On the probability that random graphs are Ramanujan

Steven J Miller, Anthony Sabelli (Brown University)
Tim Novikoff (Cornell University)

Slides and paper available at
http://www.math.brown.edu/∼sjmiller

Expanders and Ramanujan Graphs:
Construction and Applications
AMS National Meeting, San Diego, January 9, 2008.

Conjectures

Expanders and Eigenvalues

- Expanding Constant
  $h(G) := \inf\left\{\frac{|\partial U|}{\min(|U|,|V\backslash U|)} : U \subset V, |U| > 0\right\}$
- $\{G_m\}$ family of expanders if $\exists\epsilon$ with $h(G_m) \geq \epsilon$ and $|G_m| \to \infty$.
- Cheeger-Buser Inequalities
  $\frac{d-\lambda_2(G)}{2} \leq h(G) \leq 2\sqrt{2d(d-\lambda_2(G))}$
- Applications: sparse ($|E|$ grows at most linearly with $|V|$), highly connected.
  ◇ communication network theory:
  superconcentrators, nonblocking networks
  ◇ coding theory, cryptography.

Known and conjectured results for λ_2

- (Alon-Boppana, Burger, Serre) $\{G_m\}$ family of finite connected d-regular graphs, $\lim_{m\to\infty}|G_m| = \infty$:

$$\liminf_{m\to\infty} \lambda_2(G_m) \geq 2\sqrt{d-1}$$

- As $|G| \to \infty$, for $d \geq 3$ and any $\epsilon > 0$, “most" d-regular graphs G have

$$\lambda_2(G) \leq 2\sqrt{d-1} + \epsilon$$

(conjectured by Alon, proved for many families by Friedman).

Questions

For a family of d-regular graphs:

- What is the *distribution* of λ_2?
- What *percent* of the graphs are Ramanujan?

$\lambda(G) = \max(\lambda_+(G), \lambda_-(G))$, where $\lambda_\pm(G)$ are largest non-trivial positive (negative) eigenvalues. If bipartite $\lambda_-(G) = -\lambda_+(G)$. If connected $\lambda_2(G) = \lambda_+(G)$.

Families Investigated (N even)

- $\mathcal{CI}_{N,d}$: d-regular connected graphs generated by choosing d perfect matchings.

- $\mathcal{SCI}_{N,d}$: subset of $\mathcal{CI}_{N,d}$ that are simple.

- $\mathcal{CB}_{N,d}$: d-regular connected bipartite graphs generated by choosing d permutations.

- $\mathcal{SCB}_{N,d}$: subset of $\mathcal{CB}_{N,d}$ that are simple.

Tracy-Widom Distribution

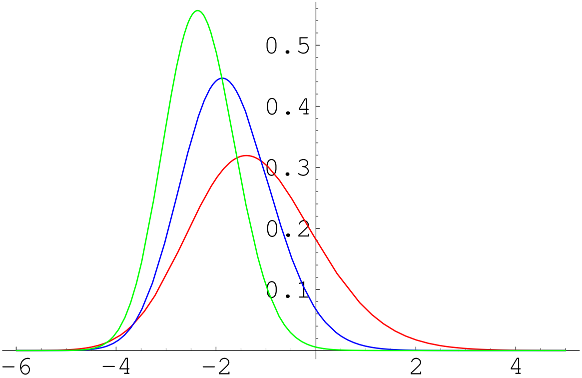
Limiting distribution of the normalized largest eigenvalues for ensembles of matrices: GOE ($\beta = 1$), GUE ($\beta = 2$), GSE ($\beta = 4$)

Applications

- Length of largest increasing subsequence of random permutations.
- Largest principle component of covariances matrices.
- Young tableaux, random tilings, queuing theory, superconductors....

Tracy-Widom Plots

Plots of the three Tracy-Widom distributions: $f_1(s)$ is red, $f_2(s)$ is blue and $f_4(s)$ is green.



