Spectra of Multi-Level Toeplitz Matrices

Random Matrices and Communication Systems (IC-30)

Olivier Roy
supervised by
Dr. Olivier Lévêque

July 14, 2005

Contents

1 Introduction

Toeplitz matrices arise naturally in the study of stationary random processes. Therefore, a careful analysis of their properties is needed in order to understand the behavior of the underlying processes. In particular, the spectral behavior of Toeplitz matrices and their close relationship with circulant matrices are of particular interest. In this report, we consider the extension to the multi-dimensional case, by means of multi-level Toeplitz and their associated multi-level circulant matrices. We also show how the asymptotic equivalences established between some classes of Toeplitz and circulant matrices can be generalized by means of low-rank corrections.

The outline of the report is as follows. In Section 2, we explain multi-level matrices and multi-level indexing. We extend some results about Toeplitz matrices to the multi-dimensional case and illustrate them with a simple example. Section 3 shows how standard results can be generalized to a larger class of Toeplitz matrices by means of a low-rank correction. We finally offer some conclusions in Section 4.

2 Multi-Level Toeplitz Matrices

In this section, we first review some concepts about (single level) Toeplitz matrices and then show how these generalize to the multi-level case. We then illustrate these theoretical results with a simple example.

2.1 Single-Level Case

An $n \times n$ Toeplitz matrix T_n is a matrix with constant coefficients along its diagonals, i.e. $(T_n)_{i,j} = t_{j-i}$ for $1 \leq i, j \leq n$ where $(T_n)_{i,j}$ denotes the element at position (i, j). It is thus completely characterized by its generating sequence $\{t_k\}_{k\in\mathbb{Z}}$. We may consider the associated (2π-periodic) Fourier series $f(x)$ defined as

$$f(x) = \sum_{k\in\mathbb{Z}} t_k e^{ikx}. \tag{1}$$

Note that we do not say anything about the convergence of the above series. In particular, if T_n is Hermitian (i.e. $f(x)$ is real-valued) and $\{t_k\} \in \ell^1(\mathbb{Z})$, then $f(x)$ exists, is continuous and bounded. In the rest of the discussion, the Toeplitz matrices under consideration will be solely characterized with respect to their associated Fourier series f and denoted $T_n(f)$. Toeplitz matrices are of particular interest in the study of stationary random processes, since the covariance matrix corresponding to any finite realization of such processes is of Toeplitz form. Unfortunately, very little is known about the eigenvalues of T_n in general. We can however alleviate this problem by considering the asymptotic eigenvalue behavior as the size n of the matrix becomes large. To this end, we first recall the two following definitions.

Definition 2.1 (Equally Distributed Sequences)**.** *Two sequences of real numbers* $\{\lambda_k^{(n)}\}_{k=1}^n$ *and* $\{\mu_k^{(n)}\}_{k=1}^n$ *are said to be equally distributed if for any continuous function* F *with bounded support*

$$\lim_{n\to\infty} \sum_{k=1}^{n} \left[F\left(\lambda_k^{(n)}\right) - F\left(\mu_k^{(n)}\right) \right] = 0. \tag{2}$$

We denote $\lambda^{(n)} \sim \mu^{(n)}$*.*

Definition 2.2 (Asymptotic Distribution)**.** *A sequence of real numbers* $\{\lambda_k^{(n)}\}_{k=1}^n$ *is said to have an asymptotic distribution if there is a Lebesgue-integrable* 2π*-periodic function* $f(x)$ *such that for any continuous function* F *with bounded support*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} F\left(\lambda_k^{(n)}\right) = \frac{1}{2\pi} \int_0^{2\pi} F\left(f(x)\right) dx. \tag{3}$$

The sequence $\{\lambda_k^{(n)}\}$ *is said to be distributed as* $f(x)$ *and we denote* $\lambda^{(n)} \sim f(x)$*.*

Definitions 2.1 and 2.2 are central to all the results considered in the sequel concerning the limiting eigenvalue distribution of Toeplitz matrices. In the rest of the discussion, we will concentrate on Hermitian Toeplitz matrices ($f(x)$ real-valued). Similar results hold for the singular values in the non-Hermitian case [1]. Note that if $f \in L^\infty$, the eigenvalues of $T_n(f)$ all belong to a common finite interval $[m, M]$ for all n. In this case, we can thus restrict the above definitions to any function F continuous on $[m, M]$.

Let us denote the eigenvalues of $T_n(f)$ by $\lambda_k^{(n)}$. In order to prove that $\lambda^{(n)} \sim f(x)$, the idea is to consider a carefully chosen circulant matrix $C_n(f)$ with eigenvalues $\mu_k^{(n)}$ such that

$$\lambda^{(n)} \sim \mu^{(n)} \quad \text{and} \quad \mu^{(n)} \sim f(x). \tag{4}$$

Circulant matrices are characterized by the fact that $(C_n)_{i,j} = c_{j-i(\mathrm{mod}\, n)}$. We write $C_n = \mathrm{circ}(c_0, c_1, \ldots, c_{n-1})$. Note that such a matrix is not necessarily unique. In [1], two types of circulant matrices are shown to both fulfill the conditions given by (4). The first type, referred to as "simple" circulant and denoted S_n, is constructed from the Toeplitz sequence $\{t_k\}_{k\in\mathbb{Z}}$ as

$$\begin{aligned} S_{2m} &= \mathrm{circ}(t_0, t_1, \ldots, t_m, t_{-m}, \ldots, t_{-1}), \\ S_{2m+1} &= \mathrm{circ}(t_0, t_1, \ldots, t_{m-1}, 0, t_{-m+1}, \ldots, t_{-1}). \end{aligned} \tag{5}$$

