Tensor Products of Random Unitary Matrices

Tomasz Tkocz[1] Marek Smaczyński[2] Marek Kuś [3] Ofer Zeitouni[4]
Karol Życzkowski[2,3]

March 16, 2012. Revised August 18, 2012.

[1]Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland. tkocz@mimuw.edu.pl

[2]The Marian Smoluchowski Institute of Physics, Jagiellonian University, Reymonta 4, 30-059 Cracow, Poland. marek.smaczynski@uj.edu.pl

[3]Center of Theoretical Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668 Warsaw, Poland. marek.kus@cft.edu.pl, karol@cft.edu.pl

[4]School of Mathematics, University of Minnesota and Faculty of Mathematics, Weizmann Institute of Science, POB 26, Rehovot 76100, Israel. zeitouni@math.umn.edu

Abstract

Tensor products of M random unitary matrices of size N from the circular unitary ensemble are investigated. We show that the spectral statistics of the tensor product of random matrices becomes Poissonian if $M = 2$, N become large or M become large and $N = 2$.

2010 Mathematics Subject Classification. 60B20, 15B52.
Key words and phrases. Random matrices, Circular Unitary Ensemble, Tensor product.

1 Introduction

In quantum mechanics a system is described with a Hamiltonian H which is a hermitian operator acting on a Hilbert space. Usually, in practical applications, this space is taken to be of finite, yet very large dimension, and the Hamiltonian H turns out to be highly complicated. Typically, we would like to find the spectrum of H (e.g. in nuclear physics it describes energy levels of the nucleus). Commonly, this is not analytically tractable. A breakthrough was achieved by E. Wigner who proposed treating H as a Gaussian random matrix, which applies very well in nuclear physics. The effectiveness of random matrices reaches far beyond nuclear physics, as for instance they are useful in analyzing generic properties of entangled states [10, 6].

Once we are given the Hamiltonian H, time evolution of the quantum system is determined by the unitary operator e^{itH}. Again, we may replace this *a priori* complicated operator by a random unitary matrix, CUE matrix in other words, and hope that generic properties of the system remain unchanged. If a physical system consists of, say two non-interacting subsystems with Hamiltonians H_1 and H_2, then the Hamiltonian of the whole system is the tensor product $H_1 \otimes H_2$. In particular the dynamics is governed by the unitary operator $e^{itH_1} \otimes e^{itH_2}$. Therefore, it is natural to ask about statistical properties of spectra of tensor products of random unitary matrices.

More generally, consider a quantum system consisting of M non-interacting subsystems. For simplicity we shall assume that each of them is described in N dimensional Hilbert space, so that any local unitary dynamics can be written as $U = U_1 \otimes \ldots \otimes U_M$, where U_j's are $N \times N$ unitary matrices. If the unitary dynamics of each subsystem is generic, the matrices U_j can be represented by random matrices from the CUE.

The main aim of the present work is to analyze properties of the tensor product of random unitary matrices. We show that when either $N = 2$ in the limit of a large number M of subsystems, or when $M = 2$ in the limit of large subsystem size N, the point process obtained from the spectrum of U, properly rescaled, becomes Poissonian, in the sense that its correlation functions converge to that of a Poisson process.

This paper is organized as follows. In section 2 we provide some definitions and introduce our main results, Theorem 1 and 2, and their corollaries; we also provide numerical simulations that confirm the results. Section 3 provides the proof of Theorem 1 and of Corollary 1, while Section 4 is devoted to the proof of Theorem 2 and of Corollary 2.

2 Spectral statistics for tensor products of random unitary matrices

The spectral statistics for two ensembles of unitary matrices will be the focal points of our investigation. The first case involves two unitary $N \times N$ matrices, whereas in the second we consider the tensor product of M two-dimensional unitary matrices. As usual, we are interested in spectral properties in the asymptotic limits of large matrices, i.e., respectively, $N \to \infty$ and $M \to \infty$.

2.1 Background and basic definitions

We recall some standard definitions and properties of some ensembles of random unitary matrices. The simplest situation is a diagonal unitary matrix with eigenvalues being independently drawn points on the unit circle. Such matrices form the **circular Poisson ensemble**, **CPE** for short. The name reflects the fact that for large matrices the number n of eigenvalues inside an interval of the length $L << 2\pi$ is approximately Poisson-distributed

$$p(L, n) \sim \frac{e^{-\lambda L}(\lambda L)^n}{n!}$$

with parameter $\lambda = N/2\pi$.

Our main interest will be in unitary matrices of size $N \times N$ drawn according to the Haar measure on the unitary group $U(N)$; such a matrix is called a matrix from the $\mathbf{CUE_N}$, where CUE stands for **circular unitary ensemble**.

Let A_N be a CUE_N matrix. Denote by $(e^{i\theta_j^N})_{j=1}^N$ its eigenvalues, where we assume that the eigenphases θ_j^N belong to the interval $[0, 2\pi)$. The random vector $(\theta_1^N, \ldots, \theta_N^N)$ possesses a density P_{CUE_N} with respect to the Lebesgue measure, which was given by Dyson in his seminal paper [3],

$$P_{\mathrm{CUE}_N}(\theta_1^N, \ldots, \theta_N^N) = C_N \prod_{1 \le k < l \le N} |e^{i\theta_k^N} - e^{i\theta_l^N}|^2. \tag{2.1}$$

This expression can be rewritten in the following form (see Paragraph 11.1 in [8])

$$P_{\mathrm{CUE}_N}(\theta_1^N, \ldots, \theta_N^N) = C_N (2\pi)^N \det \left[S_N(\theta_k^N - \theta_l^N) \right]_{k,l=1}^N,$$

where

$$S_N(x) = \frac{1}{2\pi} \frac{\sin \frac{Nx}{2}}{\sin \frac{x}{2}}. \tag{2.2}$$

In particular

$$S_N(0) = \frac{N}{2\pi}.$$

The set of eigenphases of a random unitary matrix can be seen as an example of **a point process** χ_N on the interval $[0, 2\pi)$ related to these eigenphases, by which we mean a random collection of points $\{\theta_1^N, \ldots, \theta_N^N\}$ or, in other words, an integer-valued random measure

$$\chi_N(D) = \sum_{k=1}^N \mathbf{1}_{\{\theta_k^N \in D\}}, \qquad D \subset [0, 2\pi),$$

where $\mathbf{1}_X$ denotes the indicator function of X.

A possible way to describe a point process is to give its so-called **joint intensities** or, as physicists usually say, **correlation functions** $\rho_k^N : (\mathbb{R}_+)^k \longrightarrow \mathbb{R}_+$, $k = 1, 2, \ldots$. In our case they might be defined simply as (see [1, Remark 4.2.4])

$$\rho_k^N(x_1, \ldots, x_k) = \lim_{\varepsilon \to 0} \frac{1}{(2\varepsilon)^k} \mathbb{P}\left(\exists j_1, \ldots, j_k \quad |\theta_{j_s}^N - x_s| < \varepsilon, \ s = 1, \ldots, k \right), \qquad x_i \text{ distinct.} \tag{2.3}$$

It is known [8] that the process χ_N is determinantal with joint intensities

$$\rho_k^N(x_1, \ldots, x_k) = \det \left[S_N(x_s - x_t) \right]_{s,t=1}^k. \tag{2.4}$$

(Recall that a point process is called **determinantal** with kernel K if its joint intensities can be written as $\rho_k(x_1, \ldots, x_k) = \det[K(x_i, x_j)]_{i,j=1}^k$.) For CUE_N matrices, due to the translation invariance of the measures we have that $K_N(x_i, x_j) = K_N(x_i - x_j)$, hence a kernel is given by a function $K_N(x)$ of a single variable. We refer to [1] for more background on such determinantal processes.

By definition, the joint intensity ρ_k^N equals $N!/(N-k)!$ times the k dimensional marginal distribution of the vector $(\theta_1^N, \ldots, \theta_N^N)$. Thus

$$\frac{(N-k)!}{N!} \int_{[0,2\pi)^k} \det \left[S_N(x_s - x_t) \right]_{s,t=1}^k \mathrm{d}x_1 \ldots \mathrm{d}x_k = 1. \tag{2.5}$$

If we rescale properly the eigenphases of a CUE_N matrix it turns out that they exhibit nice asymptotic behavior. Namely, it is clear that the point process $\{\frac{N}{2\pi}\theta_1^N, \dots, \frac{N}{2\pi}\theta_N^N\}$ is determinantal with the kernel $\frac{2\pi}{N} S_N\left(\frac{x}{N/2\pi}\right)$. Thanks to the fact that this function converges when $N \to \infty$, we can give a precise analytic description of the limit of the probability $\mathbb{P}\left(\frac{N}{2\pi}\theta_1^N \notin A, \dots, \frac{N}{2\pi}\theta_N^N \notin A\right)$, where $A \subset \mathbb{R}_+$ is a compact set (see Theorem 3.1.1 in [1]).

In the case of CPE matrices the situation is even simpler. The point process related to the rescaled (by the factor $\frac{N}{2\pi}$) eigenphases of a CPE_N matrix behaves for large N as a Poisson point process with the parameter $\lambda = 1$.

For point processes, related to the correlation functions is the notion of **level spacing distribution**, denoted by $P(s)$, which is defined for a point process $\{\alpha\vartheta_1, \dots, \alpha\vartheta_N\}$ of the properly rescaled eigenphases $(\vartheta_j)_{j=1}^N$ of a random N-dimensional unitary matrix by

$$P(s) := \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \frac{1}{N} \sum_{j=1}^{N} \mathbb{P}\left(s_j \in (s-\varepsilon, s+\varepsilon)\right), \tag{2.6}$$

where

$$s_1 = \alpha\left(\vartheta_1' + 2\pi - \vartheta_N'\right), \qquad s_j = \alpha(\vartheta_j' - \vartheta_{j-1}'),\ 1 < j \leq N, \tag{2.7}$$

and $(\vartheta_j')_{j=1}^N$ is the non-decreasing rearrangement of the sequence $(\vartheta_j)_{j=1}^N$. The scaling factor α is chosen so that the mean distance $\mathbb{E}s_j$ between two consecutive rescaled eigenphases is 1. In the cases of a CUE_N or CPE_N matrix, one has $\alpha = \frac{N}{2\pi}$. We should bear in mind that the level spacing distribution of the Poisson point process with the parameter $\lambda = 1$ is exponential with the density

$$P(s) = e^{-s}. \tag{2.8}$$

Moreover, it is easy to check that

$$P_{\mathrm{CPE}_N}(s) \xrightarrow[N \to \infty]{} e^{-s}.$$

Of course, the limit for the CUE_N is different.

2.2 Statement of results

We now present our main results for the two cases under consideration.

2.2.1 $M = 2$, N large

We begin by considering two independent CUE matrices A and B of size N. We are interested in the asymptotic behavior of the eigenphases of the tensor product $A \otimes B$. Our first main result is the following.

