PRECONDITIONED RANDOM TOEPLITZ OPERATORS

W. F. KE†, K. F. LAI‡ ¶, AND N. C. WONG§

Abstract. The solution of Hermitian positive definite random Toeplitz systems $Ax = b$ by the preconditioned conjugate gradient method for the Strang circulant preconditioner is studied. We established the foundation for this method by extending the work of Brown-Halmos on Toeplitz operators and Grenander-Szegö on Toeplitz form to random Toeplitz operators.

Key words. Random Toeplitz operator, random circulant matrix, preconditioned conjugate gradient method

AMS subject classifications. 65F10, 65F15, 60B20

1. Introduction. The aim of this paper is to show that we can precondition a random Toeplitz operator to yield fast convergence when conjugate gradient method is used for computation.

Preconditioned conjugate gradient (PCG) for solving finite dimensional linear system $Ax = b$ is a well established technique in numerical linear algebra (see for example [Axe 94], [GV 96]). When $A$ is a Toeplitz matrix very efficient preconditioners have been found (see [Str 86], [CJ 07]). Conjugate gradient methods also works for Hilbert spaces (see [Dan 70], [Pan 04]).

Since the innovative work of Wigner [Wig 55] random matrices with independent identically distributed (iid) random variables as entries have been intensively studied (see [BS 10], [AGZ 09], [Tao 12], [PS 11]) with many useful applications (see [Meh 04], [BCC 09]). The special class of random Toeplitz matrices, in particular their spectral measure have been studied by Bryc and others ([BDJ 06]). But there has been no work on their eigenvalue distributions in relation to the properties of a generating function as given in the classic work of Grenander and Szegö [GS 58].

On the other hand we can consider random matrices as random linear operators as given in [Sko 84]. This is the point of view we shall take. We shall establish for random Toeplitz operators the theoretical background used by Raymond Chan in his important work on circulant preconditioners for Toeplitz matrices, see in particular [ChS 89], [ChR 89], [ChR 91].

The paper is divided into three parts. In the first part we shall formulate and prove for random matrices some theorems on the distribution of eigenvalues of random matrices which are standard in the case of number matrices. We think it is useful to have these theorems written down and they will be used in subsequent papers on applications to numerical computations. In part two we first extend the work of Brown and Halmos [BH 63] on Toeplitz matrix to random Toeplitz operators. And then we extend the results of Grenander and Szegö [GS 58] on Toeplitz forms to random Toeplitz operators. These results are what we need for the extension of Raymond Chan’s results to random Toeplitz operators. In part three we apply the results of part two to show that that we can use the Strang’s circulant to precondition a random Toeplitz operator and give some numerical examples.

Part I.

†Department of Mathematics and National Center for Theoretical Sciences, National Cheng Kung University, Tainan, Taiwan (wfke@mail.ncku.edu.tw)
‡School of Mathematical Sciences, Capital Normal University, Beijing 100048, China (kinglaihonkon@gmail.com)
§Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung, Taiwan (wong@math.nsysu.edu.tw)
¶We thank the National Center of Theoretical Science for the support for a short visit during which this work is done.
2. Ordering eigenvalues. In all standard discussions on the distribution of eigenvalues of a matrix one begins with ordering the eigenvalues as

\[ \lambda_1 \leq \cdots \leq \lambda_n \]

In case of a matrix with functions as entries the "eigenvalues" are themselves functions and it may not be possible to arrange these eigenvalue-functions in an order. So we begin with this problem.

There are other possible ways to deduce some of our results but our proofs are most direct and constructive. They also set up the favours of our program.

We fix a \( \sigma \)-finite measure space \((\Omega, \mu)\). By a \( n \times n \) random matrix \( A = (a_{ij}) \) on \( \Omega \) we shall mean simply that the entries \( a_{ij} : \Omega \to \mathbb{C} \) are random variables on \( \Omega \). We shall not impose the iid condition in the beginning. In this sense we consider \( A \) as a random linear operator \( A : \Omega \times \mathbb{C}^n \to \mathbb{C}^n \) given by \( A(\omega, v) = (a_{ij}(\omega))v \).

We begin with stating a simple lemma in measure theory.

**Lemma 1.** Let \( \Omega = \bigcup_{i \geq 1} \Omega_i \) be a covering of \( \Omega \) by at most countably many measurable subsets. Set

\[ g_1 := 1_{\Omega_1}, \]

and

\[ g_i := 1_{\Omega_i}(1 - g_1 - \cdots - g_{i-1}), \quad i = 2, 3, \ldots. \]

Then

\[ 1 = \sum_i g_i \]

is a measurable partition of unity of \( \Omega \) subordinating to the (an, arbitrary but fixed, ordered) family \( \{\Omega_1, \Omega_2, \ldots\} \). More precisely, \( g_i = 1_{\Omega_i \setminus \bigcup_{j<i} \Omega_j} \), and \( g_i g_j = 0 \) for all \( i \neq j \). \( \square \)

**Lemma 2.** Let \( u_1, u_2, \ldots, u_m \) be measurable functions from \( \Omega \) into \( \mathbb{R}^n \) such that \( \{u_1(x), u_2(x), \ldots, u_m(x)\} \) is an orthonormal subset of \( \mathbb{R}^n \) for each \( x \) in \( \Omega \). Then we can find measurable functions \( u_{m+1}, \ldots, u_n \) from \( \Omega \) into \( \mathbb{R}^n \) such that \( \{u_1(x), \ldots, u_n(x)\} \) is an orthonormal basis of \( \mathbb{R}^n \) for every \( x \) in \( \Omega \).

**Proof.** Consider the standard orthonormal basis \( \{\hat{e}_1, \ldots, \hat{e}_n\} \) consisting of constant vector-valued functions \( \hat{e}_k(x) = e_k \in \mathbb{R}^n, \forall x \in X \). For each choice of \( n-m \) of them, \( \gamma := \{\hat{e}_{\gamma_1}, \ldots, \hat{e}_{\gamma_{n-m}}\} \), the subset \( X_\gamma \) of \( x \) in \( \Omega \) at which

\[ \{u_1(x), \ldots, u_m(x), \hat{e}_{\gamma_1}(x), \ldots, \hat{e}_{\gamma_{n-m}}(x)\} \]

forms a basis of \( \mathbb{R}^n \) is measurable. Indeed, it is the cozero set of the measurable function

\[ \det(u_1(x), u_2(x), \ldots, u_m(x), \hat{e}_{\gamma_1}(x), \ldots, \hat{e}_{\gamma_{n-m}}(x)). \]