The second type, referred to as "Cesàro" circulant and denoted C_n, is built as

$$C_n = \mathrm{circ}(c_0, c_1, \ldots, c_{n-1}) \tag{6}$$

where

$$\begin{aligned} c_0 &= t_0, \\ c_k &= \frac{(n-k)t_k + kt_{-n+k}}{n} \quad \text{for } k = 1, 2 \ldots, n-1. \end{aligned} \tag{7}$$

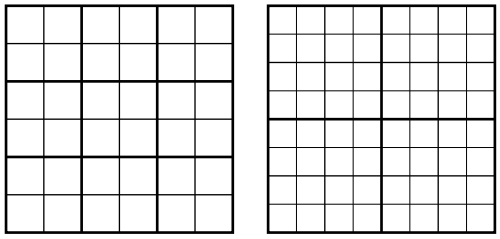


Figure 1: Multi-level matrices. The tiling of a multi-level matrix with (a) $p = 2$, $n_1 = 3$, $n_2 = 2$ and (b) $p = 3$, $n_1 = 2$, $n_2 = 2$, $n_3 = 2$.

Note that in both definitions, the idea is to build a circulant sequence out of a Toeplitz sequence. Using these results, the theorem of Szegö [2] states that for $f \in L^\infty$, the eigenvalues of $T_n(f)$ are asymptotically distributed as $f(x)$. This was extended in [1] to the case where $f \in L^2$. We state it here for future reference.

Theorem 2.3. *If $f(x) \in L^2$ is a real-valued function, then the eigenvalues of the (Hermitian) Toeplitz matrix $T_n(f)$ are distributed as $f(x)$.*

2.2 Multi-Level Case

The results presented in the single-level case can be extended to matrices with multiple levels. Multi-level matrices are matrices which are recursively partitioned into smaller blocks. We give a more precise characterization in the following definition that aims to avoid cumbersome notations.

Definition 2.4 (Multi-Level Matrix)**.** *A partitioned matrix A of size $n \times n$ is called a p-level matrix of level orders $n_1, n_2, \ldots, n_p$ if $n = n_1 \cdot n_2 \cdots n_p$ and if the matrix at level k can be obtained from the matrix at level $k-1$ by partitioning each block into n_k^2 square blocks of equal size (with the convention $A_0 = A$ and $n_0 = n$).*

Examples of multi-level tilings are shown in Figure 1. A possible way to index a multi-level matrix is with the use of multi-indices. Namely, for each element a, we define similarly its horizontal and vertical position by two vectors of indices

$$\begin{aligned} \bar{i} = &\quad (i_1, i_2, \ldots, i_p)\,, \\ \bar{j} = &\quad (j_1, j_2, \ldots, j_p)\,. \end{aligned} \tag{8}$$

The index $1 \leq i_k \leq n_k$ (resp. $1 \leq j_k \leq n_k$) corresponds to the horizontal (resp. vertical) position of the block at level k to which the element a belongs. We thus successively refine the position of the element a as we go through the vectors $\bar{i}$

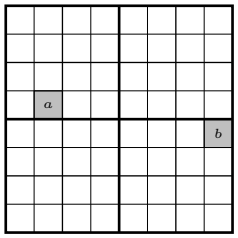


Figure 2: Multi-level indexing. The element a is indexed by the vectors $\bar{i} = (1, 2, 2)$ and $\bar{j} = (1, 1, 2)$. The element b is indexed by the vectors $\bar{i} = (2, 1, 1)$ and $\bar{j} = (2, 2, 2)$.

1	2	3	4	7	8	9	10
2	1	4	3	8	7	10	9
5	6	1	2	11	12	7	8
6	5	2	1	12	11	8	7
13	14	15	16	1	2	3	4
14	13	16	15	2	1	4	3
17	18	13	14	5	6	1	2
18	17	14	13	6	5	2	1

Figure 3: Schematic representation of a multi-level Toeplitz matrix with $p = 3$, $n_1 = 2$, $n_2 = 2$ and $n_3 = 3$.

and $\bar{j}$. Based on these vectors, the position of the element a can be expressed as

$$\begin{aligned} i = &\quad (i_1 - 1)\tilde{n}_1 + \ldots + (i_{p-1})\tilde{n}_{p-1} + i_p\,, \\ j = &\quad (j_1 - 1)\tilde{n}_1 + \ldots + (j_{p-1})\tilde{n}_{p-1} + j_p \end{aligned} \tag{9}$$

where $\tilde{n}_k = n/(n_1 \cdot n_2 \cdots n_k)$. Examples of multi-level indexing is shown in Figure 2.

