On the Spectral Norms of Pseudo-Wigner and Related Matrices

Ilya Soloveychik and Vahid Tarokh,
John A. Paulson School of Engineering and Applied Sciences, Harvard University

Abstract—**We investigate the spectral norms of symmetric $N \times N$ matrices from two pseudo-random ensembles. The first is the pseudo-Wigner ensemble introduced in "Pseudo-Wigner Matrices" by Soloveychik, Xiang and Tarokh and the second is its Sample Covariance-type analog defined in this work. Both ensembles are defined through the concept of r-independence by controlling the amount of randomness in the underlying matrices, and can be constructed from dual BCH codes. We show that when the measure of randomness r grows as N^ρ, where $\rho \in (0, 1]$ and $\varepsilon > 0$, the norm of the matrices is almost surely within $o\left(\frac{\log^{1+\varepsilon} N}{N^{\min[\rho, 2/3]}}\right)$ distance from 1. Numerical simulations verifying the obtained results are provided.**

Index Terms—**Pseudo-random matrices, spectral norm, Wigner ensemble, sample covariance matrices.**

I. INTRODUCTION

Random matrices have been a very active area of research for the last few decades and found enormous applications in various areas of modern mathematics, physics, engineering, biological modeling, and other fields [1]. In this article, we focus on two types of square symmetric matrices: 1) sign (± 1) matrices and 2) Sample Covariance Matrices (SCM) of sign vectors.

Random square symmetric sign matrices were originally examined by Wigner [2]. He proved that if the elements of the upper triangle of an $N \times N$ symmetric matrix (including the main diagonal) are independent Rademacher (± 1 with equal probabilities) random variables, then as $N \to \infty$ a properly scaled empirical spectral measure converges to the semicircular law. Wigner originally showed convergence in expectation, which was later improved to convergence in probability [3] and to almost sure weak convergence [4]. The spectral behavior of SCMs formed from p independent N dimensional vectors with independent entries and $\frac{p}{N} \to \gamma \in (0, 1)$, was for the first time rigorously investigated by Marchenko and Pastur [5]. They showed that (actually, under weaker conditions on dependencies among vector entries) the limiting spectrum converges to a non-random law.

In many engineering applications, one needs to simulate random matrices. The most natural way to generate an instance of a random $N \times N$ sign matrix is to toss a fair coin $\frac{N(N+1)}{2}$ times, fill the upper triangular part of a matrix with the outcomes and reflect the upper triangular part into the lower. Similarly, to get a random SCM matrix one would need to toss a coin $pN \approx \gamma N^2$ times. Unfortunately, for large N such approach would require a powerful source of randomness due to the independence condition [6]. In

addition, when the data is generated by a truly random source, atypical *non-random looking* outcomes have non-zero probability of showing up. Yet another issue is that any experiment involving tossing a coin would be impossible to reproduce. All these reasons stimulated researchers and engineers from different areas to seek approaches of generating *random-looking* data usually referred to as *pseudo-random* sources or sequences of binary digits [7, 8]. A wide spectrum of pseudo-random number generating algorithms have found applications in a large variety of fields including radar, digital signal processing, CDMA, error correction, cryptographic systems, and Monte Carlo simulations, navigation systems, scrambling, coding theory, etc. [7].

The term *pseudo-random* is used to emphasize that the binary data at hand is indeed generated by an entirely deterministic causal process with low algorithmic complexity, but its statistical properties resemble some of the properties of data generated by tossing a fair coin. Remarkably, most efforts were focused on one dimensional pseudo-random sequences [7, 8] due to their natural applications and to the relative simplicity of their analytical treatment. The study of pseudo-random arrays and matrices was launched around the same time [9–12]. Among the known two dimensional pseudo-random constructions the most popular are the so-called perfect maps [9, 13, 14] and two dimensional cyclic codes [11, 12]. However, none of these works considered spectral properties as the defining statistical features for their constructions.