By the Gram-Schmidt process, we can transform the linearly independent set to an orthonormal basis

\[ \{u_1(x), u_2(x), \ldots, u_m(x), f_{m+1}(x), \ldots, f_n(x)\} \]
of $\mathbb{R}^n$ for each $x$ in $\Omega$, such that the functions $f_{m+1}^\gamma(x), \ldots, f_n^\gamma(x)$ are defined and measurable on $\Omega$. Subordinating to the measurable covering $\Omega = \bigcup \Omega_\gamma$, a measurable partition of unity is given by

$$1 = \sum_\gamma g_\gamma.$$ 

Define

$$u_{m+k}(x) := \sum_\gamma g_\gamma(x)f_{m+k}^\gamma, \quad k = 1, \ldots, n - m.$$ 

We have all such $u_{m+k}$ measurable on $\Omega$, and together with $u_1, \ldots, u_m$ they form an orthonormal family we want. \[\Box\]

**Lemma 3.** Let $a_1, a_2, \ldots, a_n$ be measurable functions from $\Omega$ into $\mathbb{R}^n$. Suppose that the random matrix $A(x) = [a_1(x) \ a_2(x) \ \ldots \ a_n(x)]$ is symmetric for each $x$ in $\Omega$. Then the spectral radius $r(A(x))$ and the maximal eigenvalue $\lambda_{\max}(A(x))$ of the random matrix $A(x)$ are both measurable functions on $\Omega$.

**Proof.** We note that all matrix norms on $\mathbb{R}^{n \times n}$ are equivalent, and they give the spectral radius by

$$r(A(x)) = \limsup \|A(x)^n\|^{1/n}, \quad \forall x \in \Omega.$$ 

If we use, for example, the Hilbert-Schmidt norm $\|A\| := (\sum_{i,j} |a_{ij}|^2)^{1/2}$, then we see that $r(A(x))$ is a measurable function on $\Omega$. Note also that $r(A(x))$ equals the operator norm of $A(x)$. Consequently, the function

$$x \mapsto r(A(x) + r(A(x))I)$$

is also measurable on $\Omega$ and gives rise to the maximal eigenvalue $\lambda_{\max}(A(x) + r(A(x))I)$ of the positive matrix $A(x) + r(A(x))I$. Since $\lambda_{\max}(A(x)) = \lambda_{\max}(A(x) + r(A(x))I) - r(A(x))$, we obtain the measurability of $\lambda_{\max}(A(\Omega))$. \[\Box\]

**Lemma 4.** Let $A(x) = [a_1(x) \ a_2(x) \ \ldots \ a_n(x)]$ be a real symmetric random matrix on $\Omega$ such that $A(x)$ has rank at most one at every $x$. Then we can order the eigenvalues $\lambda_1(x) \leq \cdots \leq \lambda_n(x)$ of $A(x)$ as measurable functions on $\Omega$ with corresponding measurable eigenvector functions $u_1(x), \ldots, u_n(x)$ which form an orthonormal basis of $\mathbb{R}^n$ for every $x$ in $\Omega$.

**Proof.** Assume $A(x) \succeq 0$, i.e., positive (semi-definite), for every $x$ in $X$ first. Let

$$X_0 = \{x \in X : A(x) = 0\}.$$ 

On its complement, let

$$X_k = \{x \in X : a_k(x) \neq 0\}, \quad k = 1, \ldots, n.$$ 

Clearly, $\{\Omega_0, X_1, \ldots, \Omega_n\}$ forms a measurable covering of $\Omega$. Let $1 = \sum_{i=0}^n g_i$ be the measurable partition subordinating to this covering as in Lemma \[\Box\]. If $x \in \Omega_k$ for some $k > 0$ then $a_k(x)$ is an eigenvector of $A(x)$ for the maximal, and the unique positive, eigenvalue $\lambda_n(x)$ of $A(x)$. Define

$$u_n(x) = g_0(x)e_1(x) + \sum_{i=1}^n g_i \frac{a_i(x)}{\|a_i(x)\|}.$$
Then we have a measurable function from $\Omega$ into $\mathbb{R}^n$ such that $\|u_n(x)\| = 1$ everywhere on $\Omega$. By Lemma 2 we can enlarge it to have a measurable orthonormal basis $$\{u_1(x), u_2(x), \ldots, u_n(x)\}$$ on $X$. We set $\lambda_1 = \cdots = \lambda_{n-1} = 0$. Note that $\lambda_n$ vanishes on $\Omega_0$ and takes strictly positive values elsewhere.

In the general case, let $\Omega_+$ (resp. $\Omega_-$) be the subset of $\Omega$ of those $x$ at which the trace of $A(x)$ is non-negative (resp. non-positive). We divide $\Omega$ into a measurable union $\Omega = \Omega_+ \cup \Omega_-$ such that $A(x) \geq 0$ on $\Omega_+$, and $-A(x) \geq 0$ on $\Omega_-$. Applying above arguments separately on $\Omega_+$ and $\Omega_-$, and gluing the results together with Lemma 1 we obtain the assertion.

**Theorem 5.** Let $A(x) = [a_1(x) \ a_2(x) \ \ldots \ a_n(x)]$ be a real symmetric random matrix on $\Omega$. Then we can order the eigenvalues $\lambda_1(x) \leq \cdots \leq \lambda_n(x)$ of $A(x)$ as measurable functions on $\Omega$ with corresponding measurable eigenvector functions $u_1(x), \ldots, u_n(x)$ which form an orthonormal basis of $\mathbb{R}^n$ for every $x$ in $\Omega$.

**Proof.** By Lemma 4 the assertion holds for the case $n = 1$. Assume by induction the assertion is valid for all dimension less than $n$. Suppose also that $A(x) \geq 0$ everywhere on $\Omega$. By Lemma 3 we see that the maximal eigenvalue $\lambda_n(x)$ of $A(x)$ is a measurable function on $\Omega$. We are going to show that we can find an associated measurable eigenvector function $u_n(x)$ of $A(x)$ of norm one everywhere on $\Omega$.

For each $x$ in $\Omega$, let $$1 = P_n(x) + P_{n-1}(x) + \cdots + P_{n-m}(x)$$ be the orthogonal sum of eigenprojections of $A(x)$ for distinct eigenvalues $\lambda_n(x) > \lambda_{n-1}(x) > \cdots > \lambda_{n-m}(x)$. In particular, $P_n(x)$ is the (nonzero) orthogonal projection of $\mathbb{R}^n$ onto the eigenspace of $A(x)$ for the maximal eigenvalue $\lambda_n(x)$. Note that $m = m(x)$ depends on $x$.