Let $\bar{k} = (k_1, k_2 \ldots, k_p)$ and $\bar{n} = (n_1, n_2 \ldots, n_p)$. We define $\bar{1} \equiv (1, 1, \ldots, 1)$ and the inequality $\bar{k} \leq \bar{n}$ to mean that $k_l \leq n_l$ for $l = 1, 2, \ldots, p$. With these notations, a multi-level Toeplitz matrix can be defined similarly to the single-level case. A p-level Toeplitz matrix $T_{\bar{n}}$ of size $n \times n$ with $n = n_1 \cdot n_2 \cdots n_p$ is a matrix such that $(T_{\bar{n}})_{\bar{i},\bar{j}} = t_{\bar{j}-\bar{i}}$ for $\bar{1} \leq \bar{i}, \bar{j} \leq \bar{n}$ where $(T_{\bar{n}})_{\bar{i},\bar{j}}$ denotes the element indexed by the vectors $\bar{i}$ and $\bar{j}$. It is thus completely characterized by its generating sequence $\{t_{\bar{k}}\}_{\bar{k} \in \mathbb{Z}^p}$. A multi-level Toeplitz matrix is thus a multi-level matrix whose level k is Toeplitz with respect to level $k+1$. Such a matrix is schematically represented in Figure 3. We can naturally associate a multi-level Toeplitz matrix $T_{\bar{n}}$ with the multi-dimensional Fourier series $f(x_1, x_2, \ldots, x_p)$

defined as

$$f(x_1, x_2, \ldots, x_p) \sim \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} \cdots \sum_{k_p \in \mathbb{Z}} t_{(k_1,k_2,\ldots,k_p)} e^{i(k_1 x_1 + k_2 x_2 + \ldots + k_p x_p)} \qquad (10)$$

which is 2π-periodic with respect to each argument. In this case, we denote it $T_{\bar{n}}(f)$. Using short-hand notations, we can more compactly write

$$f(\bar{x}) \sim \sum_{\bar{k} \in \mathbb{Z}^p} t_{\bar{k}} e^{i\langle \bar{k}, \bar{x} \rangle} \qquad (11)$$

where $\bar{x} = (x_1, x_2 \ldots, x_p)$ and $\langle ., . \rangle$ denotes the Euclidian scalar product. Again, nothing is said about the convergence of the above series. At this point, we can extend Definitions 2.1 and 2.2 of Section 2.1 to the multi-level case.

Definition 2.5 (Equally Distributed Multi-Level Sequences)**.** *Two multi-level sequences of real numbers $\{\lambda_{\bar{k}}^{(\bar{n})}\}$ and $\{\mu_{\bar{k}}^{(\bar{n})}\}$ are said to be equally distributed if for any continuous function F with finite support*

$$\lim_{\bar{n} \to \infty} \sum_{\bar{1} \leq \bar{k} \leq \bar{n}} \left[F\left(\lambda_{\bar{k}}^{(\bar{n})}\right) - F\left(\mu_{\bar{k}}^{(\bar{n})}\right) \right] = 0. \qquad (12)$$

We denote $\lambda^{(\bar{n})} \sim \mu^{(\bar{n})}$.

Definition 2.6 (Asymptotic Multi-Level Distribution)**.** *A multi-level sequence of real numbers $\{\lambda_{\bar{k}}^{(\bar{n})}\}$ is said to have an asymptotic (multi-level) distribution if there is a function $f(x_1, x_2, \ldots, x_p)$, Lebesgue-integrable on the p-dimensional cube $[0, 2\pi]^p$ and 2π-periodic with respect to each argument, such that for any continuous function F with finite support*

$$\begin{aligned} &\lim_{\bar{n} \to \infty} \frac{1}{n_1 \cdot n_2 \cdots n_p} \sum_{\bar{1} \leq \bar{k} \leq \bar{n}} F\left(\lambda_{\bar{k}}^{(\bar{n})}\right) \\ &= \frac{1}{(2\pi)^p} \int_0^{2\pi} \int_0^{2\pi} \cdots \int_0^{2\pi} F\left(f(x_1, x_2, \ldots, x_p)\right) dx_1 dx_2 \cdots dx_p \,. \end{aligned} \qquad (13)$$

The sequence $\{\lambda_{\bar{k}}^{(\bar{n})}\}$ is said to be distributed as $f(x_1, x_2 \ldots, x_p)$ and we denote $\lambda^{(\bar{n})} \sim f(x_1, x_2 \ldots, x_p)$.

The notation $\bar{n} \to \infty$ means that $n_k \to \infty$ for all $1 \leq k \leq p$. Using Definitions 2.5 and 2.6, we can proceed exactly along the same lines as in Section 2.1. Let us denote the eigenvalues of $T_{\bar{n}}(f)$ by $\lambda_{\bar{k}}^{(\bar{n})}$. The main technicality is to construct a multi-level circulant matrix $C_{\bar{n}}(f)$ with eigenvalues $\mu_{\bar{k}}^{(\bar{n})}$ which fulfills

$$\lambda^{(\bar{n})} \sim \mu^{(\bar{n})} \quad \text{and} \quad \mu^{(\bar{n})} \sim f(x_1, x_2, \ldots, x_p). \qquad (14)$$

A multi-level circulant is characterized by $(C_{\bar{n}})_{\bar{i},\bar{j}} = c_{\bar{j}-\bar{i}(\mathrm{mod}\,\bar{n})}$ where $\bar{j} - \bar{i}(\mathrm{mod}\,\bar{n}) \equiv (j_1 - i_1(\mathrm{mod}\, n_1), j_2 - i_2(\mathrm{mod}\, n_2), \ldots, j_p - i_p(\mathrm{mod}\, n_p))$. In other

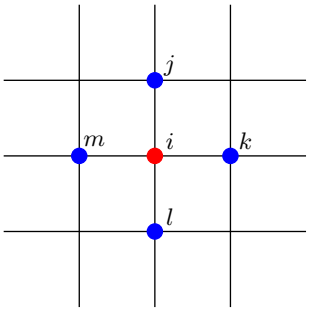


Figure 4: Nearest neighbor averaging. The pixel i is obtained by computing the average value of its nearest neighbors j,k,l and m.

words, the matrix at level k is circulant with respect to level $k+1$. As in the single-level case, "simple" and "Cesàro" multi-level circulant matrices can be built [1] and shown to both fulfill conditions (14). This allows to state the following theorem [1] which is the multi-level counterpart to Theorem 2.3.

