The spectra of random abelian G-circulant matrices

Mark W. Meckes

Department of Mathematics, Case Western Reserve University, 10900 Euclid Ave., Cleveland, Ohio 44106, U.S.A.
E-mail address: `mark.meckes@case.edu`
URL: `www.case.edu/artsci/math/mwmeckes/`

Abstract. This paper studies the asymptotic behavior of eigenvalues of random abelian G-circulant matrices, that is, matrices whose structure is related to a finite abelian group G in a way that naturally generalizes the relationship between circulant matrices and cyclic groups. It is shown that, under mild conditions, when the size of the group G goes to infinity, the spectral measures of such random matrices approach a deterministic limit. Depending on some aspects of the structure of the groups, whether the matrices are constrained to be Hermitian, and a few details of the distributions of the matrix entries, the limit measure is either a (complex or real) Gaussian distribution or a mixture of two Gaussian distributions.

1. Introduction

Given a finite group G and a function $f : G \to \mathbb{C}$, the matrix $M = \left[f(ab^{-1}) \right]_{a,b \in G}$ is called a G-circulant matrix by Diaconis (1988, 1990). This generalizes the classical notion of circulant matrices, which arise as the special case in which G is a finite cyclic group. The action of such a matrix M on the vector space $\{g : G \to \mathbb{C}\}$ is as a convolution operator: for $g : G \to \mathbb{C}$ and $a \in G$,

$$(Mg)(a) = \sum_{b \in G} f(ab^{-1})g(b) =: (f * g)(a). \qquad (1.1)$$

This paper considers the asymptotic behavior of the spectra of random G-circulant matrices, or equivalently random convolution operators on G, when G is a large abelian group. (For the rest of this paper, G will always stand for a *finite abelian* group.) Such random matrices will be generated by picking the values $f(a)$ independently, with or without imposing a constraint $f(a^{-1}) = \overline{f(a)}$ which is equivalent to insisting that the matrix M is Hermitian. This generalizes the study of random circulant matrices, whose theory has already been developed in Bose and Mitra (2002); Bose and Sen (2008); Bryc and Sethuraman (2009); Meckes

(2009); Bose et al. (2010) among many other papers, with applications discussed in Jain and Srivastava (2008); Yin et al. (2010). The richer structure of arbitrary abelian groups relative to cyclic groups leads to the appearance of some interesting phenomena which do not occur for circulant matrices, or the more familiar setting of random matrices with independent entries.

The prototypical situation (exemplified in Corollaries 3.2, 3.4, and 3.6, and Theorems 4.2 and 4.4 below) is that when the size of G grows the empirical spectral distribution of a (properly normalized) random G-circulant matrix M approaches a Gaussian distribution. When M is constrained to be Hermitian the limit will be a real Gaussian distribution; without such a constraint it will be a complex Gaussian distribution. These situations may be thought of as analogous to the semicircle law for Hermitian random matrices and circular law for non-Hermitian random matrices with independent entries, respectively. This behavior, which has already been observed for random circulant matrices in Bose and Mitra (2002); Meckes (2009), occurs in particular if only a negligible fraction of the elements of G are of order 2, and also if every nonidentity element of G is of order 2. On the other hand, if neither of these is the case then more complicated limiting distributions occur which are mixtures of two Gaussian distributions (as in Theorems 4.1 and 4.3 below).

Another perspective on these results, which is crucial in the proofs, is that they describe the distribution of values of random Fourier series on G. The supremum of such a random Fourier series is already a thoroughly studied quantity Kahane (1985); Marcus and Pisier (1981). In particular, results of Marcus and Pisier (1981) include as special cases estimates of the spectral norms of random G-circulant matrices, as pointed out in Proposition 2.4 below.

Section 2 below briefly reviews the facts about Fourier analysis on finite abelian groups which are used here and points out their immediate consequences for G-circulant matrices; some notation and conventions used in the remainder of the paper are established there. Section 3 investigates the spectra of some random G-circulant matrices whose entries are Gaussian random variables. The invariance properties of Gaussian random variables allow an easy detailed study to be undertaken which illuminates the general situation, in particular the role of the number of elements of order 2. Finally, Section 4 determines the asymptotic behavior of the spectrum for general entries with finite variances.

The cases of G-circulant matrices with heavy-tailed entries, and of random G-circulant matrices when G is a nonabelian finite group, will be investigated in future work.

Acknowledgements

2. Some Fourier analysis and notation

For a finite abelian group G, we denote by $\widehat{G}$ the family of group homomorphisms $\chi : G \to \mathbb{T}$, where $\mathbb{T}$ is the multiplicative group $\{z \in \mathbb{C} \mid |z| = 1\}$. The elements of $\widehat{G}$ are called characters of G; $\widehat{G}$ is a group under the operation of pointwise

multiplication. The multiplicative inverse of a character χ is its pointwise complex conjugate $\overline{\chi}$. From the homomorphism property it follows that for $a \in G$ and $\chi \in \widehat{G}$, $\chi(a^{-1}) = \overline{\chi}(a)$.

We denote by $\ell^2(G)$ the space of functions $f : G \to \mathbb{C}$ equipped with the inner product

$$\langle f, g \rangle = \sum_{a \in G} f(a)\overline{g(a)},$$

and $\ell^2(\widehat{G})$ is defined analogously. The Fourier transform of $f \in \ell^2(G)$ is the function $\widehat{f} \in \ell^2(\widehat{G})$ given by

$$\widehat{f}(\chi) = \langle f, \overline{\chi} \rangle = \sum_{a \in G} f(a)\chi(a).$$

This includes as special cases both the classical discrete Fourier transform (when G is cyclic) and the Walsh–Hadamard transform (when G is a product of cyclic groups of order 2). The following lemma summarizes the most important fundamental facts about the Fourier transform for our purposes.

Lemma 2.1. *Let G be a finite abelian group with $|G|$ elements.*

(1) *The functions $\left\{ \frac{1}{\sqrt{|G|}}\chi \mid \chi \in \widehat{G} \right\}$ form an orthonormal basis of $\ell^2(G)$.*

(2) *The map $f \mapsto \frac{1}{\sqrt{|G|}}\widehat{f}$ is a linear isometry of $\ell^2(G)$ onto $\ell^2(\widehat{G})$.*

(3) *If $f, g \in \ell^2(G)$, then for each $\chi \in \widehat{G}$, $\widehat{f * g}(\chi) = \widehat{f}(\chi)\widehat{g}(\chi)$ (where the convolution $f * g$ is defined in (1.1).*

Proof:

(1) See Theorem 6 on Serre (1977, p. 19).
(2) This follows easily from Proposition 7 on Serre (1977, p. 20) (which is a consequence of part (1)).
(3) This follows directly from the definitions by a straightforward computation. $\square$

Observe that contained in Lemma 2.1(1) is the fact that $|G| = |\widehat{G}|$.

We will need two additional facts about characters of finite abelian groups which are not as easily located in standard references.