Specific pseudo-random constructions usually develop from a set of properties mimicking truly random data, and attempt to come up with deterministic ways of reproducing these properties. Following this approach, in [15] we proposed a framework allowing construction of symmetric sign matrices of low Kolmogorov complexity with spectra converging to the semicircular law. Here we extend the ideas of [15] to the construction of low complexity SCMs with spectra converging to Marchenko-Pastur law. In a related work [16], the authors show that if the columns of matrices are randomly chosen from a properly designed binary code, their spectra converge to Marchenko-Pastur law as is the case for our construction. However all these works do not examine finer characteristics of the proposed matrices. In the present article we go beyond limiting spectral measures. We require more moments of the pseudo-random construction to match those of the truly random ensembles which en-

ables us to capture the behavior of the extreme eigenvalues (spectral norm). As a tradeoff, we pay a penalty for that by increased Kolmogorov complexity. We also provide an explicit construction of both Wigner-type and SCM-type ensembles from dual BCH codes and support our theoretical results by numerical simulations.

The outline of this paper is given next. Section II provides the original truly random ensembles and their properties. In Section III, we introduce our pseudo-random ensembles through the concept of r-independence and demonstrate that by increasing the amount of randomness involved in their construction, we can mimic finer properties of the true random matrices. The main results about the spectral norms are presented in Section IV followed by numerical tests in Section VI.

Notation. For a real x, $\lfloor x \rfloor$ stands for the largest integer not exceeding x. For a real random variable X, we write $F_X(x)$ for its cumulative distribution function (c.d.f.) and $f_X(x)$ for its probability density function (p.d.f.). For two real functions $g(x)$ and $f(x)$ of a real or natural argument, we say $g(x) = o(f(x))$ if $\lim\limits_{x \to +\infty} g/f = 0$ and $g(x) = O(f(x))$ if $\lim\limits_{x \to +\infty} g/f < +\infty$.

II. RANDOM MATRIX ENSEMBLES

Denote the spectrum of a symmetric real matrix $\mathbf{S}_N$ by

$$\lambda_1(\mathbf{S}_N) \leqslant \cdots \leqslant \lambda_N(\mathbf{S}_N), \tag{1}$$

the c.d.f. associated with it by

$$F_{\mathbf{S}_N}(x) = \frac{1}{N} \sum_{i=1}^{N} \theta(x - \lambda_i(\mathbf{S}_N)), \tag{2}$$

where $\theta(x)$ is the unit step function at zero, and the spectral norm by

$$\|\mathbf{S}_N\| = \max[|\lambda_1(\mathbf{S}_N)|, |\lambda_N(\mathbf{S}_N)|]. \tag{3}$$

The l-th empirical moment of $\mathbf{S}_N$ reads as

$$\int x^l dF_{\mathbf{S}_N} = \frac{1}{N} \mathrm{Tr}\left(\mathbf{S}_N^l\right). \tag{4}$$

Next, we introduce two random ensembles. We will mimic their spectral properties by pseudo-random constructions in Section III.

A. Wigner Matrices and the Semicircular Law

The first rigorous study of a random matrix ensemble was performed by Wigner in his seminal work [2, 17]. Wigner's ensemble $\mathcal{W}_N$ is the set S_N of all $N \times N$ matrices with $\pm\frac{1}{2\sqrt{N}}$ entries endowed with the uniform probability measure.

Let F_W be the c.d.f. of the standard semicircular law with the p.d.f.

$$f_W(x) = \begin{cases} \frac{2}{\pi}\sqrt{1 - x^2}, & -1 \leqslant x \leqslant 1, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

The moments of this distribution read as

$$\mu_W(s) = \int_{-\infty}^{+\infty} x^s dF_W = \begin{cases} \frac{1}{2^s}\mathbf{C}_{s/2}, & s \text{ even}, \\ 0, & s \text{ odd}, \end{cases} \tag{6}$$

where

$$\mathbf{C}_{s/2} = \frac{s!}{\left(\frac{s}{2}\right)!\left(\frac{s}{2} + 1\right)!} \tag{7}$$

are Catalan numbers. Stirling's approximation yields

$$\frac{1}{2^s}\mathbf{C}_{s/2} = \sqrt{\frac{8}{\pi s^3}}(1 + o(1)), \quad s \to +\infty. \tag{8}$$

Using the so-called method of moments, Wigner demonstrated [2] that the empirical spectral measures of matrices from $\mathcal{W}_N$ converge in expectation to the semicircular law (5). In a follow up article he improved this result to convergence in probability [17]. Almost sure weak convergence [4] and other asymptotic results were obtained later [18].

