation in Equation (7), we get

$$\mathbb{E}\mathrm{tr}A^m = \sum_{\mathbf{i}=(i_1,\cdots i_m)} \int_{\Omega^m} \prod_{j=1}^m F(x_{\phi_\mathbf{i}(j)} - x_{\phi_\mathbf{i}(j+1)})dx_1\cdots dx_{p(\mathbf{i})},$$

We then reorder the terms. We consider the equivalence relation in $\Sigma_{m,p}$, $\phi \sim \phi'$ if there exists a permutation σ of $\{1,\cdots,p\}$ such that $\sigma \circ \phi = \phi'$. The value of $\int_{\Omega^m} \prod_{j=1}^m F(x_{\phi(j)} - x_{\phi(j+1)})dx_1\cdots dx_p$ is constant on each equivalence class. Let $\phi \in \Sigma_{m,p}$, the numbers of indices $\mathbf{i}$ such that $\phi_\mathbf{i} \sim \phi$ is equal to $n!/(n-p)!$ (if $n \geq p$ and 0 otherwise). Since there are $p!$ surjective mappings in the class of equivalence of ϕ, we deduce Equation (6). $\square$

Lemma 6 *For each m,*

$$\mathbb{E}\mu_n(P_m) = \mu(P_m) + \frac{1}{n}\left(\sum_{q=1}^{m-1} q\mu(P_q)\mu(P_{m-q}) - \frac{m(m-1)}{2}\mu(P_m)\right) + o\left(\frac{1}{n}\right).$$

Proof. We apply Lemma 5 and identify the coefficients in n^m and n^{m-1} in Equation (6). We first consider the term in n^m, such a term comes from $p = m$:

$$\frac{n!}{(n-m)!}\int_{\Omega^m}\prod_{j=1}^m F(x_j - x_{j+1})dx_1\cdots dx_m,$$

By induction, we easily obtain that

$$\int_{\Omega^m}\prod_{j=1}^m F(x_j - x_{j+1})dx_1\cdots dx_m = \int_\Omega F^{*m}(0)dx_1 = F^{*m}(0),$$

where $*$ denotes the convolution operator: $F*G(y) = \int_\Omega F(y-x)G(x)dy$ and F^{*m} is $F*F\cdots*F$ (m times). We recall the two properties: $\widehat{F*G}(k) = \hat{F}(k)\hat{G}(k)$ and $F(0) = \sum_{k\in\mathbb{Z}^d}\hat{F}(k)$, in order to get:

$$\int_{\Omega^m}\prod_{j=1}^m F(x_j - x_{j+1})dx_1\cdots dx_m = \sum_{k\in\mathbb{Z}^d}\hat{F}(k)^m = \mu(P_m).$$

We thus deduce that:

$$\lim_n \mathbb{E}\mu_n(P_m) = \mu(P_m).$$

It remains to identify the terms in n^{m-1} in Equation (6). This term comes from two contributions $p = m$ and $p = m-1$. Since $n!/(n-m)! = n^m - n^{m-1}\sum_{i=0}^{m-1} i + o(n^{m-1})$, the term in n^{m-1} in $p = m$ is equal to:

$$-\frac{m(m-1)}{2}\mu(P_m). \tag{8}$$

The leading term for $p = m-1$ is

$$\frac{n!}{(n-m+1)!(m-1)!}\sum_{\phi\in\Sigma_{m,m-1}}\int_{\Omega^{m-1}}\prod_{j=1}^m F(x_{\phi(j)} - x_{\phi(j+1)})dx_1\cdots dx_{m-1}, \tag{9}$$

Now if $\phi \in \Sigma_{m,m-1}$, $\phi^{-1}(i)$ is not reduced to a single point for a unique index i_ϕ. Since the value of $\int_{\Omega^{m-1}}\prod_{j=1}^m F(x_{\phi(j)} - x_{\phi(j+1)})dx_1\cdots dx_{m-1}$ is invariant under permutations of the indices, without loss of generality, we may assume that $i_\phi = 1$, $\phi^{-1}(1) = \{1, q+1\}$ with $q \in \{1,\cdots,m-1\}$ and $\phi(j) = j$ if $j \le q$ and $\phi(j) = j+1$ if $j > q+1$. For such ϕ, integrating over $x_2,\cdots,x_q,x_{q+1},\cdots,x_{m-1}$,

$$\int_{\Omega^m}\prod_{j=1}^m F(x_{\phi(j)} - x_{\phi(j+1)})dx_1\cdots dx_{m-1} = \int_\Omega F^{*(q)}(0)F^{*(m-q)}(0)dx_1$$
$$= \mu(P_q)\mu(P_{m-q}).$$

Finally, for each q, there are $(m-q)\times(m-1)!$ surjective mappings such that, up to a permutation of the indices, $\phi^{-1}(1) = \{1, q+1\}$ and $\phi(j) = j$ if $j \le q$ and $\phi(j) = j+1$ if $j > q+1$. Indeed for

such ϕ, there are $(m-q)$ possible pairs (i_1, i_1+q), $1 \le i_1 \le m-q$ such that $\phi(i_1) = \phi(i_1+q)$. Therefore Equation (9) can be written as:

$$\frac{n!(m-1)!}{(n-m+1)!(m-1)!} \sum_{q=1}^{m-1} (m-q)\mu(P_q)\mu(P_{m-q}) = n^{m-1} \sum_{q=1}^{m-1} q\mu(P_q)\mu(P_{m-q}) + o(n^{m-1}). \quad (10)$$

Adding this last term with the term (8), we get the stated formula. $\qquad\square$

We may now prove Theorem 1.

Proof of Theorem 1.

We fix n and for each $m \ge 1$, we define the functional:

$$Q_m(\mathcal{X}_n) = \frac{1}{n^{m-1}}\mathrm{tr}A^m = n\mu_n(P_m).$$

If $\mathbf{x}, \mathbf{y} \in \Omega^n$, let $d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \mathbb{1}(x_i \ne y_i)$ denote the Hamming distance. The functional Q_m is Lipschitz for the Hamming distance d. Indeed, define $\mathbf{x^l} = (x_j^l)_{1 \le j \le n}$ by $x_j^l = x_j$ for $j \ne l$ and $x_l^l \ne x_l$, we have:

$$\left| Q_m(\mathbf{x}) - Q_m(\mathbf{x^l}) \right| = \frac{1}{n^{m-1}} \left| \sum_{i_1,\cdots,i_m} \prod_{j=1}^m F(x_{i_j} - x_{i_{j+1}}) - \prod_{j=1}^m F(x_{i_j}^l - x_{i_{j+1}}^l) \right|$$

$$\le 2m\|f\|_\infty^m,$$

indeed $|\prod_{j=1}^m F(x_{i_j} - x_{i_{j+1}}) - \prod_{j=1}^m F(x_{i_j}^l - x_{i_{j+1}}^l)|$ is at most $2\|f\|_\infty^m$ and it is non zero only if there exists a index i_j such that $i_j = l$. It follows easily that Q_m is $2m\|f\|_\infty^m$-Lipschitz for the Hamming distance d.