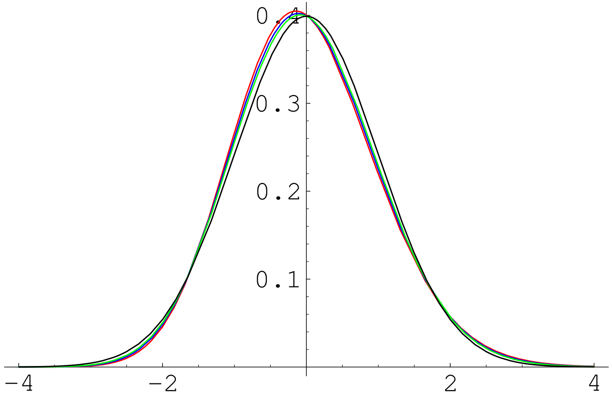
Tracy-Widom Distributions

Parameters for the Tracy-Widom distributions. F_β is the cumulative distribution function for f_β, and $F_\beta(\mu_\beta)$ is the mass of f_β to the left of its mean.

	Mean μ	Std Dev σ	$F_\beta(\mu_\beta)$
TW($\beta = 1$)	-1.21	1.268	0.5197
TW($\beta = 2$)	-1.77	0.902	0.5150
TW($\beta = 4$)	-2.31	0.720	0.5111
Std Normal	0.00	1.000	0.5000

Normalized Tracy-Widom Plots

Plots normalized to have mean 0 and variance 1: $f_1^{\rm norm}(s)$ is red, $f_2^{\rm norm}(s)$ is blue, $f_4^{\rm norm}(s)$ is green, standard normal is black.



Conjectures

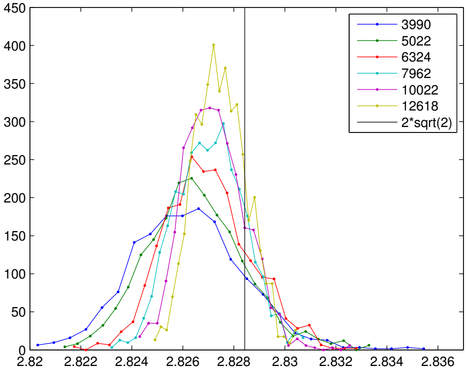
Conjectures

- The distribution of $\lambda_\pm(G)$ converges to the $\beta = 1$ Tracy-Widom distribution as $N \to \infty$ in all studied families.
- For non-bipartite families, $\lambda_\pm(G)$ are independent.
- The percent of the graphs that are Ramanujan approaches 52% as $N \to \infty$ (resp., 27%) in bipartite (resp., non-bipartite) families.

Evidence weaker for $\mathcal{CB}_{N,d}$ (d-regular connected bipartite graphs, not necessarily simple).

Distribution of $\lambda_+(G)$

Distribution of $\lambda_+(G)$ for 1000 graphs randomly chosen from $\mathcal{CI}_{N,3}$ for various N (vertical line is $2\sqrt{2}$).



Statistical evidence for conjectures

- Well-modeled by Tracy-Widom with $\beta = 1$.
- Means approach $2\sqrt{d-1}$ according to power law.
- Variance approach 0 according to power law.
- Comparing the exponents of the power laws, see the number of standard deviations that $2\sqrt{d-1}$ falls to the right of the mean goes to 0 as $N \to \infty$.
- $\lambda_{\pm}(G)$ appear independent in non-bipartite families.
- As $N \to \infty$ the probability that a graph is Ramanujan is the mass of the Tracy-Widom distribution to the left of its mean (52%) if bipartite (27% otherwise).

Power law exponents of means and standard deviations

- Means: $\mu_{\mathcal{F}_{N,d}} \approx 2\sqrt{d-1} - c_{\mu,N,d} N^{m(\mathcal{F}_{N,d})}$
- Standard Deviations: $\sigma_{\mathcal{F}_{N,d}} \approx c_{\sigma,N,d} N^{s(\mathcal{F}_{N,d})}$
- Thus $2\sqrt{d-1} \approx \mu_{\mathcal{F}_{N,d}} + \frac{c_{\mu,N,d}}{c_{\sigma,N,d}} N^{m(\mathcal{F}_{N,d})-s(\mathcal{F}_{N,d})} \sigma_{\mathcal{F}_{N,d}}$

Ramanujan Threshold

As $N \to \infty$, if $m(\mathcal{F}_{N,d}) < s(\mathcal{F}_{N,d})$ then $2\sqrt{d-1}$ falls zero standard deviations to the right of the mean.