Here we focus on a series of results obtained by Soshnikov and Sinai [19–21]. These papers developed a combinatorial technique enabling exact quantification of the high-order expected moments of Wigner matrices, and led to the proof of universality of the joint distribution of their largest eigenvalues.

Lemma 1 (Corollary of Main Theorem from [19]). *Let* $\mathbf{W}_N \in \mathcal{W}_N$ *and* $s_N = o\left(N^{2/3}\right)$, *then*

$$\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{W}_N^{s_N}\right)\right] = \begin{cases} \sqrt{\frac{8}{\pi s_N^3}}N(1 + o(1)), & s_N \text{ even}, \\ 0, & s_N \text{ odd}, \end{cases} \tag{9}$$

as $N \to +\infty$, *and the random variables*

$$\mathrm{Tr}\left(\mathbf{W}_N^{s_N}\right) - \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{W}_N^{s_N}\right)\right] \tag{10}$$

converge in distribution to the normal law $\mathcal{N}\left(0, \frac{1}{\pi}\right)$.

This result in particular implies almost sure weak convergence of the empirical spectra of matrices from Wigner's ensemble to the semicircular law [18]. Below we also use the following variation of a result proven in [21].

Lemma 2 (Corollary of Theorem 2 from [21]). *Let* $\mathbf{W}_N \in \mathcal{W}_N$, *then for any sequence* $s_N = O\left(N^{2/3}\right)$,

$$\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{W}_N^{s_N}\right)\right] \leqslant c_W(s_N)N, \tag{11}$$

where $c_W(s_N)$ *is bounded uniformly over* N.

B. Sample Covariance Matrices and Marchenko-Pastur Law

Let $\mathcal{X}_{N,p}$ be the set $M_{N,p}$ of $N \times p$ matrices with $\pm\frac{1}{\sqrt{N}}$ entries endowed with the uniform probability measure. Below we consider a setting where the dimensions N and $p_N = p(N)$ grow such that the limit

$$\gamma = \lim_{N \to \infty} \frac{p_N}{N}, \tag{12}$$

exists. The spectra of the SCMs $\mathbf{X}_N^\top \mathbf{X}_N$ with $\mathbf{X}_N \in \mathbf{M}_{N,p}$ are invariant under the replacement of $\mathbf{X}_N$ with $\mathbf{X}_N^\top$ up to zero eigenvalues, therefore, without loss of generality we

assume $\gamma \leqslant 1$. The Marchenko-Pastur distribution is defined through its p.d.f. as

$$f_{MP}(x) = \begin{cases} \frac{1}{2\pi\gamma x}\sqrt{(b-x)(x-a)}, & a \leqslant x \leqslant b, \\ 0, & \text{otherwise,} \end{cases} \quad (13)$$

where

$$a = (1 - \sqrt{\gamma})^2, \qquad b = (1 + \sqrt{\gamma})^2. \quad (14)$$

The moments of this distribution read as

$$\mu_W(s) = \int_{-\infty}^{+\infty} x^s f_W dx = \sum_{k=1}^{s} \gamma^k \boldsymbol{N}(s_N, k), \quad (15)$$

where

$$\boldsymbol{N}(s,k) = \frac{1}{s}\binom{k}{s}\binom{k-1}{s} \quad (16)$$

are Narayana numbers. Stirling's approximation gives [22]

$$\sum_{k=1}^{s} \gamma^k \boldsymbol{N}(s_N, k) = \frac{\gamma^{1/4}}{2\sqrt{\pi}} \frac{N(1+\sqrt{\gamma})^{2s+1}}{s^{3/2}}(1 + o(1)).$$

Marchenko and Pastur proved in [5] that the spectrum of the product $\mathbf{X}_N^\top \mathbf{X}_N$ converges almost surely weakly to the limiting distribution (13). Later this result was strengthened in [23] and other works.

Péché proved [24] the universality of the joint distribution of top eigenvalues of SCM for a rich family of marginal distributions by developing a tight bound on the expected high-order moments. Adapted to our setup their main technical result reads as follows.