Let M_m denote the median of Q_m. We may apply a Talagrand's Concentration Inequality (see for example Proposition 2.1 of Talagrand [16]),

$$\mathbb{P}(|Q_m - M_m| > t) \le 4\exp(-\frac{t^2}{4m^2\|f\|_\infty^{2m}n}),$$

integrating over all t we deduce:

$$|n\mathbb{E}\mu(P_m) - M_m| \le \mathbb{E}|\mu(P_m) - M_m| \le C_m\sqrt{n},$$

for some constant C_m and it follows, that for all $s > C_m/\sqrt{n}$:

$$\mathbb{P}(|\mu_n(P_m) - \mathbb{E}\mu(P_m)| > s) \le 4\exp(-n\frac{(s - C_m/\sqrt{n})^2}{4m^2\|f\|_\infty^{2m}}),$$

Using the Borel Cantelli Lemma and Lemma 6, a.s. $\lim_n \mu_n(P_m) = \mu(P_m)$. $\qquad\square$

3 Limit Spectral Measure of Scaled ERM

3.1 Proof of Theorem 2

The study of the first model was simplified by the absence of boundary effects with Ω. So in order to prove Theorem 2, we will first discard them in the second model. We define F_δ as the 1-periodic extension of f_δ: for all $x \in \mathbb{R}^d$, there exists a unique couple (y, u) such that $x = y+u$, with $u \in \mathbb{Z}^d$ and $y \in \Omega$, and we set $F_\delta(x) = f_\delta(y)$.

We now introduce a matrix and its spectral empirical measure:

$$\tilde{B}_n = (F_{\delta_n}(X_i - X_j))_{1 \leq i \leq j \leq n} \quad \text{and} \quad \tilde{\nu}_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\tilde{\lambda}_i(n)},$$

where $(\tilde{\lambda}_1(n), \cdots, \tilde{\lambda}_n(n))$ is the spectrum of $\tilde{B}_n$. The next lemma states that the limiting spectral measures of $\tilde{\nu}_n$ and ν_n are equal.

Lemma 7 *For the topology of the weak convergence of (signed) measures, a.s. $\nu_n - \tilde{\nu}_n$ converges as n goes to infinity to the null measure.*

Proof. It is sufficient to prove that for all $m \geq 1$, a.s. $\lim_n \nu_n(P_m) - \tilde{\nu}_n(P_m) = 0$. To this end, we notice that if $x, y \in \Omega$, $f_\delta(x - y) = F_\delta(x - y)$ unless $x \in \Omega \backslash (1 - \delta)\Omega$ and $y \in B(x, \delta)$. We write:

$$\begin{aligned}
\left| \nu_n(P_m) - \tilde{\nu}_n(P_m) \right| &\leq \frac{1}{n}\sum_{i_1, \cdots, i_m} \left| \prod_{j=1}^{m} f_{\delta_n}^m(X_{i_j} - X_{i_{j+1}}) - \prod_{j=1}^{m} F_{\delta_n}(X_{i_j} - X_{i_{j+1}}) \right| \\
&\leq \frac{1}{n}\sum_{i_1, \cdots, i_m} 2\|f\|_\infty^m \mathbb{1}(X_{i_1} \in \Omega \backslash (1 - \delta_n)\Omega) \prod_{j=2}^{m} \mathbb{1}(X_{i_j} \in B(X_{i_1}, m\delta_n)) \\
&\leq \frac{2}{n}\|f\|_\infty^m N_n(\Omega \backslash (1 - m\delta_n)\Omega)^m,
\end{aligned}$$

where N_n is the counting measure $N_n(\cdot) = \#\{i \in \{1, \cdots, n\} : X_i \in \cdot\}$. Note that $\mathbb{P}(X_1 \in \Omega \backslash (1 - m\delta_n)\Omega) \leq C\delta_n$. By the strong law of large numbers, it follows easily that that $N_n(\Omega \backslash (1 - m\delta_n)\Omega)/n$ converges almost surely to 0. $\square$

By Lemma 7, we may focus on $\tilde{B}_n$ and $\tilde{\nu}_n$. In order to keep the notations as light as possible we drop the "~" in $\tilde{B}_n$ and $\tilde{\nu}_n$.

We first prove that,

$$\nu_n \text{ converges in probability to a measure } \nu_\gamma \text{ for the weak convergence.} \tag{11}$$

By Lemma 5, if $m \geq 1$,

$$\mathbb{E}\nu_n(P_m) = \frac{1}{n}\sum_{p=1}^{m}\binom{n}{p}\sum_{\phi \in \Sigma_{m,p}}\int_{\Omega^p}\prod_{j=1}^{m} F_{\delta_n}(x_{\phi(j)} - x_{\phi(j+1)})dx_1 \cdots dx_p. \tag{12}$$

We begin with an elementary lemma.

Lemma 8 *If $\phi \in \Sigma_{m,p}$, $p > 1$ the value of*

$$\int_{\Omega^{p-1}}\prod_{j=1}^{m} F(x_{\phi(j)} - x_{\phi(j+1)})dx_2 \cdots dx_p$$

does not depend on x_1.