Theorem 1. *Let $(\theta_j)_{j=1}^N$ and $(\phi_j)_{j=1}^N$ be the eigenphases of two independent* CUE_N *matrices A and B. Define the point process σ_N of rescaled eigenphases of the matrix $A \otimes B$ as*

$$\sigma_N(D) := \sum_{k,l=1}^{N} \mathbf{1}_{\left\{\frac{N^2}{2\pi}(\theta_k + \phi_l \mod 2\pi) \in D\right\}}, \qquad \textit{for any compact set } D \subset \mathbb{R}_+. \tag{2.9}$$

Let ρ_k^N, $k = 1, 2, \ldots$ be the intensities of the process σ_N. Then

$$\rho_k^N \xrightarrow[N\to\infty]{} 1, \tag{2.10}$$

uniformly on any compact subset of $(\mathbb{R}_+)^k$.

Thus, Theorem 1 relates the statistical properties of a properly rescaled phase-spectrum of a large $\mathrm{CUE}_N \otimes \mathrm{CUE}_N$ matrix to those of a Poisson point process. A (not immediate) corollary of the convergence of intensities is the following.

Corollary 1. *For the point process σ_N defined in* (2.9)*,*

$$\begin{aligned} &\mathbb{P}\left(\sigma_N \text{ has no rescaled eigenphase in the interval } [0,s]\right) \\ &\quad = \mathbb{P}\left(\sigma_N([0,s]) = 0\right) \xrightarrow[N\to\infty]{} e^{-s}, \qquad s > 0. \end{aligned} \tag{2.11}$$

In particular

$$P_{\mathrm{CUE}_N \otimes \mathrm{CUE}_N}(s) \xrightarrow[N\to\infty]{} e^{-s}, \tag{2.12}$$

where the level spacing distribution $P_{\mathrm{CUE}_N \otimes \mathrm{CUE}_N}(s)$ is defined by (2.6)

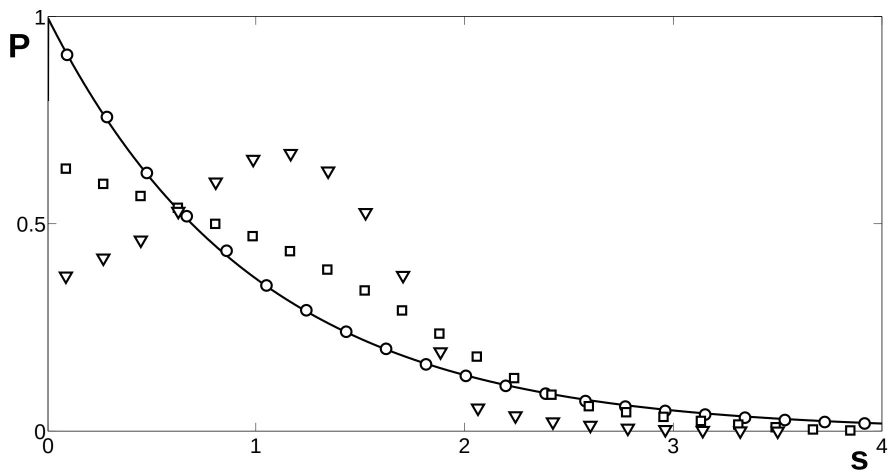

Figure 2.1: The level spacing distributions $P(s)$ for the tensor products of random unitary matrices $CUE_N \otimes \mathrm{CUE}_N$ for $N = 2$ ($\triangledown$), $N = 3$ ($\square$), $N = 20$ ($\circ$). The symbols denote the numerical results respectively obtained for $2^{17}, 2^{16}, 2^{13}$ independent matrices, while the solid line represents the exponential distribution (2.8).

Our numerical results support (2.10), i.e. the level spacing distribution of the tensor product of two random unitary matrices of size N is described asymptotically by the Poisson ensemble. The numerical data presented in Figure 2.1 reveals that $P_{\mathrm{CUE}_N \otimes \mathrm{CUE}_N}(s)$ and $P_{\mathrm{CPE}_N}(s)$ are close already for $N = 20$.

2.2.2 $N = 2$, M large

We next consider M independent CUE_2 matrices $A_1, \ldots, A_M$ and study the asymptotic properties of the phase-spectrum of a matrix $A_1 \otimes \ldots \otimes A_M$. Our main result is as follows.

Theorem 2. *Let θ_j^1, θ_j^2, $j = 1, \ldots, M$ be the eigenphases of independent CUE_2 matrices $A_1, \ldots, A_M$. Define the point process τ_M of the rescaled eigenphases of a matrix $A_1 \otimes \ldots \otimes A_M$ as*

$$\tau_M(D) := \sum_{\epsilon=(\epsilon_1,\ldots,\epsilon_M)\in\{1,2\}^M} \mathbf{1}_{\left\{\frac{2M}{2\pi}\left(\theta_1^{\epsilon_1}+\ldots+\theta_M^{\epsilon_M} \mod 2\pi\right)\in D\right\}}, \qquad \text{for any compact set } D \subset \mathbb{R}_+. \tag{2.13}$$

Then, for each k there exists a continuous function $\delta_k : \mathbb{R}_+ \to \mathbb{R}_+$ with $\delta_k(0) = 0$ so that for any mutually disjoint intervals $I_1, \ldots, I_k \subset \mathbb{R}_+$

$$\limsup_{M\to\infty} \frac{\mathbb{P}\left(\tau_M(I_1) > 0, \ldots, \tau_M(I_k) > 0\right)}{|I_1| \cdot \ldots \cdot |I_k|} \leq \left(1 + \delta_k(\max_j |I_j|)\right),$$
$$\liminf_{M\to\infty} \frac{\mathbb{P}\left(\tau_M(I_1) > 0, \ldots, \tau_M(I_k) > 0\right)}{|I_1| \cdot \ldots \cdot |I_k|} \geq \left(1 - \delta_k(\max_j |I_j|)\right).$$

Note that the statement of Theorem 2 is weaker than that of Theorem 1. This is due to the fact that stronger correlations exist in the point process τ_M, which prevent us from discussing the convergence of its intensities to those of a Poisson process. The mode of convergence is however strong enough to deduce interesting information, including the weak convergence of the processes. We exhibit this by considering the behavior of the level spacings when M tends to infinity.

Corollary 2. *For the point process τ_M defined in (2.13) we have*

$$\begin{aligned} &\mathbb{P}\left(\tau_M \text{ has no eigenphase in the interval } [0, s]\right) \\ &= \mathbb{P}\left(\tau_M([0, s]) = 0\right) \xrightarrow[M\to\infty]{} e^{-s}, \qquad s > 0. \end{aligned} \tag{2.14}$$

In particular

$$P_{\mathrm{CUE}_2^{\otimes M}}(s) \xrightarrow[M\to\infty]{} e^{-s}, \tag{2.15}$$

where the level spacing distribution $P_{\mathrm{CUE}_2^{\otimes M}}(s)$ is defined by (2.6).

The relevant numerical results which confirm (2.15) are shown in Figure 2.2. Again we may observe that it is enough to take relatively small M in order to get a good approximation of the spectrum of a matrix $\mathrm{CUE}_2^{\otimes M}$ by the Poisson ensemble.

2.3 Discussion

The convergence exhibited in Theorems 1 and 2, and in their corollaries, is arguably not surprising: taking the tensor product introduces so many eigenphases (N^2 in the case of Theorem 1, 2^M in the case of Theorem 2) that, after appropriate scaling, the local correlations between adjacent eigenphases are not influenced by the long range correlation that is present due to the tensorization. One should however be careful in carrying this heuristic too far:

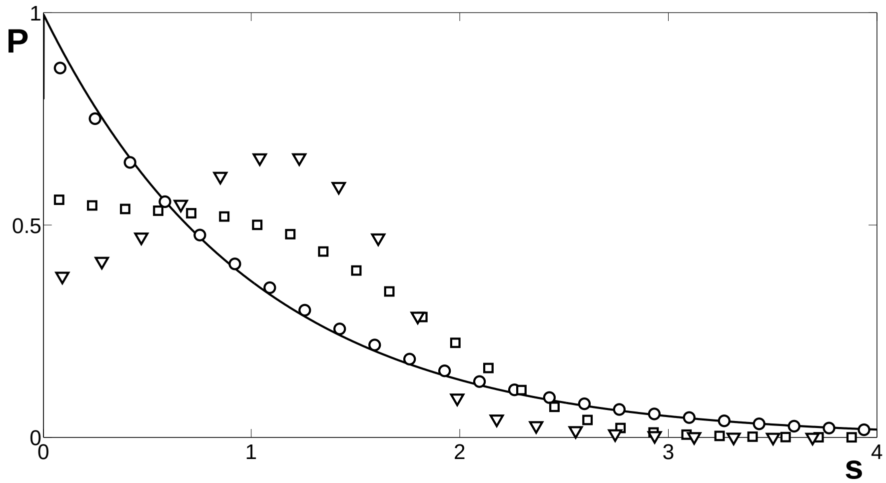

Figure 2.2: Level spacing distributions $P(s)$ for the tensor products of random unitary matrices $CUE_2^{\otimes M}$ for $M = 2$ ($\triangledown$), $M = 3$ ($\square$), $M = 8$ ($\circ$). The symbols denote the numerical results respectively obtained for $2^{17}, 2^{16}, 2^{14}$ independent matrices, while the solid line represents the exponential distribution (2.8).

well known superposition and interpolation relations, see [4] and the discussion in [1, Section 2.5.5], show that the point process obtained by the union of eigenvalues of, say, a GOE_N and GOE_{N+1} independent matrices, is closely related to that obtained from of a GUE_N matrix, and thus definitely not Poissonian. This phenomenon had been also discussed in the physics literature [9]. Compared to that, the tensorization operation appears to strongly decorrelate eigenphases on the local level.

It is natural to try to generalize Theorems 1 and 2 to other situations, where either N or M are finite but not necessarily equal to 2, or both N and M go to infinity. While we expect similar methods to apply and yield similar decorrelation results, there are several technical issues to control, and we do not discuss such extensions here. We pose the following conjecture.