Theorem 2.7. *If $f(x_1, x_2, \ldots, x_p) \in L^2$ is a real-valued function, then the eigenvalues of the p-level (Hermitian) Toeplitz matrix $T_{\bar{n}}(f)$ are distributed as $f(x_1, x_2, \ldots, x_p)$.*

2.3 Example: Nearest Neighbor Averaging

We now illustrate the theory explained previously by means of a simple example. We define the 2-level Hermitian Toeplitz matrix $T_{\bar{n}}$ of size $m^2 \times m^2$ as

$$(T_{\bar{n}})_{i_1,i_2,j_1,j_2} = \begin{cases} 1 & \text{if } |i_1 - j_1| = 1 \text{ and } i_2 = j_2, \\ 1 & \text{if } |i_2 - j_2| = 1 \text{ and } i_1 = j_1, \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

for $1 \leq i_1, i_2, j_1, j_2 \leq m$. The parameters of this multi-level matrix are thus $p = 2$ and $n_1 = n_2 = m$. A possible scenario where we might encounter such a matrix is the following. Consider an image of size $m \times m$ where the pixel i has coordinate (i_1, i_2). Suppose that we represent this image by a vector of length m^2 whose coefficients are obtained by scanning the image row-wise. Applying the matrix $T_{\bar{n}}$ on this vector amounts (up to a normalization factor) to replace the value of the pixel i by the average value of its nearest neighbors, as shown in Figure 4. In other words, it amounts to filter the image with a two-dimensional low-pass filter whose frequency response is given by the two-dimensional Fourier series $f(x_1, x_2)$ derived hereafter. The corresponding two-level circulant matrix amounts to consider the nearest neighbor averaging of an image which has been wrapped around both axis. We show the effect of this averaging process on a real image in Figure 5. The matrix $T_{\bar{n}}$ is generated by the sequence $t_{\bar{k}} = t_{k_1,k_2}$

(a) (b) (c)

Figure 5: Nearest neighbor averaging on a real image. (a) The original 64×64 image. (b) The image after the first transformation. (c) The image after the second transformation.

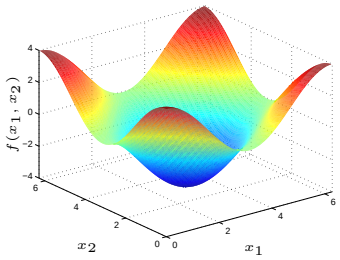


Figure 6: Fourier series $f(x_1, x_2)$ corresponding to the 2-level Hermitian Toeplitz matrix $T_{\bar{n}}$.

where

$$t_{k_1,k_2} = \begin{cases} 1 & \text{if } (k_1,k_2) \in \{(-1,0),(0,-1),(1,0),(0,1)\}, \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Thus, the corresponding two-dimensional Fourier series $f(x_1, x_2)$ exists, is continuous and bounded. It can easily be computed as

$$\begin{aligned} f(x_1,x_2) &= \sum_{k_1\in\mathbb{Z}} \sum_{k_2\in\mathbb{Z}} t_{k_1,k_2} e^{i(k_1x_1+k_2x_2)} \\ &= 2\cos x_1 + 2\cos x_2. \end{aligned} \tag{17}$$

We plot it over one period in Figure 6. Denoting the eigenvalues of $T_{\bar{n}}(f)$ by

$\lambda_{\bar{j}}^{\bar{n}}$, we can compute the moments m_k of the limiting eigenvalue distribution as

$$\begin{aligned}
m_k &= \lim_{\bar{n}\to\infty} \frac{1}{n_1 \cdot n_2} \sum_{\bar{1}\le\bar{j}\le\bar{n}} \left(\lambda_{\bar{j}}^{\bar{n}}\right)^k \\
&\overset{(1)}{=} \frac{1}{(2\pi)^2} \int_0^{2\pi}\int_0^{2\pi} f(x_1,x_2)^k dx_1 dx_2 \\
&= \frac{1}{(2\pi)^2} \int_0^{2\pi}\int_0^{2\pi} \left(2\cos x_1 + 2\cos x_2\right)^k dx_1 dx_2 \\
&\overset{(2)}{=} \frac{2^k}{(2\pi)^2} \int_0^{2\pi}\int_0^{2\pi} \sum_{l=0}^{k} \binom{k}{l} \cos^l(x_1)\cos^{k-l}(x_2)\, dx_1 dx_2 \\
&\overset{(3)}{=} \frac{2^k}{(2\pi)^2} \sum_{l=0}^{k} \binom{k}{l} \int_0^{2\pi} \cos^l(x_1) dx_1 \int_0^{2\pi} \cos^{k-l}(x_2) dx_2
\end{aligned}$$

where (1) follows from Theorem 2.7 using $F(x) = x^k$, (2) from the binomial formula, (3) from the linearity of the integrals. When k is odd, the above summation is easily seen to be zero owing to the parity of the cosine function. Thus $m_{2k+1} = 0$ for all $k \in \mathbb{Z}$. In the even case, m_{2k} can be computed as