Proof. We consider the change of variable, for $j > 1$, $x_j' = x_j - x_1$. The Jacobian of this change of variable is 1. If we set $x_1' = 0$, we obtain $\int_{\Omega^{p-1}}\prod_{j=1}^{m} F(x_{\phi(j)} - x_{\phi(j+1)})dx_2 \cdots dx_p = \int_{\Omega^{p-1}}\prod_{j=1}^{m} F(x'_{\phi(j)} - x'_{\phi(j+1)})dx_2' \cdots dx_p'$. $\square$

8

Assume $m \geq 2$, by Lemma 8, we have:

$$
\begin{aligned}
\mathbb{E}\nu_n(P_m) &= F_{\delta_n}(0)^m + \frac{1}{n}\sum_{p=2}^{m}\binom{n}{p}\sum_{\phi\in\Sigma_{m,p}}\int_{\Omega^{p-1}}\prod_{j=1}^{m}F_{\delta_n}(x_{\phi(j)} - x_{\phi(j+1)})dx_2\cdots dx_p \\
&= f(0)^m + \frac{1}{n}\sum_{p=2}^{m}\binom{n}{p}\sum_{\phi\in\Sigma_{m,p}}\Delta(\phi) + \int_{\Omega^{p-1}}\prod_{j=1}^{m}f_{\delta_n}(x_{\phi(j)} - x_{\phi(j+1)})dx_2\cdots dx_p (13)
\end{aligned}
$$

where $\Delta(\phi) = \int_{\Omega^{p-1}}\prod_{j=1}^{m}F_{\delta_n}(x_{\phi(j)} - x_{\phi(j+1)}) - \prod_{j=1}^{m}f_{\delta_n}(x_{\phi(j)} - x_{\phi(j+1)})dx_2\cdots dx_p$. Since the support of f_δ is included in $\delta\Omega$, if $f_\delta(x_{\phi(j)} - x_{\phi(j+1)}) \neq F_\delta(x_{\phi(j)} - x_{\phi(j+1)})$ then $x_{\phi(j)}, x_{\phi(j+1)} \in \Omega\setminus(1-\delta)\Omega$. Moreover notice that if $\prod_{j=1}^{m}F(x_{\phi(j)} - x_{\phi(j+1)}) \neq 0$ then $x_2,\cdots x_p \in B(x_1, (m-1)\delta)$. By Lemma 8, from now on, we can assume without loss of generality:

$$
x_1 = 0,
$$

and then $\Delta(\phi) = 0$ for $\delta < 1/(2m)$.

Considering the change of variable $y_i = x_i/\delta_n$ in the integrands of Equation (13), we obtain, for $\delta < 1/(2m)$, with $y_1 = 0$,

$$
\mathbb{E}\nu_n(P_m) = f(0)^m + \sum_{p=2}^{m}\frac{\delta_n^{d(p-1)}}{n}\binom{n}{p}\sum_{\phi\in\Sigma_{m,p}}\int_{(\delta_n^{-1}\Omega)^{p-1}}\prod_{j=1}^{m}f(y_{\phi(j)} - y_{\phi(j+1)})dy_2\cdots dy_p. \quad (14)
$$

Finally, since $\binom{n}{p} \sim n^p/p!$ as n goes to infinity, we deduce that, for $m \geq 2$,

$$
\lim_{n\to\infty}\mathbb{E}\nu_n(P_m) = f(0)^m + \sum_{p=2}^{m}\frac{\gamma^{p-1}}{p!}\sum_{\phi\in\Sigma_{m,p}}\int_{(\mathbb{R}^d)^{p-1}}\prod_{j=1}^{m}f(y_{\phi(j)} - y_{\phi(j+1)})dy_2\cdots dy_p. \quad (15)
$$

(For $m \leq 1$, we have $\nu_n(P_0) = 1$ and $\nu_n(P_1) = f(0)$).

We check easily that the right hand side of Equation (15) is bounded by $(Cm)^m$ for some constant C not depending on m. Therefore, by Carleman's Condition, there is exists a unique measure ν_γ such that $\lim_{n\to\infty}\mathbb{E}\nu_n(P_m) = \nu_\gamma(P_m)$. In particular, the sequence $(\nu_n)_{n\in\mathbb{N}}$ is tight and we have proved (11).

The continuity of $\gamma \mapsto \nu_\gamma$ follows from the comtinuity of $\gamma \mapsto \nu_\gamma(P_m)$. Indeed, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence converging to $\gamma < \infty$. Since $\sup_n \nu_{\gamma_n}(P_2) < \infty$, the sequence $(\nu_{\gamma_n})_{n\in\mathbb{N}}$ is tight. Hence for all $\epsilon > 0$, there exists a compact set K such that for all n $\nu_{\gamma_n}(K^c) \leq \epsilon$. Now, let h be a continuous function with compact support, we need to prove that $\lim_{n\to\infty}\nu_{\gamma_n}(h) = \nu_\gamma(h)$. Fix ϵ, there exists a polynomial P such that $\sup_{x\in K}|h(x)-P(x)| \leq \epsilon$, we deduce that $|\nu\gamma_n(h) - \nu_\gamma(h)| \leq |\nu_{\gamma_n}(h)-\nu_{\gamma_n}(P)|+|\nu_{\gamma_n}(P)-\nu_\gamma(P)|+|\nu_\gamma(P)-\nu_\gamma(h)| \leq 2\epsilon(1+\|h\|_\infty)+|\nu_{\gamma_n}(P)-\nu_\gamma(P)|$. Letting n tends to infinity, since ϵ is arbitrary small and $\gamma \mapsto \nu_\gamma(P)$ is continuous, we obtain: $\lim_{n\to\infty}\nu_{\gamma_n}(h) = \nu_\gamma(h)$.

It remains to prove the almost sure convergence of ν_n. We will prove that for each $m \geq 1$, there exists a constant C and

$$
\mathbb{E}\left(\operatorname{tr}B_n^m - \mathbb{E}\operatorname{tr}B_n^m\right)^4 \leq Cn^2. \quad (16)
$$

This last equation implies $\mathbb{E}\left(\nu_n(P_m) - \mathbb{E}\nu_n(P_m)\right)^4 \leq C/n^2$ and by Borel Cantelli Lemma, we deduce that $\nu_n(P_m)$ converges almost surely toward $\nu_\gamma(P_m)$.

It remains to prove Inequality (16). A circuit in $\{1, \cdots, n\}$ of length m is a mapping $\pi : \mathbb{Z} \to \{1, \cdots n\}$ such that for all integer r, $\pi(m + r) = \pi(r)$. Following Bryc, Dembo and Jiang [3], we introduce the new notation:

$$F_\pi = \prod_{i=1}^{m} F_{\delta_n}(X_{\pi(i)} - X_{\pi(i+1)}).$$

We then write:

$$\mathbb{E}\left(\operatorname{tr}B_n^m - \mathbb{E}\operatorname{tr}B_n^m\right)^4 = \mathbb{E}\left(\sum_\pi F_\pi - \mathbb{E}F_\pi\right)^4 = \sum_{\pi_1, \cdots, \pi_4} \mathbb{E}\left[\prod_{l=1}^{4} F_{\pi_l} - \mathbb{E}F_{\pi_l}\right], \qquad (17)$$

where the sums are over all circuits in $\{1, \cdots, n\}$ of length m.

