The Tracy–Widom law for some sparse random matrices

Sasha Sodin*

January 19, 2011

Abstract

Consider the random matrix obtained from the adjacency matrix of a random d-regular graph by multiplying every entry by a random sign. The largest eigenvalue converges, after proper scaling, to the Tracy–Widom distribution.

1 The result

Let $G = (V, E)$ be an element of $\mathcal{G}(N, d)$, i.e. a random uniformly chosen d-regular graph with N vertices, and let $S : E \to \{-1, +1\}$ be independent random signs:

$$\mathbb{P}\big\{S(u, v) = 1\big\} = \mathbb{P}\big\{S(u, v) = -1\big\} = 1/2 \ .$$

Consider the $N \times N$ (random) Hermitian matrix M,

$$M_{uv} = \begin{cases} S(u, v) \ , & (u, v) \in E \\ 0, & (u, v) \notin E \end{cases}$$

with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$.

*Tel Aviv University, sodinale@post.tau.ac.il. Supported in part by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities and by the ISF.

Theorem. *Assume that M is as above, with $3 \leq d \ll N^{2/3}$. The scaled largest eigenvalue*

$$2 \left(\frac{d(d-1)}{(d-2)^2} N \right)^{2/3} \left\{ \frac{\lambda_N}{2\sqrt{d-1}} - 1 \right\} \tag{1}$$

converges in distribution to the Tracy–Widom law TW_1. The same is true for the scaled smallest eigenvalue

$$-2 \left(\frac{d(d-1)}{(d-2)^2} N \right)^{2/3} \left\{ \frac{\lambda_1}{2\sqrt{d-1}} + 1 \right\} . \tag{2}$$

The proof is based on the combinatorial arguments from [2], which can be seen as a modification of Soshnikov's approach [9] to the asymptotic distribution of the extreme eigenvalues of random Hermitian matrices with independent entries. Another crucial ingredient is an estimate of McKay [4] on the number of subgraphs of a random graph.

2 Connection to non-backtracking walks

Consider the matrices

$$M^{(n)} = (d-1)^{n/2} \left\{ U_n \left(\frac{M}{2\sqrt{d-1}} \right) - \frac{1}{d-1} U_{n-2} \left(\frac{M}{2\sqrt{d-1}} \right) \right\} , \tag{3}$$

where

$$U_n(\cos\theta) = \frac{\sin((n+1)\theta)}{\sin\theta}$$

are the Chebyshev polynomials of the second kind, and formally

$$U_{-2} \equiv U_{-1} \equiv 0 .$$

It is not hard to see (cf. [8]) that, for any $u_0, u_n \in V$,

$$M^{(n)}_{u_0 u_n} = \sum{}' M_{u_0 u_1} M_{u_1 u_2} \cdots M_{u_{n-1} u_n} ,$$

where the sum is over all paths $p_n = u_0 u_1 u_2 \cdots u_n$ such that

(a) $(u_j, u_{j+1}) \in E$ for any $0 \leq j \leq n-1$;

(b) $u_{j+2} \neq u_j$ for $0 \leq j \leq n-2$ (p_n is *non-backtracking*.)

Therefore

$$\operatorname{tr} M^{(n)} = \sum\nolimits'' M_{u_0 u_1} M_{u_1 u_2} \cdots M_{u_{n-1} u_n} \ ,$$

where the sum is over paths p_n satisfying (a), (b), and

(c) $u_n = u_0$.

Taking the expectation (over S) and noting that $\mathbb{E}_S M_{uv}^k = ((-1)^k + 1)/2$, we deduce:

Lemma 1. *The expectation $\mathbb{E}_S \operatorname{tr} M^{(n)}$ is equal to the number $P_n(G)$ of paths p_n that satisfy (a), (b), (c), and*

(d) *every edge $(u, v) \in E$ appears an even number of times on p_n.*

In particular, $\mathbb{E} \operatorname{tr} M^{(2n+1)} = 0$, we shall therefore study $\mathbb{E} \operatorname{tr} M^{(2n)} = \mathbb{E}_G P_{2n}(G)$, where $\mathbb{E}_G$ denotes expectation over the random choice of the graph G.

3 Diagrams

Following [2], we group the paths p_{2n} satisfying (a)-(d) into (topological) equivalence classes, which are in one-to-one correspondence with *diagrams*:

Definition 2. Let $\beta \in \{1, 2\}$.