Let $X_0 = \{x \in X : \lambda_n(x) = 0\}$, which is a measurable set. Let $x \in X \setminus X_0$. For each $k = 1, 2, \ldots, n$, write $$e_k = P_n(x)e_k + P_{n-1}(x)e_k + \cdots + P_{n-m}(x)e_k.$$ Applying $A(x)$ on both sides repeatedly, we get $$A(x)^q e_k = \lambda_n(x)^q P_n(x)e_k + \lambda_{n-1}(x)^q P_{n-1}(x)e_k + \cdots + \lambda_{n-m}(x)^q P_{n-m}(x)e_k,$$ for $q = 1, 2, \ldots$. Consequently, $$\frac{A(x)^q e_k}{\lambda_n(x)^q} \to P_n(x)e_k, \quad \text{as } q \to \infty.$$ As a pointwise limit of measurable functions, $P_n(x)e_k$ is measurable from $X \setminus \Omega_0$ into $\mathbb{R}^n$. Consequently, its cozero set $\Omega_k = \{x \in \Omega \setminus \Omega_0 : P_n(x)e_k \neq 0\}$ is measurable for $k = 1, 2, \ldots, n$. Clearly, $\Omega = \Omega_0 \cup \Omega_1 \cup \cdots \cup \Omega_n$ is a measurable covering of $\Omega$. Using Lemma 1 we have an associated measurable partition of unity

$$1 = g_0 + g_1 + \cdots + g_n.$$ We can then define a measurable function $$u_n(x) = \frac{g_0(x)e_1 + \sum_{k=1}^n g_k(x)P_n(x)e_k}{\|g_0(x)e_1 + \sum_{k=1}^n g_k(x)P_n(x)e_k\|}.$$
which is of norm one everywhere on \( \Omega \).

By Lemma 2 we can enlarge \( \{u_n(x)\} \) to a measurable orthonormal random basis
\[
\{u_1(x), u_2(x), \ldots, u_n(x)\}
\]
of \( \mathbb{R}^n \) everywhere on \( \Omega \). Define an orthogonal matrix \( U = [u_1(x) \ u_2(x) \ \cdots \ u_n(x)] \). Then
\[
U^t(x)A(x)U(x) = \begin{pmatrix} A'(x) & 0 \\ 0 & \lambda_n(x) \end{pmatrix}, \quad \forall x \in \Omega.
\]

Here, \( A'(x) \) is an \((n-1) \times (n-1)\) positive semi-definite real matrix. By the induction hypothesis, we can order its \( n-1 \) eigenvalues
\[
\lambda_1(x) \leq \cdots \leq \lambda_{n-1}(x), \quad \forall x \in \Omega,
\]
as measurable functions on \( \Omega \). Clearly, \( \lambda_{n-1}(x) \leq \lambda_n(x) \) everywhere on \( \Omega \).

In general, we consider the random matrix \( B(x) = A(x) + r(A(x))I \geq 0 \). Let the ordered measurable eigenfunctions of \( B \) be \( g_1(x), \ldots, g_n(x) \). Then, the eigenvalue functions of \( A \) are given by \( \lambda_k(x) = g_k(x) - r(A(x)) \) for \( k = 1, \ldots, n \). Moreover, \( A \) and \( B \) share the same set of measurable eigenvector functions. \( \Box \)

We remark that with slight modification the above arguments also work for the complex case.

By Theorem 5 we can arrange the (all real) eigenvalues of an \( n \times n \) random Hermitian matrix \( A(x) \) in an increasing order
\[
\lambda_{\min}(x) = \lambda_1(x) \leq \lambda_2(x) \leq \cdots \leq \lambda_n(x) = \lambda_{\max}(x).
\]

We write \( \lambda_k(A(x)) \) if the random matrix \( A \) is to be emphasized.

The following two lemmas are direct applications of Theorem 5 and the theorems of Rayleigh-Ritz, Courant-Fischer, and Weyl.

It is now clear that the standard proofs of the theorems of Rayleigh-Ritz, Courant-Fischer and Weyl in [HJ 90], [GV 96] extend to random matrices. Together with Theorem 5 we obtain the following lemmas.

**Lemma 6.** Let \( A(x) \) be an \( n \times n \) Hermitian random matrix on \( \Omega \). Then
\[
\lambda_{\max}(x) = \max_v \frac{v(x)^*A(x)v(x)}{v(x)^*v(x)} = \max_{v(x)^*v(x)=1} v(x)^*A(x)v(x),
\]
\[
\lambda_{\min}(x) = \min_v \frac{v(x)^*A(x)v(x)}{v(x)^*v(x)} = \min_{v(x)^*v(x)=1} v(x)^*A(x)v(x), \quad \forall x \in \Omega.
\]

Here, \( v \) runs through all non-vanishing random vectors, i.e., measurable functions from \( \Omega \) into \( \mathbb{C}^n \).

In general, for \( 1 \leq k \leq n \) we have
\[
\lambda_k(x) = \min_{w_1, \ldots, w_{n-k}} \max_{v(x) \perp w_1(x), \ldots, w_{n-k}(x)} \frac{v(x)^*A(x)v(x)}{v(x)^*v(x)}, \quad \forall x \in \Omega,
\]
\[
= \max_{w_1, \ldots, w_{n-k}} \min_{v(x) \perp w_1(x), \ldots, w_{n-k}(x)} \frac{v(x)^*A(x)v(x)}{v(x)^*v(x)}, \quad \forall x \in \Omega.
\]
have a sequence \( \{ f \} \) uniformly bounded everywhere on \( \Omega \), and (3.1) holds for all \( \omega \neq 0 \) and \( j \).

**Lemma 7.** Let \( A, B \) be \( n \times n \) Hermitian random matrices on \( \Omega \). Let the eigenvalue functions \( \lambda_j(A(x)), \lambda_j(B(x)) \) and \( \lambda_j(A(x) + B(x)) \) be arranged in increasing order. Then

\[
\lambda_k(A(x)) + \lambda_1(B(x)) \leq \lambda_k(A(x) + B(x)) \leq \lambda_k(A(x)) + \lambda_B(B(x)), \quad \forall x \in X.
\]

\[ \square \]

With these lemmas we can see immediately that Cauchy Interlacing theorem for random matrices also holds. This will allow us to extends the proofs of Chan in [ChR 89].