Lemma 2.2. *The number of elements $a \in G$ such that $a^2 = 1$ is equal to the number of characters $\chi \in \widehat{G}$ such that $\chi = \overline{\chi}$.*

Proof: For $a \in G$, define $\delta_a : G \to \mathbb{C}$ by $\delta_a(b) = \delta_{a,b}$, where the latter is the Kronecker delta function, and observe that $\{\delta_a \mid a \in G\}$ is an orthonormal basis of $\ell^2(G)$. Then $\widehat{\delta}_a(\chi) = \chi(a)$ for each $\chi \in \widehat{G}$. By Lemma 2.1(2), the number of $a \in G$ such that $a^2 = 1$ is equal to

$$\sum_{a \in G} \langle \delta_a, \delta_{a^{-1}} \rangle = \frac{1}{|G|} \sum_{a \in G} \left\langle \widehat{\delta}_a, \widehat{\delta_{a^{-1}}} \right\rangle = \frac{1}{|G|} \sum_{a \in G} \sum_{\chi \in \widehat{G}} \chi(a)\overline{\chi}(a^{-1})$$

$$= \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \sum_{a \in G} \chi(a)^2 = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \langle \chi, \overline{\chi} \rangle,$$

which by Lemma 2.1(1) is equal to the number of $\chi \in \widehat{G}$ such that $\chi = \overline{\chi}$. $\square$

Lemma 2.2 says that G and $\widehat{G}$ have equal numbers of elements of order 2. A much stronger fact is also true: G and $\widehat{G}$ are isomorphic groups. However, this isomorphism is noncanonical, depends on the classification of finite abelian groups, and in any case is not useful here.

Lemma 2.3. *Let H be a subgroup of a finite abelian group H. Then each character on H extends to a character on G in precisely $|G| / |H|$ distinct ways.*

Proof: It is easy to check that restriction to H defines a homomorphism $\widehat{G} \to \widehat{H}$. Since each coset of this homomorphism's kernel has the same size, it suffices to prove that that it is surjective, or equivalently that each character on H extends to a character on G at all. For a proof of this fact see, e.g., Apostol (1976, p. 134). $\square$

From (1.1) and Lemma 2.1(3) it follows that the Fourier transform diagonalizes G-circulant matrices. In particular, if $M = [f(ab^{-1})]_{a,b \in G}$ for $f \in \ell^2(G)$, then the eigenvalues of M are precisely the values $\{\widehat{f}(\chi) \mid \chi \in \widehat{G}\}$ of the Fourier transform of f, and the characters of G are eigenvectors of M. (For generalizations of these facts for nonabelian G, see Diaconis (1988, 1990).) Observe that every G-circulant matrix is normal, but that M is Hermitian if and only if $f(a^{-1}) = \overline{f(a)}$ for each $a \in G$.

Given a family of random variables $\{Y_a \mid a \in G\}$, define the random function $f \in \ell^2(G)$ by $f(a) = \frac{1}{\sqrt{|G|}} Y_a$. (We are avoiding using X to name random variables because of its typographical similarity to χ.) The corresponding G-circulant matrix is the random matrix $M = [Y_{ab^{-1}}]_{a,b \in G}$. Its eigenvalues, indexed by $\chi \in \widehat{G}$, are given by

$$\lambda_\chi = \widehat{f}(\chi) = \frac{1}{\sqrt{|G|}} \sum_{a \in G} Y_a \chi(a), \qquad (2.1)$$

and the empirical spectral distribution of M is

$$\mu = \frac{1}{|\widehat{G}|} \sum_{\chi \in \widehat{G}} \delta_{\lambda_\chi} = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \delta_{\lambda_\chi},$$

where δ_z here denotes the point mass at $z \in \mathbb{C}$.

The Fourier transform $\widehat{f}$ is a random trigonometric polynomial on G, of the kind studied extensively by Marcus and Pisier (1981). From (2.1) it follows in particular that $\|M\| = \|\widehat{f}\|_\infty$, where the former norm is the spectral norm of M. The following result is thus a special case of Marcus and Pisier (1981, Theorem 1.4), which also applies to infinite compact abelian groups.

Proposition 2.4. *Suppose that $\{Y_a \mid a \in G\}$ are independent (except possibly for a constraint $Y_{a^{-1}} = \overline{Y_a}$ for each $a \in G$) and mean 0 with finite second moments. Then*

$$c \left(\min_{a \in G} \mathbb{E} |Y_a| \right) \leq \frac{\mathbb{E} \|M\|}{\sqrt{\log |G|}} \leq C \sqrt{\max_{a \in G} \mathbb{E} |Y_a|^2},$$

where $c, C > 0$ are constants, independent of G and the distributions of the Y_a.

The rest of this paper deals mainly with infinite sequences of finite abelian groups $G^{(n)}$, always assumed to satisfy $|G^{(n)}| \to \infty$. For each n a family of random

variables $\{Y_g^{(n)} \mid g \in G^{(n)}\}$ will be used to construct a random $G^{(n)}$-circulant matrix

$$M^{(n)} = \left[\frac{1}{\sqrt{|G^{(n)}|}} Y_{ab^{-1}}^{(n)} \right]_{a,b \in G^{(n)}}$$

with empirical spectral measure $\mu^{(n)}$. As mentioned earlier, an important role will be played by the quantity

$$p_2^{(n)} = \frac{\left| \{ a \in G^{(n)} \mid a^2 = 1 \} \right|}{\left| G^{(n)} \right|} = \frac{\left| \{ \chi \in \widehat{G} \mid \chi = \overline{\chi} \} \right|}{\left| \widehat{G}^{(n)} \right|}.$$

The standard real Gaussian measure is denoted $\gamma_{\mathbb{R}}$, and the standard complex Gaussian distribution, normalized such that $\mathbb{E}|Z|^2 = 1$ when Z is a standard complex Gaussian random variable, is denoted $\gamma_{\mathbb{C}}$. For $\alpha \in [0,1]$, γ_α denotes the Gaussian measure on $\mathbb{C} \cong \mathbb{R}^2$ with covariance $\frac{1}{2} \begin{bmatrix} 1+\alpha & 0 \\ 0 & 1-\alpha \end{bmatrix}$, so that in particular $\gamma_0 = \gamma_{\mathbb{C}}$ and $\gamma_1 = \gamma_{\mathbb{R}}$.

The integral of a function f with respect to a measure ν will be denoted by $\nu(f)$.

3. Gaussian matrix entries

The following is an immediate consequence of Lemma 2.1(2) and the rotation-invariance of the standard Gaussian distribution. The special case of this result for classical circulant matrices (that is, when G is a cyclic group) was observed in Meckes (2009).

Proposition 3.1. *Let G be a finite abelian group and let $\{Y_a \mid a \in G\}$ be independent, standard complex Gaussian random variables. Then the eigenvalues $\{\lambda_\chi \mid \chi \in \widehat{G}\}$ of M given by (2.1) are independent, standard complex Gaussian random variables.*

The random matrix ensemble in Proposition 3.1 is the G-circulant analogue of the complex Ginibre ensemble X, which consists of a square matrix with independent, standard complex Gaussian entries.