Conjecture. *Let $\theta_j^1, \dots, \theta_j^N$, $j = 1, \dots, M$ be the eigenphases of independent* CUE_N *matrices $A_1, \dots, A_M$. Define the point process $\tau_{M,N}$ of the rescaled eigenphases of a matrix $A_1 \otimes \dots \otimes A_M$ as*

$$\tau_{M,N}(D) := \sum_{\epsilon=(\epsilon_1,\dots,\epsilon_M)\in\{1,\dots,N\}^M} \mathbf{1}_{\left\{\frac{N^M}{2\pi}\left(\theta_1^{\epsilon_1}+\dots+\theta_M^{\epsilon_M} \mod 2\pi\right)\in D\right\}}, \quad \text{for any compact set } D \subset \mathbb{R}_+. \tag{2.16}$$

Then, for each k there exists a continuous function $\delta_k : \mathbb{R}_+ \to \mathbb{R}_+$ with $\delta_k(0) = 0$ so that for *any mutually disjoint intervals* $I_1, \ldots, I_k \subset \mathbb{R}_+$

$$\limsup \frac{\mathbb{P}\left(\tau_{M,N}(I_1) > 0, \ldots, \tau_{M,N}(I_k) > 0\right)}{|I_1| \cdot \ldots \cdot |I_k|} \leq \left(1 + \delta_k(\max_j |I_j|)\right),$$
$$\liminf \frac{\mathbb{P}\left(\tau_{M,N}(I_1) > 0, \ldots, \tau_{M,N}(I_k) > 0\right)}{|I_1| \cdot \ldots \cdot |I_k|} \geq \left(1 - \delta_k(\max_j |I_j|)\right)$$

with fixed $N > 2$ *and* $M \to \infty$*, or* $N \to \infty$ *and fixed* $M > 2$*.*

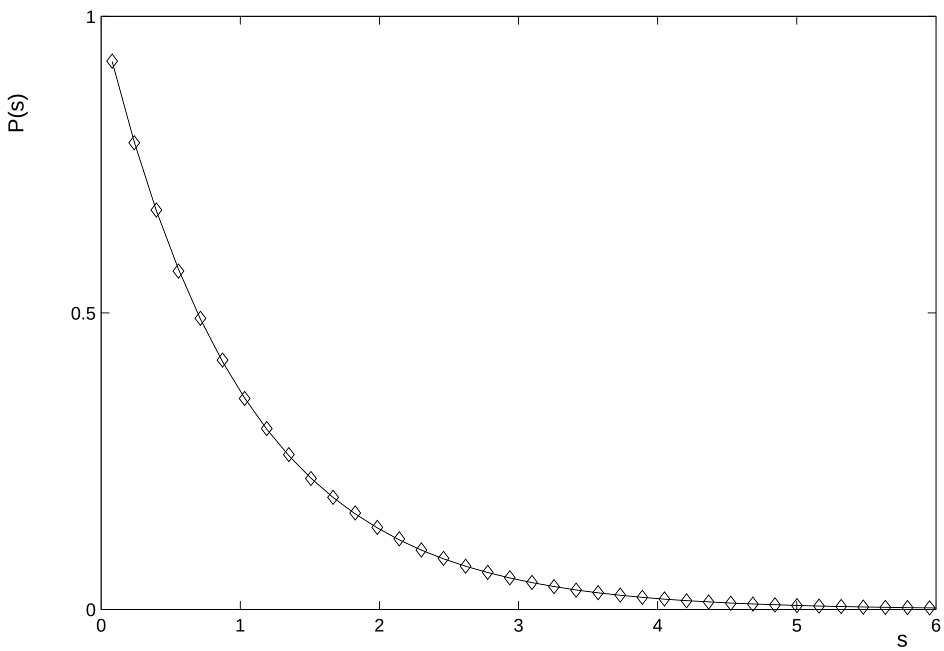

Figure 2.3: The symbols ($\Diamond$) show the level spacing distribution $P(s)$ obtained numerically for the tensor products of random unitary matrices $\mathrm{CUE}_N^{\otimes M}$ drawn 10^3 times when $M = 4$ and $N = 8$. The solid line represents the exponential distribution (2.8).

We offer some numerical evidence to support the expected consequence of the conjecture concerning spacings, see Figure 2.3. Moreover, we believe the conjecture holds true in the real case as well, i.e. if we replace CUE matrices with COE matrices.

3 Tensor product of two $N \times N$ unitary matrices

We prove in this section Theorem 1 and Corollary 1, that correspond to the case $M = 2$ and N large. We start with an elementary observation. Recall the kernel S_N, see (2.2).

Lemma 1. *For any $N \geq 1$*

$$\sup_{x\in\mathbb{R}} |S_N(x)| = \frac{N}{2\pi}. \tag{3.1}$$

Proof. Clearly there are many ways to see the lemma. We choose an elementary argument. First, we show inductively that

$$|\sin(nu)| \leq n|\sin u|, \qquad \text{for } u \in \mathbb{R}, n \geq 1.$$

Hence

$$|S_N(x)| = \frac{1}{2\pi}\left|\frac{\sin\left(N\frac{x}{2}\right)}{\sin\frac{x}{2}}\right| \leq \frac{N}{2\pi}.$$

Then, for $x = 0$ we have equality, which completes the proof. □

Proof of Theorem 1. We begin with setting $\tilde{x}_1, \ldots, \tilde{x}_k \geq 0$ and recalling that by definition

$$\rho_k^N(\tilde{x}_1, \ldots, \tilde{x}_k) = \lim_{\tilde{\varepsilon}\to 0} \frac{1}{(2\tilde{\varepsilon})^k}\mathbb{P}\Bigg(\exists \begin{array}{c}(i) = (i_1, \ldots, i_k) \in \{1, \ldots, N\}^k \\ (j) = (j_1, \ldots, j_k) \in \{1, \ldots, N\}^k\end{array} \ \forall s = 1, \ldots, k$$
$$\frac{N^2}{2\pi}(\theta_{i_s} + \phi_{j_s} \mod 2\pi) \in (\tilde{x}_s - \tilde{\varepsilon}, \tilde{x}_s + \tilde{\varepsilon})\Bigg).$$

Let us first of all get rid of the addition modulo 2π noticing that the event, probability of which we want to compute, is the sum of 2^k mutually exclusive events occurring when $\theta_{i_s} + \phi_{j_s}$ is in the interval $[0, 2\pi)$ or $[2\pi, 4\pi)$. Thus we can write the sought after probability as

$$\sum_{(\eta)=(\eta_1,\ldots,\eta_k)\in\{0,1\}^k} \mathbb{P}\left(\exists^{(i)}_{(j)} \ \forall s \ \theta_{i_s} + \phi_{j_s} \in (\eta_s \cdot 2\pi + x_s - \varepsilon, \eta_s \cdot 2\pi + x_s + \varepsilon)\right), \tag{3.2}$$

where we denote $x_s = \frac{2\pi}{N^2}\tilde{x}_s$ and $\varepsilon = \frac{2\pi}{N^2}\tilde{\varepsilon}$. Let us now concentrate solely on the first term corresponding to the index $(\eta) = (\eta_1, \ldots, \eta_k) = (0, \ldots, 0)$ (the other terms can be dealt with in the same manner). In order to take advantage of the independence we write explicitly (3.2) in terms of a convolution and thus observe that the considered quantity equals

$$\lim_{K\to\infty} \sum_{\substack{\ell_1,\ldots,\ell_k=1, \\ 2\pi\ell_s/K<x_s}}^{K} \mathbb{P}\left(\exists^{(i)}_{(j)} \ \forall s \ \begin{array}{c}\theta_{i_s} \in (2\pi\ell_s/K - \pi/K, 2\pi\ell_s/K + \pi/K) \\ \phi_{j_s} \in (x_s - 2\pi\ell_s/K - \varepsilon, x_s - 2\pi\ell_s/K + \varepsilon)\end{array}\right),$$

where the constrains $2\pi\ell_s/K < x_s$ are the result of the fact that $\theta_{i_s} + \phi_{j_s} \in (0\cdot 2\pi + x_s - \varepsilon, 0\cdot 2\pi + x_s + \varepsilon)$, for $(\eta) = 0$, so, in particular, that $\theta_{i_s} < x_s + \varepsilon$. Exploiting the independence we obtain that the last expression equals

$$\sum_{\substack{\ell_1,\ldots,\ell_k=1 \\ 2\pi\ell_s/K<x_s}}^{K} \mathbb{P}\left(\exists(i)\forall s \ \ \theta_{i_s} \in (2\pi\ell_s/K - \pi/K, 2\pi\ell_s/K + \pi/K)\right)$$
$$\cdot\, \mathbb{P}\left(\exists(j)\forall s \ \ \phi_{j_s} \in (x_s - 2\pi\ell_s/K - \varepsilon, x_s - 2\pi\ell_s/K + \varepsilon)\right). \tag{3.3}$$

Now observe that for a determinantal point process $\{\alpha_j\}_{j=1}^N$ with a kernel K and fixed numbers $u_1, \dots, u_k$ we have

$$\begin{aligned}
&\mathbb{P}\left(\exists (i) \in \{1, \dots, N\}^k \ \forall s = 1, \dots, k \quad \alpha_{i_s} \in (u_s - \delta, u_s + \delta)\right) \\
&= \sum_{p=1}^{k} \sum_{\pi \in \mathfrak{S}(k,p)} \lambda_\pi(u_1, \dots, u_k) \left((2\delta)^p \det\left[K(u_{\pi(s,1)}, u_{\pi(t,1)})\right]_{s,t=1}^p + o(\delta^p)\right),
\end{aligned} \tag{3.4}$$

where $\mathfrak{S}(k,p)$ is the collection of all partitions into p non-empty pairwise disjoint subsets of the set $\{1, \dots, k\}$. By this we mean that if π is such a partition then

$$\pi = \{\{\pi(1,1), \dots, \pi(1, \sharp\pi(1))\}, \dots, \{\pi(p,1), \dots, \pi(p, \sharp\pi(p))\}\},$$

where $\sharp\pi(q)$ is cardinality of the q-th block of the partition π. Moreover, to compactify the notation, we attach to a partition π the function $\lambda_\pi \colon \mathbb{R}^k \longrightarrow \{0,1\}$, defined as

$$\lambda_\pi(u_1, \dots, u_k) = \mathbf{1}_{\{u_{\pi(1,1)} = \ldots = u_{\pi(1,\sharp\pi(1))}, \ldots, u_{\pi(p,1)} = \ldots = u_{\pi(p,\sharp\pi(p))}\}}(u_1, \dots, u_k).$$

Applying this to formula (3.3) we obtain

$$\begin{aligned}
&\sum_{\substack{\ell_1, \dots, \ell_k = 1 \\ 2\pi\ell_s/K < x_s}}^{K} \sum_{p_1, p_2 = 1}^{k} \sum_{\substack{\pi_1 \in \mathfrak{S}(k,p_1) \\ \pi_2 \in \mathfrak{S}(k,p_2)}} \lambda_{\pi_1}\left((2\pi\ell_s/K)_{s=1}^k\right) \lambda_{\pi_2}\left((x_s - 2\pi\ell_s/K)_{s=1}^k\right) \\
&\quad \cdot \left(\left(\frac{2\pi}{K}\right)^{p_1} \det\left[S_N(2\pi\ell_{\pi_1(s,1)}/K - 2\pi\ell_{\pi_1(t,1)}/K)\right]_{s,t=1}^{p_1} + o(1/K^{p_1})\right) \\
&\quad \cdot \left((2\varepsilon)^{p_2} \det\left[S_N\left(x_{\pi_2(s,1)} - 2\pi\ell_{\pi_2(s,1)}/K - x_{\pi_2(t,1)} + 2\pi\ell_{\pi_2(t,1)}/K\right)\right]_{s,t=1}^{p_2} + o(\varepsilon^{p_2})\right) .
\end{aligned}$$