**Theorem 8.** Let \( \hat{A}(x) \) be an \( n \times n \) Hermitian random matrix with eigenvalue functions \( \{ \hat{\lambda}_j(x) \} \) arranged in increasing order. Let \( A(x) \) be a principal random submatrix of \( \hat{A}(x) \) of order \( n-1 \) with eigenvalue functions \( \{ \lambda_j(x) \} \) arranged in increasing order. Then

\[
\hat{\lambda}_1(x) \leq \lambda_1(x) \leq \hat{\lambda}_2(x) \leq \lambda_2(x) \leq \cdots \leq \lambda_{n-2}(x) \leq \hat{\lambda}_{n-1}(x) \leq \lambda_{n-1}(x) \leq \hat{\lambda}_n(x)
\]

for all \( x \) in \( \Omega \). In general, if \( A(x) \) is an \( r \times r \) principal random submatrix of \( \hat{A}(x) \) obtained by deleting \( n-r \) rows and the corresponding columns from \( \hat{A}(x) \). Then

\[
\lambda_k(\hat{A}(x)) \leq \lambda_k(A(x)) \leq \lambda_{k+n-r}(\hat{A}(x)), \quad \forall x \in X, \ 1 \leq k \leq r.
\]

\[ \square \]

**3. Almost equidistribution.** We extend the important concept of equidistribution in analytic number theory to almost equidistribution.

Let \( (a^{(n)}_{j})_{1 \leq j \leq n+1}, (b^{(n)}_{j})_{1 \leq j \leq n+1} \) be two infinite lower triangular matrices of real random variables on \( \Omega \) such that for almost all \( \omega \) in \( \Omega \) the entries in the lower triangular matrices are uniformly bounded, i.e., there is some constant \( K(\omega) > 0 \) such that \( a^{(n)}_{j}(\omega), b^{(n)}_{j}(\omega) \) are in \( [-K(\omega), K(\omega)] \) for all admissible indices \( n, j \). We say that they are *equally distributed* if

\[
\lim_{n \to \infty} \frac{1}{n+1} \sum_{j=1}^{n+1} (f(a^{(n)}_{j}(\omega)) - f(b^{(n)}_{j}(\omega))) = 0, \ a.e. \ on \ \Omega,
\]

for any continuous function \( f \) on \( \mathbb{R} \).

**Proposition 9.** The following are equivalent conditions for a.e. uniformly bounded real infinite random lower triangular matrices \( (a^{(n)}_{j})_{1 \leq j \leq n+1} \) and \( (b^{(n)}_{j})_{1 \leq j \leq n+1} \) to be equally distributed.

(a) \( (3.1) \) holds for \( f(t) = t^k \) for all \( k = 0, 1, 2, \ldots \).

(b) \( (3.1) \) holds for \( f(t) = \log(1 + x_n t) \) for a real sequence \( x_n \) converging to zero.

**Proof.** (a) Forgetting a set of measure zero, we can assume both the random matrices are uniformly bounded everywhere on \( \Omega \), and \( (3.1) \) holds for all \( f(t) = t^k \) everywhere on \( \Omega \) for all \( k = 0, 1, 2, \ldots \). By the Stone-Weierstrass theorem, for every real continuous function \( f \) in \( C(\mathbb{R}) \) we have a sequence \( \{ p_m(f) \} \) of polynomial such that

\[
\text{sup}\{|f(x) - p_m(f)(x)| : x \in [-m, m]| < 1/m, \quad m = 1, 2, \ldots
\]
For every fixed $\omega$ in $\Omega$, we have

\[
\lim_{n \to \infty} \frac{1}{n+1} \sum_{j=1}^{n+1} (f(a_j^{(n)}(\omega)) - f(b_j^{(n)}(\omega))) = 0, \quad \forall \omega \in \Omega.
\]

(b) For (b), we recall that in [GS 58] the assumption should be that $f(x) = \log(1 + xt)$ for all real $x$ and $|x| < K^{-1}$. Indeed, to utilize the proof there, i.e. to use Vitali’s theorem for holomorphic functions ([Tit 60] p.168) we simply need (3.1) to hold for a sequence of real $x_n$ with $|x_n| < K^{-1}$ and $\{x_n\}$ has a cluster point $x$ with $|x| < K^{-1}$. In this case, the proof in page 63 of [GS 58] works.

For the sake of completeness, we include the proof here. Assume that for each occasion $z = x_m$, where $x_m$ is a sequence of real numbers with $|x_m| \leq R < K^{-1}$ and having a cluster point, we have

\[
\lim_{n \to \infty} \sum_{j=1}^{n+1} (f(a_j^{(n)}(\omega)) - f(b_j^{(n)}(\omega))) = 0
\]

holds all $\omega$ in $\Omega$ except a measurable subset of measure zero. Fix $\omega$, the quotient following the limit sign in the above displayed formula is a single-valued and analytic function of the complex variable $z$ provided that $|z| < 1/K$. It is uniformly bounded in $z$ and $n$ provided that $|z| \leq R_1$ with $R < R_1 < 1/K$. Since the sequence converges on a set $\{x_1, x_2, \ldots\}$ with cluster point in the complex open disk $B(0; R_1)$ centered at zero with radius $R_1$, Vitali’s theorem ensures that the sequence indeed converges uniformly on $B(0; R)$. Applying the Cauchy integral formula, we see that

\[
\lim_{n \to \infty} \frac{\sum_{j=1}^{n+1} (\log(1 + z a_j^{(n)}(\omega)) - \log(1 + z b_j^{(n)}(\omega)))}{n+1} = 0
\]

Then we can apply (a).

A useful special case of the previous proposition is when $K$ is constant. The obvious statements are left to the readers.

Part II.

Given a sequence $\{c_n : -\infty < n < \infty\}$ of complex numbers. We can form an $(n+1) \times (n+1)$ matrix $T_n$ whose $(i,j)$-th entry is $c_{i-j}$.

On the other hand a random Toeplitz matrix is like the $T_n$ given above except now $c_n$ is a sequence of independent identically distributed (iid) random variables with Gaussian distributions. See [BDJ 06].
Let us at the moment ignore the iid condition. Suppose we are given a sequence \( \{c_n : -\infty < n < \infty\} \) of real valued random variables on a probability space \((\Omega, \mu)\). Suppose there is a function \( f(\omega, x) \) on \( \Omega \times S^1 \) such that it has a Fourier series expansion at almost everywhere on \( \Omega \) that

\[
f(\omega, z) = \sum_{n=-\infty}^{\infty} c_n(\omega)e^{inz}, \quad \forall z \in S^1.
\]

We say that the above Fourier series is uniformly summable on \( \Omega \) if for every \( \epsilon > 0 \) there is a positive integer \( N \) such that outside a subset of \( \Omega \) of zero measure we have

\[
\sum_{|n| > N} |c_n(\omega)| < \epsilon.
\]

We shall establish the results of Szegő necessary for the analysis of Chan in [ChS 89] and [ChR 89].