3-Regular Graphs

Experiments: Comparisons with Tracy-Widom Distribution

- Each set is 1000 random 3-regular graphs from $\mathcal{CI}_{N,3}$ normalized to have mean 0 and variance 1.
- 19 degrees of freedom, critical values 30.1435 ($\alpha = .05$) and 36.1908 ($\alpha = .01$).
- Only showing subset of data.

χ^2-Tets of $\lambda_+(G)$ for $\mathcal{CI}_{N,3}$ versus Tracy-Widom Distributions

Critical values: 30.1 ($\alpha = .05$), 36.2 ($\alpha = .01$).

N	$\mathrm{TW}_1^{\mathrm{norm}}$	$\mathrm{TW}_2^{\mathrm{norm}}$	$\mathrm{TW}_4^{\mathrm{norm}}$	$\mathrm{N}(0,1)$
26	52.4	43.7	36.8	30.3
100	72.1	41.3	28.9	13.2
796	3.7	4.9	7.0	19.3
3168	17.4	19.6	24.0	61.3
6324	20.8	19.8	21.4	28.6
12618	9.9	9.3	10.6	17.2
20000	37.4	41.1	41.4	71.2
mean (all)	32.5	27.2	24.9	49.1
median (all)	20.0	19.1	18.0	25.2
mean (last 10)	22.3	24.9	29.1	66.7
median (last 10)	21.2	21.8	22.2	64.5

χ^2-Tests of $\lambda_+(G)$ against $\beta = 1$ Tracy-Widom

Critical values: 30.1 ($\alpha = .05$), 36.2 ($\alpha = .01$).

N	$\mathcal{CI}_{N,3}$	$\mathcal{SCI}_{N,3}$	$\mathcal{CB}_{N,3}$	$\mathcal{SCB}_{N,3}$
26	52.4	111.6	142.7	14.3
100	72.1	19.8	23.4	18.5
796	3.7	14.9	20.9	19.6
3168	17.4	22.2	70.6	25.4
12618	9.9	13.1	36.9	13.7
20000	37.4	14.9	27.4	12.1
mean (all)	32	21	78	19
standard deviation (all)	42	18	180	7
mean (last 10)	22	17	44	17
standard deviation (last 10)	8	5	37	8
mean (last 5)	22	17	32	14
standard deviation (last 5)	10	4	23	1

Experiment: Mass to the left of the mean for $\lambda_+(G)$

- Each set of 1000 3-regular graphs.
- mass to the left of the mean of the Tracy-Widom distributions:
  - ◇ 0.519652 ($\beta = 1$)
  - ◇ 0.515016 ($\beta = 2$)
  - ◇ 0.511072 ($\beta = 4$)
  - ◇ 0.500000 (standard normal).
- two-sided z-test: critical thresholds: 1.96 (for $\alpha = .05$) and 2.575 (for $\alpha = .01$).

Experiment: Mass to the left of the mean for $\mathcal{CI}_{N,3}$

Critical values: 1.96 ($\alpha = .05$), 2.575 ($\alpha = .01$).

N	Obs mass	$z_{\mathrm{TW},1}$	$z_{\mathrm{TW},2}$	$z_{\mathrm{TW},4}$	z_{StdNorm}
26	0.483	-2.320	-2.026	-1.776	-1.075
100	0.489	-1.940	-1.646	-1.396	-0.696
796	0.521	0.085	0.379	0.628	1.328
6324	0.523	0.212	0.505	0.755	1.455
20000	0.526	0.402	0.695	0.944	1.644
μ (last 10)	0.518	0.473	0.531	0.655	1.202
$\widetilde{\mu}$ (last 10)	0.523	0.411	0.537	0.755	1.455
μ (last 5)	0.517	0.591	0.532	0.630	1.050
$\widetilde{\mu}$ (last 5)	0.515	0.421	0.695	0.700	0.949

Experiment: Mass left of mean: 3-Regular, sets of 100,000

Discarded: Matlab's algorithm didn't converge.
Critical values: 1.96 ($\alpha = .05$), 2.575 ($\alpha = .01$).