Corollary 3.2. *Suppose that for each n, $\{Y_a^{(n)} \mid a \in G^{(n)}\}$ are independent, standard complex Gaussian random variables. Then $\mathbb{E}\mu^{(n)} = \gamma_{\mathbb{C}}$ for each n, and $\mu^{(n)} \to \gamma_{\mathbb{C}}$ weakly in probability. Furthermore, if $|G^{(n)}| = \Omega(n^\varepsilon)$ for some $\varepsilon > 0$, then $\mu^{(n)} \to \gamma_{\mathbb{C}}$ weakly almost surely.*

Proof: For each, say, Lipschitz $f : \mathbb{C} \to \mathbb{R}$,

$$(\mathbb{E}\mu)(f) := \mathbb{E}\big(\mu(f)\big) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \mathbb{E}f(\lambda_\chi),$$

where the (n) superscripts are omitted for simplicity. By Proposition 3.1, each λ_χ is distributed according to $\gamma_{\mathbb{C}}$, and so $(\mathbb{E}\mu)(f) = \gamma_{\mathbb{C}}(f)$. Thus $\mathbb{E}\mu = \gamma_{\mathbb{C}}$.

By the concentration properties of Gaussian measure (see Ledoux (2001)), since the λ_χ are distributed as independent standard complex Gaussian random variables, if f is 1-Lipschitz, then

$$\mathbb{P}\big[|\mu(f) - \gamma_{\mathbb{C}}(f)| \geq t\big] \leq 2e^{-|G|t^2}$$

for each $t > 0$. If $\left|G^{(n)}\right| = \Omega(n^\varepsilon)$, then the Borel–Cantelli lemma implies that $\mu^{(n)}(f) \to \gamma_\mathbb{C}(f)$ almost surely. Applying this to a countable dense family of f, it follows that $\mu^{(n)} \to \gamma_\mathbb{C}$ weakly almost surely.

In the general case, since $\left|G^{(n)}\right| \to \infty$, each subsequence of $\mu^{(n)}$ has a subsequence $\mu^{(n_j)}$ for which, say, $\left|G^{(n_j)}\right| \geq j$, so that by the above argument $\mu^{(n_j)}$ converges to $\gamma_\mathbb{C}$ almost surely as $j \to \infty$. It follows that $\mu^{(n)}$ converges to $\gamma_\mathbb{C}$ in probability. $\qquad\square$

The next proposition deals with the G-circulant analogue of the Gaussian Unitary Ensemble (GUE), which, up to a choice of normalization, is distributed as $2^{-1/2}(X+X^*)$, where X is the complex Ginibre ensemble mentioned above. Equivalently, the diagonal entries of the GUE are standard real Gaussian random variables, the off-diagonal entries are standard complex Gaussian random variables, and the entries are independent except for the constraint that the matrix is Hermitian. It is worth noting explicitly that while each entry of the GUE has (complex) variance 1, the variance of a diagonal entry and the real part of an off-diagonal entry differ by a factor of 2. (Again, the special case for classical circulant matrices was observed earlier in Meckes (2009).)

Proposition 3.3. *Let G be a finite abelian group and let $\{Y_a \mid a \in G\}$ be random variables which are independent except for the constraint $Y_{a^{-1}} = \overline{Y_a}$, and such that*

$$Y_a \sim \begin{cases} \gamma_\mathbb{R} & \text{if } a^2 = 1, \\ \gamma_\mathbb{C} & \text{if } a^2 \neq 1. \end{cases}$$

Then the eigenvalues $\{\lambda_\chi \mid \chi \in \widehat{G}\}$ of M given by (2.1) are independent, standard real Gaussian random variables.

Proof: Let $\{Z_a \mid a \in G\}$ be independent, standard complex Gaussian random variables. Then $\{Y_a \mid a \in G\}$ are distributed as $\left\{2^{-1/2}\left(Z_a + \overline{Z_{a^{-1}}}\right) \mid a \in G\right\}$. Thus the eigenvalues λ_χ of M in the present proposition are jointly distributed as $\sqrt{2}$ times the real parts of the eigenvalues of the random matrix defined in Proposition 3.1, and are thus independent real standard normal random variables. $\qquad\square$

Observe that in the "G-circulant GUE" of Proposition 3.3, every element $a \in G$ with $a = a^{-1}$ corresponds to a "diagonal" of M in which the entries are constrained to be real.

The following corollary follows from Proposition 3.3 in the same way that Corollary 3.2 follows from Proposition 3.1.

Corollary 3.4. *Suppose that for each n, $\left\{Y_a^{(n)} \mid a \in G^{(n)}\right\}$ are real and complex Gaussian random variables as described in Proposition 3.3. Then $\mathbb{E}\mu^{(n)} = \gamma_\mathbb{R}$ for each n, and $\mu^{(n)} \to \gamma_\mathbb{R}$ weakly in probability. Furthermore, if $\left|G^{(n)}\right| = \Omega(n^\varepsilon)$ for some $\varepsilon > 0$, then $\mu^{(n)} \to \gamma_\mathbb{R}$ weakly almost surely.*

The real Ginibre ensemble X consists of a square matrix with independent, real standard Gaussian random variables. The Gaussian Orthogonal Ensemble (GOE) is distributed as $2^{-1/2}(X + X^t)$. Equivalently, the diagonal entries of the GOE are distributed as $\mathcal{N}(0,2)$ and the off-diagonal entries are distributed as $\mathcal{N}(0,1)$. In general the analogues of Propositions 3.1 and 3.3 for matrices with real entries

are less elegant. In the nonsymmetric case the eigenvalues have a Gaussian joint distribution in a $|G|$-dimensional real subspace of $\mathbb{C}^{|G|}$, and in the symmetric case the $|G|$ eigenvalues are not independent in general. We will not state such results in general, but will note for future reference that in the "G-circulant GOE", every element $a \in G$ with $a = a^{-1}$ corresponds to a diagonal of M in which the variance of the entries is 2 instead of 1. (See Theorem 4.3 below and the discussion following it.)

On the other hand, the analogous results *are* simple in the case in which the characters $\chi \in \widehat{G}$ are all real-valued, so that the Fourier transform defines an isometry (up to scaling) between the *real* ℓ^2 spaces on G and $\widehat{G}$. By Lemma 2.2, this is the case precisely when every $a \in G$ satisfies $a^2 = 1$, or in other words, when $G \cong (\mathbb{Z}_2)^n$ for some n. In this case a G-circulant matrix is automatically symmetric, so that there is no difference (except for scaling) between the "G-circulant real Ginibre ensemble" and the "G-circulant GOE". The following results are proved in the same way as Proposition 3.1 and Corollary 3.2.