Lemma 3 (Corollary from Propositions 2.4 and 2.5 from [24]). *Let* $\mathbf{X}_N \in \mathcal{X}_{N,p_N}$ *and* $s_N = o\left(\sqrt{N}\right)$, *then*

$$\mathbb{E}\left[\text{Tr}\left(\left(\frac{\mathbf{X}_N^\top \mathbf{X}_N}{(1+\sqrt{\gamma})^2}\right)^{s_N}\right)\right]$$
$$= \frac{\gamma^{1/4}(1+\sqrt{\gamma})}{2\sqrt{\pi}} \frac{N}{s_N^{3/2}}(1 + o(1)), \quad (17)$$

as $N \to +\infty$, *and the random variables*

$$\text{Tr}\left(\left(\frac{\mathbf{X}_N^\top \mathbf{X}_N}{(1+\sqrt{\gamma})^2}\right)^{s_N}\right) - \mathbb{E}\left[\text{Tr}\left(\left(\frac{\mathbf{X}_N^\top \mathbf{X}_N}{(1+\sqrt{\gamma})^2}\right)^{s_N}\right)\right]$$

converge in distribution to the normal law $\mathcal{N}\left(0, \frac{1}{\pi}\right)$.

Below we utilize the following result from [24].

Lemma 4 (Corollary of Theorem 3.1 from [24]). *Let* $\mathbf{X}_N \in \mathcal{X}_{N,p_N}$, *then for any sequence* $s_N = O\left(N^{2/3}\right)$,

$$\mathbb{E}\left[\text{Tr}\left(\left(\frac{\mathbf{X}_N^\top \mathbf{X}_N}{(1+\sqrt{\gamma})^2}\right)^{s_N}\right)\right] \leqslant c_{MP}(\gamma, s_N)N,$$

where $c_{MP}(\gamma, s_N)$ *is bounded uniformly over* N.

III. PSEUDO-RANDOM ENSEMBLES

A. Definitions

In this section, we recall some definitions from [15] and introduce a family of pseudo-Marchenko-Pastur (pseudo-MP) ensembles analogous to the pseudo-Wigner matrices.

Definition 1 ([15]). *Let* $\mathbf{x} = \{X_i\}_{i=1}^{N}$ *be a sequence of sign-valued random variables.* $\mathbf{x}$ *is* r-*independent if any* r *of its elements* $X_{i_1}, \ldots, X_{i_r}$ *are statistically independent,*

$$\mathbb{P}\left[X_{i_1} = b_1, \ldots, X_{i_r} = b_r\right] = \prod_{l=1}^{r} \mathbb{P}\left[X_{i_l} = b_l\right], \quad (18)$$

for any $i_1 \neq \cdots \neq i_r$ *in the range* $[1, N]$ *and* $b_i \in \{\pm 1\}$.

Definition 2 ([15]). *Let a subset* $\mathcal{A}_N^r \subset S_N$ *be endowed with the uniform measure. We say that it is an* r-*independent pseudo-Wigner ensemble of order* N *if the elements of the upper triangular (including the main diagonal) parts of its matrices form an* r-*independent sequence w.r.t. (with respect to) the measure induced on them by* $\mathcal{A}_N^r$.

Definition 3 (r-independent Pseudo-MP Ensemble of order N). *Let a subset* $\mathcal{Y}_{N,p}^r \subset M_{N,p}$ *be endowed with the uniform measure. We say that the ensemble of matrices*

$$\{\mathbf{Y}_N \mathbf{Y}_N^\top \mid \mathbf{Y}_N \in \mathcal{Y}_{N,p}^r\} \quad (19)$$

is an r-**independent pseudo-MP ensemble of order** N *if the elements of the matrices* $\mathbf{Y}_N$ *form an* r-*independent sequence w.r.t. the measure induced on them by* $\mathcal{Y}_{N,p}^r$.

Below, whenever probability measure over $\mathcal{A}_N^r$ or $\mathcal{Y}_{N,p}^r$ are considered, they are always assumed to be uniform as in Definitions 2 and 3.

The last definition is justified by the following result.