Performing the limit $K \to \infty$ we notice that only the terms corresponding to $p_2 = k$ do not vanish, for, otherwise, λ_{π_2} would give nontrivial relations for (ℓ) which altogether with λ_{π_1} make the sum over (ℓ) of at most $O(K^{p_1 - 1})$ terms. Recall that $\varepsilon/\tilde{\epsilon} = 2\pi/N^2$. Thus, taking the limit $\tilde{\varepsilon} \to 0$, the extra factor $(2\pi/N^2)^k$ is produced, so we finally find that the considered term contributes

$$\begin{aligned}
\sum_{p=1}^{k} \frac{1}{N^{k-p}} \sum_{\pi \in \mathfrak{S}(k,p)} \frac{1}{(2\pi)^p} \int_{\substack{[0,2\pi)^k \\ y_s < x_s}} \lambda_\pi(y_1, \dots, y_k) \det\left[\frac{2\pi}{N} S_N(y_{\pi(s,1)} - y_{\pi(t,1)})\right]_{s,t=1}^p \\
\cdot \det\left[\frac{2\pi}{N} S_N(x_s - y_s - x_t + y_t)\right]_{s,t=1}^k \mathrm{d}\mathcal{H}_p(y_1, \dots, y_k)
\end{aligned}$$

to $\rho_k^N(x_1, \dots, x_k)$, where $\mathcal{H}_p$ denotes the p-dimensional Hausdorff measure in $\mathbb{R}^k$. As already mentioned the other terms in (3.2) can be calculated in a similar way, only the limits of the integration have to be changed. Summing up, we get

$$\rho_k^N(x_1,\dots,x_k) = \sum_{(\eta)\in\{0,1\}^k} \sum_{p=1}^{k} \frac{1}{N^{k-p}} \sum_{\pi\in\mathfrak{S}(k,p)} \frac{1}{(2\pi)^p} \int_{A_{(\eta)}} \Bigg(\lambda_\pi(y_1,\dots,y_k) \cdot \det\left[\frac{2\pi}{N} S_N(y_{\pi(s,1)} - y_{\pi(t,1)})\right]_{s,t=1}^{p} \cdot \det\left[\frac{2\pi}{N} S_N(2\pi\eta_s + x_s - y_s - 2\pi\eta_t - x_t + y_t)\right]_{s,t=1}^{k} \Bigg) \mathrm{d}\mathcal{H}_p(y_1,\dots,y_k), \tag{3.5}$$

where the subset $A_{(\eta)}$ of $[0,2\pi)^k$ is the set of all $(y_1,\dots,y_k)$ such that either $y_s < x_s$ if $\eta_s = 0$, or $y_s \geq x_s$ if $\eta_s = 1$ for $s = 1,\dots,k$.

To proceed we have to investigate the asymptotic behavior of the integrand in (3.5). We will do it again only for $(\eta) = (0,\dots,0)$, observing that an adaptation to other terms is straightforward. We start with the term $p = k$. Then the integrand is a product of two determinants of matrices of size k, so applying to each of them the permutation definition and extracting the term referring to the trivial permutations, we find it equals

$$\left(\frac{2\pi}{N} S_N(0)\right)^{2k} + \sum_{\sigma\neq\mathrm{id}\text{ or }\tau\neq\mathrm{id}} \operatorname{sgn}\sigma \operatorname{sgn}\tau \prod_{i=1}^{k} \frac{2\pi}{N} S_N(y_i - y_{\sigma(i)}) \prod_{j=1}^{k} \frac{2\pi}{N} S_N(x_j - y_j - x_{\tau(j)} + y_{\tau(j)}), \tag{3.6}$$

where the summation involves all permutations σ and τ of k indices. The first term $(2\pi S_N(0)/N)^{2k} = 1$, after substituting in (3.5), gives simply

$$\frac{1}{(2\pi)^k} \sum_{(\eta)\in\{0,1\}^k} \int_{A_{(\eta)}} \left(\frac{2\pi}{N} S_N(0)\right)^{2k} = 1.$$

We will show that the second term in (3.6) after being put into (3.5) vanishes in the limit. We consider here only the case $k = 2$ to explain the main idea. The terms involving more factors can be treated along the same lines. The sum over σ and τ reduces to

$$-\left(\frac{2\pi}{N} S_N(0)\right)^2 \left(\left(\frac{2\pi}{N} S_N(y_1 - y_2)\right)^2 + \left(\frac{2\pi}{N} S_N(x_1 - y_1 - x_2 + y_2)\right)^2 \right) + \left(\frac{2\pi}{N} S_N(y_1 - y_2)\right)^2 \left(\frac{2\pi}{N} S_N(x_1 - y_1 - x_2 + y_2)\right)^2. \tag{3.7}$$

Let us for instance deal with the last term in equation (3.7). Putting it into (3.5) we arrive at

$$\frac{1}{(2\pi)^2} \sum_{(\eta)} \int_{A_{(\eta)}} \left(\frac{2\pi}{N} S_N(y_1 - y_2)\right)^2 \left(\frac{2\pi}{N} S_N(x_1 - y_1 - x_2 + y_2)\right)^2.$$

Taking a quick look at the integrand we see that the above expression goes to 0 when $N \to \infty$ by Lebesgue's dominated convergence theorem, for $\frac{1}{N} S_N(u) \xrightarrow{N\to\infty} 0$, when $u \neq 0$, and the appropriate bound (3.1) follows from Lemma 1.

For the terms corresponding to $k < p$, we easily notice that thanks to the factor $\frac{1}{N^{k-p}}$ they converge to 0. The proof is now complete. $\square$

Remark 1. By virtue of formula (3.5) the joint intensities ρ_k^N can be estimated as

$$\sup_{\mathbb{R}^k} |\rho_k^N| \leq \frac{1}{N^k} \sup_{u_1,\dots,u_k \in \mathbb{R}} \det \left[\frac{2\pi}{N} S_N(u_s - u_t) \right]_{s,t=1}^k \cdot \sum_{p=1}^{k} \sharp\mathfrak{S}(k,p) \int_{[0,2\pi)^p} \det \left[S_N(y_s - y_t) \right]_{s,t=1}^p \mathrm{d}y_1 \dots \mathrm{d}y_p,$$

where $\sharp X$ denotes cardinality of a set X. Using Hadamard's inequality (see, e.g. (3.4.6) in [1]) for the first term, the observation (2.5) for the second one, and finally (3.1) we obtain

$$\sup_{\mathbb{R}^k} |\rho_k^N| \leq \frac{1}{N^k} \left(\sup \left| \frac{2\pi}{N} S_N \right| \right)^k k^{k/2} \sum_{p=1}^{k} \sharp\mathfrak{S}(k,p) \frac{N!}{(N-k)!} = k^{k/2} \frac{1}{N^k} \sum_{p=1}^{k} \sharp\mathfrak{S}(k,p) \frac{N!}{(N-k)!}.$$

Due to the well-known combinatorial fact that

$$\sum_{p=1}^{k} \sharp\mathfrak{S}(k,p) x(x-1) \cdot \ldots (x-p+1) = x^k,$$

($\sharp\mathfrak{S}(k,p)$ is the Stirling number of the second kind, consult e.g. [5]) we may conclude with a useful bound

$$\sup_{\mathbb{R}^k} |\rho_k^N| \leq k^{k/2}. \quad \square \tag{3.8}$$

Proof of Corollary 1. For the proof of (2.11) we have to calculate the probability of the event that there is no rescaled eigenphase in a given interval. This is done in the following lemma.

Lemma 2. *Let χ be a point process related to the eigenphases, possibly rescaled, of a CUE_N matrix A_N with the joint intensities ρ_k, $k = 1, 2, \dots$ (so $\rho_\ell \equiv 0$, for $\ell > N$). Then for any compact set D*

$$\mathbb{P}\left(\chi(D) = 0\right) = 1 + \sum_{\ell=1}^{\infty} \frac{(-1)^\ell}{\ell!} \int_{D^\ell} \rho_\ell. \tag{3.9}$$

Proof. Clearly, we have

$$\mathbb{P}\left(\chi(D) = 0\right) = 1 - \sum_{k=1}^{N} \mathbb{P}\left(\chi(D) = k\right).$$

One way to compute the probability $\mathbb{P}\left(\chi(D) = k\right)$ is to use the notion of Jánossy densities $j_{D,k}(x_1, \dots, x_k)$ (see Definition 4.2.7 in [1]). They can be expressed in terms of the joint intensities as

$$j_{D,k}(x_1, \dots, x_k) = \sum_{r=0}^{\infty} \frac{1}{r!} (-1)^r \rho_{k+r}(x_1, \dots, x_k, \underbrace{D, \dots, D}_{r}), \tag{3.10}$$

where

$$\rho_{k+r}(x_1,\ldots,x_k,\underbrace{D,\ldots,D}_{r}) = \int_{D^r} \rho_{k+r}(x_1,\ldots,x_k,y_1,\ldots,y_r)dy_1\cdots dy_r. \tag{3.11}$$

They exist whenever

$$\sum_k \int_{D^k} \frac{k^r \rho_k(x_1,\ldots,x_k)}{k!} dx_1\cdots dx_k < \infty, \tag{3.12}$$

which is clearly fulfilled in our case, as $\rho_\ell \equiv 0$ for $\ell > N$. Moreover, the vanishing of ρ_ℓ for large enough ℓ makes every sum in the following finite so we will not have troubles with interchanging the order of summations.

In terms of the Jánossy intensities, the probability $\mathbb{P}(\chi(D) = k)$ reads as (see Equation (4.2.7) of [1])

$$\mathbb{P}(\chi(D) = k) = \frac{1}{k!}\int_{D^k} j_{D,k}(x_1,\ldots,x_k)dx_1\cdots dx_k, \tag{3.13}$$

and, consequently,

$$\begin{aligned}
\mathbb{P}(\chi(D) = 0) &= 1 - \sum_{k=1}^{N} \frac{1}{k!}\int_{D^k} j_{D,k} \\
&= 1 - \sum_{k=1}^{n} \frac{1}{k!}\int_{D^k} \sum_{r\geq 0} \frac{(-1)^r}{r!} \rho_{k+r}(x_1,\ldots,x_k,\underbrace{D,\ldots,D}_{r})\mathrm{d}x_1\ldots\mathrm{d}x_k \\
&= 1 - \sum_{k\geq 1}\sum_{r\geq 0} \frac{1}{k!}\frac{(-1)^r}{r!}\int_{D^{k+r}} \rho_{k+r} = 1 - \sum_{k\geq 1}\sum_{\ell\geq k} \frac{1}{k!}\frac{(-1)^{\ell-k}}{(\ell-k)!}\int_{D^\ell}\rho_\ell \\
&= 1 - \sum_{\ell\geq 1}\left[\sum_{k\geq 1}\binom{\ell}{k}(-1)^k\right]\frac{(-1)^\ell}{\ell!}\int_{D^\ell}\rho_\ell = 1 + \sum_{\ell\geq 1}\frac{(-1)^\ell}{\ell!}\int_{D^\ell}\rho_\ell.
\end{aligned}$$

□

Lemma 2 applied to the process σ_N yields

$$\mathbb{P}(\sigma_N([0,s]) = 0) = 1 + \sum_{\ell\geq 1}\frac{(-1)^\ell}{\ell!}\int_{[0,s]^\ell}\rho_\ell^N.$$

To pass to the limit $N \to \infty$ we need an appropriate bound on the intensities ρ_ℓ^N. In Remark 1 we showed that $|\rho_\ell^N| \leq \ell^{\ell/2}$ (see (3.8)). Therefore, by Lebesgue's dominated convergence theorem, we get

$$\lim_{N\to\infty}\mathbb{P}(\sigma_N([0,s]) = 0) = 1 + \sum_{\ell\geq 1}\frac{(-1)^\ell}{\ell!}\int_{[0,s]^\ell}\lim_{N\to\infty}\rho_\ell^N = 1 + \sum_{\ell\geq 1}\frac{(-1)^\ell}{\ell!}s^\ell = e^{-s}.$$

This completes the proof of (2.11).