4. Random Laurent operators. Let \((\Omega, \mu)\) be a measure space, \(H\) be a Hilbert space and \(\mathcal{L}(H)\) be the \(C^*\) algebra of bounded linear operators on \(H\). By a random linear operator we mean a map \( \Phi : \Omega \to \mathcal{L}(H) \) such that for every pair \( x, y \in H \) the map \( \Omega \to \mathbb{C} : \omega \mapsto \langle \Phi(\omega)x, y \rangle \) is measurable. See [Sko 84].

We begin with a construction of random Laurent operators following Brown and Halmos [BH 63].

Let \((\Omega, \mu)\) be a probability space. Let \( S^1 \) be the unit circle in the complex plane. \( L^2(S^1) \) denotes the Hilbert space of square integrable complex valued functions on \( S^1 \) with respect to the Lebesgue measure on the Borel sets.

The space of bounded linear operators on a Hilbert space \( H \) is denoted by \( \mathcal{L}(H) \). But we shall write \( L(\mathcal{L}(S^1)) \) for \( \mathcal{L}(L^2(S^1)) \).

Let \( \varphi : \Omega \times S^1 \to \mathbb{C} \) be measurable with respect to the product measure. Then the function \( \varphi(\omega, \cdot) \) defines a measurable function \( \varphi_\omega : S^1 \to \mathbb{C} \) which we shall assume to be bounded. By the Laurent operator defined by \( \varphi \) we mean the map \( L = L^{\varphi} : \Omega \to \mathcal{L}(S^1) \) defined by multiplication

\[
L_\omega f = \varphi_\omega f, \quad f \in L^2(S^1), \omega \in \Omega.
\]

**Proposition 10.** Given \( f, g \in L^2(S^1) \) the map \( \Omega \to \mathbb{C} \) taking \( \omega \) to \( (L_\omega f, g) \) is measurable.

**Proof.** Note that every measurable function can be written as a pointwise limit of finite linear sums of simple functions. We can use simple functions of the form as the indicator functions \( 1_{A \times B}(\omega, z) = 1_A(\omega)1_B(z) \) of measurable squares \( A \times B \). Such functions give rise the measurability of the map, and so do the finite linear sums of them. Taking pointwise limit of a sequence we verify the asserted measurability. \( \square \)

**Lemma 11.** Suppose \( A : \Omega \to \mathcal{L}(S^1) \) is a random linear operator. Then the functions \( \varphi : \Omega \times S^1 \to \mathbb{C} \) defined by \( \varphi(\omega, z) := (A(\omega)e_0)(z) \) and \( \varphi_\omega : S^1 \to \mathbb{C} \) defined by \( \varphi_\omega(z) = \varphi(\omega, z) \), for \( \omega \in \Omega \) are measurable.

**Proof.** Let \( \{e_1, \ldots\} \) be an orthonormal basis of the separable Hilbert space \( L^2(S^1) \). The measurability of \( A \) implies that the maps

\[
(\omega, z) \mapsto \sum_{j=1}^{n} (A(\omega)e_0, e_j)e_j(z)
\]
are measurable on $\Omega \times S^1$ for $n = 1, 2, \ldots$. Letting $n \to \infty$, we see that the map

$$\varphi(\omega, z) = (A(\omega)e_0)(z) = \sum_{j=1}^{\infty} (A(\omega)e_0, e_j)(z)$$

is measurable on $\Omega \times S^1$. By the definition of product measures, we see that $\varphi_\omega$ is measurable on $S^1$ for all $\omega$ in $\Omega$.

**Lemma 12.** Let $\phi$ be a measurable function in $L^2(S^1)$. Define a linear map $B : L^2(S^1) \to L^2(S^1)$ by $f \mapsto \phi f$. Then the domain of $B$ is dense and $B$ is a closed operator.

**Proof.** Since $\phi g \in L^2(S^1)$ for all $g$ in $C(S^1)$, the domain $D(B)$ of $B$ contains the dense subspace $C(S^1)$ of $L^2(S^1)$. Suppose that $f_n \to f$ and $\phi f_n \to g$ in $L^2$-norm. Then both $f_n \to f$ and $\phi f_n \to g$ in measure. Therefore, we have a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that both $f_{k_h} \to f$ and $\phi f_{k_h} \to g$ almost everywhere on $S^1$. Consequently, $\phi f = g$ and $B$ has a closed graph.

For an integer $n$ and $z \in S^1$, we write $e_n(z) = z^n$. We introduce the shift operator $W$ on $L^2(S^1)$ as multiplication by $e_1$, i.e. $Wf(z) = (e_1f)(z) = zf(z)$. Clearly $We_n = e_{n+1}$ for all $n$ and $W^n f = e_n f$ for all $n \geq 0$. We are interested in the centralizer $Z(W)$ of the shift operator $W$ in the space $\mathcal{R}$ of all random linear operators from $\Omega$ to $L^2(S^1)$.

$$Z(W) = \{ A \in \mathcal{R} : A(\omega) W = W A(\omega), \, \omega \in \Omega \}$$

**Theorem 13.** We have $Z(W) = \{ L^2 \}$ where $\varphi : \Omega \times S^1 \to \mathbb{C}$ with $\varphi_\omega$ bounded for $\omega \in \Omega$.

**Proof.** Since multiplication operators commutes it follows that $\{ L^2 \} \subset Z(W)$.

Conversely take $A \in Z(W)$. Define $\varphi : \Omega \times S^1 \to \mathbb{C}$ defined by $\varphi(\omega, z) = (A(\omega)e_0)(z)$ Then for $n \geq 0$, we have

$$A(\omega)e_n = A(\omega)W^n e_0 = W^n A(\omega)e_0 = W^n \varphi_\omega = e_n \cdot \varphi_\omega = \varphi_\omega \cdot e_n$$

Let $B(\omega)$ be the operator defined by multiplication by $\varphi_\omega$. Then the above equations that $A(\omega)e_n = B(\omega)e_n$ for $n \geq 0$. Fuglede’s theorem ([Fug 50]) says that $A(\omega)$ commutes with $W$ implies that $A(\omega)$ commutes with $W^*$. The same argument that gives $A(\omega)e_n = B(\omega)e_n$ for all $n$. Thus $A(\omega) = B(\omega)$ on $L^2(S^1)$.

That $\varphi_\omega$ is bounded follow by a norm argument.