Proposition 1. *Let* $q < e$, *then for* $r \leqslant q \log_2 N$ *and any* $\alpha \in (\frac{q}{e}, 1)$ *there exists* N_0 *such that for any* $N \geqslant N_0$, *with probability at least* $1 - \frac{r}{N^{2(1-\alpha)}}$ *a matrix* $\mathbf{Y}_N$ *chosen uniformly from* M_{N,p_N}^{2r} *satisfies*

$$\left| F_{\mathbf{Y}_N^\top \mathbf{Y}_N}(x) - F_{MP}(x) \right| \leqslant \frac{1}{r}, \quad \forall x \in \mathbb{R}. \quad (20)$$

Proof. The proof from [15] applies with minor changes. $\square$

B. High-Order Moments

Lemma 5. *Let* $\{\beta_N\}$, $\{r_N\}$, $\{s_N\} \subset \mathbb{N}$ *be such that* $s_N = o\left(N^{2/3}\right)$ *with* $s_N \leqslant \beta_N r_N$, *and* $\mathbf{A}_N$ *be chosen uniformly from* $\mathcal{A}_N^{\beta_N r_N}$, *then for the expected moments we have*

$$\mathbb{E}\left[\text{Tr}\left(\mathbf{A}_N^{s_N}\right)\right] = \begin{cases} \sqrt{\frac{8}{\pi s_N^3}}N(1 + o(1)), & s_N \text{ even,} \\ 0, & s_N \text{ odd,} \end{cases} \quad (21)$$

as $N \to \infty$. *In addition, the first* $p = 1, \ldots, 2\beta_N$ *moments of the random variable*

$$\text{Tr}\left(\mathbf{A}_N^{s_N}\right) - \mathbb{E}\left[\text{Tr}\left(\mathbf{A}_N^{s_N}\right)\right] \quad (22)$$

converge to the moments of the normal law $\mathcal{N}\left(0, \frac{1}{\pi}\right)$.

Proof. The proof follows that of Main Theorem of [19]. $\square$

Lemma 6. *Let* $\{r_N\}$, $\{s_N\} \subset \mathbb{N}$ *be such that* $s_N = O\left(N^{2/3}\right)$ *with* $s_N \leqslant r_N$, *and* $\mathbf{A}_N$ *be chosen uniformly from* $\mathcal{A}_N^{r_N}$, *then*

$$\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{A}_N^{s_N}\right)\right] \leqslant c_W(s_N)N, \tag{23}$$

where $c_W(s_N)$ *is bounded uniformly over* N.

Proof. The proof is analogous to that of Theorem 2 from [21]. $\square$

Lemma 7. *Let* $\{\beta_N\}$, $\{r_N\}$, $\{s_N\} \subset \mathbb{N}$ *be such that* $s_N = o\left(\sqrt{N}\right)$ *with* $s_N \leqslant \beta_N r_N$, *and* $\mathbf{Y}_N$ *be chosen uniformly from* $\mathcal{Y}_{N,p_N}^{\beta_N r_N}$, *then*

$$\mathbb{E}\left[\mathrm{Tr}\left(\left(\frac{\mathbf{Y}_N^\top \mathbf{Y}_N}{(1+\sqrt{\gamma})^2}\right)^{s_N}\right)\right]$$
$$= \frac{\gamma^{1/4}(1+\sqrt{\gamma})}{2\sqrt{\pi}}\frac{N}{s_N^{3/2}}(1+o(1)), \quad (24)$$

as $N \to \infty$. *In addition, the first* $p = 1, \ldots, 2\beta_N$ *moments of the random variable*

$$\mathrm{Tr}\left(\left(\frac{\mathbf{Y}_N^\top \mathbf{Y}_N}{(1+\sqrt{\gamma})^2}\right)^{s_N}\right) - \mathbb{E}\left[\mathrm{Tr}\left(\left(\frac{\mathbf{Y}_N^\top \mathbf{Y}_N}{(1+\sqrt{\gamma})^2}\right)^{s_N}\right)\right]$$

converge to the moments of the normal law $\mathcal{N}\left(0, \frac{1}{\pi}\right)$.

Proof. The proof is analogous to those of Propositions 2.4 and 2.5 from [24]. $\square$

Lemma 8. *Let* $\{r_N\}$, $\{s_N\} \subset \mathbb{N}$ *be such that* $s_N = O\left(N^{2/3}\right)$ *with* $s_N \leqslant r_N$, *and* $\mathbf{Y}_N$ *be chosen uniformly from* $\mathcal{Y}_{N,p_N}^{r_N}$, *then*

$$\mathbb{E}\left[\mathrm{Tr}\left(\left(\frac{\mathbf{Y}_N^\top \mathbf{Y}_N}{N(1+\sqrt{\gamma})^2}\right)^{s_N}\right)\right] \leqslant c_{MP}(\gamma, s_N)N,$$

where $c_{MP}(\gamma, s_N)$ *is bounded uniformly over* N.