Notice that $\mathbb{E}\left[\prod_{l=1}^{4} F_{\pi_l} - \mathbb{E}F_{\pi_l}\right] = 0$ if there exists a circuit π_k, $1 \le k \le 4$ such that the image of π_k has an empty intersection with the union of the images of π_l, $l \ne k$. Indeed, due to the independence of the variables $(X_i)_{1 \le i \le n}$, $F_{\pi_k} - \mathbb{E}F_{\pi_k}$ is then independent of $\prod_{l \ne k} F_{\pi_l} - \mathbb{E}F_{\pi_l}$.

Two circuits π_1 and π_2 in $\{1, \cdots, n\}$ of length m_1 and m_2 with a non empty intersection of their images may be concatenated into a circuit in $\{1, \cdots, n\}$ of length $m_1 + m_2$ as follows. Assume that $\pi_1(i_0) = \pi_2(j_0)$, we define the circuit $\pi_{1,2}$ of length $m_1 + m_2$ by for $i \in \{1, \cdots, m_1 + m_2\}$

$$\pi_1 . \pi_2(i) = \begin{cases} \pi_1(i) & \text{if} \quad 1 \le i \le i_0 \\ \pi_2(i - i_0 + j_0) & \text{if} \quad i_0 + 1 \le i \le i_0 + m_2 \\ \pi_1(i - m - i_0) & \text{if} \quad i_0 + m + 1 \le i \le m_1 + m_2 \end{cases}$$

We have:

$$F_{\pi_1} F_{\pi_2} = F_{\pi_1 \cdot \pi_2}.$$

Using the same reasoning as for Equation (14), we get

$$\mathbb{E}F_{\pi_1} F_{\pi_2} = \delta_n^{d(q-1)} \int_{(\delta_n^{-1}\Omega)^{q-1}} \prod_{j=1}^{2m} f(y_{\pi_1 . \pi_2(j)} - y_{\pi_1 . \pi_2(j+1)}) dy_{i_2} \cdots dy_{i_q},$$

where is $q = q(\pi_1, \pi_2)$ is the cardinal of the union of the images of π_1 and π_2 and $(y_{i_1}, \cdots, y_{i_q})$ is the image of $\pi_1 \cdot \pi_2$ and $y_{i_1} = 0$.

If $N(\pi_1, \pi_2)$ is the cardinal of the intersection of the images of π_1 and π_2, if $N(\pi_1, \pi_2) \ge 1$, we obtain

$$\mathbb{E}|F_{\pi_1} F_{\pi_2}| \le Cn^{-q(\pi_1, \pi_2)+1}. \qquad (18)$$

Otherwise, $N(\pi_1, \pi_2) = 0$, if $q(\pi_i)$ is the cardinal of of the image of π_i,

$$\mathbb{E}|F_{\pi_1} F_{\pi_2}| = \mathbb{E}|F_{\pi_1}|\mathbb{E}|F_{\pi_2}| \le Cn^{-q(\pi_1)-q(\pi_2)+2} = Cn^{-q(\pi_1, \pi_2)+2}. \qquad (19)$$

Similarly assume that $N(\pi_1, \pi_2) \ge 1$, $N(\pi_1 . \pi_2, \pi_3) \ge 1$, if $q(\pi_1, \pi_2, \pi_3)$ is the cardinal of the union if the images of π_1, π_2, π_3 then $F_{\pi_1} F_{\pi_2} F_{\pi_3} = F_{(\pi_1 . \pi_2) . \pi_3}$ and we deduce similarly

$$\mathbb{E}|F_{\pi_1} F_{\pi_2} F_{\pi_3}| \le Cn^{-q(\pi_1, \pi_2, \pi_3)+1}.$$

Finally assume that $N(\pi_1, \pi_2) \ge 1$, $N(\pi_1 . \pi_2, \pi_3) \ge 1$, $N(\pi_1 . \pi_2 . \pi_3, \pi_4) \ge 1$, if $q(\pi_1, \pi_2, \pi_3)$ is the cardinal of the union if the images of $\pi_1, \pi_2, \pi_3, \pi_4$, we obtain:

$$\mathbb{E}|F_{\pi_1} F_{\pi_2} F_{\pi_3} F_{\pi_4}| \le Cn^{-q(\pi_1, \pi_2, \pi_3, \pi_4)+1}. \qquad (20)$$

By (17), it remains to decompose:

$$4! \sum_{(\pi_1,\cdots,\pi_4)\in S\cup S'} \mathbb{E}\left[\prod_{l=1}^{4} F_{\pi_l} - \mathbb{E}F_{\pi_l}\right]$$

where S is the set of quadruples of circuits such that $N(\pi_1,\pi_2) \geq 1$, $N(\pi_1.\pi_2,\pi_3) \geq 1$, $N(\pi_1.\pi_2.\pi_3,\pi_4) \geq 1$ and S' is the set of quadruples of circuits such that $N(\pi_1,\pi_2) \geq 1$ and $N(\pi_3,\pi_4) \geq 1$ and otherwise for $i < j$, $N(\pi_i,\pi_j) = 0$.

The decomposition of the $\mathbb{E}\left[\prod_{l=1}^{4} F_{\pi_l} - \mathbb{E}F_{\pi_l}\right]$ gives rise to four types of terms:

1. $\sum_{(\pi_1,\cdots,\pi_4)\in S\cup S'} \prod_{l=1}^{4} \mathbb{E}F_{\pi_l}$,

2. $\sum_{(\pi_1,\cdots,\pi_4)\in S\cup S'} \mathbb{E}\prod_{l=1}^{4} F_{\pi_l}$,

3. $\sum_{(\pi_1,\cdots,\pi_4)\in S\cup S'} \mathbb{E}F_{\pi_{l_1}} F_{\pi_{l_2}} \mathbb{E}F_{\pi_{l_3}} F_{\pi_{l_4}}$,

4. $\sum_{(\pi_1,\cdots,\pi_4)\in S\cup S'} \mathbb{E}F_{\pi_{l_1}} \mathbb{E}F_{\pi_{l_2}} F_{\pi_{l_3}} F_{\pi_{l_4}}$,

where (l_1, l_2, l_3, l_4) is a permutation of $(1, 2, 3, 4)$. We will apply successively the same method to bound these terms.