Proposition 3.5. *Let $G \cong (\mathbb{Z}_2)^n$ and let $\{Y_a \mid a \in G\}$ be independent, standard real Gaussian random variables. Then the eigenvalues $\{\lambda_\chi \mid \chi \in \widehat{G}\}$ of M given by* (2.1) *are independent, standard real Gaussian random variables.*

Corollary 3.6. *Suppose that for each n, $G^{(n)} \cong (\mathbb{Z}_2)^n$ and $\{Y_a^{(n)} \mid a \in G^{(n)}\}$ are independent, standard real Gaussian random variables. Then $\mathbb{E}\mu^{(n)} = \gamma_\mathbb{R}$ for each n, and $\mu^{(n)} \to \gamma_\mathbb{R}$ weakly almost surely.*

4. General matrix entries

Our main results are stated under a Lindeberg-type condition on the random variables $Y_a^{(n)}$ used to generate the random matrices:

$$\forall \varepsilon > 0 : \quad \lim_{n \to \infty} \frac{1}{|G^{(n)}|} \sum_{a \in G^{(n)}} \mathbb{E}\left(|Y_a^{(n)}|^2 \mathbb{1}_{|Y_a^{(n)}| \geq \varepsilon \sqrt{|G^{(n)}|}}\right) = 0. \qquad (4.1)$$

The usual remarks apply about the sufficiency of identical distribution or a Lyapunov-type condition: (4.1) holds in the settings of Theorems 4.1 and 4.2 if all the $Y_a^{(n)}$ are identically distributed, or have uniformly bounded $(2 + \delta)$ moments; it holds in the settings of Theorems 4.3 and 4.4 if all the random variables with a given variance assumption satisfy such assumptions.

We now state our main results, deferring the proofs until the end of the section.

Theorem 4.1. *Let $\alpha \in [0, 1]$. Suppose that for each n, $\{Y_a^{(n)} \mid a \in G^{(n)}\}$ are independent; that*

$$\mathbb{E}Y_a^{(n)} = 0, \quad \mathbb{E}|Y_a^{(n)}|^2 = 1, \quad and \quad \mathbb{E}(Y_a^{(n)})^2 = \alpha$$

for every $a \in G^{(n)}$; and that (4.1) *holds. Suppose further that $\lim_{n \to \infty} p_2^{(n)} = p$ exists. Then $\mu^{(n)}$ converges, in mean and in probability, to $(1 - p)\gamma_\mathbb{C} + p\gamma_\alpha$.*

One of the main special cases of interest in Theorem 4.1 is when $\alpha = 1$, that is, when the matrix entries are all real. In that case, the limiting spectral distribution of $M^{(n)}$ is complex Gaussian if the number of a with $a^2 = 1$ is negligible for large n. On the other hand, if the fraction of such a is asymptotically constant then, due

to the presence of many real-valued characters χ, the limiting spectral distribution will be a mixture of $\gamma_{\mathbb{C}}$ and $\gamma_{\mathbb{R}}$.

The other main special case of interest is when $\alpha = 0$, so that the matrix entries have uncorrelated real and imaginary parts. In that case, which generalizes the setting of Corollary 3.2, one can remove the assumption that $p_2^{(n)}$ approaches a limit.

Theorem 4.2. *Suppose that for each n, $\{Y_a^{(n)} \mid a \in G^{(n)}\}$ are independent; that*

$$\mathbb{E}Y_a^{(n)} = 0, \quad \mathbb{E}\big|Y_a^{(n)}\big|^2 = 1, \quad and \quad \mathbb{E}\big(Y_a^{(n)}\big)^2 = 0$$

for every $a \in G^{(n)}$; and that (4.1) holds. Then $\mu^{(n)}$ converges, in mean and in probability, to $\gamma_{\mathbb{C}}$.

The special case of Theorem 4.2 for classical circulant matrices (that is, when the $G^{(n)}$ are cyclic groups) was proved by the author in Meckes (2009).

Theorem 4.3. *Let $\alpha \in [0,1]$, $\beta > 0$. Suppose that for each n, $\{Y_a^{(n)} \mid a \in G^{(n)}\}$ are mean 0 and independent except for the constraint $Y_{a^{-1}}^{(n)} = \overline{Y_a^{(n)}}$; that*

$$\mathbb{E}Y_a^{(n)}Y_b^{(n)} = \begin{cases} 1 & \text{if } a = b^{-1} \neq a^{-1}, \\ \alpha & \text{if } a = b \neq a^{-1}, \\ \beta & \text{if } a = b = a^{-1}, \\ 0 & \text{otherwise,} \end{cases}$$

for $a, b \in G^{(n)}$; and that (4.1) holds. Assume further that $\lim_{n \to \infty} p_2^{(n)} = p$ exists. Then $\mu^{(n)}$ converges, in mean and in probability, to

$$(1-p)\mathcal{N}\big(0, 1 + p(\beta - \alpha - 1)\big) + p\mathcal{N}\big(0, 1 + \alpha + p(\beta - \alpha - 1)\big).$$

if $p < 1$ and to $\mathcal{N}\big(0, \beta\big)$ if $p = 1$.

Observe that by Lagrange's theorem on orders of subgroups, $1/p_2^{(n)}$ is an integer, which implies that if $p < 1$ then in fact $p \leq 1/2$, and therefore the stated variances of the normal distributions named above are indeed positive.

The most obvious (though not necessarily, as we shall see, the most natural) special case of interest in Theorem 4.3 is when the $Y_a^{(n)}$ are real and i.i.d. (except for the symmetry constraint), so that $\alpha = \beta = 1$. In that case the limiting spectral distribution is the mixture distribution

$$(1-p)\mathcal{N}(0, 1-p) + p\mathcal{N}(0, 2-p). \tag{4.2}$$

Two other special cases are suggested by considering the analogy with the GOE and GUE. The G-circulant analogue of the GOE, as discussed in the previous section, would have real entries such that $\alpha = 1$ and $\beta = 2$, and thus the limiting spectral distribution

$$(1-p)\mathcal{N}(0, 1) + p\mathcal{N}(0, 2). \tag{4.3}$$

The slightly simpler nature of this limiting distribution (note that the parameter p plays only one role in (4.3), as opposed to two roles in (4.2)) reflects that a "GOE-like" normalization of entries is more natural than equal variances. However, this phenomenon is only evident when $0 < p < 1$. In the classical case of Wigner matrices it is well known that in order for the semicircle law to hold, no variance

assumption need be made on the diagonal entries of the matrix. The situation described above emphasizes that this is the case precisely because the number of diagonal entries in a Wigner matrix is negligible.

Finally, when the second moments are the same as for the "G-circulant GUE" of Proposition 3.3, then $\alpha = 0$ and $\beta = 1$ and, as in Corollary 3.4, the limiting spectral distribution is simply the standard real Gaussian distribution, even regardless of the value of p. Thus for G-circulant matrices, a constraint to be complex Hermitian appears to be somehow more natural than a constraint to be real symmetric. As in Theorem 4.2, the assumption that p_2 approaches a limit can even be removed in this situation.

Theorem 4.4. *Suppose that for each n, $\left\{ Y_a^{(n)} \mid a \in G^{(n)} \right\}$ are mean 0 and independent except for the constraint $Y_{a^{-1}}^{(n)} = \overline{Y_a^{(n)}}$; that $\mathbb{E}\left|Y_a^{(n)}\right|^2 = 1$ for every $a \in G^{(n)}$; that $\mathbb{E}\left(Y_a^{(n)}\right)^2 = 0$ if $a \neq a^{-1}$; and that (4.1) holds. Then $\mu^{(n)}$ converges, in mean and in probability, to $\gamma_\mathbb{R}$.*

The special case of Theorem 4.4 for classical circulant matrices (with more restrictive assumptions on the distributions of the matrix entries) was proved by Bose and Mitra (2002).