Proof. The proof is analogous to that of Theorem 3.1 from [24]. $\square$

IV. SPECTRAL NORMS

Here we present the main results of the article.

A. Pseudo-Wigner Matrices

Proposition 2. *Let* $\mathbf{A}_n \in \mathcal{A}_N^{r_N}$ *with* $\liminf \frac{r_N}{N^\rho} > 0$ *for some* $\rho \in (0, 1]$, *then for any* $\varepsilon > 0$

$$\|\mathbf{A}_N\| = 1 + o\left(\frac{\log^{1+\varepsilon}N}{N^{\min[\rho, 2/3]}}\right), \quad a.s. \tag{25}$$

Proof. For simplicity, let us start with the case $\rho \leqslant \frac{2}{3}$. Given $\varepsilon > 0$, set

$$q_N = 2\left\lfloor \frac{1}{2}\frac{N^\rho}{\log^{\varepsilon/2}N}\right\rfloor. \tag{26}$$

Using Markov's inequality we obtain the following chain of bounds,

$$\mathbb{P}\left\{\|\mathbf{A}_N\| \geqslant 1 + \frac{\log^{1+\varepsilon}N}{N^\rho}\right\} \tag{27}$$
$$\leqslant \mathbb{P}\left\{\mathrm{Tr}\left(\mathbf{A}_N^{q_N}\right) \geqslant \left(1 + \frac{\log^{1+\varepsilon}N}{N^\rho}\right)^{q_N}\right\}$$
$$= \mathbb{P}\left\{\mathrm{Tr}\left(\mathbf{A}_N^{q_N}\right) \geqslant \left(1 + \frac{\log^{1+\varepsilon}N}{N^\rho}\right)^{2\left[\frac{1}{2}\frac{N^\rho}{\log^{\varepsilon/2}N}\right]}\right\}$$
$$\leqslant \mathbb{P}\left\{\mathrm{Tr}\left(\mathbf{A}_N^{q_N}\right) \geqslant \frac{1}{2}\exp\left(\log^{1+\varepsilon/2}N\right)\right\}$$
$$\leqslant \frac{\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{A}_N^{q_N}\right)\right]}{\frac{1}{2}\exp\left(\log^{1+\varepsilon/2}N\right)} = O\left(N\exp\left(-\log^{1+\varepsilon/2}N\right)\right),$$

where the last line follows from Lemma 6. This implies

$$\sum_{N=1}^\infty \mathbb{P}\left\{\|\mathbf{A}_N\| \geqslant 1 + \frac{\log^{1+\varepsilon}N}{N^\rho}\right\} < +\infty. \tag{28}$$

It now follows from Borel-Cantelli lemma that

$$\|\mathbf{A}_N\| \leqslant 1 + \frac{\log^{1+\varepsilon}N}{N^\rho}, \quad a.s. \tag{29}$$

In order to get the opposite direction inequality, note that Lemma 5 together with the linear algebraic relation

$$\|\mathbf{A}_N\| \leqslant \mathrm{Tr}\left(\mathbf{A}_N^q\right)^{1/q} \leqslant N^{1/q}\|\mathbf{A}_N\|, \tag{30}$$

give

$$\mathbb{E}\left[\|\mathbf{A}_N\|\right] = 1 + o\left(\frac{1}{N^\kappa}\right), \tag{31}$$

for any fixed positive κ and therefore,

$$\|\mathbf{A}_N\| \geqslant 1 + \frac{\log^{1+\varepsilon}N}{N^\rho}, \quad a.s. \tag{32}$$

which together with (29) implies the desired statement.

Assume now that $\rho > \frac{2}{3}$. We know from [21] that Lemma 3 is no longer valid in this case and the expected traces can grow faster that $O(N)$. Therefore, to keep the first ratio in the last line of (27) bounded by a summable sequence, the largest (in order) possible choice for q_N is

$$q_N = 2\left\lfloor \frac{1}{2}\frac{N^{2/3}}{\log^{\varepsilon/2}N}\right\rfloor. \tag{33}$$

Now the same reasoning as above together with Lemma 6 complete the proof. $\square$