We begin with the terms of type 1, we have: $\prod_{l=1}^{4} \mathbb{E}F_{\pi_l} \leq Cn^{-\sum_{l=1}^{4} q(\pi_l)+4}$. Since $(\pi_1, \cdots, \pi_4) \in S \cup S'$, $q(\pi_1, \pi_2, \pi_3, \pi_4) \leq \sum_{l=1}^{4} q(\pi_l) - 2$, hence:

$$\prod_{l=1}^{4} \mathbb{E}F_{\pi_l} \leq Cn^{-q(\pi_1,\pi_2,\pi_3,\pi_4)+2}.$$

There are at most Cn^q quadruples of circuits such that $q(\pi_1, \pi_2, \pi_3, \pi_4) = q$, therefore the terms of type 1 may be bounded as by

$$\sum_{(\pi_1,\cdots,\pi_4)\in S\cup S'} \prod_{l=1}^{4} \mathbb{E}|F_{\pi_l}| \leq Cn^2.$$

We now deal with the terms of type 2. By (20), if $(\pi_1, \cdots, \pi_4) \in S$, $\mathbb{E}\prod_{l=1}^{4}|F_{\pi_l}| \leq Cn^{-q(\pi_1,\pi_2,\pi_3,\pi_4)+1}$ otherwise $(\pi_1, \cdots, \pi_4) \in S'$ and, by (19), $\mathbb{E}\prod_{l=1}^{4}|F_{\pi_l}| \leq Cn^{-q(\pi_1,\pi_2,\pi_3,\pi_4)+2}$. There are at most Cn^q quadruples of mappings such that $q(\pi_1, \pi_2, \pi_3, \pi_4) = q$. Hence

$$\sum_{(\pi_1,\cdots,\pi_4)\in S\cup S'} \mathbb{E}\prod_{l=1}^{4}|F_{\pi_l}| \leq Cn^2.$$

We turn to the terms of type 3: $\sum_{(\pi_1,\cdots,\pi_4)S\cup S'} \mathbb{E}F_{\pi_{l_1}} F_{\pi_{l_2}} \mathbb{E}F_{\pi_{l_3}} F_{\pi_{l_4}}$. Assume first that the quadruple $(\pi_1, \cdots, \pi_4) \in S'$. If $l_1 = 1$, $l_2 = 3$, $l_3 = 2$, $l_4 = 4$, then $\mathbb{E}F_{\pi_{l_1}} F_{\pi_{l_2}} \mathbb{E}F_{\pi_{l_3}} F_{\pi_{l_4}} = \prod_{l=1}^{4} \mathbb{E}F_{\pi_l}$ and we obtain the same bound that the terms of type 1. The other cases reduce to the case $l_1 = 1$, $l_2 = 2$, $l_3 = 3$, $l_4 = 4$ and by (18), $\mathbb{E}F_{\pi_1} F_{\pi_2} \leq Cn^{-q(\pi_1,\pi_2)+1}$. There are at most $Cn^{q+q'}$ quadruples such that $q(\pi_1,\pi_2) = q$ and $q(\pi_3,\pi_4) = q'$. We deduce that $\sum_{(\pi_1,\cdots,\pi_4)\in S'} \mathbb{E}F_{\pi_1} F_{\pi_2} \mathbb{E}F_{\pi_3} F_{\pi_4} \leq Cn^2$.

Assume now that that $(\pi_1, \cdots, \pi_4) \in S$. We have:

$$\mathbb{E}F_{\pi_{l_1}} F_{\pi_{l_2}} \mathbb{E}F_{\pi_{l_3}} F_{\pi_{l_4}} \leq Cn^{-q(\pi_{l_1},\pi_{l_2})-q(\pi_{l_3},\pi_{l_4})+2+1(N(\pi_{l_1},\pi_{l_2})=0)+1(N(\pi_{l_3},\pi_{l_4})=0)}.$$

If $N(\pi_{l_1}, \pi_{l_2}) = 0$, then $N(\pi_{l_3}, \pi_{l_4}) \geq 1$ and there are at most $Cn^{q+q'-2}$ quadruples such that $q(\pi_{l_1}, \pi_{l_2}) = q$ and $q(\pi_{l_3}, \pi_{l_4}) = q'$. Indeed, since $(\pi_1, \cdots, \pi_4) \in S$, the cardinal of the intersection of the images of (π_{l_1}, π_{l_2}) and (π_{l_3}, π_{l_4}) is at least 2. The other cases reduce to the case, $N(\pi_{l_1}, \pi_{l_2}) \geq 1$ and $N(\pi_{l_3}, \pi_{l_4}) \geq 1$, for such cases, we notice that there are at most $Cn^{q+q'}$ quadruples such that $q(\pi_{l_1}, \pi_{l_2}) = q$ and $q(\pi_{l_3}, \pi_{l_4}) = q'$. In all cases, we conclude that: $\sum_{(\pi_1, \cdots, \pi_4) \in S} \mathbb{E} F_{\pi_1} F_{\pi_2} \mathbb{E} F_{\pi_3} F_{\pi_4} \leq Cn^2$. Hence,

$$\sum_{(\pi_1, \cdots, \pi_4) \in S \cup S'} \mathbb{E} F_{\pi_{l_1}} F_{\pi_{l_2}} \mathbb{E} F_{\pi_{l_3}} F_{\pi_{l_4}} \leq Cn^2.$$

It remains to treat the terms of type 4. Assume that $(\pi_1, \cdots, \pi_4) \in S \cup S'$, we have:

$$\mathbb{E} F_{\pi_{l_1}} \mathbb{E} F_{\pi_{l_2}} \mathbb{E} F_{\pi_{l_3}} F_{\pi_{l_4}} \leq Cn^{-q(\pi_{l_1})+1-q(\pi_{l_2},\pi_{l_3},\pi_{l_4})+\epsilon(\pi)}, \tag{21}$$

where $\epsilon(\pi) \in \{1,2\}$, $\epsilon(\pi) = 2$ if there exists $j \in \{2,3,4\}$ such that $N(\pi_{l_j}, \pi_{l_k}) = 0$ for all $k \in \{2,3,4\}\backslash\{j\}$, otherwise, $\epsilon(\pi) = 1$.