The formula (2.12) follows now from a relation connecting the probability $E(0, s)$ that a randomly chosen interval of length s is free from eigenphases with the level spacing distribution $P(s)$, (2.6) (see equation (6.1.16a) in [8]),

$$P(s) = \frac{\mathrm{d}^2}{\mathrm{d}s^2} E(0; s). \tag{3.14}$$

We have just showed that $\lim_{N\to\infty} \mathbb{P}\left(\sigma_N([0, s]) = 0\right) = E(0; s) = e^{-s}$. Thus, indeed

$$\lim_{N\to\infty} P_{\mathrm{CUE}_N \otimes \mathrm{CUE}_N}(s) = \frac{\mathrm{d}^2}{\mathrm{d}s^2} e^{-s} = e^{-s}.$$

□

4 Tensor product of M unitary matrices of size 2×2

Now we will prove Theorem 2. In the course of the proof we will need three lemmas. Let us start with them.

Lemma 3. *Fix a positive integer s and a number $\gamma \in (0, 1/s)$. For each positive integer n let us define the set $\mathcal{L}_n = \{\ell = (\ell_1, \ldots, \ell_s) \mid \mathbb{Z} \ni \ell_j \geq 0, \sum_{j=1}^{s} \ell_j = n\}$. Then*

$$\sum_{\ell\in\mathcal{L}_n, \exists j\ \ell_j/n \leq \gamma} \frac{1}{s^n} \frac{n!}{\ell!} = 1 - \sum_{\ell\in\mathcal{L}_n, \forall j\ \ell_j/n > \gamma} \frac{1}{s^n} \frac{n!}{\ell!} \xrightarrow[n\to\infty]{} 0. \tag{4.1}$$

Here, we adopt the convention that $\ell! = \ell_1! \cdot \ldots \cdot \ell_s!$.

Proof. First observe that

$$\sum_{\ell, \exists j\ \ell_j/n \leq \gamma} \frac{1}{s^n} \frac{n!}{\ell!} \leq s \sum_{\ell, \ell_1/n \leq \gamma} \frac{1}{s^n} \frac{n!}{\ell!} = s \sum_{\ell_1=0}^{\lfloor \gamma n \rfloor} \frac{1}{s^n} \frac{n!}{\ell_1!(n-\ell_1)!} \sum_{\ell_2+\ldots+\ell_s \leq n-\ell_1} \frac{(n-\ell_1)!}{\ell_2! \cdot \ldots \cdot \ell_s!}$$
$$= s \sum_{\ell_1=0}^{\lfloor \gamma n \rfloor} \frac{1}{s^n} \binom{n}{n-\ell_1} (s-1)^{n-\ell_1} = s \sum_{k=n-\lfloor \gamma n \rfloor}^{n} \binom{n}{k} \left(1 - \frac{1}{s}\right)^k \left(\frac{1}{s}\right)^{n-k}.$$

Let $X_1, X_2, \ldots$ be i.i.d. Bernoulli random variables such that $\mathbb{P}(X_1 = 0) = 1/s = 1 - \mathbb{P}(X_1 = 1)$. Denote $S_n = X_1 + \ldots + X_n$. Then the last expression equals $s\mathbb{P}(S_n \geq n - \lfloor \gamma n \rfloor)$ and can be estimated from above as follows

$$s\mathbb{P}(S_n \geq n - \gamma n) = s\mathbb{P}\left(\frac{S_n - \mathbb{E}S_n}{n} \geq \frac{1}{s} - \gamma\right) \leq s \exp\left(-2n(1/s - \gamma)^2\right) \xrightarrow[n\to\infty]{} 0,$$

where the last inequality follows for instance from Hoeffding's inequality (see [7]). □

Lemma 4. *Let X be a random vector in $\mathbb{R}^n$ with a bounded density. Let $A\colon \mathbb{R}^n \longrightarrow \mathbb{R}^k$ be a linear mapping of rank r. Then there exists a constant C such that for any intervals $I_1, \ldots, I_k \subset \mathbb{R}$ of finite length we have*

$$\mathbb{P}(AX \in I_1 \times \ldots I_k) \leq C|I_{i_1}| \cdot \ldots \cdot |I_{i_r}|,$$

where $1 \leq i_1 < \ldots < i_r \leq k$ are indices of those rows of the matrix A which are linearly independent.

Proof. Let $a_1, \ldots, a_k \in \mathbb{R}^n$ be rows of the matrix A. We know there are r of them, say $a_1, \ldots, a_r$, which are linearly independent. Thus there exists an invertible $r \times r$ matrix U such that

$$U \begin{bmatrix} a_1 \\ \vdots \\ a_r \end{bmatrix} = \begin{bmatrix} e_1 \\ \vdots \\ e_r \end{bmatrix} =: E,$$

where $e_i \in \mathbb{R}^n$ is the i-th vector of the standard basis of $\mathbb{R}^n$. Notice that

$$\begin{aligned} \mathbb{P}\left(AX \in I_1 \times \ldots \times I_k\right) \leq \mathbb{P}\left(U^{-1}EX \in I_1 \times \ldots \times I_r\right) = \mathbb{P}\left((X_1, \ldots, X_r) \in U(I_1 \times \ldots I_r)\right) \\ \leq C|U(I_1 \times \ldots \times I_k)| = C|\det U| \cdot |I_1| \cdot \ldots \cdot |I_r|, \end{aligned}$$

for the vector $(X_1, \ldots, X_r)$ also has a bounded density on $\mathbb{R}^r$. This finishes the proof. □

Lemma 5. *Let A be a matrix of dimension $k \times j$, with entries in $\{0, 1\}$, and satisfying the following conditions*

(i) no two columns are equal.

(ii) no two rows are equal.

(iii) no zero row

Then, the rank of A is at least $\min(k, \lfloor \log_2 j \rfloor + 1)$.

Proof. (Due to Dima Gourevitch) Denote $r = \text{rank}A$. The assertion of the lemma is equivalent to the statement that $2^r \geq j$ and if $2^r = j$ then $r = k$.

We may assume without loss of generality that the first r rows of A are linearly independent and the others are their linear combinations. Under this assumption, if two columns are identical in the first r coordinates then they are identical in all coordinates. By condition *(i)*, such columns do not exist. Therefore the $r \times j$ submatrix B which consists of the first r rows has distinct columns. As a result $j \leq 2^r$.

Now suppose $j = 2^r$. If $k > r$, consider the $r + 1$ row of A. It is a linear combination of the first r rows. Since the columns of B include the column $e_i = (0, .., 0, 1, 0, .., 0)$ for all $i = 1, \ldots, r$, the coefficient of each row is either 0 or 1. B includes also a column of all 1s, thus there is at most one nonzero coefficient (if there were more than one, a certain entry would be greater than 1). Consequently, the coefficient of exactly one row is 1, and all other coefficients vanish, because if all coefficients were zero, the $r + 1$ row would be zero which contradicts *(iii)*. Thus, the $r+1$-th row is identical to one of the first r rows - in contradiction to condition *(ii)*. □

Proof of Theorem 2. Fix an integer $k \geq 1$ and finite intervals $I_1, \ldots, I_k \subset \mathbb{R}_+$ which are mutually disjoint. We need to compute the probability of the event $\{\tau_M(I_j) > 0, j = 1, \ldots, k\}$ which means that in each interval I_j there is a rescaled eigenphase. Such eigenphase is of the form $\frac{2^M}{2\pi}\left(\theta_1^{\epsilon_1} + \ldots + \theta_M^{\epsilon_M} \mod 2\pi\right)$ for some $\epsilon = (\epsilon_1, \ldots, \epsilon_M) \in \{1, 2\}^M$. Therefore

$$\{\tau_M(I_j) > 0, j = 1, \ldots, k\} = \bigcup_{\epsilon} A_\epsilon,$$

where

$$A_\epsilon = \left\{ \sum_{i=1}^{M} \theta_i^{\epsilon_i^j} \mod 2\pi \in \underbrace{\frac{2\pi}{2^M} I_j}_{J_j}, j = 1, \dots, k \right\}, \tag{4.2}$$

and ϵ runs over the set

$$\mathcal{E} = \left\{ [\epsilon_i^j]_{i=1,\dots,M}^{j=1,\dots,k} \mid \epsilon_i^j \in \{1,2\},\ \epsilon^u \neq \epsilon^v, \text{for } u \neq v, u, v = 1, \dots, k \right\} \tag{4.3}$$

of all $k \times M$ matrices with entries $1, 2$ which have pairwise distinct rows $\epsilon^j = (\epsilon_1^j, \dots, \epsilon_M^j) \in \{1,2\}^M$, $j = 1, \dots, k$ (j-th row ϵ^j describes the j-th eigenphase and since intervals are disjoint we assume the rows are distinct). Column vectors are denoted by $\epsilon_i = [\epsilon_i^1, \dots, \epsilon_i^k]^T$, $i = 1, \dots, M$.

We say that ϵ is *bad* if the collection of its vector columns $\{\epsilon_i, i = 1, \dots, M\}$ is less than 2^k. Otherwise ϵ is called *good*.

Obviously,

$$\mathbb{P}\left(\bigcup_{\text{good } \epsilon\text{'s}} A_\epsilon \right) \leq \mathbb{P}\left(\bigcup_{\epsilon} A_\epsilon \right) \leq \mathbb{P}\left(\bigcup_{\text{good } \epsilon\text{'s}} A_\epsilon \right) + \mathbb{P}\left(\bigcup_{\text{bad } \epsilon\text{'s}} A_\epsilon \right).$$

The strategy is to show that the contribution of bad ϵ's vanishes for large M while good ϵ's essentially provide the desired result $\prod_j |I_j|$ when M goes to infinity. So the proof will be divided into several parts.

Good ϵ's. The goal here is to prove

$$\lim_{\max_j |I_j| \to 0} \lim_{M\to\infty} \frac{1}{|I_1| \cdot \ldots \cdot |I_k|} \mathbb{P}\left(\bigcup_{\text{good } \epsilon\text{'s}} A_\epsilon \right) = 1, \tag{4.4}$$

with the required uniformity in the choice of the disjoint intervals I_j. By virtue of

$$\sum_{\text{good } \epsilon\text{'s}} \mathbb{P}(A_\epsilon) - \sum_{\substack{\text{good } \epsilon, \widetilde{\epsilon} \\ \epsilon \neq \widetilde{\epsilon}}} \mathbb{P}(A_\epsilon \cap A_{\widetilde{\epsilon}}) \leq \mathbb{P}\left(\bigcup_{\text{good } \epsilon\text{'s}} A_\epsilon \right) \leq \sum_{\text{good } \epsilon\text{'s}} \mathbb{P}(A_\epsilon)$$

it suffices to prove that

$$\lim_{M\to\infty} \sum_{\text{good } \epsilon\text{'s}} \mathbb{P}(A_\epsilon) = \prod |I_j| \tag{4.5}$$

uniformly, and that the correlations between two different good ϵs does not matter

$$\limsup_{\max_j |I_j| \to 0} \limsup_{M\to\infty} \frac{1}{\prod |I_j|} \sum_{\substack{\text{good } \epsilon, \widetilde{\epsilon} \\ \epsilon \neq \widetilde{\epsilon}}} \mathbb{P}(A_\epsilon \cap A_{\widetilde{\epsilon}}) = 0. \tag{4.6}$$

Let us now prove (4.5). The proof of (4.6) is deferred to the very end as we will need the ideas developed here as well as in the part devoted to bad ϵ's.