$$\begin{aligned}
m_{2k} &= \frac{2^{2k}}{(2\pi)^2} \sum_{l=0}^{k} \binom{2k}{2l} \int_0^{2\pi} \cos^{2l}(x_1) dx_1 \int_0^{2\pi} \cos^{2(k-l)}(x_2) dx_2 \\
&\overset{(1)}{=} \sum_{l=0}^{k} \binom{2k}{2l} \binom{2l}{l} \binom{2k-2l}{k-l} \\
&= \sum_{l=0}^{k} \frac{(2k)!}{(l!)^2[(k-l)!]^2} \\
&\overset{(2)}{=} \binom{2k}{k}^2
\end{aligned}$$

where (1) follows from the fact that [3, Homework 7]

$$\frac{1}{2\pi} \int_0^{2\pi} \cos^{2l}(x_1) dx_1 = \frac{1}{2^{2l}} \binom{2l}{l}$$

and (2) from the formula

$$\sum_{l=0}^{k} \binom{k}{l}^2 = \binom{2k}{k}.$$

The moments m_k are thus given by

$$m_k = \begin{cases} \binom{2k}{k}^2 & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases} \tag{18}$$

Numerical evaluation shows that the first six even moments are: 1, 4, 36, 400, 4900 and 63504. We can also try to derive the corresponding limiting eigenvalue distribution $F(t)$ which can be computed as

$$F(t) = \frac{1}{(2\pi)^2} \int_{f(x_1,x_2)\leq t} dx_1 dx_2 = \frac{1}{\pi^2} \int_{f(x_1,x_2)\leq t} dx_1 dx_2 \qquad (19)$$

for $t \in [-4,4]$ where in the second integral $(x_1,x_2) \in [0,\pi]^2$. The second equality follows from obvious arguments about the symmetry of $f(x_1,x_2) = 2\cos x_1 + 2\cos x_2$, as can be inferred from Figure 6. Assume first that $t \in [-4,0]$. For all $x_1 \in [\arccos(1+t/2), \pi]$, there exists a unique $x_2 \in [\arccos(t/2-\cos x_1), \pi]$ such that $f(x_1,x_2) = t$. We can thus write

$$\begin{aligned} \frac{1}{\pi^2} \int_{f(x_1,x_2)\leq t} dx_1 dx_2 &= \frac{1}{\pi^2} \int_{\arccos(1+t/2)}^{\pi} \int_{\arccos(t/2-\cos x_1)}^{\pi} dx_2 dx_1 \\ &= \frac{1}{\pi^2} \int_{\arccos(1+t/2)}^{\pi} \left(\pi - \arccos\left(\frac{t}{2} - \cos x_1\right)\right) dx_1. \end{aligned}$$

Unfortunately, the above integral does not seem to be solvable in closed-form. Nevertheless, we can derive the underlying probability density function (pdf) $p(t)$ using the derivation formula

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(t,s) ds = f(t,b(t))b'(t) - f(t,a(t))a'(t) + \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(t,s) ds.$$

This yields

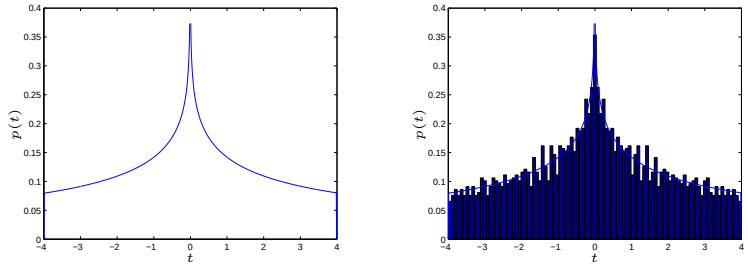
$$\begin{aligned} p(t) &= \frac{d}{dt} F(t) \\ &= \frac{1}{\pi^2} \int_{\arccos(1+t/2)}^{\pi} \frac{1}{2\sqrt{1-\left(\frac{t}{2}-\cos x_1\right)^2}} dx_1 \\ &= \frac{1}{2\pi^2} \int_{-1-t/4}^{1+t/4} \frac{1}{\sqrt{\left(1-\left(\frac{t}{4}-x_1\right)^2\right)\left(1-\left(\frac{t}{4}+x_1\right)^2\right)}} dx_1 \end{aligned}$$

where the last equality follows from the change of variable $x_1 \rightarrow \cos(x_1) - t/4$. Using tabulated formulas [4], the above integration yields for $t \in (-4,0)$

$$p(t) = -\frac{4}{\pi^2(t-4)} K\left(\frac{(t+4)^2}{(t-4)^2}\right)$$

where K denotes the complete elliptic integral of the first kind given by

$$K(t) = \int_0^{\pi/2} \frac{1}{\sqrt{(1-t\sin^2\theta)}} d\theta.$$



(a) (b)

Figure 7: Probability density function $p(t)$ of the limiting eigenvalue distribution of the 2-level Hermitian Toeplitz matrix $T_{\bar{n}}(f)$. (a) Theoretical result. (b) Experimental result for $m = 50$.

In particular, $\lim_{t\to 0} K(t) = \pi/2$ and $\lim_{t\to 1} K(t) = \infty$, hence

$$\lim_{t\to -4^+} p(t) = \frac{1}{4\pi}$$

and

$$\lim_{t\to 0^-} p(t) = \infty.$$

The pdf thus exhibits a discontinuity in -4 and 0. Note that $T_{\bar{n}}(f)$ has the same eigenvalues as $-T_{\bar{n}}(f)$. The pdf is thus symmetric around zero and we can deduce that for all $t \in (-4, 0) \cup (0, 4)$,

$$p(t) = -\frac{4}{\pi^2(|t|-4)} K\left(\frac{(|t|+4)^2}{(|t|-4)^2}\right). \tag{20}$$

The pdf $p(t)$ is shown in Figure 7. Note that the moments m_k correspond to the squared moments of an arcsin law whose pdf is given by

$$p_\nu(t) = \frac{1}{\pi\sqrt{4-t^2}}\, 1_{\{|t|<2\}}\,. \tag{21}$$

Since these moments satisfy Carleman's condition, our density $p(t)$ corresponds to the pdf of the random variable $Z = XY$, where X and Y are two independent random variables distributed according to p_ν.