**Proposition 14.** Let $\Omega$ be a measure space. Let $S^1$ be the unit circle in the complex plane $\mathbb{C}$. Let $f : \Omega \times S^1 \to \mathbb{R}$ (or $\mathbb{C}$) such that

- for each fixed $\omega$ in $\Omega$, the map $z \mapsto f(\omega, z)$ is continuous, and
- for each fixed $z$ in $S^1$, the map $\omega \mapsto f(\omega, z)$ is measurable.

Then there are measurable functions $g_n(\omega)$ on $\Omega$ such that

$$f(\omega, z) = \sum_{n \in \mathbb{Z}} g_n(\omega) z^n, \quad \forall \omega \in \Omega, \forall z \in S^1.$$  

**Proof.** Since $C(S^1)$ is a separable Banach space, its weak* compact convex dual ball is metrizable. In particular, the norm one linear functional

$$h(z) \mapsto \frac{1}{2\pi} \int_{S^1} h(z)z^{-n}dz, \quad h \in C(S^1),$$
is a limit of a sequence of convex combinations of point masses.

For a point mass $\delta_t$ with $t$ in $S^1$, the function

$$\omega \mapsto \delta_t(f(\omega, z)) = f(\omega, t)$$

is measurable on $\Omega$. As a pointwise limit of a sequence of measurable functions, the function $g_n(x)$ defined by

$$\omega \mapsto \frac{1}{2\pi} \int_{S^1} f(\omega, z)z^{-n}dz$$

is measurable on $\Omega$ for each $n$ in $\mathbb{Z}$.

It is plain that $f$ carries the stated form (by Fourier transform).

The matrix coefficients of a random linear operator $A : \Omega \to \mathcal{L}(S^1)$ with respect to $\{e_j\}$ are defined to be

$$a_{ij}(\omega) = \langle A(\omega)e_j, e_i \rangle$$

for $i,j \in \mathbb{Z}$.

**Lemma 15.** Given a measurable function $\varphi : \Omega \times S^1 \to \mathbb{C}$ with $\varphi_\omega$ bounded. Let $\varphi_\omega = \sum_i c_i(\omega)e_i$ be the Fourier expansion of $\varphi$. Let $L = L^\varphi$ be the random linear operator defined by multiplication by $\varphi$. Then the matrix coefficients of $L$ are given by

$$\ell_{ij}(\omega) = c_{i-j}(\omega)$$

**Theorem 16.** A random linear operator $A : \Omega \to \mathcal{L}(S^1)$ is a Laurent random operator if and only if its matrix coefficients satisfy

$$a_{i+1,j+1}(\omega) = a_{ij}(\omega)$$

for $\omega \in \Omega$.

**Proof.** After the preceding lemma it remains to show that the condition is sufficient which will follow if we show that $A \in Z(W)$. But $a_{i+1,j+1}(\omega) = a_{ij}(\omega)$ implies

$$\langle A(\omega)W e_j, e_i \rangle = \langle WA(\omega)e_j, e_i \rangle.$$

\[ \blacksquare \]

5. Random Toeplitz operators. Let $\mathcal{H}^2(S^1)$ be the space of square integrable analytic functions on $S^1$. Write $P : \mathcal{L}^2(S^1) \to \mathcal{H}^2(S^1)$ for the projection.

Let $\varphi : \Omega \times S^1 \to \mathbb{C}$ be measurable with respect to the product measure. Write $\varphi_\omega : S^1 \to \mathbb{R}$ for the function $\varphi(\omega, \cdot)$. We assume that $\varphi_\omega$ are bounded. The random Toeplitz operator defined by $\varphi$ is the map $T = T^\varphi : \Omega \times \mathcal{H}^2(S^1) \to \mathcal{H}^2(S^1)$ to be given by

$$T_\omega(u) = P(\varphi_\omega \cdot u), \quad u \in \mathcal{H}^2(S^1).$$

That is $T^\varphi f = PL^\varphi f$ for $f \in \mathcal{H}^2(S^1)$.

We can compute the matrix coefficients of $T^\varphi$ with respect to $\{e_j : j \geq 0\}$ as follows

$$\langle T^\varphi(e_j, \cdot) \rangle e_j = \langle P L^\varphi(e_j, \cdot) \rangle e_j = \langle L^\varphi(e_j, \cdot) \rangle e_j = \langle L^\varphi(e_{j+1}, \cdot) \rangle e_{j+1}$$

$$= \langle P L^\varphi(e_{j+1}, \cdot) \rangle e_{j+1} = \langle T^\varphi(e_{j+1}, \cdot) \rangle e_{j+1}$$
Thus we see that the matrix coefficients $t_{ij}(\omega) = \langle T^\varphi(\omega)e_j, e_i \rangle$ of random Toeplitz operator defined by $\varphi$ satisfy

$$t_{i+1,j+1}(\omega) = t_{ij}(\omega).$$

**Lemma 17.** Let $A_n : \Omega \to \mathcal{L}(H)$ be a sequence of random linear operators. Suppose that for each $\omega \in \Omega$ the sequence of operators $A_n(\omega)$ has a weak limit $A^\omega_\infty$. Put $A_\infty(\omega) = A^\omega_\infty$. Then $A_\infty : \Omega \to \mathcal{L}(H)$ is a random linear operator.

**Proof.** We put on $\mathcal{L}(H)$ the weak operator topology, which is defined by the seminorms $T \mapsto \langle Tx, y \rangle$ where $x, y$ are vectors in $H$. In this case, for every $\omega$ in $\Omega$, and $x, y$ in $H$ we have

$$\langle A_n(\omega)x, y \rangle \longrightarrow \langle A_\infty(\omega)x, y \rangle, \text{ as } n \to \infty.$$  

This says exactly the map $\omega \mapsto \langle A_\infty(\omega)x, y \rangle$ is the pointwise limit of a sequence of measurable functions. So it is measurable. $\Box$

**Theorem 18.** A random linear operator $A : \Omega \to \mathcal{L}(\mathcal{H}^2(S^1))$ is a random Toeplitz operator if and only if its matrix coefficients satisfy

$$a_{i+1,j+1}(\omega) = a_{ij}(\omega)$$

for $\omega \in \Omega$ and $i, j \geq 0$.

Given a sequence $\{c_n : -\infty < n < \infty\}$ of complex valued random variables on $\Omega$. We can form an infinite matrix $T$ whose $(i, j)$-th entry is $t_{ij} = c_{i-j}$. We can write down the first $(n+1) \times (n+1)$ submatrix as

$$T_n = \begin{pmatrix}
  c_0 & c_{-1} & c_{-2} & \ldots & c_{-n} \\
  c_1 & c_0 & c_{-1} & \ldots & c_{-(n-1)} \\
  c_2 & c_1 & c_0 & \ldots & c_{-(n-2)} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  c_{n-2} & c_{n-3} & c_{n-4} & \ldots & c_{-2} \\
  c_{n-1} & c_{n-2} & c_{n-3} & \ldots & c_{-1} \\
  c_n & c_{n-1} & c_{n-2} & \ldots & c_0
\end{pmatrix}$$

We shall always assume that $T_n$ is hermitian.