B. Pseudo-Wishart Matrices

Proposition 3. *Let* $\mathbf{Y}_N$ *be chosen uniformly from* $\mathcal{Y}_{N,p_N}^{r_N}$ *for some* $\rho \in (0, 1]$, *then for any* $\varepsilon > 0$

$$\left\|\frac{\mathbf{Y}_N^\top \mathbf{Y}_N}{(1+\sqrt{\gamma})^2}\right\| = 1 + o\left(\frac{\log^{1+\varepsilon}N}{N^{\min[\rho, 2/3]}}\right), \quad a.s. \tag{34}$$

Proof. The proof of Proposition 2 works verbatim with Lemmas 7 and 8 replacing Lemmas 5 and 6, correspondingly. $\square$

V. A Construction from Dual BCH codes

Next we provide an explicit constructions of the pseudo-Wigner and pseudo-MP ensembles from dual BCH codes. The idea was presented in [15] for the r-independent pseudo-Wigner matrices with r of the order of $\log_2 N$. Here we focus on higher levels of independence with $r \propto N^\rho$, $\rho > 0$.

For $m \in \mathbb{N}$, a primitive narrow-sense binary BCH code $\mathcal{C}_m^\delta$ of length $n = 2^m - 1$ and designed minimum distance $\delta \geqslant 3$ is a cyclic code generated by the lowest degree binary polynomial having roots $\alpha, \alpha^2, \ldots, \alpha^{\delta-1}$, where α is a primitive element of $GF(2^m)$.

Lemma 9 (Theorem 9.1.1, Theorem 9.2.6 from [25]). *A primitive narrow-sense binary BCH code $\mathcal{C}_m^\delta$ of length $n = 2^m - 1$ and designed distance δ has*

- *minimum distance d such that $\delta \leqslant d \leqslant 2\delta - 1$, and*
- *dimension at least $n - mt$.*

Under the same assumptions as in Lemma 9, the dual BCH code is a cyclic code of dimension $k^\perp \leqslant mt$ [25].

Lemma 10 (Lemma 3.2 from [16]). *If a code $\mathcal{C}$ has minimum distance d, then its dual code $\mathcal{C}^\perp$ is $(d-1)$-independent (see Definition 2) w.r.t. to the uniform measure over its codewords.*

Given these results, the pseudo-Wigner matrices are built as explained in Section as IV of [15]. Pseudo-MP matrices are constructed analogously, by first packing the codewords of the dual BCH code row by row into rectangular $N \times p$ matrices $\mathbf{Y}_N$ scaled by $\frac{1}{\sqrt{N}}$. Then the desired SCMs are obtained as $\mathbf{Y}_N^\top \mathbf{Y}_N$.

VI. Numerical Simulations

To illustrate the results obtained in Section IV, we constructed a BCH code of length $n = 2^{14} - 1 = 16383$ and minimum distance 15 (the generating polynomial was computed by calling `bchgenpoly(16383,16173)` function of MATLAB). Using the obtained polynomial, we calculated the generating polynomial of the dual code as explained in [15] and randomly chose 10^5 words from the dual code. These codewords were packed into 180×180 symmetric sign matrices as described in Section as IV of [15]. In Figure 1 the empirical distribution of the spectral norms of the obtained pseudo-Wigner matrices (dBCH curve in the picture) is compared to the theoretical limit for the truly random matrices, the so-called Tracy-Widom distribution [26].

VII. Conclusions

In this article, we extend the framework of pseudo-Wigner matrices introduced in [15] to a new family of pseudo-Marchenko-Pastur ensembles. The definitions of both classes of matrices are based on the concept of r-independence of the matrix entries to mimic the behavior of the truly random Wigner and sample covariance ensembles, correspondingly. The designed properties of these pseudo-random ensembles
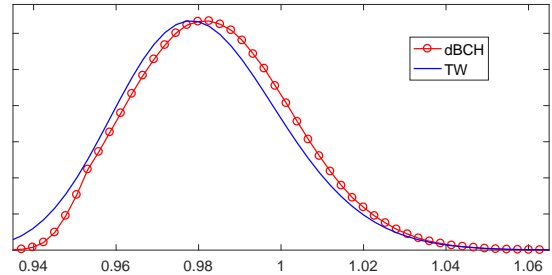

Fig. 1: Distribution of norms of pseudo-Wigner matrices constructed from a dual BCH code, $N = 180$, $m = 14$, $d = 15$ versus Tracy-Widom law.

allow us to derive approximations of the expected moments similar to those for corresponding truly random matrices, which further enables us to achieve bounds on the spectral norms of the pseudo-Wigner and pseudo-MP ensembles as functions of the level of independence r. We also provide explicit constructions of pseudo-Wigner and pseudo-MP ensembles from dual BCH codes.