$\mathcal{CI}_{N,3}$	$z_{\mathrm{TW},1}$	$z_{\mathrm{TW},2}$	$z_{\mathrm{TW},4}$	z_{StdNorm}	Discarded
1002	0.239	3.173	5.667	12.668	0
2000	-0.128	2.806	5.300	12.301	0
5022	1.265	4.198	6.692	13.693	0
10022	0.391	3.324	5.819	12.820	0
40000	2.334	5.267	7.761	14.762	0

$\mathcal{SCI}_{N,3}$	$z_{\mathrm{TW},1}$	$z_{\mathrm{TW},2}$	$z_{\mathrm{TW},4}$	z_{StdNorm}	Discarded
1002	-1.451	1.483	3.978	10.979	0
2000	-0.457	2.477	4.971	11.972	0
5022	-0.042	2.891	5.386	12.387	1

Experiment: Mass left of mean: 3-Regular, sets of 100,000

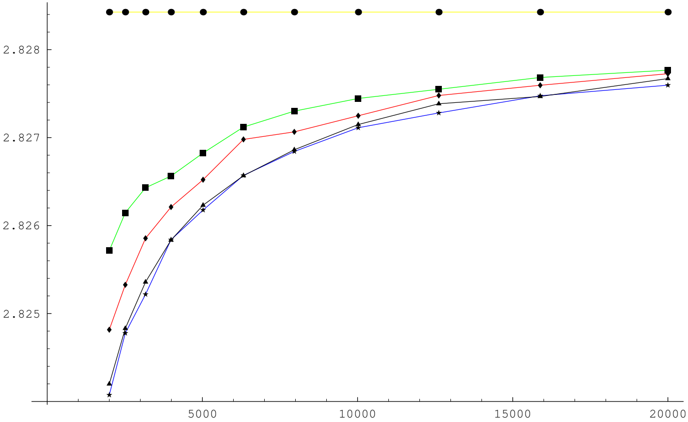
Critical values: 1.96 ($\alpha = .05$), 2.575 ($\alpha = .01$).

$\mathcal{CB}_{N,3}$	$z_{\mathrm{TW},1}$	$z_{\mathrm{TW},2}$	$z_{\mathrm{TW},4}$	z_{StdNorm}	Discarded
1002	3.151	6.083	8.577	15.577	0
2000	3.787	6.719	9.213	16.213	1
5022	3.563	6.495	8.989	15.989	4
10022	2.049	4.982	7.476	14.477	0
12618	3.701	6.634	9.127	16.128	0
15886	2.999	5.931	8.425	15.426	0
20000	2.106	5.039	7.533	14.534	0
40000	1.853	4.786	7.280	14.281	0

$\mathcal{SCB}_{N,3}$	$z_{\mathrm{TW},1}$	$z_{\mathrm{TW},2}$	$z_{\mathrm{TW},4}$	z_{StdNorm}	Discarded
1002	-1.963	0.971	3.465	10.467	0
2000	-0.767	2.167	4.661	11.663	2
5022	-0.064	2.869	5.364	12.365	4

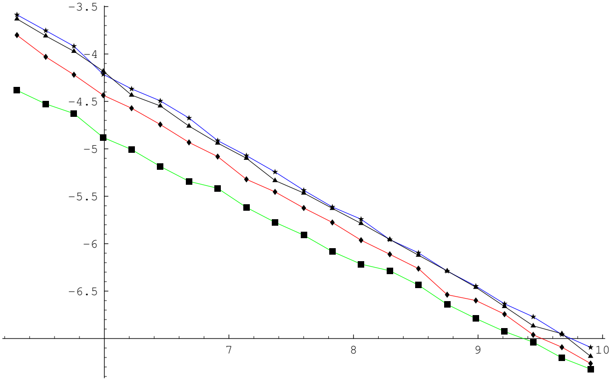
3-regular graphs: Sample means of $\lambda_+(G)$

Sets of 1000 random 3-regular graphs. $\mathcal{CI}_{N,3}$ is red, $\mathcal{SCI}_{N,3}$ is blue, $\mathcal{CB}_{N,3}$ is green, $\mathcal{SCB}_{N,3}$ is black; the solid yellow line is $2\sqrt{2} \approx 2.8284$.