Given $\epsilon \in \mathcal{E}$ and a vector $\alpha = [\alpha_1 \ldots \alpha_k]^T \in \{1,2\}^k$ we count how many column vectors of ϵ equals α and call this number ℓ_α. Then $\sum_\alpha \ell_\alpha = M$. Note that ϵ is good iff all ℓ_αs are nonzero. The crucial observation is that the probability of the event A_ϵ does depend only on the vector $\ell = (\ell_\alpha)_{\alpha\in\{1,2\}^k}$ associated with ϵ as described before. Indeed, the sum $\sum_{i=1}^M [\theta_i^{\epsilon_i^1} \ldots \theta_i^{\epsilon_i^k}]^T \bmod 2\pi$ is identically distributed as the random vector $\sum_\alpha \psi(\alpha, \ell_\alpha) \bmod 2\pi$, where

$$\psi(\alpha,\ell_\alpha) = \begin{bmatrix} \psi_1(\alpha,\ell_\alpha) \\ \vdots \\ \psi_k(\alpha,\ell_\alpha) \end{bmatrix} = \begin{bmatrix} \theta_{i_1}^{\alpha_1} \\ \vdots \\ \theta_{i_1}^{\alpha_k} \end{bmatrix} + \ldots + \begin{bmatrix} \theta_{i_{\ell_\alpha}}^{\alpha_1} \\ \vdots \\ \theta_{i_{\ell_\alpha}}^{\alpha_k} \end{bmatrix} \bmod 2\pi \tag{4.7}$$

is a sum modulo 2π of i.i.d. vectors. Note that the distribution of $\psi(\alpha, \ell_\alpha)$ does not depend on the choice of indices $i_1, \ldots, i_{\ell_\alpha}$ but only on α and ℓ_α. Consequently, denoting by $\mathcal{E}_\ell$ the set of all ϵ's such that there are exactly ℓ_α indices $1 \le i_1 < \ldots < i_{\ell_\alpha} \le M$ for which $\epsilon_{i_1} = \ldots = \epsilon_{i_{\ell_\alpha}} = \alpha$, we have that the value of $\mathbb{P}(A_\epsilon)$ is the same for all $\epsilon \in \mathcal{E}_\ell$. Clearly $\sharp\mathcal{E}_\ell = \frac{M!}{\ell!}$, whence

$$\sum_{\text{good } \epsilon\text{'s}} \mathbb{P}(A_\epsilon) = \sum_{\text{good } \ell\text{'s}} \frac{M!}{\ell!} \mathbb{P}\left(\sum_{\alpha\in\{1,2\}^k} \psi(\alpha,\ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k \right). \tag{4.8}$$

The idea is to identify those terms which will sum up to $\prod |I_i|$ and the rest which will be neglected in the limit of large M. To do this, set a positive parameter $\gamma < 1/2^k$ and let us call a good ℓ *very good* (*v.g.* for short) if $\ell_\alpha > \gamma M$ for every α and *quite good* (*q.g.* for short) otherwise. We claim that

$$\mathbb{P}\left(\sum \psi(\alpha,\ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right) \le C \prod |J_j|, \qquad \textit{for a good } \ell, \tag{C1}$$

and

$$\mathbb{P}\left(\sum \psi(\alpha,\ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right) = \frac{\prod |J_j|}{(2\pi)^k}\left(1 + \frac{r_\ell}{\sqrt{M}}\right), \qquad |r_\ell| \le C, \quad \textit{for a very good } \ell, \tag{C2}$$

where C is a constant (from now on in this proof we adopt the convention that C is a constant depending only on k which may differ from line to line).

Let us postpone the proofs and see how to conclude (4.5). Notice that $\frac{\prod |J_j|}{(2\pi)^k} = \frac{1}{2^{kM}} \prod |I_j|$. Thus applying (C1) we obtain

$$\sum_{\text{q.g. } \ell\text{'s}} \mathbb{P}\left(\sum \psi(\alpha,\ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right) \le \prod |I_j| \cdot C \sum_{\text{q.g. } \ell\text{'s}} \frac{1}{2^{kM}} \frac{M!}{\ell!}.$$

By Lemma 3 it vanishes when $M \to \infty$. Now we deal with very good ℓ's writing with the aid of (C2) that

$$\sum_{\text{v.g. } \ell\text{'s}} \mathbb{P}\left(\sum \psi(\alpha,\ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right) = \prod |I_j| \Bigg(\sum_{\text{v.g. } \ell\text{'s}} \frac{1}{2^{kM}} \frac{M!}{\ell!} + \sum_{\text{v.g. } \ell\text{'s}} \frac{1}{2^{kM}} \frac{M!}{\ell!} \frac{r_\ell}{\sqrt{M}} \Bigg).$$

The first term in the bracket approaches 1 in the limit $M \to \infty$ due to Lemma 3, while the second one approaches 0 as it is bounded above by $C\frac{1}{\sqrt{M}}$.

Proof of (C1). Let us define the vectors

$$e_j = (\underbrace{2,\ldots,2}_{j-1}, 1, \underbrace{2,\ldots,2}_{k-j}) \in \{1,2\}^k, \qquad j = 1,\ldots,k.$$

Since ℓ is good, in particular we have that $\ell_{e_j} > 0$, so denoting the random vector $\psi(e_j, \ell_{e_j})$ by Ψ^j we have $\sum_\alpha \psi(\alpha,\ell_\alpha) = (\Psi^1 + \ldots + \Psi^k) + \sum_{\alpha\notin\{e_1,\ldots,e_k\}} \psi(\alpha,\ell_\alpha)$. By independence it is enough to show that the random vector $\Psi = \Psi^1 + \ldots + \Psi^k \bmod 2\pi$ has a bounded density on $[0,2\pi)^k$. Equation (4.7) yields that

$$\Psi^j = (\underbrace{Y_j,\ldots,Y_j}_{j-1}, X_j, \underbrace{Y_j,\ldots,Y_j}_{k-j}),$$

where (X_j, Y_j) are independent random vectors on $[0,2\pi)^2$ with the same distributions as the vectors $(\theta_1^1 + \ldots + \theta_{\ell_{e_j}}^1 \bmod 2\pi, \theta_1^2 + \ldots + \theta_{\ell_{e_j}}^2 \bmod 2\pi)$ respectively. Clearly, the vector (X_j, Y_j) has a bounded density on $[0,2\pi)^2$ because the vector (θ_1^1, θ_1^2) has a bounded density. Therefore the vector $(X_1, Y_1, \ldots, X_k, Y_k)$ has a bounded density on $[0,2\pi)^{2k}$. A certain linear transformation with determinant 1 maps this vector to $(\Psi^1 + \ldots + \Psi^k, Y_1, \ldots, Y_k)$ which consequently also has a bounded density. One projects it to the first k coordinates and then takes care of addition modulo 2π obtaining that Ψ has a bounded density, which finishes the proof. □

Proof of (C2). Given a vector $\alpha \in \{1,2\}^k$ let Θ^α denote the random vector in $[0,2\pi)^k$ identically distributed as the vector $(\theta_1^{\alpha_1}, \ldots, \theta_1^{\alpha_k})$. Take its independent copies $\Theta_1^\alpha, \Theta_2^\alpha, \ldots$ such that the family $\{\Theta_1^\alpha, \Theta_2^\alpha, \ldots\}_{\alpha\in\{1,2\}^k}$ also consists of independent random vectors. Then $\mathbb{E}\Theta^\alpha = [\pi,\ldots,\pi]^T$, and

$$\begin{aligned} p_{\ell,M} &= \mathbb{P}\left(\sum_\alpha \psi(\alpha,\ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right) = \mathbb{P}\left(\sum_\alpha \sum_{l=1}^{\ell_\alpha} \Theta_l^\alpha \bmod 2\pi \in J_1 \times \ldots \times J_k\right) \\ &= \sum_{i_1,\ldots,i_k=0}^{M-1} \mathbb{P}\left(\sum_\alpha \sum_{l=1}^{\ell_\alpha} \Theta_l^\alpha \in (J_1 + 2\pi i_1) \times \ldots \times (J_k + 2\pi i_k)\right) \\ &= \sum_i \mathbb{P}\left(\sum_\alpha \sum_{l=1}^{\ell_\alpha} \frac{\Theta_l^\alpha - \mathbb{E}\Theta_l^\alpha}{\sqrt{M}} \in \frac{1}{\sqrt{M}}(J_1 + 2\pi(i_1 - M/2)) \times \ldots \times \frac{1}{\sqrt{M}}(J_k + 2\pi(i_k - M/2))\right). \end{aligned}$$

To ease the notation we introduce new indices

$$j = \left(i_1 - \frac{M}{2}, \ldots, i_k - \frac{M}{2}\right) \in \left\{-\frac{M}{2}, -\frac{M}{2}+1, \ldots, \frac{M}{2}-1\right\}^k,$$

sets

$$K_{j,M} = \frac{1}{\sqrt{M}}(J_1 + 2\pi j_1) \times \ldots \times \frac{1}{\sqrt{M}}(J_k + 2\pi j_k),$$

and the vector

$$S_M = \sum_\alpha \sum_{l=1}^{\ell_\alpha} \frac{\Theta_l^\alpha - \mathbb{E}\Theta_l^\alpha}{\sqrt{M}}.$$

Now we intend to use the local Central Limit Theorem of [2]. Indeed, due to independence such a theorem should hopefully yield that S_M has a normal distribution for large M. To be more precise, let us consider the matrix $\operatorname{Cov} S_M = \sum_\alpha \frac{\ell_\alpha}{M} \operatorname{Cov} \Theta^\alpha$ and its eigenvalues. Since for any $x \in \mathbb{R}^k$

$$x^T(\operatorname{Cov} S_M)x = \sum_\alpha \frac{\ell_\alpha}{M} x^T(\operatorname{Cov}\Theta^\alpha)x \leq \underbrace{\max_\alpha \|\operatorname{Cov}\Theta^\alpha\|^{1/2}}_{C} |x|^2,$$

it is clear that the largest eigenvalues are uniformly (i.e. with respect to M) bounded by C, which depends only on k. To provide an uniform bound for the smallest eigenvalues let us observe that (recall that e_i is the vector $(2, \ldots, 2, 1, 2, \ldots, 2)$)

$$x^T(\operatorname{Cov} S_M)x \geq \sum_{j=1}^k \frac{\ell_{e_j}}{M} x^T(\operatorname{Cov}\Theta^{e_j})x > \gamma x^T\Big(\sum_{j=1}^k \operatorname{Cov}\Theta^{e_j}\Big)x \geq \gamma \cdot \frac{\pi^2}{3}|x|^2,$$

where the second inequality is because ℓ is very good.