If $\epsilon(\pi) = 1$ then since there are at most $Cn^{q+q'}$ quadruples such that $q(\pi_{l_1}) = q$ and $q(\pi_{l_2}, \pi_{l_3}, \pi_{l_4}) = q'$, we deduce that $\sum_{(\pi_1, \cdots, \pi_4) \in S \cup S'} \mathbf{1}(\epsilon(\pi) = 1) \mathbb{E} F_{\pi_{l_1}} F_{\pi_{l_2}} \mathbb{E} F_{\pi_{l_3}} F_{\pi_{l_4}} \leq Cn^2$.

If $\epsilon(\pi) = 2$, then, without loss of generality, we may assume $N(\pi_{l_2}, \pi_{l_k}) = 0$ for $k \in \{3,4\}$. It implies that $q(\pi_{l_2}, \pi_{l_3}, \pi_{l_4}) = q(\pi_{l_2}) + q(\pi_{l_3}, \pi_{l_4})$. Since $(\pi_1, \cdots, \pi_4) \in S \cup S'$, $N(\pi_{l_2}, \pi_{l_1}) \geq 1$, therefore $q(\pi_{l_1}) + q(\pi_{l_2}) \geq q(\pi_{l_1}, \pi_{l_1}) + 1$ and by Inequality (21),

$$\mathbb{E} F_{\pi_{l_1}} \mathbb{E} F_{\pi_{l_2}} \mathbb{E} F_{\pi_{l_3}} F_{\pi_{l_4}} \leq Cn^{-q(\pi_{l_1})+1-q(\pi_{l_2},\pi_{l_3},\pi_{l_4})+2} \leq Cn^{-q(\pi_{l_1},\pi_{l_2})-q(\pi_{l_3},\pi_{l_4})+2}.$$

Finally, we notice that there are at most $Cn^{q+q'}$ quadruples such that $q(\pi_{l_1}, \pi_{l_2}) = q$ and $q(\pi_{l_3}, \pi_{l_4}) = q'$, it follows that:

$$\sum_{(\pi_1, \cdots, \pi_4) \in S \cup S'} \mathbb{E} F_{\pi_{l_1}} \mathbb{E} F_{\pi_{l_2}} F_{\pi_{l_3}} F_{\pi_{l_4}} \leq Cn^2.$$

Inequality (16) is proved.

3.2 Proof of Theorem 3

By Equation (15), for $m \geq 2$, we have (with $y_1 = 0$):

$$\nu_\gamma(P_m) = f(0)^m + \sum_{p=2}^{m} \frac{\gamma^{p-1}}{p!} \sum_{\phi \in \Sigma_{m,p}} \int_{(\mathbb{R}^d)^{p-1}} \prod_{j=1}^{m} f(y_{\phi(j)} - y_{\phi(j+1)}) dy_2 \cdots dy_p. \tag{22}$$

The leading term in γ is of order γ^{m-1}. Taking $p = m$ in the above expression gives:

$$\nu_\gamma(P_m) \sim \gamma^{m-1} \int_{(\mathbb{R}^d)^{m-1}} \prod_{j=1}^{m} f(y_j - y_{j+1}) dy_2 \cdots dy_m.$$

A direct iteration leads to:

$$\int_{(\mathbb{R}^d)^{m-1}} \prod_{j=1}^{m} f(y_j - y_{j+1}) dy_2 \cdots dy_m = f^{*m}(0),$$

where $f * g(y) = \int_{\mathbb{R}^d} f(x) g(y-x) dx$, $f^{*1}(x) = f(x)$, and for $m \geq 2$, $f^{*m} = f^{*(m-1)} * f$.

12

Hence $\int_{(\mathbb{R}^d)^{m-1}} \prod_{j=1}^m f(y_j - y_{j+1}) dy_2 \cdots dy_m = \int_{\mathbb{R}^d} \hat{f}^m(\xi) d\xi = \int t^m \psi(t) dt$ and for all $m \geq 2$,

$$\nu_\gamma(P_m) \sim \gamma^{m-1} \int t^m \psi(t) dt = \int t^m \gamma^{-2} \psi(\frac{t}{\gamma}) dt.$$

Since $\int t\psi(t) dt = \int \hat{f}(\xi) d\xi = f(0)$, this formula is still valid for $m = 1$. Now, let $h(t) = \sum_{m \geq 1} h_m t^m$ with $\sum_{m \geq 1} |h_m| t^m$ finite for all t, then since, $|\nu_\gamma(P_m)| \leq m\gamma^{m-1} C^m$, using Fubini's Theorem, the conclusion follows. $\qquad\square$

Remark. We can easily identify the next term in the asymptotic of $\nu_\gamma(P_m)$. The second leading term in Equation (22) is of order γ^{m-2}. As in (10), it is equal to

$$I_m = \gamma^{m-2} \sum_{p=1}^{m-1} p \int_\mathbb{R} u^p \psi(u) du \int_\mathbb{R} v^{m-p} \psi(v) dv.$$

Since if $u \neq v$, $\sum_{p=1}^{m-1} pu^p v^{m-p} = uv(u-v)^{-2}((m-1)u^m - mu^{m-1}v + v^m)$, we deduce that:

$$I_m = \gamma^{m-2} \int_{\mathbb{R}^2} uv \frac{(m-1)u^m - mu^{m-1}v + v^m}{(u-v)^2} \psi(u)\psi(v) du dv.$$

3.3 Proof of Proposition 4

Let D_n denote the $n \times n$ matrix with entry i, j equal to: $\mathbb{1}(\|X_i - X_j\| \leq \delta_n)$, if I_n denotes the $n \times n$ identity matrix, $D_n - I_n$ is the adjacency matrix of the random geometric graph $\mathcal{G}(\mathcal{X}_n, \delta_n)$ where there is an edge between $i \neq j$ if $\|X_i - X_j\| \leq \delta_n$. We have component wise:

$$-\|f\|_\infty D_n \leq B_n \leq \|f\|_\infty D_n.$$

Since the spectral radius $\rho(B_n)$ of B_n is upper bounded by $\max_{1 \leq i \leq n} |\sum_{j=1}^n (B_n)_{ij}|$, we deduce that:

$$\rho(B_n) \leq \|f\|_\infty (1 + \Delta_n),$$

where Δ_n is the maximal degree of the graph $\mathcal{G}(\mathcal{X}_n, \delta_n)$. Then, the proposition follows from Theorem 6.6 of Penrose [15]. $\qquad\square$