VIII. Acknowledgment

References

[1] G. Akemann, J. Baik, and P. Di Francesco, "The Oxford handbook of random matrix theory," *Oxford University Press*, 2011.

[2] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Annals of Mathematics*, vol. 62, no. 3, pp. 548–564, 1955.

[3] U. Grenander, "Probabilities on algebraic structures," *John Wiley & Sons, Inc., New York-London*, 1963.

[4] L. Arnold, "On the asymptotic distribution of the eigenvalues of random matrices," *Journal of Mathematical Analysis and Applications*, vol. 20, no. 2, pp. 262–268, 1967.

[5] V. A. Marchenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Matematicheskii Sbornik*, vol. 114, no. 4, pp. 507–536, 1967.

[6] J. E. Gentle, "Random number generation and Monte Carlo methods," *Springer Science & Business Media*, 2013.

[7] H.-J. Zepernick and A. Finger, "Pseudo random signal processing: theory and application," *John Wiley & Sons*, 2013.

[8] S. W. Golomb *et al.*, "Shift register sequences," *Aegean Park Press*, 1982.

[9] I. Reed and R. Stewart, "Note on the existence of perfect maps," *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 10–12, 1962.

[10] F. J. MacWilliams and N. J. A. Sloane, "Pseudo-random sequences and arrays," *Proceedings of the IEEE*, vol. 64, no. 12, pp. 1715–1729, 1976.

[11] H. Imai, "A theory of two-dimensional cyclic codes," *Information and Control*, vol. 34, no. 1, pp. 1–21, 1977.

[12] S. Sakata, "On determining the independent point set for doubly periodic arrays and encoding two-dimensional cyclic codes and their duals," *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 556–565, 1981.

[13] K. G. Paterson, "Perfect maps," *IEEE Transactions on Information Theory*, vol. 40, no. 3, pp. 743–753, 1994.

[14] T. Etzion, "Constructions for perfect maps and pseudorandom arrays," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 1308–1316, 1988.

[15] I. Soloveychik, Y. Xiang, and V. Tarokh, "Pseudo-Wigner matrices," *arXiv:1701.05544*, 2017.

[16] B. Babadi and V. Tarokh, "Spectral distribution of random matrices from binary linear block codes," *IEEE Transactions of Information Theory*, vol. 57, no. 6, pp. 3955–3962, 2011.

[17] E. P. Wigner, "On the distribution of the roots of certain symmetric matrices," *Annals of Mathematics*, pp. 325–327, 1958.

[18] G. W. Anderson, A. Guionnet, and O. Zeitouni, "An introduction to random matrices," *Cambridge university press*, vol. 118, 2010.

[19] Y. Sinai and A. Soshnikov, "Central limit theorem for traces of large random symmetric matrices with independent matrix elements," *Boletim da Sociedade Brasileira de Matemática*, vol. 29, no. 1, pp. 1–24, 1998.

[20] ——, "A refinement of Wigner's semicircle law in a neighborhood of the spectrum edge for random symmetric matrices," *Functional Analysis and Its Applications*, vol. 32, no. 2, pp. 114–131, 1998.

[21] A. Soshnikov, "Universality at the edge of the spectrum in Wigner random matrices," *Communications in Mathematical Physics*, vol. 207, no. 3, pp. 697–733, 1999.

[22] ——, "A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices," *Journal of Statistical Physics*, vol. 108, no. 5, pp. 1033–1056, 2002.

[23] I. M. Johnstone, "On the distribution of the largest eigenvalue in principal components analysis," *Annals of statistics*, pp. 295–327, 2001.

[24] S. Péché, "Universality results for the largest eigenvalues of some sample covariance matrix ensembles," *Probability Theory and Related Fields*, vol. 143, no. 3, pp. 481–516, 2009.

[25] F. J. MacWilliams and N. J. A. Sloane, "The theory of error correcting codes," *Elsevier*, 1977.

[26] C. A. Tracy and H. Widom, "Level-spacing distributions and the Airy kernel," *Communications in Mathematical Physics*, vol. 159, no. 1, pp. 151–174, 1994.