Let $(\Omega, \mu)$ be a probability space. Let $f(\omega, x)$ such that $f(\omega, \cdot)$ is in $L^1(S^1)$ for a.e. in $\omega \in \Omega$. We can consider its Fourier coefficient

$$c_n(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} f(\omega, x) dx.$$ 

We use these coefficients to build the finite Toeplitz forms

$$T_n(f)(\omega) = \sum_{0 \leq j,k \leq n} c_{k-j}(\omega)u_j\bar{u}_k$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} |u_0 + u_1e^{ix} + \cdots + u_ne^{inx}|^2 f(\omega, x) dx.$$
- this is a hermitian form in the variables $u$.

The eigenvariables of the Hermitian form $T_n(f)(\omega)$ are defined as the roots of the characteristic equation $\det(T_n(f) - \lambda) = 0$; we denote them by

$$\lambda_1^{(n)}(\omega), \ldots, \lambda_{n+1}^{(n)}(\omega)$$

These are real valued random variables defined a.e. in $\Omega$.

If we assume that (1) $m \leq f(x, w) \leq M$ for all $x \in S^1$ and a.e. $\omega$ and that (2)

$$I = \sum_{p=0}^{n} |u_p|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |u_0 + u_1 e^{ix} + \cdots + u_n e^{inx}|^2 dx = 1$$

Then it follows from the definition of $T_n(f)(\omega)$ that

$$m \leq T_n(f)(\omega) \leq M.$$ 

To each eigenvariable $\lambda_j^{(n)}(\omega)$ we have a nonvanishing eigenvector $u = (u_j^0, \ldots, u_j^n)$ determined up to a scalar, such that

$$T_n(f)(\omega) = \lambda_j^{(n)}(\omega) \sum_{p=0}^{n} |u_p|^2 = \lambda_j^{(n)}(\omega)$$

for our choice of $u$. Thus we see that

$$m \leq \lambda_j^{(n)}(\omega) \leq M.$$

**Theorem 19.** The matrices $(\lambda_j^{(n)}(\omega))$ and $(f(\omega, -\pi + \frac{2j\pi}{n+1}))$ are equally distributed.

**Proof.** If we take $F(t) = \log t$, $t > 0$, and make use of a Riemann sum, then we need to show

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{j=1}^{n+1} \log \lambda_j^{(n)}(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\omega, x) dx$$

Suppose $n > 0$. Let $D_n(f)$ be the determinant of $T_n(f)$. Then

$$D_n(f)(\omega) = \lambda_1^{(n)}(\omega) \cdots \lambda_{n+1}^{(n)}(\omega).$$

So we need to show

$$\lim_{n \to \infty} (D_n(f)(\omega))^{\frac{1}{n+1}} = \exp(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\omega, x) dx).$$

But this follows from the minimal value theorem

$$\lim_{n \to \infty} \frac{D_n(f)(\omega)}{D_{n-1}(f)(\omega)} = \exp(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\omega, x) dx).$$

The proof of the minimal value theorem for orthogonal polynomials given in [GS 58 §2.2] extends easily to our situation. 

Part III.
6. **Circulant preconditioner.** We want to solve efficiently a linear system of the form $T_n x = b$. The method of pre-conditioning means that we find a matrix $S_n$ such that it is efficient to solve the linear system $S_n^{-1} T_n x = S_n^{-1} b$. The conjugate gradient method used to solve the above linear system says that the more the eigenvalues of the coefficient matrix $S_n^{-1} T_n$ are clustered together the faster the convergence rate. A sequence of matrices $\{A_n\}_1^\infty$ is said to have clustered spectra around 1 if for any given $\epsilon > 0$ there exist positive integers $n_1$ and $n_2$ such that for all $n > n_1$, at most $n_2$ eigenvalues of the matrix $A_n - I_n$ have absolute value larger than $\epsilon$.

This leads to the problem of the distribution of the eigenvalues of the matrix $S_n^{-1} T_n$ when $S_n$ is the Strang circulant in the paper [ChR 89] of R. Chan. The basic technique is due to Szegő which assumes that there is a function $f$ on the circle $S^1$ such that the given sequence $\{c_n\}$ is the sequence of Fourier coefficients of $f$ and then the method of orthogonal polynomials is applied. The function $f$ is then called the generating function of the sequence $\{c_n\}$.

Let $(\Omega, \mu)$ be a probability space. Let $f(\omega, x)$ be a real-valued function on $\Omega \times S^1$ such that $f(\omega, \cdot)$ is in $L^1(S^1)$ for a.e. in $\omega \in \Omega$. We can consider its Fourier coefficient

$$c_n(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} f(\omega, x) dx.$$ 

We use these coefficients to build the $(n+1) \times (n+1)$ Toeplitz matrix $T_n$ whose $(k, j)$ entry $(T_n)_{k,j}$ is $c_{k-j}$. We have written $c_{k-j}$ for the function $c_{k-j}(w)$. We shall consider the case that he $T_n$ are Hermitian positive definite a.e. in $\omega$.

The Strang preconditioner $S_n = (s_{k,j})$ of $T_n$ is the Hermitian circulant defined as follows in two cases.

If $n = 2m + 1$ then $s_{k,j} = s_{k-j}$ are given by

$$s_\ell = \begin{cases} 
    c_\ell & \text{for } 0 \leq \ell \leq m, \\
    c_{\ell-n} & \text{for } m < \ell \leq n-1
\end{cases}$$

If $n = 2m$ then $s_{k,j} = s_{k-j}$ are given by

$$s_\ell = \begin{cases} 
    c_\ell & \text{for } 0 \leq \ell \leq m-1, \\
    0 & \text{for } \ell = m, \\
    \frac{1}{2}(c_m + c_{-m}) & \text{for } m < \ell \leq n-1, \\
    \bar{s}_{-\ell} & \text{for } 0 < -\ell \leq n-1.
\end{cases}$$

**Theorem 20.** Write the sup-norm as $\|c_n\|_\infty = \sup_{\omega \in \Omega} |c_n(\omega)|$. Suppose $f$ is positive and $\sum_{n} \|c_n\|_\infty$ is finite (Wiener class). Then for all $\epsilon > 0$, there exist $M$ and $N > 0$ such that for all $n > N$, at most $M$ eigenvalue functions of $S_n - T_n$ have sup norms exceeding $\epsilon$.