Let us now consider another matrix $R_{\bar{n}}$ obtained by independently choosing the non-zero coefficients of $T_{\bar{n}}$ to be equal to 1 with probability p and -1 with probability $1-p$. We nevertheless force the matrix to be Hermitian. For $p = 0$, $R_{\bar{n}} = T_{\bar{n}}$ and for $p = 1$, $R_{\bar{n}} = -T_{\bar{n}}$. Thus, for both cases, the pdf is well known and is given by (20). However, for a general value of p, $R_{\bar{n}}$ is not a multi-level Toeplitz and the corresponding pdf seems difficult to derive analytically. We

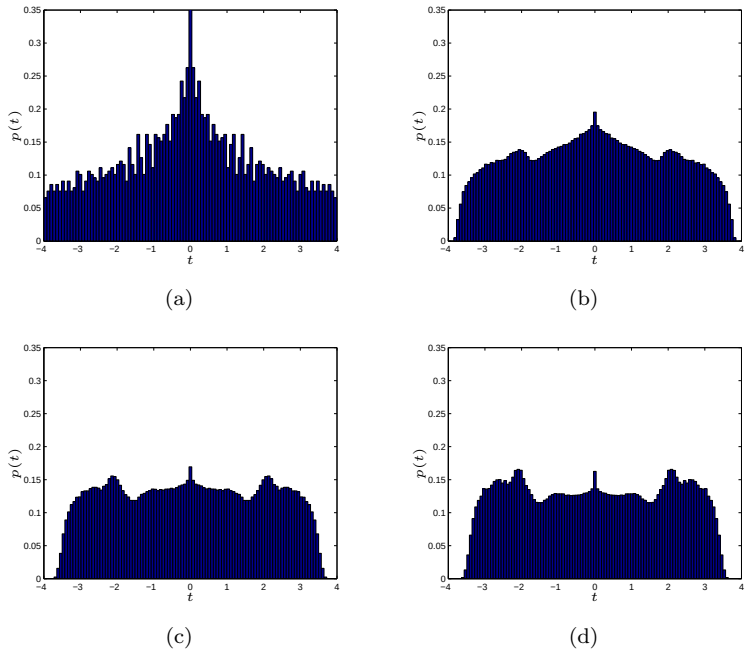
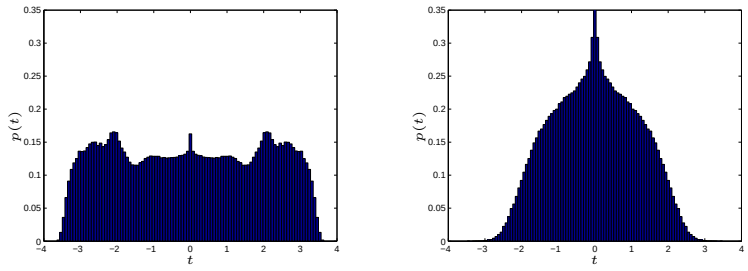


Figure 8: Probability density function $p(t)$ of the limiting eigenvalue distribution of the 2-level random Hermitian matrix $R_{\bar{n}}$. Experimental result for $m = 50$ and (a) $p = 0$, (b) $p = 0.1$ (c) $p = 0.2$ and (d) $p = 0.5$. The graphs have been averaged over 1000 realizations.



(a) (b)

Figure 9: Probability density function $p(t)$ of the limiting eigenvalue distribution of the 2-level random Hermitian matrix $R_{\bar{n}}$. Experimental result for $m = 50$ and (a) a binomial distribution of parameter $p = 1/2$ and (b) a zero mean Gaussian distribution with variance $1/3$. The graphs have been averaged over 1000 realizations.

show in Figure 8 the empirical pdfs obtained for different values of p. Note that $R_{\bar{n}}$ exhibits the same behavior as $T_{\bar{n}}$, i.e. $R_{\bar{n}}$ and $-R_{\bar{n}}$ have the same eigenvalues. Hence, the pdfs obtained with p and $1-p$ will be the same and we can restrict our analysis to $p \in [0, 0.5]$. It also interesting to compare the graphs obtained in the case where the non-zero elements of the matrix are distributed according to two different distributions. In Figure 9, we plot the pdfs corresponding to the binomial distribution considered so far, with $p = 1/2$, and a zero mean Gaussian distribution with variance $1/3$. With these parameters, the two distributions have mean 0 and variance $1/3$. The different curves obtained suggest that there is no universal law, such as the Wigner's law, for this type of matrices.

3 Low-Rank Correction

The results presented in Section 2 are based on the assumption that the multi-dimensional Fourier series f belongs to L^2. In this section, we show [5] that these results still hold if f is in L^1. In itself, this extension is only of formal purpose. What is interesting in the L^1 case however, is that the approach adopted to prove the claim is based on low-rank correction matrices. More precisely, our goal is to prove that, given a p-level Toeplitz matrix $T_{\bar{n}}(f)$ with $f \in L^1$, the conditions (14) are satisfied for some p-level circulant matrix $C_{\bar{n}}(f)$.