4 Further properties of the Euclidean Random Matrices

4.1 Eigenvectors of Euclidean Random Matrices

As it is pointed by Mézard, Parisi and Zee [13], if $U_i = (\Phi_{k,n})_i = e^{2i\pi k.X_i}$ we have:

$$(A\Phi_{k,n})_i = \left(\sum_j F(X_i - X_j) e^{-2i\pi k.(X_i - X_j)} \right) (\Phi_{k,n})_i, \qquad (23)$$

In particular, if $F(x) = e^{2i\pi k.x}$, then n is an eigenvalue with $\Phi_{k,n}$ as eigenvector and the rank of A is 1. Note also by the Strong Law of Large Numbers that for all i, a.s.

$$\lim_{n\to\infty} \frac{1}{n} \sum_j F(X_i - X_j) e^{-2i\pi k.(X_i - X_j)} = \hat{F}(k).$$

if $A_n = A/n$, by Equation (23), for all i, a.s.:

$$\lim_{n\to\infty} (A_n \Phi_{k,n})_i = \hat{F}(k)(\Phi_{k,n})_i.$$

This last equation is consistent with Theorem 1: a.s. for n large enough there exists an eigenvalue of A_n close to $\hat{F}(k)$. It is possible to strengthen this last convergence as follows:

Proposition 9 *For $p \geq 1$, let $\|U\|_p = \left(\sum_{i \geq 1} |U_i|^p \right)^{1/p}$ and $\|U\|_\infty = \sup_{i \geq 1} |U_i|$. For all $p \in (2, \infty]$, a.s. for all $k \in \mathbb{Z}^d$,*

$$\lim_{n \to \infty} \|A_n \Phi_{k,n} - \hat{F}(k) \Phi_{k,n}\|_p = 0.$$

Moreover, $\lim_{n \to \infty} \mathbb{E}\|A_n \Phi_{k,n} - \hat{F}(k) \Phi_{k,n}\|_2^2 = \|f\|_2^2 - |\hat{F}(k)|^2 = \sum_{l \neq k} |\hat{F}(l)|^2$.

Proof. To simplify notation, we write $\Phi = \Phi_{k,n}$ and $f_{ij} = F(X_i - X_j)$.

$$
\begin{aligned}
\mathbb{P}\left(\|A_n \Phi_{k,n} - \hat{F}(k)\Phi_{k,n}\|_p > \epsilon \right) &= \mathbb{P}\left(\sum_{i=1}^{n} \left| \frac{1}{n} \sum_{j=1}^{n} f_{ij}\Phi_j - \hat{F}(k)\Phi_i \right|^p > \epsilon^p \right) \\
&\leq n\mathbb{P}\left(\left| \frac{1}{n} \sum_{j=1}^{n} f_{1j}\Phi_j - \hat{F}(k)\Phi_1 \right|^p > \frac{\epsilon^p}{n} \right) \\
&\leq n\mathbb{P}\left(\left| \sum_{j=1}^{n} f_{1j}\Phi_j - n\hat{F}(k)\Phi_1 \right| > \epsilon n^{1-1/p} \right).
\end{aligned}
$$

From Equation (23), $|\sum_{j=1}^{n} f_{1j}\Phi_j - n\hat{F}(k)\Phi_1| = |\sum_{j=1}^{n} F(X_1 - X_j)e^{-2i\pi k.(X_1 - X_j)} - n\hat{F}(k)|$. Hence:

$$
\begin{aligned}
\mathbb{P}\left(\|A_n\Phi_{k,n} - \hat{F}(k)\Phi_{k,n}\|_p > \epsilon \right) &\leq n\mathbb{P}\left(|\sum_{j=1}^{n} F(X_1 - X_j)e^{-2i\pi k.(X_1-X_j)} - n\hat{F}(k)| > \epsilon n^{1-1/p} \right) \\
&\leq n\mathbb{E}\left[\mathbb{P}\left(|\sum_{j=2}^{n} F(X_1 - X_j)e^{-2i\pi k.(X_1-X_j)} - (n-1)\hat{F}(k)| \right.\right. \\
&\qquad\qquad\qquad \left.\left. > \epsilon n^{1-1/p} - |F(0)| - |\hat{F}(k)| \,\Big|\, X_1 \right) \right] \\
&\leq 2n \exp\left(-\frac{\max(0, (\epsilon n^{1-1/p} - |F(0)| - |\hat{F}(k)|))^2}{\|F\|_\infty(n-1)} \right),
\end{aligned}
$$

where the last equation is Hoeffding's Inequality. We then apply Borel Cantelli Lemma.

It remains to prove the statement of the proposition for $p = 2$. Similarly, we obtain:

$$\mathbb{E}\|A_n\Phi_{k,n} - \hat{F}(k)\Phi_{k,n}\|_2^2 = \frac{1}{n}\mathbb{E}|\sum_j F(X_1 - X_j)e^{-2i\pi k.(X_1-X_j)} - n\hat{F}(k)|^2.$$

We then write $\mathbb{E}|\sum_j F(X_1-X_j)e^{-2i\pi k.(X_1-X_j)} - n\hat{F}(k)|^2 = \mathbb{E}[\mathbb{E}[|\sum_j \left(F(X_1-X_j)e^{-2i\pi k.(X_1-X_j)} - \hat{F}(k) \right)|^2|X_1]] = |F(0) - \hat{F}(k)|^2 + (n-1)\int_\Omega |F(x)e^{-2i\pi k.x} - \hat{F}(k)|^2 dx$. The statement follows. □

4.2 Correlation of the Eigenvalues

In this paragraph, we state an elementary lemma on the m-correlation of the eigenvalues of A ($m \leq n$):

$$M_m = 1/\binom{n}{m} \mathbb{E} \sum_{\{i_1, \cdots, i_m\} \subset \{1, \cdots n\}} \prod_{j=1}^{m} (\lambda_{i_j} - F(0)),$$

where the sum is over all subsets of $\{1, \cdots, n\}$ of cardinal m. Note that $M_1 = 0$ and that M_m is related to the factorial moment measure $\rho_m(dz_1, \cdots, dz_m)$ (also called the joint intensity

14

measure, refer to Daley and Vere-Jones [6]) of the point process $\{\lambda_1 - F(0), \cdots, \lambda_n - F(0)\}$ as follows:

$$M_m = \int_{\mathbb{C}^m} \prod_{j=1}^{m} z_j \rho_m(dz_1, \cdots, dz_m),$$

Heuristically, $\rho_m(dz_1, \cdots, dz_m)$ is the infinitesimal probability of having an eigenvalue at $F(0) + z_i$ for each $i \in \{1, \cdots, m\}$. We define $\bar{A} = A - F(0)I$, where I is the $n \times n$ identity matrix (note that "$\bar{\cdot}$" is not the complex conjugate of the matrix A). $\bar{A}(x_1, \cdots, x_m)$ is the $m \times m$ matrix where the coefficient i, j is equal to $F(x_i - x_j) - \delta_{ij} F(0)$.