We will not attempt to deal thoroughly with the question of when the convergence in probability in the results above can be strengthened to almost sure convergence. However, the following result gives some sufficient conditions. Each of the conditions stated automatically implies the Lindeberg-type condition (4.1); for the first part this follows from exponential tail decay which is implied by a Poincaré inequality (see Ledoux (2001, Corollary 3.2)), and for the other parts it is elementary.

Theorem 4.5. *In the setting of Theorem 4.1, 4.2, 4.3, or 4.4, suppose in addition that $\left|G^{(n)}\right| = \Omega(n^\varepsilon)$ for some $\varepsilon > 0$ and that one of the following conditions holds:*

(1) *There is a constant $K > 0$ such that for every n and every $a \in G^{(n)}$, $Y_a^{(n)}$ satisfies a Poincaré inequality with constant K. That is,*

$$\operatorname{Var} f\left(Y_a^{(n)}\right) \leq K\mathbb{E}\left|\nabla f\left(Y_a^{(n)}\right)\right|^2$$

for every smooth $f : \mathbb{R}^2 \to \mathbb{R}$.

(2) *There is a constant $K > 0$ such that $\left|Y_a^{(n)}\right| \leq K$ a.s. for every n and every $a \in G^{(n)}$.*

(3) *For some $\delta \in (0, 1]$, $\sup_{n \in \mathbb{N}} \max_{a \in G^{(n)}} \mathbb{E}\left|Y_a^{(n)}\right|^{2+\delta} < \infty$, and $\sum_{n=1}^\infty \left|G^{(n)}\right|^{-\delta/2} < \infty$.*

(4) *For some $\delta \in (0, 1]$, $\sup_{n \in \mathbb{N}} \max_{a \in G^{(n)}} \mathbb{E}\left|Y_a^{(n)}\right|^{2+\delta} < \infty$, and $p_2^{(n)} \to p > 0$.*

Then $\mu^{(n)}$ converges to the stated limit almost surely.

We now turn to the proofs of our main results. Unsurprisingly, generalizing the results of the last section to non-Gaussian matrix entries is achieved by using an appropriate version of the central limit theorem to show that the eigenvalues λ_χ are approximately distributed like uncorrelated Gaussian random variables. Even to prove asymptotic results, it is necessary here to apply some *quantitative* version of the central limit theorem, in order to achieve suitably uniform control over the λ_χ.

The approach taken here (and earlier in Meckes (2009)) generalizes and extends the method used by Bose and Mitra (2002), which applied a multivariate version of the Berry–Esseen theorem and thus required the matrix entries to have uniformly bounded third moments. Here a quantitative, multivariate version of Lindeberg's theorem is applied.

If $f : \mathbb{R}^d \to \mathbb{R}$ is bounded and Lipschitz with Lipschitz constant $|f|_L$, its bounded Lipschitz norm may be defined by

$$\|f\|_{BL} = \max\{\|f\|_\infty, |f|_L\}.$$

The bounded Lipschitz distance between random vectors X and Y in $\mathbb{R}^d$ is defined by

$$d_{BL}(X, Y) = \sup_{\|f\|_{BL} \leq 1} |\mathbb{E}f(X) - \mathbb{E}f(Y)|.$$

It is well known (see e.g. Dudley (2002, section 11.3)) that the class of bounded Lipschitz functions is a convergence-determining class. The subclass of compactly supported such functions is furthermore separable with respect to the sup norm Dudley (2002, Corollary 11.2.5). Thus to show that a sequence $\nu^{(n)}$ of probability measures on $\mathbb{R}^d$ converges weakly to ν in mean, in probability, or almost surely, it suffices to show that for each bounded Lipschitz function f, $\nu^{(n)}(f) \to \nu(f)$ in the same sense.

The following is a special case of Bhattacharya and Ranga Rao (1986, Theorem 18.1) (cf. the proof of Bhattacharya and Ranga Rao (1986, Corollary 18.2)).

Proposition 4.6. *Suppose that $X_1, \ldots, X_k$ are independent mean 0 random vectors in $\mathbb{R}^d$ such that $\frac{1}{k}\sum_{j=1}^k \mathrm{Cov}(X_j) = I_d$. For $\varepsilon > 0$ let*

$$\theta(\varepsilon) = \frac{1}{k}\sum_{j=1}^k \mathbb{E}\left(\|X_j\|^2 \, \mathbb{1}_{\|X_j\| > \varepsilon\sqrt{k}}\right).$$

Then

$$d_{BL}\left(\frac{1}{\sqrt{k}}\sum_{j=1}^k X_j, Z\right) \leq C_d \inf_{0 \leq \varepsilon \leq 1}(\varepsilon + \theta(\varepsilon)),$$

where Z is a standard Gaussian random vector in $\mathbb{R}^d$, and $C_d > 0$ depends only on d.

Proof of Theorem 4.1: Let $f : \mathbb{C} \to \mathbb{R}$ with $\|f\|_{BL} \leq 1$. Observe that

$$\mathbb{E}\mu(f) = \frac{1}{|G|}\sum_{\chi \in \widehat{G}} \mathbb{E}f(\lambda_\chi) = \frac{1}{|G|}\sum_{\chi \in \widehat{G}} \mathbb{E}f\left(\frac{1}{\sqrt{|G|}}\sum_{a \in G} \chi(a)Y_a\right), \qquad (4.4)$$

where (n) superscripts have been omitted for simplicity. We consider λ_χ as a sum of independent random vectors in $\mathbb{R}^2 \cong \mathbb{C}$. The relevant covariances are

$$\mathrm{Cov}(\chi(a)Y_a) = \begin{bmatrix} \mathbb{E}(\mathrm{Re}\,\chi(a)Y_a)^2 & \mathbb{E}(\mathrm{Re}\,\chi(a)Y_a)(\mathrm{Im}\,\chi(a)Y_a) \\ \mathbb{E}(\mathrm{Re}\,\chi(a)Y_a)(\mathrm{Im}\,\chi(a)Y_a) & \mathbb{E}(\mathrm{Im}\,\chi(a)Y_a)^2 \end{bmatrix}.$$

The identities

$$\begin{aligned}
(\mathrm{Re}\,w)(\mathrm{Re}\,z) &= \tfrac{1}{2}\,\mathrm{Re}\big[(w + \overline{w})z\big], \\
(\mathrm{Im}\,w)(\mathrm{Im}\,z) &= \tfrac{1}{2}\,\mathrm{Re}\big[(\overline{w} - w)z\big], \\
(\mathrm{Re}\,w)(\mathrm{Im}\,z) &= \tfrac{1}{2}\,\mathrm{Im}\big[(w - \overline{w})z\big],
\end{aligned} \qquad (4.5)$$

will be useful.