Let $\|A\|_1$, $\|A\|_2$ and $\|A\|_F$ denote respectively the L^1, L^2 and Frobenius

norms of the $n \times n$ matrix A. These norms are defined as

$$\|A\|_1 = \sup_{x \in \mathbb{C}^n : x \neq 0} \frac{\|Ax\|}{\|x\|}, \tag{22}$$

$$\|A\|_2 = \sqrt{\frac{1}{n} \operatorname{tr}(A^* A)}, \tag{23}$$

$$\|A\|_F = \sqrt{\operatorname{tr}(A^* A)} = \sqrt{n} \|A\|_2 . \tag{24}$$

The eigenvalues of $T_{\bar{n}}(f)$ and $C_{\bar{n}}(f)$ are denoted $\lambda_{\bar{k}}^{(\bar{n})}$ and $\mu_{\bar{k}}^{(\bar{n})}$, respectively. The condition $\mu^{(\bar{n})} \sim f(x_1, x_2, \ldots, x_p)$ has been proved for $f \in L^1$ in [6] using a Cesàro multi-level circulant matrix. It was also shown in [6] that the following holds

$$f \in L^2 \quad \Rightarrow \quad \|T_{\bar{n}} - C_{\bar{n}}\|_F^2 = o(n) \quad \Rightarrow \quad \lambda^{(\bar{n})} \sim \mu^{(\bar{n})}. \tag{25}$$

Unfortunately, the properties (25) are no longer valid when $f \in L^1$. In other words, the multi-level Toeplitz $T_{\bar{n}}(f)$ is not close in norm to the multi-level circulant matrix $C_{\bar{n}}(f)$. The derivation that follows aims at showing that the above properties can be saved if we consider the closeness a bit differently. The next theorem [7] shows that $\lambda^{(\bar{n})} \sim \mu^{(\bar{n})}$ also holds for $f \in L^1$ if $T_{\bar{n}}(f)$ and $C_{\bar{n}}(f)$ are close after a "low-rank correction", i.e. $\|T_{\bar{n}} - C_{\bar{n}} + \Delta_{\bar{n}}\|_F^2 = o(n)$ for some correction matrix $\Delta_{\bar{n}}$ with $\operatorname{rank}(\Delta_{\bar{n}}) = o(n)$. Here again, we will assume that f is real-valued. The complex-valued case follows by splitting the matrix into real and imaginary parts and can be found in [5].

Theorem 3.1. *Suppose that for any $\epsilon > 0$, there exists a matrix $\Delta_{\bar{n}}(\epsilon)$ with rank smaller than ϵn such that*

$$\|T_{\bar{n}} - C_{\bar{n}} + \Delta_{\bar{n}}(\epsilon)\|_F^2 \leq \epsilon n$$

for all $\bar{n}$ with sufficiently large components. Then,

$$\lambda^{(\bar{n})} \sim \mu^{(\bar{n})}.$$

According to Theorem 3.1, it thus remains to show that such a low-rank correction matrix exists in our case. To this end, we prove the following theorem [5] by skipping some technical details in order to gain in clarity.

Theorem 3.2. *Let $f \geq 0$ and $f \in L^1$. Then, for an arbitrary $\epsilon > 0$, there exists a matrix $\Delta_{\bar{n}}(\epsilon)$ with rank smaller than ϵn such that*

$$\|T_{\bar{n}} - C_{\bar{n}} + \Delta_{\bar{n}}(\epsilon)\|_F^2 \leq \epsilon n$$

for all $\bar{n}$ with sufficiently large components.

Proof: For convenience, we will write $f(x_1, x_2, \ldots, x_p)$ as $f(\bar{x})$. The key idea of this proof is to split the generating Fourier series $f(\bar{x}) \in L^1$ as

$$f_M(\bar{x}) = \begin{cases} f(\bar{x}) & \text{for } f(\bar{x}) \leq M, \\ M & \text{for } f(\bar{x}) > M \end{cases}$$

and $r_M(\bar{x}) = f(\bar{x}) - f_M(\bar{x})$ for some $M > 0$. Since $f_M(\bar{x}) \in L^\infty$, the associated multi-level Toeplitz $T_{\bar{n}}(f_M)$ and circulant $C_{\bar{n}}(f_M)$ are close in norm, i.e.

$$\|T_{\bar{n}}(f_M) - C_{\bar{n}}(f_M)\|_F^2 \leq \epsilon n \tag{26}$$

for all $\bar{n}$ with sufficiently large components. Let us consider $r_M(\bar{x})$. We have that

$$\begin{aligned}
\|r_M\|_1 &= \int_{\{\bar{x}\in[0,2\pi]^p : f(\bar{x})>M\}} |f(\bar{x}) - M| d\bar{x} \\
&\leq \int_{\{\bar{x}\in[0,2\pi]^p : f(\bar{x})>M\}} |f(\bar{x})| d\bar{x} \\
&= \int_{[0,2\pi]^p} |f(\bar{x})| 1_{\{f(\bar{x})>M\}} d\bar{x}.
\end{aligned}$$

Now, $|f(\bar{x})|1_{\{f(\bar{x})>M\}} \leq |f(\bar{x})|$ with $f \in L^1$, thus by the dominated convergence theorem

$$\lim_{M\to\infty} \int_{[0,2\pi]^p} |f(\bar{x})| 1_{\{f(\bar{x})>M\}} d\bar{x} = \int_{[0,2\pi]^p} \lim_{M\to\infty} |f(\bar{x})| 1_{\{f(\bar{x})>M\}} d\bar{x} = 0.$$