- A *diagram* is an (undirected) multigraph $\bar{G} = (\bar{V}, \bar{E})$, together with a circuit $\bar{p} = \bar{u}_0 \bar{u}_1 \cdots \bar{u}_0$ on $\bar{G}$, such that
  - $\bar{p}$ is *non-backtracking* (meaning that no edge is followed by its reverse, unless the edge is a loop $\bar{u}\bar{u}$);
  - For every $(\bar{u}, \bar{v}) \in \bar{E}$,
  $$\# \{j \,|\, \bar{u}_j = \bar{u}, \, \bar{u}_{j+1} = \bar{v}\} + \# \{j \,|\, \bar{u}_j = \bar{v}, \, \bar{u}_{j+1} = \bar{u}\} = 2 \ ;$$
  - the degree of $\bar{u}_0$ in $\bar{G}$ is 1; the degrees of all the other vertices are equal to 3.
- A *weighted diagram* is a diagram $\bar{G}$ together with a weight function $\bar{w} : \bar{E} \to \{-1, 0, 1, 2, \cdots\}$.

Informally, if $(u_1, u_2), (u_2, u_3), \cdots, (u_{k-1}, u_k)$ is a sequence of edges with

$$\deg u_1 = \deg u_k = 3, \quad \deg u_2 = \cdots = \deg u_{k-1} = 2 \ ,$$

we replace $u_1 u_2 \cdots u_k$ with a single edge $\bar{e} = (\bar{u_1}, \bar{u_k})$, to which we assign the weight $\bar{w}(\bar{e}) = k - 2$. Thus, the path

1.2.3.4.5.6.7.4.6.5.7.4.3.2.1

corresponds to the diagram in Figure 1 (left), whereas

1.2.3.4.5.6.7.8.9.7.8.9.6.5.10.11.12.13.14.15.16.12.13.14.15.16.12.11.10.5.4.3.2.1

corresponds to Figure 1 (right).

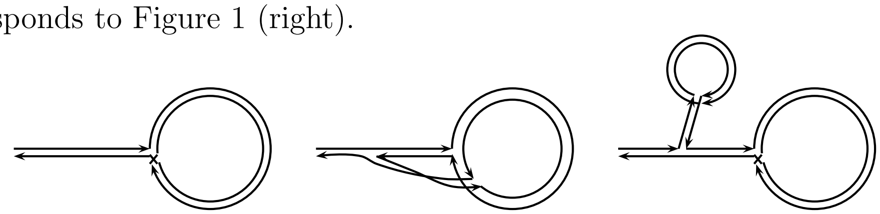

Figure 1: Some diagrams: $s = 1$ (left), $s = 2$ (center, right)

More complicated paths (with vertices of degree greater than 3) also correspond to weighted diagrams; the formal correspondence is described in [2, II.1].

The following lemma summarises the results of [2, II.1-II.2] that we shall use:

Lemma 3.

1. *(Claim II.2.2.) All the diagrams can be classified according to a parameter $s = 1, 2, \cdots$ (see Figure 1.) A diagram with parameter s has $\#\bar{V} = 2s$ vertices and $\#\bar{E} = 3s - 1$ edges.*

2. *(Proposition II.2.3) The number $D_1(s)$ of diagrams corresponding to a given value s satisfies*

$$(s/C)^s \leq D_1(s) \leq C^{s-1} s^s \ ,$$

where $C > 1$ is a numerical constant.

3. *For a diagram δ with parameter s, let $W(\delta)$ be the number of associated weighted diagrams, and $\widetilde{W}(\delta)$ – the number of associated weighted diagrams with positive weights. Then*

$$\binom{n-3s}{3s-2} \leq \widetilde{W}(\delta) \leq W(\delta) \leq \binom{n+3s-2}{3s-2} .$$

In particular, if $s = o(\sqrt{n})$,

$$\widetilde{W}(\delta), W(\delta) = (1+o(1)) \frac{n^{3s-2}}{(3s-2)!} .$$

4. *Every path corresponding to δ has at most $n-s+1$ distinct vertices. If the associated weighted diagram has positive weights, the path has exactly $n-s+1$ distinct vertices; of these, 1 is of degree 1, $n-3s+1$ are of degree 2, and $2s-1$ are of degree 3. A general path is obtained from one as above by identifying several pairs of connected vertices, and erasing the connecting edges.*

5. *(Claim II.1.4) The number of different paths corresponding to a given weighted diagram is at most N^{n-s+1}. If the weights are positive, the number is exactly*

$$N(N-1)\cdots(N-n+s) .$$

4 Counting the paths

Fix a diagram δ with parameter s. We shall now estimate the expected number of paths p_{2n} associated to δ in a random graph G. We shall use the following special case of [4, Theorem 2.10].