It is a matter of a direct computation to see the last inequality as for $k \geq 2$ we have $\sum_{j=1}^k \operatorname{Cov}\Theta^{e_j} = \big((k-2)\pi^2/3 - 2\big)[1\ldots1]^T[1\ldots1] + \operatorname{diag}(2 + 2\pi^2/3, \ldots, 2 + 2\pi^2/3)$ and for $k = 1$ the sum equals $\pi^2/3$. Therefore, with the matrix B_M given by

$$B_M^2 = (\operatorname{Cov} S_M)^{-1}$$

it holds that

$$\frac{1}{C}|x| \leq |B_M x| \leq C|x|.$$

Therefore the assumptions of [2, Corollary 19.4] are satisfied (for the family of independent random vectors $\{\Theta_1^\alpha, \Theta_2^\alpha, \ldots\}_{\alpha\in\{1,2\}^k}$), so the vector $B_M S_M$ possesses a density q_M and

$$\sup_{x\in\mathbb{R}^k} \big(1 + |x|^{k+2}\big)\left(q_M(x) - \phi(x) - \frac{1}{\sqrt{M}}P_M(x)\phi(x)\right) = O(M^{-k/2}),$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}^k} e^{-|x|^2/2}$ is the density of the standard normal distribution in $\mathbb{R}^k$ and P_M is a polynomial of degree $k-1$ whose coefficients depends on the cumulants of the vectors

$B_M\Theta^\alpha$. We may put it differently, i.e.

$$q_M(x) = \phi(x) + \frac{1}{\sqrt{M}}\bigg(\underbrace{P_M(x)\phi(x) + \frac{f_M(x)}{1+|x|^{k+2}}}_{h_M(x)}\bigg),$$

for some functions f_M uniformly bounded $\sup_M \sup_{x\in\mathbb{R}^k} |f_M(x)| = C < \infty$. Therefore, denoting $L_{j,M} = B_M K_{j,M}$,

$$\begin{aligned} p_{\ell,M} &= \sum_j \mathbb{P}\left(S_M \in K_{j,M}\right) = \sum_j \mathbb{P}\left(B_M S_M \in B_M K_{j,M}\right) = \sum_j \int_{L_{j,M}} q_M \\ &= \sum_j \int_{L_{j,n}} \phi + \frac{1}{\sqrt{M}} \sum_j \int_{L_{j,n}} h_M = a_M + \frac{1}{\sqrt{M}} b_M. \end{aligned} \tag{4.9}$$

Let us firstly deal with the error term b_M. Denoting

$$\kappa = \frac{|J_1| \cdot \ldots \cdot |J_k|}{(2\pi)^k}$$

we are to show that

$$|b_M| \leq C\kappa. \tag{4.10}$$

To do this we estimate the integrated function

$$|h_M(x)| \leq |P_M(x)|\phi(x) + \frac{C}{1+|x|^{k+2}} =: h(x).$$

Then $|b_M| \leq \sum_j \int_{L_{j,M}} h$. Introduce *full* boxes

$$F_{j,M} = B_M\left(\frac{1}{\sqrt{M}}([0,2\pi) + 2\pi j_1) \times \ldots \times \frac{1}{\sqrt{M}}([0,2\pi) + 2\pi j_k)\right)$$

and observe that

$$\int_{L_{j,M}} h = \frac{|L_{j,M}|}{|F_{j,M}|} |F_{j,M}| \frac{1}{|L_{j,n}|} \int_{L_{j,M}} h \leq \kappa |F_{j,M}| \sup_{L_{j,M}} h \leq \kappa |F_{j,M}| \sup_{F_{j,M}} h.$$

Since $\mathrm{diam} F_{j,M} \leq C\frac{2\pi\sqrt{k}}{\sqrt{M}} \xrightarrow[M\longrightarrow\infty]{} 0$, the sets $F_{j,M}$ are pairwise disjoint and sum up to $B_M[-\pi\sqrt{M}, \pi\sqrt{M})^k$, we can infer that the sum $\sum_j |F_{j,M}| \sup_{F_{j,M}} h$ converges to $\int_{\mathbb{R}^k} h = C < \infty$. Hence, this sum is bounded by C and we get (4.10).

Now we handle the main term a_M. We prove it equals κ up to another error $\kappa\frac{C}{\sqrt{M}}$. Let $A_{j,M} : \mathbb{R}^k \longrightarrow \mathbb{R}^k$ be the linear isomorphism mapping $F_{j,M}$ onto $L_{j,M}$. It equals $B_M \tilde{A}_{j,M} B_M^{-1}$, where $\tilde{A}_{j,M}$ is the linear mapping transforming the box $B_M^{-1}F_{j,M}$ onto the box $B_M^{-1}L_{j,M}$, whence $|\det A_{j,M}| = \kappa$. Thus, changing the variable we obtain

$$\int_{L_{j,M}} \phi(x)\mathrm{d}x = \kappa \int_{F_{j,M}} \phi(A_{j,M}x)\mathrm{d}x.$$

Notice that $A_{j,M}x$ is close to x, whenever $x \in F_{j,M}$, for

$$|A_{j,n}x - x| \leq \mathrm{diam} F_{j,M}, \qquad x \in F_{j,M}.$$

Consequently, on $F_{j,M}$, $\phi(A_{j,M}x)$ is close to $\phi(x)$. Strictly, we use the mean value theorem and get

$$\int_{L_{j,M}} \phi(x)\mathrm{d}x = \kappa \int_{F_{j,M}} \phi(x)\mathrm{d}x + \kappa \int_{F_{j,M}} \nabla\phi_V(\eta_x) \cdot (A_{j,M}x - x)\mathrm{d}x,$$

for some mean points $\eta_x \in [x, A_{j,M}x]$. This results in

$$\begin{aligned} a_M &= \sum_j \int_{L_{j,M}} \phi(x)\mathrm{d}x = \kappa \sum_j \int_{F_{j,M}} \phi + \kappa \int_{\bigcup_j F_{j,M}} \nabla\phi(\eta_x) \cdot (A_{j,M}x - x)\mathrm{d}x \\ &= \kappa\Big(1 \underbrace{- \int_{\mathbb{R}^k \setminus B_M[-\pi\sqrt{M}, \pi\sqrt{M})^k} \phi}_{c_M} + \underbrace{\sum_j \int_{F_{j,M}} \nabla\phi_V(\eta_x) \cdot (A_{j,M}x - x)\mathrm{d}x}_{d_M}\Big). \end{aligned}$$

We are almost done. Clearly c_M converges to 0 faster that $1/\sqrt{M}$, so $|c_M| \leq C/\sqrt{M}$. For d_M we use the Schwarz inequality and integrability of $|\nabla\phi(\eta_x)|$

$$|d_M| \leq \sum_j \int_{F_{j,M}} |\nabla\phi(\eta_x)||A_{j,M}x - x|\mathrm{d}x \leq \mathrm{diam} F_{j,M} \int_{\bigcup F_{j,M}} |\nabla\phi(\eta_x)|\mathrm{d}x \leq \frac{C}{\sqrt{M}}.$$

This completes the proof of (C2). □

We have proved claims (C1) and (C2), so the proof of the part concerning good ϵ's is now complete. Let us proceed to tackle bad ϵ's.

Bad ϵ's. The goal here is to show that

$$\lim_{M\to\infty} \mathbb{P}\left(\bigcup_{\text{bad } \epsilon\text{'s}} A_\epsilon\right) = 0, \tag{4.11}$$

again, with the required uniformity. Obviously it suffices to show that $\sum_{\text{bad } \epsilon\text{'s}} \mathbb{P}(A_\epsilon) \xrightarrow[M\to\infty]{} 0$. Let $\mathcal{F}_j$ be the set of those bad ϵ's for which the cardinality of the set $\{\epsilon_i, i = 1, \ldots, M\}$ equals j. Observe that $\sharp\mathcal{F}_j \leq j^M$. With the aid of Lemma 5 we will show that

$$\forall \epsilon \in \mathcal{F}_j \ \mathbb{P}(A_\epsilon) \leq C \cdot 2^{-M(1+\lfloor \log_2 j \rfloor)} \cdot O\big((\max_j |I_j|)^{1+\lfloor \log_2 j \rfloor}\big), \quad \text{when } \max_j |I_j| \longrightarrow 0. \tag{4.12}$$

This will finish the proof, for

$$\begin{aligned} \sum_{\text{bad } \epsilon\text{'s}} \mathbb{P}(A_\epsilon) &\leq C \cdot O(\max_j |I_j|) \sum_{j=1}^{2^k-1} j^M \cdot 2^{-M(1+\lfloor \log_2 j \rfloor)} \\ &= C \cdot O(\max_j |I_j|) \sum_{j=1}^{2^k-1} 2^{-M(1+\lfloor \log_2 j \rfloor - \log_2 j)} \xrightarrow[M\to\infty]{} 0. \end{aligned} \tag{4.13}$$

For the proof of (4.12) fix $\epsilon \in \mathcal{F}_j$. We have seen that

$$\mathbb{P}(A_\epsilon) = \mathbb{P}\left(\sum \psi(\alpha, \ell_\alpha) \bmod 2\pi \in J_1 \times \ldots \times J_k\right)$$

and we know that there are exactly j numbers ℓ_α which are nonzero, say those which correspond to vectors $\alpha^1, \ldots, \alpha^j \in \{1,2\}^k$. Denote $\Psi^j = \psi(\alpha^i, \ell_{\alpha^i})$, $i = 1, \ldots, j$ and consider the random vector $S_j = \Psi^1 + \ldots + \Psi^j$ in $\mathbb{R}^k$. As in the proof of Claim (C1) we observe that S_j is a linear image of the vector $(X_1, Y_1, \ldots, X_j, Y_j)$. This mapping is given by the matrix $A = [a_{uv}]$ where

$$a_{2i-1,v} = \begin{cases} 1, & \alpha_v^i = 1 \\ 0, & \alpha_v^i = 2 \end{cases}, \qquad a_{2i,v} = \begin{cases} 0, & \alpha_v^i = 1 \\ 1, & \alpha_v^i = 2 \end{cases}.$$

By Lemma 4 we obtain

$$\mathbb{P}(S_j \bmod 2\pi \in J_1 \times \ldots \times J_k) \le C \max\left(|J_{i_1}| \cdot \ldots \cdot |J_{i_r}|\right) = C \cdot O(\max_j |I_j|) \cdot 2^{-Mr}, \quad (4.14)$$

where $r = \operatorname{rank} A$. The number r does not change if we replace the $2i$-th column of A with the vector e with 1 at each its entry, as the sum of $2i-1$-th and $2i$-th columns is just e. Now taking only the columns $1, 2, 3, 5, \ldots, 2j-1$ we get the matrix B which has the same rank as A. It has $j+1$ columns and fulfills the assumptions of Lemma 5 (it has no zero row as the second column consists of all 1s). Thus $r \ge \min(1 + \lfloor \log_2(1+j) \rfloor, k)$ and when $j < 2^k - 1$ this minimum equals $1 + \lfloor \log_2(1+j) \rfloor \ge 1 + \lfloor \log_2 j \rfloor$. If $j = 2^k - 1$ in the matrix A there must be two identical columns, one with even, say $2u$, and one with odd, say $2v-1$ index, which means that the u-th and the v-th column of B add up to e, so the v-th column may be erased and the rank of B does not change. Therefore we apply the lemma to the matrix B with erased the v-th column which is of size $k \times j$ and get again $r \ge \min(1 + \lfloor \log_2 j \rfloor, k) = 1 + \lfloor \log_2 j \rfloor$. This completes the proof of (4.12).