This is theorem 2 of [ChR 89]; see also theorem 4 §8 of [ChS 89]. The proof of Chan works for we have Cauchy Interlacing theorem in our case.

7. **Numerical results.** To test our preconditioners we shall compute using random phases generated by random sequences in the following manner. Let $\mathbb{N}$ be the set of all integers $\geq 0$. The generating function will be $f : \mathbb{N} \times S^1 \to \mathbb{C}$ with Fourier series

$$f = \sum_{k=-\infty}^{\infty} c_k(t) e^{-ikt\theta}.$$
We shall take a simple example of a random variable by choosing $c_k(t) = a_k e^{i \phi_k(t)}$. For $a_k$ we shall take the example as given in [ChR 89] namely

$$a_k = \begin{cases} \sqrt{\frac{1+\sqrt{1+k^2}}{(1+k)^2}}, & k > 0, \\ 2, & k = 0 \\ \bar{a}_{-k}, & k < 0. \end{cases}$$

For each $k > 0$ we take random sequences $\{\phi_k(t) : t = 0, 1, 2, \ldots \}$. Note that $f$ remains in the Wiener class. The $c_k(t)$ are used to build the Toeplitz matrices $T_n(t)$.

We choose the circulant preconditioner $S_n$ as above and run the PCG for $T_n x = b$ as given in [ChS 89] p. 106, namely, start from $x_0 = 0$ and $r_0 = b$. Solve

$$S z_{j-1} = r_{j-1}$$

$$\beta_j = z_{j-1}^T r_{j-1} / z_{j-2}^T r_{j-2}$$

$$d_j = z_{j-1} + \beta_j d_{j-1}$$

$$\alpha_{j-1} = z_{j-1}^T r_{j-1} / d_{j}^T T d_{j}$$

$$x_j = x_{j-1} + \alpha_j d_j$$

$$r_j = r_{j-1} - \alpha_j T d_j.$$

We input random sequences $\phi_k(t)$ which are Gaussian $N(0, 1)$. Write $r_j$ for the residue after $j$ iterations. We calculate the number $I(n, t)$ of iterations required in order that Strang’s circulant PCG applied to $T_n(t)$ achieve the residue ratios

$$\frac{||r_j||_2}{||r_j||_0} < 10^{-10}.$$
In Figure 7.1 we display the results for $n = 65$ i.e. a $65 \times 65$ Toeplitz matrix $T_{65}(t)$. The X-axis shows the value of $t$ and the Y-axis shows the number of iterations. The graph above is for the usual conjugate gradient method for $T_n$ and the graph below is that of $I(65, t)$ for the preconditioned $S^{-1}_n T_n$ which shows a mean value of around 20 iterations.

Figure 7.2 shows the variation of the average number of required iterations with the size of the Toeplitz matrices. Here the X axis is $n$ to indicate that the Toeplitz matrix is $n \times n$. We take $n$ to be odd $n = 2m + 1$ and we run for $10 \leq m \leq 120$. The Y axis gives the average number of required iterations. The lower graph shows the average $\frac{1}{50} \sum_{t=1}^{50} I(n, t)$.
Just to confirm the clustering of the eigenvalues we show the eigenvalues of $S_n^{-1}T_n$ in Figure 7.3. The eigenvalues $x + \sqrt{-1}y$ are displayed as $(x, y)$ for $t = 1, 2, \ldots, 6$.

REFERENCES

[AGZ 09] G. Anderson, A. Guionnet, O. Zeitouni, An introduction to random matrices, Cambridge University Press, (2009).

[Axe 85] O. Axelsson, A survey of preconditioned iterative methods for linear systems of algebraic equations, BIT Numerical Mathematics 25, (1985) 165-187.

[Axe 94] O. Axelsson, Iterative solution methods, Cambridge University Press, (1994).

[BCC 09] Z. Bai, Y. Chen, Y. Liang, Random matrix theory and its applications, World Scientific Pub. Co. (2009).

[BS 10] Z. Bai, J. Silverstein, Spectral analysis of large dimensional random matrices, Springer Verlag, (2010).

[BDJ 06] W. Bryc, A. Dembo, T. Jiang, Spectral measure of large random Hankel, Markov and Toeplitz matrices, Annals of Prob. 34, (2006) 1-38.

[BH 63] A. Brown, P. Halmos, Algebraic properties of Toeplitz operators, J. Reine Angew. Math. 213, (1963) 89-102.

[ChS 89] R. Chan, G. Strang, Toeplitz equations by conjugate gradients with circulant preconditioner, SIAM J. Sci. Stat. Comput., 10, (1989) 104-119.

[ChR 89] R. Chan, Circulant preconditioners for Hermitian Toeplitz systems, SIAM J. Matrix Anal. Appl., 10, (1989) 542-550.

[ChR 91] R. Chan, Toeplitz preconditioners for Toeplitz systems with nonnegative generating functions, IMA J. Numer. Anal. 11, (1991) 333-345.

[CJ 07] R. Chan, X. Jin, An introduction to iterative Toeplitz solvers, SIAM, (2007).

[Dan 70] J. Daniel, The approximate minimization of functionals, Prentice Hall, (1970).

[Fug 50] B. Fuglede, A commutativity theorem for normal operators, Proc. N. A. S., 36, (1950) 35-40.

[GV 96] G. Golub, C. Van Loan, Matrix computations, Johns Hopkins University Press, (1996).

[GS 58] U. Grenander, G. Szegö, Toeplitz forms and their applications, University of California Press, (1958).

[HJ 90] R. Horn, C. Johnson, Matrix analysis, Cambridge University Press, (1990).

[Meh 04] M. Mehta, Random matrices, North Holland Pub., (2004).

[Pan 04] I. Panayotov, Conjugate gradient in Hilbert spaces, McGill University Master thesis, (2004).

[PS 11] L. Pastur, M. Shcherbina, Eigenvalue distribution of large random matrices, Amer Math Society, (2011).

[Phi 03] G. Phillips, Interpolation and approximation by polynomials, Springer Verlag, (2003).

[Sko 84] A. Skorohod, Random linear operators, Kluwer Academic Publishers, (1984).

[Str 86] G. Strang, A proposal for Toeplitz matrix calculations, Stud. Appl. Math., 74 (1986), pp. 171-176.

[Tao 12] T. Tao, Topics in random matrix theory, Amer Math Society, (2012).

[Tit 60] E. Titchmarsh, The Theory of Functions, Oxford University Press, (1960).

[Wig 55] E. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, Annals of Math 62, (1955) 548-564.