Let L be a subgraph of the complete (labelled) graph K_N; let ℓ_u be the degree of u in L, $E_G = \#E(G) = dN/2$, and $E_L = \#E(L) = \sum \ell_u/2$. Finally, denote $a^{[b]} = a!/(a-b)!$

Proposition 4 (McKay)**.** *If $E_G - E_L \geq 3d(d+1)$, then the probability F_L that a random d-regular graph G contains L satisfies*

$$\xi(L) \, \frac{\prod_u d^{[\ell_u]}}{2^{E_L} E_G^{[E_L]}} \leq F_L \leq \Xi(L) \, \frac{\prod_u d^{[\ell_u]}}{2^{E_L} E_G^{[E_L]}} ,$$

where

$$\xi(L) = \left\{ \frac{1 - \frac{d(d+1)}{2(E_G - E_L - 2d(d+1))}}{1 + \frac{d^2}{2(E_G - 2d^2 - \frac{e-1}{e} E_L)}} \right\}^{E_L} \frac{E_G^{[E_L]}}{(E_G - 1)^{[E_L]}} ,$$

$$\Xi(L) = \frac{E_G^{[E_L]}}{(E_G - 2d^2)^{[E_L]}} .$$

We shall deduce the following estimates:

Lemma 5.

1. *For $n \leq N$,*

$$\mathbb{E}P_{2n}(G) \leq n(d-2)(d-1)^{n-1} \exp\left\{ \frac{Cn^{3/2}}{N^{1/2}} (1 + d/\sqrt{nN}) \right\} , \qquad (4)$$

   where $C > 0$ is a numerical constant.

2. *For $d^2/N \ll n \ll \min(\sqrt{n}, N/d)$,*

$$\mathbb{E}P_{2n}(G) = (1 + o(1)) \, (d-1)^n \sum_{s \geq 1} \frac{(d-2)^{2s-1}}{d^{s-1}(d-1)^s N^{s-1}} \frac{n^{3s-2}}{(3s-2)!} D_1(s) . \qquad (5)$$

Proof. Let p_{2n} be a path associated to a diagram δ with parameter s, and let L be the induced (undirected) graph:

$$E(L) = \{(u_j, u_{j+1})\} .$$

Then, by item 4 of Lemma 3,

$$\begin{aligned} \frac{\prod_u d^{[\ell_u]}}{2^{E_L} E_G^{[E_L]}} &\leq \frac{d^{n-s+1}(d-1)^{n-s}(d-2)^{2s-1}}{2^n (dN/2)^{[n]}} \\ &= \frac{(d-2)^{2s-1}}{d^{s-1}(d-1)^s} \left(\frac{d-1}{N} \right)^n \frac{(dN/2)^n}{(dN/2)^{[n]}} \\ &\leq \exp\left\{ \frac{C_1 n^2}{dN} \right\} \left(\frac{d-1}{N} \right)^n \frac{(d-2)^{2s-1}}{d^{s-1}(d-1)^s} , \end{aligned} \qquad (6)$$

where $C_1 > 0$ is a numerical constant. Also,

$$\Xi(L) = \frac{(dN/2)^{[n]}}{(dN/2 - 2d^2)^{[n]}} \leq \exp\left\{C_2 nd/N\right\} \ ;$$

therefore

$$F_L \leq \exp\left\{\frac{C_1 n^2}{dN} + \frac{C_2 nd}{N}\right\} \left(\frac{d-1}{N}\right)^n \frac{(d-2)^{2s-1}}{d^{s-1}(d-1)^s} \ .$$

According to item 3, the number of ways to choose the weights on δ is at most

$$\binom{n+3s-2}{3s-2} \leq \frac{n^{3s-2}}{(3s-2)!} \ ,$$

and, according to item 4, the number of ways to choose the vertices is at most N^{n-s+1}. Finally, the number of diagrams with parameter s is at most $C^{s-1}s^s$. Therefore

$$\begin{aligned}
\mathbb{E}P_{2n}(G) &\leq \sum_{s\geq 1} C^{s-1} s^s N^{n-s+1} \frac{n^{3s-2}}{(3s-2)!} \\
&\qquad \times \exp\left\{\frac{C_1 n^2}{dN} + \frac{C_2 nd}{N}\right\} \left(\frac{d-1}{N}\right)^n \frac{(d-2)^{2s-1}}{d^{s-1}(d-1)^s} \\
&\leq n(d-2)(d-1)^{n-1} \exp\left\{\frac{C_1 n^2}{dN} + \frac{C_2 nd}{N}\right\} \qquad (7) \\
&\qquad \times \sum_{s\geq 1} \frac{(C_3 n^3/N)^{s-1}}{(2(s-1))!} \\
&\leq n(d-2)(d-1)^{n-1} \exp\left\{\frac{C_4 n^{3/2}}{N^{1/2}}(1 + d/\sqrt{nN})\right\} \ .
\end{aligned}$$

This proves the first statement.