Setting $w = z = \chi(a)Y_a$ for a fixed $\chi \in \widehat{G}$,

$$\sum_{a \in G} \mathbb{E}(\operatorname{Re} \chi(a)Y_a)^2 = \sum_{a \in G} \left[\frac{1}{2} \operatorname{Re} \mathbb{E} \left(\chi(a)^2 Y_a^2 + |\chi(a)|^2 |Y_a|^2 \right) \right]$$

$$= \frac{1}{2} \left(|G| + \alpha \sum_{a \in G} \chi^2(a) \right) = \frac{|G|}{2} (1 + \alpha \mathbb{1}_{\chi = \overline{\chi}}).$$

In the last step we have used that unless χ is real-valued, χ and $\overline{\chi}$ are distinct characters, and hence orthogonal in $\ell^2(G)$. In similar fashion, we find that

$$\operatorname{Cov}(\lambda_\chi) = \frac{1}{|G|} \sum_{a \in G} \operatorname{Cov}(\chi(a)Y_a) = \frac{1}{2} \left(I_2 + \mathbb{1}_{\chi = \overline{\chi}} \left[\begin{smallmatrix} \alpha & 0 \\ 0 & -\alpha \end{smallmatrix} \right] \right).$$

Observe in particular that if $\alpha = 1$ and χ is real-valued, then λ_χ is almost surely real, with variance 1; in that case we treat λ_χ as a random variable in $\mathbb{R}$, as opposed to a random vector in $\mathbb{R}^2$. Proposition 4.6 and (4.1) (recalling that $|\chi(a)| = 1$ always) now imply that there is a sequence δ_n decreasing to 0 such that for each $\chi \in \widehat{G}$,

$$\left| \mathbb{E}f \left(\frac{1}{\sqrt{|G|}} \sum_{a \in G} \chi(a)Y_a \right) - \gamma_\alpha(f) \right| \leq \delta_n$$

if χ is real-valued, and

$$\left| \mathbb{E}f \left(\frac{1}{\sqrt{|G|}} \sum_{a \in G} \chi(a)Y_a \right) - \gamma_{\mathbb{C}}(f) \right| \leq \delta_n$$

otherwise. Writing $\nu^{(n)} = (1 - p_2^{(n)})\gamma_{\mathbb{C}} + p_2^{(n)}\gamma_\alpha$, by (4.4) it follows that

$$\begin{aligned} |\mathbb{E}\mu(f) - \nu(f)| &= \left| \frac{1}{|G|} \sum_{\chi = \overline{\chi}} \mathbb{E}f \left(\frac{1}{\sqrt{|G|}} \sum_{a \in G} \chi(a)Y_a \right) - p_2\gamma_\alpha(f) \right. \\ &\quad \left. + \frac{1}{|G|} \sum_{\chi \neq \overline{\chi}} \mathbb{E}f \left(\frac{1}{\sqrt{|G|}} \sum_{a \in G} \chi(a)Y_a \right) - (1 - p_2)\gamma_{\mathbb{C}}(f) \right| \\ &\leq p_2\delta_n + (1 - p_2)\delta_n = \delta_n, \end{aligned} \quad (4.6)$$

where as above the subscripts (n) are omitted. Since $p_2^{(n)} \to p$, it follows that $\nu^{(n)} \Rightarrow (1 - p)\gamma_{\mathbb{C}} + p\gamma_\alpha$, and so $\mathbb{E}\mu^{(n)} \Rightarrow (1 - p)\gamma_{\mathbb{C}} + p\gamma_\alpha$.

Next observe that

$$\mathbb{E}(\mu(f))^2 = \frac{1}{|G|^2} \sum_{\chi_1, \chi_2 \in \widehat{G}} \mathbb{E}f(\lambda_{\chi_1})f(\lambda_{\chi_2}) = \frac{1}{|G|^2} \sum_{\chi_1, \chi_2 \in \widehat{G}} \mathbb{E}F((\lambda_{\chi_1}, \lambda_{\chi_2})), \quad (4.7)$$

where $F : \mathbb{C}^2 \to \mathbb{R}$ is defined by $F(w, z) = f(w)f(z)$, so that $\|F\|_{BL} \leq 2$. We now consider $(\lambda_{\chi_1}, \lambda_{\chi_2})$ as a sum of independent random vectors in $\mathbb{R}^4$. The upper-left and lower-right 2×2 blocks of $\operatorname{Cov}((\lambda_{\chi_1}, \lambda_{\chi_2}))$ are of course just $\operatorname{Cov}(\lambda_{\chi_1})$ and $\operatorname{Cov}(\lambda_{\chi_2})$, computed above. For the off-diagonal blocks, we use $w = \chi_1(a)Y_a$ and

$z = \chi_2(a)Y_a$ in (4.5) to obtain for example

$$\sum_{a \in G} \mathbb{E}(\operatorname{Re}\chi_1(a)Y_a)(\operatorname{Re}\chi_2(a)Y_a) = \sum_{a \in G}\left[\frac{1}{2}\operatorname{Re}\mathbb{E}\left(\chi_1(a)\chi_2(a)Y_a^2 + \overline{\chi_1(a)}\chi_2(a)|Y_a|^2\right)\right]$$

$$= \frac{1}{2}\left(\alpha\sum_{a \in G}\chi_1(a)\chi_2(a) + \sum_{a \in G}\overline{\chi_1(a)}\chi_2(a)\right)$$

$$= \frac{|G|}{2}\left(\alpha\mathbb{1}_{\chi_1 = \overline{\chi_2}} + \mathbb{1}_{\chi_1 = \chi_2}\right).$$

Similarly, it follows that the off-diagonal blocks of $\operatorname{Cov}\left((\lambda_{\chi_1}, \lambda_{\chi_2})\right)$ are 0 unless $\chi_1 = \chi_2$ or $\chi_1 = \overline{\chi_2}$.

Assume for now that $\chi_1 \neq \chi_2$ and $\chi_1 \neq \overline{\chi_2}$. Applying Proposition 4.6, we now obtain that there is a sequence δ_n' decreasing to 0 such that whenever $\|f\|_{BL} \leq 1$,

$$\left|\mathbb{E}f(\lambda_{\chi_1})f(\lambda_{\chi_2}) - \gamma_\alpha(f)^2\right| \leq \delta_n'$$

if χ_1 and χ_2 are both real-valued,

$$|\mathbb{E}f(\lambda_{\chi_1})f(\lambda_{\chi_2}) - \gamma_{\mathbb{C}}(f)\gamma_\alpha(f)| \leq \delta_n'$$

if exactly one of χ_1 and χ_2 is real-valued, and

$$\left|\mathbb{E}f(\lambda_{\chi_1})f(\lambda_{\chi_2}) - \gamma_{\mathbb{C}}(f)^2\right| \leq \delta_n'$$

if neither χ_1 nor χ_2 is real-valued. (Note that Proposition 4.6 may be applied in the case of nonidentity covariance via a linear change of coordinates. For $\alpha < 1$, the determinant of the covariance is bounded away from zero, whereas for $\alpha = 1$ the variables are real.) Given χ_1, note that there are at most 2 characters χ_2 which are unaccounted for. By (4.7), it now follows that

$$\left|\mathbb{E}\mu^{(n)}(f)^2 - \nu^{(n)}(f)^2\right| \leq \delta_n' + \frac{2}{\left|G^{(n)}\right|}. \tag{4.8}$$