Hence, $\lim_{M\to\infty} \|r_M\|_1 = 0$ and $\lim_{M\to\infty} |\{\bar{x} \in [0,2\pi]^p : f(\bar{x}) > M\}| = 0$. It can be shown [5] that under these conditions, for large enough M, we can write $T_{\bar{n}}(r_M)$ and $C_{\bar{n}}(r_M)$ as

$$\begin{aligned}
T_{\bar{n}}(r_M) &= T_{1,\bar{n}}(r_M) + T_{2,\bar{n}}(r_M), \\
C_{\bar{n}}(r_M) &= C_{1,\bar{n}}(r_M) + C_{2,\bar{n}}(r_M)
\end{aligned}$$

where

$$\begin{aligned}
\max\{\|T_{1,\bar{n}}(r_M)\|_2, \|C_{1,\bar{n}}(r_M)\|_2\} &\leq \epsilon, &(27)\\
\max\{\operatorname{rank} T_{2,\bar{n}}(r_M), \operatorname{rank} C_{2,\bar{n}}(r_M)\} &\leq \epsilon n. &(28)
\end{aligned}$$

The way we perform this split is quite technical and we will not discuss this matter here. The important point is that both $T_{\bar{n}}(r_M)$ and $C_{\bar{n}}(r_M)$ can be divided into one low-norm component (indexed by 1) and one low-rank component (indexed by 2). The low-rank components allow us to define our small-rank correction matrix $\Delta_{\bar{n}}(\epsilon)$ as

$$\Delta_{\bar{n}}(\epsilon) \equiv C_{2,\bar{n}}(r_M) - T_{2,\bar{n}}(r_M),$$

which satisfies $\operatorname{rank} \Delta_{\bar{n}} \leq 2\epsilon n$ by equation (28). We can thus write

$$\begin{aligned}
&\|T_{\bar{n}}(f) - C_{\bar{n}}(f) + \Delta_{\bar{n}}(\epsilon)\|_F^2 \\
&\quad= \| \left(T_{\bar{n}}(f_M) - C_{\bar{n}}(f_M)\right) + \left(T_{1,\bar{n}}(r_M) - C_{1,\bar{n}}(r_M)\right) \|_F^2 \\
&\quad\leq 2\|T_{\bar{n}}(f_M) - C_{\bar{n}}(f_M)\|_F^2 + 2\|T_{1,\bar{n}}(r_M) - C_{1,\bar{n}}(r_M)\|_F^2 \\
&\quad\leq 2\epsilon n + 4\epsilon^2 n.
\end{aligned}$$

The last inequality follows from equation (26) and the fact that, using the triangle inequality and equation (27), we have

$$\begin{aligned}\|T_{1,\bar{n}}(r_M) - C_{1,\bar{n}}(r_M)\|_F^2 &= n\|T_{1,\bar{n}}(r_M) - C_{1,\bar{n}}(r_M)\|_2^2 \\ &\leq n\|T_{1,\bar{n}}(r_M)\|_2^2 + n\|C_{1,\bar{n}}(r_M)\|_2^2 \\ &\leq 2n \max\left\{\|T_{1,\bar{n}}(r_M)\|_2^2, \|C_{1,\bar{n}}(r_M)\|_2^2\right\} \\ &\leq 2n\epsilon^2.\end{aligned}$$

The proof follows. □

Theorem 3.2 can be easily extended to any real-valued f by appropriate splitting into its positive and negative parts. In the light of the results obtained, we can thus conclude this discussion with the following theorem.

Theorem 3.3. *If $f(x_1, x_2, \ldots, x_p) \in L^1$ is a real-valued function, then the eigenvalues of the p-level (Hermitian) Toeplitz matrix $T_{\bar{n}}(f)$ are distributed as $f(x_1, x_2, \ldots, x_p)$.*

4 Conclusions

In this report, we have shown how the notions of Toeplitz and circulant matrices can be extended to the multi-dimensional case by considering multi-level matrices and multi-level indexing. We have provided a simple example to illustrate these concepts. Finally, the idea of low-rank correction was explained in order to prove asymptotic equivalence results between Toeplitz and circulant matrices for a larger class of Fourier series.

References

[1] E. E. Tyrtyshnikov, "A unifying approach to some old and new theorems on distribution and clustering," *Linear Algebra and Its Applications*, no. 232, pp. 1–43, 1996.

[2] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. Chelsea Publishing Company, New-York, 2nd ed., 1984.

[3] O. Lévêque, "Random matrices and communication systems." Lecture Notes, EPFL, 2005.

[4] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press, London, 5th ed., 1994.

[5] E. E. Tyrtyshnikov and N. L. Zamarashkin, "Spectra of multilevel toeplitz matrices: Advanced theory via simple matrix relationships," *Linear Algebra and Its Applications*, no. 270, pp. 15–27, 1998.

[6] E. E. Tyrtyshnikov, "New theorems on the distribution of eigen and singular values of multilevel Toeplitz matrices (in russian)," *Dokl. Akad. Nauk.*, vol. 3, no. 333, pp. 300–302, 1999.

[7] E. E. Tyrtyshnikov, "Influence of matrix operations on the distribution of eigenvalues and singular values of Toeplitz matrices," *Linear Algebra and Its Applications*, no. 207, pp. 225–249, 1994.