Lemma 10

$$M_m = \int_{\Omega^m} \det \bar{A}(x_1, \cdots, x_m) dx_1 \cdots dx_m.$$

For $m = 2$ we get:

$$M_2 = -\int_{\Omega} F(x)^2 dx,$$

the point process of eigenvalues is thus repulsive.

Proof. The characteristic polynomial of $\bar{A}$ is $\chi_{\bar{A}}(t) = \det(\bar{A} - tI) = \prod_{i=1}^{n}(\lambda_i - f(0) - t) = \sum_{m=0}^{n} a_m(-t)^{n-m}$, where, $a_m = \sum_{\{i_1,\cdots,i_m\}} \prod_{j=1}^{m}(\lambda_{i_j} - f(0))$. However, by Newton formula, we also have, $a_m = \sum_{\{i_1,\cdots,i_m\}} det\bar{A}_{\{i_1,\cdots,i_m\}}$, where for a set of indices $\mathbf{i} = \{i_1, \cdots, i_m\}$, $\bar{A}_\mathbf{i}$ is the $m \times m$ extracted matrix obtained from $\bar{A}$ by keeping the raws and columns $\{i_1, \cdots, i_m\}$ (i.e. $\bar{A}_\mathbf{i}$ is a principal minor). Taking expectation, we deduce that $\mathbb{E}a_m = \binom{n}{m} \int_{\Omega^m} \det \bar{A}(x_1, \cdots, x_m) dx_1 \cdots dx_m$. □

In Lemma 10, we have computed the mean value of the symmetric polynomials:

$$\alpha_m(x_1, \cdots, x_n) = \sum_{\{i_1,\cdots,i_m\} \subset \{1,\cdots n\}} \prod_{j=1}^{m} x_j$$

for the vector $\bar{\lambda} = (\lambda_1 - F(0), \cdots, \lambda_n - F(0))$. Actually, it is possible to compute the mean value of the symmetric polynomials:

$$\alpha_{m,k}(x_1, \cdots, x_n) = \sum_{\{i_1,\cdots,i_m\} \subset \{1,\cdots n\}} \prod_{j=1}^{m} x_j^k.$$

for the vector $\bar{\lambda}$. To this end simply consider, $\chi_{\bar{A}^k}(t) = \det(\bar{A}^k - tI) = \prod_{i=1}^{n}((\lambda_i - f(0))^k - t)$. We obtain similarly:

$$1/\binom{n}{m} \mathbb{E}\alpha_{m,k}(\bar{\lambda}) = \int_{\mathbb{C}^n} \prod_{j=1}^{m} z_j^k \rho_m(dz_1, \cdots, dz_m) = \int_{\Omega^n} \det \bar{A}_\mathbf{m}^k(x_1, \cdots, x_n) dx_1 \cdots dx_n,$$

where $\mathbf{m} = \{1, \cdots, m\}$ and for a set of indices $\mathbf{i} = \{i_1, \cdots, i_m\}$, $\bar{A}_\mathbf{i}$ is the $m \times m$ extracted matrix obtained from $\bar{A}$ by keeping the raws and columns $\{i_1, \cdots, i_m\}$.

Acknowledgments

References

[1] Z. D. Bai. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica*, 9(3):611–677, 1999. With comments by G. J. Rodgers and Jack W. Silverstein; and a rejoinder by the author.

[2] E. Bogomolny, O. Bohigas, and C. Schmidt. Spectral properties of distance matrices. *Journal of Physics A: Mathematical and General*, 36:3595–3616, 2003.

[3] W. Bryc, A. Dembo, and T. Jiang. Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.*, 34(1):1–38, 2006.

[4] F. Chung, L. Lu, and V. Vu. Eigenvalues of random power law graphs. *Ann. Comb.*, 7(1):21–33, 2003.

[5] F. Chung, L. Lu, and V. Vu. The spectra of random graphs with given expected degrees. *Internet Math.*, 1(3):257–275, 2004.

[6] D.J. Daley and D. Vere-Jones. *An introduction to the Theory of Point Processes*. Springer Series in Statistics. Springer-Verlag, New-York, 1988.

[7] M. Draief and A. Ganesh. Efficient routeing in Poisson small-world networks. *J. Appl. Probab.*, 43(3):678–686, 2006.

[8] M. Draief, A. Ganseh, and L. Massoulié. Thresholds for virus spread on networks. *to appear in Ann. Appl. Probab.*

[9] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[10] C. Hammond and S. Miller. Distribution of eigenvalues for the ensemble of real symmetric Toeplitz matrices. *J. Theoret. Probab.*, 18(3):537–566, 2005.

[11] V. Koltchinskii and E. Giné. Random matrix approximation of spectra of integral operators. *Bernoulli*, 6(1):113–167, 2000.

[12] L. Lovász. Random walks on graphs: a survey. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 353–397. János Bolyai Math. Soc., Budapest, 1996.

[13] M. Mézard, G. Parisi, and A. Zee. Spectra of Euclidean random matrices. *Nuclear Phys. B*, 559(3):689–701, 1999.

[14] C. Offer and B. D. Simons. Field theory of Euclidean matrix ensembles. *J. Phys. A*, 33(42):7567–7583, 2000.

[15] M. Penrose. *Random Geometric Graphs*. Oxford Studies in Probability. Oxford Univeristy Press, Oxford, 2003.

[16] M. Talagrand. A new look at independence. *Ann. Probab.*, 24(1):1–34, 1996.

[17] A. M. Vershik. Random metric spaces and universality. *Russian Math. Surveys*, 59(2):259–295, 2004.

[18] E. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67:325–327, 1958.