Finally,

$$\mathbb{E}\left|\mu^{(n)}(f) - \nu^{(n)}(f)\right|^2 = \left[\mathbb{E}\mu^{(n)}(f)^2 - \nu^{(n)}(f)^2\right] - 2\nu^{(n)}(f)\left[\mathbb{E}\mu^{(n)}(f) - \nu^{(n)}(f)\right]$$

$$\leq \left|\mathbb{E}\mu^{(n)}(f)^2 - \nu^{(n)}(f)^2\right| + 2\left|\mathbb{E}\mu^{(n)}(f) - \nu^{(n)}(f)\right|,$$

so by (4.6) and (4.8),

$$\mu^{(n)}(f) \to \left[(1-p)\gamma_{\mathbb{C}} + p\gamma_\alpha\right](f)$$

in L^2, and hence in probability. $\qquad\square$

Proof of Theorem 4.2: The proof is analogous to that of Theorem 4.1, setting $\alpha = 0$. In that case $\operatorname{Cov}(\lambda_\chi)$ no longer depends on whether χ is real-valued, which makes it unnecessary to assume that $p_2^{(n)}$ approaches a limit. $\qquad\square$

Proof of Theorem 4.3: We omit (n) superscripts as before. We will assume that $p < 1$; the case $p = 1$ (which implies that in fact $p_2 = 1$ for sufficiently large n) is similar and slightly simpler. Let $A = \{a \in G \mid a = a^{-1}\}$. Since G is abelian, A is a

subgroup of G. The restriction of a character of G to A is a character on A, which is necessarily real-valued on A. It follows that for $\chi_1, \chi_2 \in \widehat{G}$,

$$
\begin{aligned}
|G|\, \mathbb{E}\lambda_{\chi_1}\lambda_{\chi_2} &= \sum_{a,b \in G} \chi_1(a)\chi_2(b)\mathbb{E}Y_a Y_b \\
&= \sum_{a \in G} \chi_1(a)\big[\big(\overline{\chi_2(a)} + \alpha\chi_2(a)\big)\mathbb{1}_{a \neq a^{-1}} + \beta\chi_2(a)\mathbb{1}_{a=a^{-1}}\big] \\
&= \sum_{a \in G \backslash A} \chi_1(a)\overline{\chi_2(a)} + \alpha \sum_{a \in G \backslash A} \chi_1(a)\chi_2(a) + \beta \sum_{a \in A} \chi_1(a)\chi_2(a) \\
&= \sum_{a \in G} \chi_1(a)\overline{\chi_2(a)} + \alpha \sum_{a \in G} \chi_1(a)\chi_2(a) + (\beta - \alpha - 1)\sum_{a \in A} \chi_1(a)\chi_2(a) \\
&= |G|\big(\mathbb{1}_{\chi_1=\chi_2} + \alpha\mathbb{1}_{\chi_1=\overline{\chi_2}}\big) + |A|\,(\beta - \alpha - 1)\mathbb{1}_{\chi_1|_A=\overline{\chi_2}|_A} \\
&= |G|\big(\mathbb{1}_{\chi_1=\chi_2} + \alpha\mathbb{1}_{\chi_1=\overline{\chi_2}} + p_2(\beta - \alpha - 1)\mathbb{1}_{\chi_1|_A=\chi_2|_A}\big).
\end{aligned}
$$
(4.9)

In particular, for $\chi \in \widehat{G}$,

$$
\mathrm{Var}(\lambda_\chi) = 1 + \alpha\mathbb{1}_{\chi=\overline{\chi}} + p_2(\beta - \alpha - 1).
$$

Denoting

$$
\nu^{(n)} = (1 - p_2^{(n)})\mathcal{N}\big(0, 1 + p_2^{(n)}(\beta - \alpha - 1)\big) + p_2^{(n)}\mathcal{N}\big(0, 1 + \alpha + p_2^{(n)}(\beta - \alpha - 1)\big),
$$

it follows as in the proof of Theorem 4.1 that $d_{BL}(\mathbb{E}\mu^{(n)}, \nu^{(n)}) \to 0$, and thus that

$$
\mathbb{E}\mu^{(n)} \Rightarrow (1 - p)\mathcal{N}\big(0, 1 + p(\beta - \alpha - 1)\big) + p\mathcal{N}\big(0, 1 + \alpha + p(\beta - \alpha - 1)\big).
$$

In this situation just the 1-dimensional case of Proposition 4.6 is necessary. Observe also that the variances $\mathrm{Var}(\lambda_\chi)$ are uniformly bounded away from 0 (cf. the comments following the statement of the theorem.) This is necessary so that Proposition 4.6 may be applied for nonidentity covariance, via a linear change of coordinates, and still yield error bounds δ_n which are uniform in f with $\|f\|_{BL} \leq 1$.

By (4.9), if $\chi_1 \neq \chi_2$ and $\chi_1 \neq \overline{\chi_2}$, then

$$
\mathrm{Cov}\big((\lambda_{\chi_1}, \lambda_{\chi_2})\big) = \big(1 + p_2(\beta - \alpha - 1)\big)I_2 + \alpha \begin{bmatrix} \mathbb{1}_{\chi_1=\overline{\chi_1}} & 0 \\ 0 & \mathbb{1}_{\chi_2=\overline{\chi_2}} \end{bmatrix}
$$

$$
+ p_2(\beta - \alpha - 1)\mathbb{1}_{\chi_1|_A=\chi_2|_A} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.
$$

We consider separately the cases $p = 0$ and $p > 0$. If $p = 0$, then when $\chi_1 \neq \chi_2$ and $\chi_1 \neq \overline{\chi_2}$, we have

$$
\mathrm{Cov}\big((\lambda_{\chi_1}, \lambda_{\chi_2})\big) = \begin{bmatrix} 1 + \alpha\mathbb{1}_{\chi_1=\overline{\chi_1}} & 0 \\ 0 & 1 + \alpha\mathbb{1}_{\chi_2=\overline{\chi_2}} \end{bmatrix} + o(1).
$$

From here the argument is completed as in the proof of Theorem 4.1.