Suppose $d^2/N \ll n \ll \min(\sqrt{N}, N/d)$. Choose s_0 such that

$$\begin{cases} s_0 = 1 \ , & n \leq N^{1/4} \\ n^{3/2}/N^{1/2} \ll s_0 \ll \sqrt{n} \ , & n > N^{1/4} \ , \end{cases}$$

For $s \leq s_0$ and positive weights, the inequalities above are asymptotic equalities. Namely,

$$F_L = (1 + o(1)) \left(\frac{d-1}{N}\right)^n \frac{(d-2)^{2s-1}}{d^{s-1}(d-1)^s} \ ,$$

$$\widetilde{W}(\delta) = (1 + o(1)) \frac{n^{3s-2}}{(3s-2)!} \ .$$

Hence the contribution of paths with positive weights and $s \leq s_0$ is

$$(1+o(1)) \sum_{1 \leq s \leq s_0} D_1(s) N^{n-s+1} \frac{n^{3s-2}}{(3s-2)!} \left(\frac{d-1}{N} \right)^n \frac{(d-2)^{2s-1}}{d^{s-1}(d-1)^s} \ .$$

The contribution of paths with $s \leq s_0$ and not all weights positive is negligible, since

$$W(\delta) - \widetilde{W}(\delta) = o(1) \frac{n^{3s-2}}{(3s-2)!} \ .$$

The contribution of paths with $s > s_0$ is also negligible, by the estimates in the proof of item 1 of this lemma (see (7) above.) Thus

$$\begin{aligned} \mathbb{E} P_{2n}(G) &= (1+o(1)) \sum_{1 \leq s \leq s_0} D_1(s) N^{n-s+1} \frac{n^{3s-2}}{(3s-2)!} \left(\frac{d-1}{N} \right)^n \frac{(d-2)^{2s-1}}{d^{s-1}(d-1)^s} \\ &= (1+o(1)) \sum_{1 \leq s < \infty} D_1(s) N^{n-s+1} \frac{n^{3s-2}}{(3s-2)!} \left(\frac{d-1}{N} \right)^n \frac{(d-2)^{2s-1}}{d^{s-1}(d-1)^s} \ . \end{aligned}$$

□

5 Conclusion of the proof

Applying Lemma 1 and using the definition of $M^{(2n)}$, we immediately deduce:

Corollary 6.

1. *For $n \leq N$,*

$$\mathbb{E} \operatorname{tr} U_{2n} \left\{ \frac{M}{2\sqrt{d-1}} \right\} \leq n \exp \left\{ \frac{C_4 n^{3/2}}{N^{1/2}} (1 + d/\sqrt{nN}) \right\} \ ,$$

   where $C > 0$ is a numerical constant.

2. *For $d^2/N \ll n \ll \min(\sqrt{N}, N/d)$,*

$$\mathbb{E} \operatorname{tr} U_{2n} \left\{ \frac{M}{2\sqrt{d-1}} \right\} = (1 + o(1))\, n \sum_{s \geq 1} \left(\frac{(d-2)^2 n^3}{d(d-1)N} \right)^{s-1} \frac{D_1(s)}{(3s-2)!} \ .$$

Now, by [2, Theorem I.5.3, Proposition II.2.5], an $N' \times N'$ random matrix A_{inv} from the Gaussian Orthogonal Ensemble GOE(N') satisfies, for $n = o(\sqrt{N'})$,

$$\mathbb{E} U_{2n}\left(A_{\mathrm{inv}}/(2\sqrt{N'-2})\right) = (1 + o(1))\, n \sum_{s \geq 1} (n^3/N')^{s-1} \frac{D_1(s)}{(3s-2)!} \;.$$

Hence, for $N' = \lfloor \frac{d(d-1)}{(d-2)^2} N \rfloor$, the second item of Corollary 6 implies:

$$\mathbb{E}\,\mathrm{tr}\, U_{2n}\left\{\frac{M}{2\sqrt{d-1}}\right\} = (1 + o(1))\mathbb{E}\,\mathrm{tr}\, U_{2n}\left\{\frac{A_{\mathrm{inv}}}{2\sqrt{N-2}}\right\} \;.$$

Applying the arguments of [2, II.3], this is extended to expectations of products of several such traces.