Pairs of good ϵ's, i.e. the proof of (4.6)**.** We denote by $\Theta_i(\epsilon)$ the random vector $(\theta_i^{\epsilon_i^1}, \ldots, \theta_i^{\epsilon_i^k})$. By the definition of A_ϵ we may write

$$A_\epsilon \cap A_{\widetilde{\epsilon}} = \left\{ \sum_{i=1}^M \begin{bmatrix} \Theta_i(\epsilon) \\ \Theta_i(\widetilde{\epsilon}) \end{bmatrix} \bmod 2\pi \in \begin{matrix} J_1 \times \ldots \times J_k \\ J_1 \times \ldots \times J_k \end{matrix} \right\}. \quad (4.15)$$

Since the intervals J_u and J_v are disjoint for $u \neq v$, we may restrict ourselves to those ϵ and $\widetilde{\epsilon}$ for which $\epsilon^u \neq \widetilde{\epsilon}^v$ whenever $u \neq v$, $u, v = 1, \ldots, k$ as otherwise the event $A_\epsilon \cap A_{\widetilde{\epsilon}}$ is impossible. However it might happen that $\epsilon^u = \widetilde{\epsilon}^u$. Let us count for how many u's it takes place, i.e. given $s \in \{1, \ldots, k\}$ let $\mathcal{P}_s$ be the set of all considered unordered pairs $\{\epsilon, \widetilde{\epsilon}\}$ for which there are exactly $k-s$ indices $1 \le u_1 < \ldots < u_{k-s} \le k$ such that $\epsilon^{u_j} = \widetilde{\epsilon}^{u_j}$, $j = 1, \ldots, k-s$. The value $s = 0$ is excluded as $\epsilon \neq \widetilde{\epsilon}$. We have

$$\sum_{\epsilon \neq \widetilde{\epsilon}} \mathbb{P}(A_\epsilon \cap A_{\widetilde{\epsilon}}) = \sum_{s=1}^k \sum_{\{\epsilon, \widetilde{\epsilon}\} \in \mathcal{P}_s} \mathbb{P}(A_\epsilon \cap A_{\widetilde{\epsilon}}).$$

Thus we fix s and prove that $\limsup_{\max_j |I_j| \to 0} \limsup_{M \to \infty} \frac{1}{\prod |I_j|} \sum_{\{\epsilon, \widetilde{\epsilon}\} \in \mathcal{P}_s} \mathbb{P}(A_\epsilon \cap A_{\widetilde{\epsilon}}) = 0$. There are two cases. A pair $\{\epsilon, \widetilde{\epsilon}\} \in \mathcal{P}_s$ can be *good* which means $\sharp \left\{ \left[\begin{smallmatrix} \epsilon_i \\ \widetilde{\epsilon}_i \end{smallmatrix}\right],\ i = 1, \ldots, M \right\} \ge$

2^{k+s}, or, otherwise we call it *bad*. We obtain a decomposition $\mathcal{P}_s = \mathcal{P}_s^{\text{good}} \cup \mathcal{P}_s^{\text{bad}}$. Now for a good pair, applying the reasoning already used for bad ϵ's, i.e. combining lemmas 4 and 5, we get the estimate

$$\mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) \leq C|J_1| \cdot \ldots \cdot |J_k| \big(\max_{j=1,\ldots,k} |J_j|\big)^s = \frac{C}{2^{(k+s)M}} \left(\prod |I_j|\right) \big(\max_j |I_j|\big)^s.$$

But $\sharp \mathcal{P}_s^{\text{good}} \leq \sharp \mathcal{P}_s \leq \binom{k}{s} \cdot 2^{(k+s)M}$, so

$$\limsup_{\max_j |I_j| \to 0} \limsup_{M\to\infty} \frac{1}{\prod |I_j|} \sum_{\{\epsilon,\widetilde{\epsilon}\} \in \mathcal{P}_s^{\text{good}}} \mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) = 0.$$

For a bad pair $\{\epsilon, \widetilde{\epsilon}\}$ we know that there are $k+s$ different rows and at most $2^{k+s}-1$ different columns in the matrix $\left[\begin{smallmatrix}\epsilon \\ \widetilde{\epsilon}\end{smallmatrix}\right]$. Hence we repeat the argument of the part concerning bad ϵ's. Namely, first exactly in the same manner as in that part we use Lemma 5 in order to establish an appropriate inequality in the spirit of (4.12). Then we follow the estimate of (4.13) and conclude that

$$\lim_{M\to\infty} \sum_{\{\epsilon,\widetilde{\epsilon}\} \in \mathcal{P}_s^{\text{bad}}} \mathbb{P}\left(A_\epsilon \cap A_{\widetilde{\epsilon}}\right) = 0.$$

This finishes the proof of Theorem 2. □

Proof of Corollary 2. Fix Δ small so that s/Δ is an integer and divide the interval $[0, s]$ into consecutive intervals of length Δ, denoted I_i. Let $Z_i = \tau_M(I_i)$ and $\bar{Z}_i = \mathbf{1}_{\{Z_i > 0\}}$. Of course, $\tau_M([0, s]) = \sum_{i=1}^{s/\Delta} Z_i$. Our goal is to show that $\tau_M([0, s])$ becomes Poissonian in the limit of large M, from which the statement of the corollary follows immediately.

The proof of Theorem 2 yields the following facts. There exist a sequence $\delta_{M,\Delta,k}$ with

$$\limsup_{\Delta\to 0} \limsup_{M\to\infty} \delta_{M,\Delta,k} = 0\,,$$

and a universal constant C such that the following hold.

$$\mathbb{P}\left(Z_i \neq \bar{Z}_i\right) \leq C\Delta^2\,, \tag{4.16}$$

$$\mathbb{E}\left(\prod_{i\in J_k} \bar{Z}_i\right) = \Delta^k (1 + O(\delta_{M,\Delta,k})) \tag{4.17}$$

where J_k denotes an arbitrary subset of k distinct integers in $\{1, \ldots, s/\Delta\}$.

Indeed, to justify (4.16) notice that

$$\mathbb{P}\left(Z_i \neq \bar{Z}_i\right) = \mathbb{P}\left(\tau_M(I_i) \geq 2\right) \leq \mathbb{P}\left(\bigcup_{\epsilon\neq\widetilde{\epsilon}} A_\epsilon \cap A_{\widetilde{\epsilon}}\right),$$

where $\epsilon, \widetilde{\epsilon} \in \{1,2\}^M$ and A_ϵ is the event that there is an eigenphase described by ϵ in the interval I_i (see (4.2)). The probability of the event $A_\epsilon \cap A_{\widetilde{\epsilon}}$ can be estimated by $C \cdot 2^{-2M} \cdot |I_i|^2 = 2^{-2M} \cdot C\Delta^2$. To see this, recall (4.15) and follow the same argument which led to estimate (4.14) (in this case the relevant matrix has the rank no less than 2). It suffices, as $\mathbb{P}\left(\bigcup_{\epsilon\neq\tilde{\epsilon}} A_\epsilon \cap A_{\tilde{\epsilon}}\right) \le \binom{2^M}{2} \cdot 2^{-2M} \cdot C\Delta^2 \le C\Delta^2$. For (4.17), observe that

$$\mathbb{E}\prod_{i\in J_k} \bar{Z}_i = \mathbb{E}\mathbf{1}_{\{Z_i>0, i\in J_k\}} = \mathbb{P}\left(\tau_M(I_i) > 0, i \in J_k\right),$$

and apply Theorem 2 (with its uniformity statement).

Let Y_i be i.i.d. Bernoulli random variables with $\mathbb{P}(Y_1 = 1) = 1 - \mathbb{P}(Y_1 = 0) = \Delta$. By (4.17) we have that for any integer ℓ,

$$\limsup_{\Delta\to 0}\limsup_{M\to\infty}\left|\mathbb{E}\left(\sum_{i=1}^{s/\Delta}\bar{Z}_i\right)^\ell - \mathbb{E}\left(\sum_{i=1}^{s/\Delta} Y_i\right)^\ell\right| = 0\,.$$

Since $\sum_{i=1}^{s/\Delta} Y_i$ converges to a Poisson random variable of parameter s as $\Delta \to 0$, it follows that $\sum_{i=1}^{s/\Delta} \bar{Z}_i$ converges in distribution to a Poisson variable of parameter s, when first $M \to \infty$ and then $\Delta \to 0$. On the other hand, using (4.16) we have that

$$\mathbb{P}\left(\sum_{i=1}^{s/\Delta}\bar{Z}_i \neq \sum_{i=1}^{s/\Delta} Z_i\right) \le Cs\Delta\,,$$

and therefore, one concludes that also $\sum_{i=1}^{s/\Delta} Z_i$ converges in distribution to a Poisson variable of parameter s, when first $M \to \infty$ and then $\Delta \to 0$. This yields the corollary. □

Acknowledgements.

TT was partially supported by NCN Grant no. 2011/01/N/ST1/05960. MS, MK and KZ were supported by the SFB Transregio-12 project der Deutschen Forschungsgemeinschaft and a grant financed by the Polish National Science Centre under the contract number DEC-2011/01/M/ST2/00379. OZ was supported by NSF grant DMS-0804133 and by a grant from the Israel Science Foundation.

Part of the work was done while the first named author was participating in The Kupcinet-Getz International Summer Science School at the Weizmann Institute of Science in Rehovot, Israel. We are grateful to the WIS for financial support making this possible.

We thank S. Jain and A. Pandey for comments related to the discussion in [9]. Finally, we thank Dima Gourevitch for both providing and allowing us to use his proof of Lemma 5.

References

[1] G. Anderson, A. Guionnet and O. Zeitouni, An introduction to random matrices, *Cambridge Studies in Advanced Mathematics **118** (Cambridge: Cambridge University Press)*, 2010.

[2] R. N. Bhattacharya and R. Ranga Rao, Normal approximation and asymptotic expansions. Classics in Applied Mathematics; 64, *Philadelphia : Society for Industrial and Applied Mathematics, cop.*, 2010.

[3] F. J. Dyson, Statistical Theory of the Energy Levels of Complex Systems, *J. Math. Phys.* **3** (1962), 140–156.

[4] P. J. Forrester and E. M. Rains, Interrelationships between orthogonal, unitary and symplectic matrix ensembles. In *Random Matrix Models and their Applications, Math. Sci. Res. Inst. Publ.* **40**, 171–207, *Cambridge, Cambridge University Press*, 2001.

[5] R. L. Graham, D. E. Knuth and O. Patashnik, Concrete Mathematics, *Addison-Wesley* , 1998.

[6] P. Hayden, D. W. Leung and A. Winter, Aspects of generic entanglement, *Commun. Math. Phys.* **265** (2006), 95–117.

[7] Zhengyan Lin and Zhidong Bai, Probability inequalities, *Science Press Beijing, Beijing; Springer, Heidelberg*, 2010.

[8] M. L. Mehta, Random matrices. Third edition, *Amsterdam: Elsevier/Academic Press*, 2004.

[9] N. Rosenzweig and C. E. Porter, Repulsion of energy levels in complex atomic spectra, *Phys. Review* **120** (1960), 1698–1714.

[10] K. Życzkowski and H-J. Sommers, Induced measures in the space of mixed quantum states, *J. Phys.* **A 34** (2001), 7111–7125.