Suppose now that $p > 0$. Given $\chi_1 \in \widehat{G}$, by Lemma 2.3 there are exactly $\frac{1}{p_2}$ values of $\chi_2 \in \widehat{G}$ with $\chi_1|_A = \chi_2|_A$. Therefore,

$$
\mathrm{Cov}\big((\lambda_{\chi_1}, \lambda_{\chi_2})\big) = \big(1 + p_2(\beta - \alpha - 1)\big)I_2 + \alpha \begin{bmatrix} \mathbb{1}_{\chi_1=\overline{\chi_1}} & 0 \\ 0 & \mathbb{1}_{\chi_2=\overline{\chi_2}} \end{bmatrix}
$$

for all but a negligible fraction of pairs $\chi_1, \chi_2 \in \widehat{G}$. The argument is again completed as in the proof of Theorem 4.1. $\qquad\square$

Proof of Theorem 4.4: The proof is analogous to that of Theorem 4.3, setting $\alpha = 0$ and $\beta = 1$. In that case $\mathrm{Cov}(\lambda_{\chi_1}, \lambda_{\chi_2}) = \mathbb{1}_{\chi_1 = \chi_2}$, so it is unnecessary to assume that $p_2^{(n)}$ approaches a limit. $\qquad\square$

Proof of Theorem 4.5: (1) The Poincaré inequality assumption and independence imply an exponential concentration property for the family of eigenvalues $\{\lambda_\chi \mid \chi \in \widehat{G}^{(n)}\}$. In particular, combining Corollaries 5.7 and 3.2 of Ledoux (2001), it follows that for each L-Lipschitz $F : \ell^2(G^{(n)}) \to \mathbb{R}$,

$$\mathbb{P}\left[\left|F(Y^{(n)}) - \mathbb{E}F(Y^{(n)})\right| \geq t\right] \leq 2e^{-ct/\sqrt{K}L}$$

for each $t > 0$, where $c > 0$ is some absolute constant and $Y^{(n)}$ is shorthand for $(Y_a^{(n)})_{a \in G^{(n)}}$. Now for a 1-Lipschitz $f : \mathbb{C} \to \mathbb{R}$ and $k \in \mathbb{N}$,

$$\left|\frac{1}{k}\sum_{j=1}^{k} f(w_j) - \frac{1}{k}\sum_{j=1}^{k} f(z_j)\right| \leq \frac{1}{k}\sum_{j=1}^{k}|w_j - z_j| \leq \sqrt{\frac{1}{k}\sum_{j=1}^{k}|w_j - z_j|^2}$$

by the Cauchy–Schwarz inequality. Combining this with Lemma 2.1(2) it follows that $\mu^{(n)}(f)$ is $\left|G^{(n)}\right|^{-1/2}$-Lipschitz as a function of $Y^{(n)}$, and so

$$\mathbb{P}\left[\left|\mu^{(n)}(f) - \mathbb{E}\mu^{(n)}(f))\right| \geq t\right] \leq 2e^{-ct\sqrt{|G^{(n)}|/K}}.$$

Combined with the already known convergence in mean and the Borel–Cantelli lemma, this implies almost sure convergence of $\mu(f)$.

(2) The proof is similar to the previous part, using instead Talagrand's convex-distance concentration inequality for independent bounded random variables Talagrand (1995, Theorem 4.1.1) (see e.g. Meckes (2004, Corollary 4) for an explicit statement of a version that applies directly to complex random variables), cf. the proof of Meckes (2009, Theorem 2).

(3) The stated Lyapunov-type assumption yields upper bounds on all the δ_n quantities in the proofs above of order $\left|G^{(n)}\right|^{-\delta/2}$ for $0 < \delta \leq 1$ (cf. Bhattacharya and Ranga Rao (1986, Corollary 18.3)). Thus the assumption that $\sum_{n=1}^{\infty}\left|G^{(n)}\right|^{-\delta/2}$ allows the Borel–Cantelli lemma to be applied again.

(4) The assumption that $p > 0$ implies that $\left|G^{(n)}\right|$ actually grows exponentially: since $p_2^{(n)}$ is always the reciprocal of an integer (by Lagrange's theorem about the orders of subgroups of finite groups), $p_2^{(n)} \to p > 0$ implies that $p_2^{(n)}$ is eventually constant. By the classification of finite abelian groups,

$$G \cong \left(\prod_{j=1}^{m} \mathbb{Z}_{2^{k_j}}\right) \times H,$$

where $m \geq 0$, $k_j \geq 1$ for each j, and each nonidentity element of H has odd order. (For simplicity of notation, we are again suppressing the dependence of all these on n.) In this notation, the number of $a \in G$ such that $a = a^{-1}$ is 2^m, so that $|G| = 2^m/p_2$. The hypothesis that $\left|G^{(n)}\right|$ is strictly increasing thus implies that m is eventually strictly increasing, and hence $\left|G^{(n)}\right|$ is

eventually *exponentially* increasing. Therefore the previous part of the theorem applies. ☐

References

T. M. Apostol. *Introduction to analytic number theory*. Springer-Verlag, New York (1976). Undergraduate Texts in Mathematics. MR0434929.

R. N. Bhattacharya and R. Ranga Rao. *Normal approximation and asymptotic expansions*. Robert E. Krieger Publishing Co. Inc., Melbourne, FL (1986). ISBN 0-89874-690-6. Reprint of the 1976 original, MR855460.

A. Bose and J. Mitra. Limiting spectral distribution of a special circulant. *Statist. Probab. Lett.* **60** (1), 111–120 (2002). MR1945684.

A. Bose and A. Sen. Another look at the moment method for large dimensional random matrices. *Electron. J. Probab.* **13**, no. 21, 588–628 (2008). MR2399292.

A. Bose, R.t Subhra Hazra and K. Saha. Spectral norm of circulant type matrices with heavy tailed entries. *Electron. Commun. Probab.* **15**, 299–313 (2010). MR2670197.

W. Bryc and S. Sethuraman. A remark on the maximum eigenvalue for circulant matrices. In *High dimensional probability V: the Luminy volume*, volume 5 of *Inst. Math. Stat. Collect.*, pages 179–184. Inst. Math. Statist., Beachwood, OH (2009). MR2797947.

P. Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA (1988). ISBN 0-940600-14-5. MR964069.

P. Diaconis. Patterned matrices. In *Matrix theory and applications (Phoenix, AZ, 1989)*, volume 40 of *Proc. Sympos. Appl. Math.*, pages 37–58. Amer. Math. Soc., Providence, RI (1990). MR1059483.

R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (2002). ISBN 0-521-00754-2. Revised reprint of the 1989 original. MR1932358.

S. R. Jain and S. C. L. Srivastava. Random cyclic matrices. *Phys. Rev. E* **78**, 036213 (2008).

J.-P. Kahane. *Some random series of functions*, volume 5 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition (1985). ISBN 0-521-24966-X; 0-521-45602-9. MR833073.

M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI (2001). ISBN 0-8218-2864-9. MR1849347.

M. B. Marcus and G. Pisier. *Random Fourier Series with Applications to Harmonic Analysis*, volume 101 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, N.J. (1981). MR630532.

M. W. Meckes. Concentration of norms and eigenvalues of random matrices. *J. Funct. Anal.* **211** (2), 508–524 (2004). MR2057479.

M. W. Meckes. Some results on random circulant matrices. In *High dimensional probability V: the Luminy volume*, volume 5 of *Inst. Math. Stat. Collect.*, pages 213–223. Inst. Math. Statist., Beachwood, OH (2009). MR2797949.

J.-P. Serre. *Linear representations of finite groups*. Springer-Verlag, New York (1977). ISBN 0-387-90190-6. Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42. MR0450380.

M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.* (81), 73–205 (1995). MR1361756.

W. Yin, S. Morgan, J. Yang and Y. Zhang. Practical compressive sensing with Toeplitz and circulant matrices (2010). Rice University CAAM Technical Report TR10-01.