As in [2, I.5] this implies that the random point process

$$\left\{ 2\left(\frac{d(d-1)}{(d-2)^2}\, N\right)^{2/3} \left[\frac{\lambda_j}{2\sqrt{d-1}} - 1\right] \right\}_{j=1}^{N}$$

converges (in the appropriate topology, [2, I.2]) to the Airy point process $\mathfrak{Ai}_1$, and hence the distribution of (1) converges to TW_1.

$\square$

6 Some remarks

1. This note continues the study of the distribution of the scaled largest eigenvalues of random matrices with independent entries. For the Gaussian Orthogonal Ensemble, the limiting distribution was identified by Tracy and Widom, building on several earlier works; see [10] and references therein. This result was extended to a rather general class of matrices with independent entries by Soshnikov [9]. Important further results have been obtained by Ruzmaikina [7], and Khorunzhiy and Vengerovsky [3]. A different proof of Soshnikov's theorem was given in [2]; it is the basis of the present argument.

2. In the physical literature, the spectrum of random matrices similar to M has been recently studied by Oren, Godel, and Smilansky [6], who have focused on the bulk of the spectrum (rather than the extreme eigenvalues).

3. In the published version of this paper [1], we stated that the conclusion of the theorem remains valid for more general weights $S(u, v)$, such that

[1] J. Stat. Phys. 136 (2009), no. 5, 834-841.

(A1) The distribution of every $S(u, v)$ is symmetric;

(A2) $\mathbb{E}S(u, v)^{2k} \le (C_0 k)^k$;

(A3$_1$) $\mathbb{E}S(u, v)^2 = 1$.

Guilhem Semerjian and Victor Bapst have kindly brought to our attention that this can not be true, and in fact, for any distribution of the entries that differs from the one we consider, the random matrix M has Lifshitz tails.

4. One may consider complex weights $S(u, v)$ that are uniformly distributed on the unit circle. Then the scaled extreme eigenvalues (1), (2) converge in distribution to the Tracy–Widom law TW_2.

5. If all the weights $S(u, v)$ are equal to 1, the matrix M is the adjacency matrix A_G of G. Obviously, the largest eigenvalue of A_G is equal to d. Numerical evidence due to Miller, Novikoff, and Sabelli [5] indicates that the second largest eigenvalue λ_2 of A_G converges to TW_1 after proper scaling. If true, this would have important consequences, see [5] and references therein.

6. Our result can be reformulated as a bound for the "new" eigenvalues of a random 2-lift of a random d-regular graph. We refer the reader to the work of Bilu and Linial [1] for the definitions.

Acknowledgment. I am grateful to my supervisor Vitali Milman for his support and encouragement. I thank Nathan Linial, Steven J. Miller, Idan Oren, and Uzy Smilansky for their remarks on a preliminary version of this note. I thank Guilhem Semerjian and Victor Bapst for spotting the error in Remark 3.

References

[1] Y. Bilu, N. Linial, *Lifts, discrepancy and nearly optimal spectral gaps*, Combinatorica, 26(2006) 495–519.

[2] O. N. Feldheim, S. Sodin, *A universality result for the smallest eigenvalues of certain sample covariance matrices*, to appear in Geom. Funct. Anal., preprint: arXiv:0812.1961.

[3] O. Khorunzhiy, V. Vengerovsky, *Even Walks and Estimates of High Moments of Large Wigner Random Matrices*, preprint: arXiv:0806.0157

[4] B. D. McKay, *Subgraphs of random graphs with specified degrees*, Proceedings of the Twelfth Southeastern Conference on Combinatorics, Graph Theory and Computing, Vol. II (Baton Rouge, La., 1981). Congr. Numer. 33 (1981), 213–223.

[5] S. J. Miller, T. Novikoff, A. Sabelli, *The distribution of the largest non-trivial eigenvalues in families of random regular graphs*, Experiment. Math. 17 (2008), no. 2, 231–244.

[6] I. Oren, A.Godel, U. Smilansky, *Trace Formulae and Spectral Statistics for Discrete Laplacians on Regular Graphs*, in preparation

[7] A. Ruzmaikina, *Universality of the edge distribution of eigenvalues of Wigner random matrices with polynomially decaying distributions of entries*, Comm. Math. Phys. 261 (2006), no. 2, 277–296.

[8] S. Sodin, *Random matrices, nonbacktracking walks, and orthogonal polynomials*, J. Math. Phys. 48 (2007), no. 12.

[9] A. Soshnikov, *Universality at the edge of the spectrum in Wigner random matrices*, Comm. Math. Phys. 207 (1999), no. 3, 697–733.

[10] C. Tracy, H. Widom, *On orthogonal and symplectic matrix ensembles*, Comm. Math. Phys. 177 (1996), no. 3, 727–754.