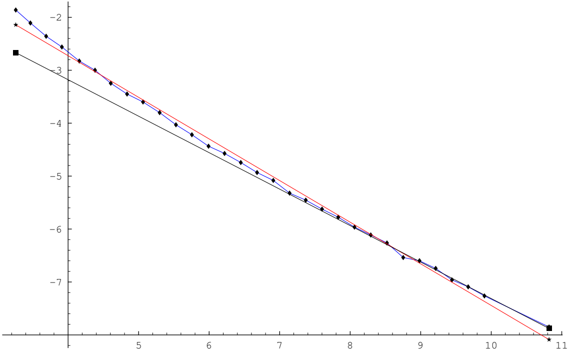
3-regular graphs: Sample means of $\lambda_+(G)$

Sets of 1000 random 3-regular graphs. $\mathcal{CI}_{N,3}$ is red, $\mathcal{SCI}_{N,3}$ is blue, $\mathcal{CB}_{N,3}$ is green, $\mathcal{SCB}_{N,3}$ is black.



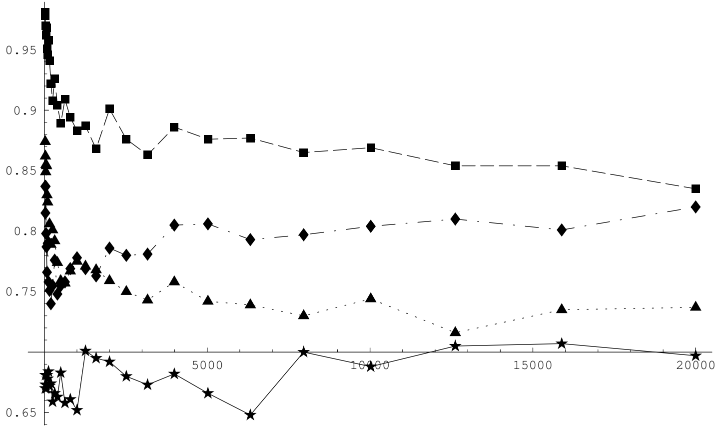
3-regular graphs: best fit means of $\lambda_+(G)$

Logarithm of the mean on $\log\left(c_{\mu,N,3} N^{m(\mathcal{CI}_{N,3})}\right)$ on N. Blue: data; red: best fit (all); black: best fit (last 10).



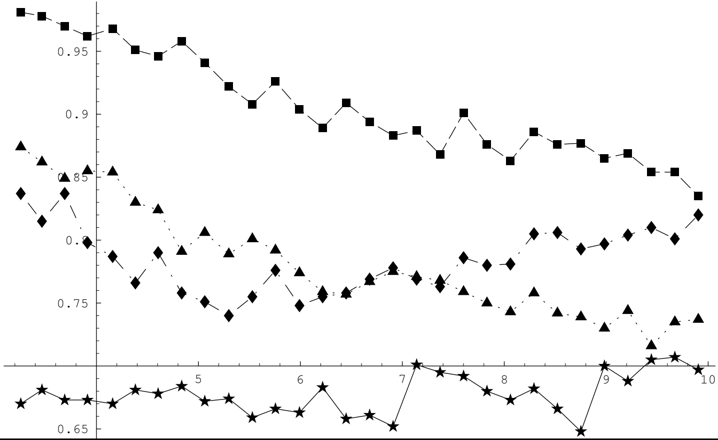
3-regular graphs: percent Ramanujan

Each set is 1000 random 3-regular graphs with N vertices. $\mathcal{CI}_{N,3}$ are stars, $\mathcal{SCI}_{N,3}$ are diamonds, $\mathcal{CB}_{N,3}$ are triangles, $\mathcal{SCB}_{N,3}$ are boxes.



3-regular graphs: percent Ramanujan

Each set is 1000 random 3-regular graphs with N vertices. $\mathcal{CI}_{N,3}$ are stars, $\mathcal{SCI}_{N,3}$ are diamonds, $\mathcal{CB}_{N,3}$ are triangles, $\mathcal{SCB}_{N,3}$ are boxes.



Best-fit exponents ($d = 3$) for $\lambda_+(G)$

First table means $m(\mathcal{F})$, second std devs $s(\mathcal{F})$.

Bold entries: $m(\mathcal{F}) > s(\mathcal{F})$.

N	$\mathcal{CI}_{N,3}$	$\mathcal{SCI}_{N,3}$	$\mathcal{CB}_{N,3}$	$\mathcal{SCB}_{N,3}$
$\{26,\dots,20000\}$	-0.795	-0.828	-0.723	- 0.833
$\{80,\dots,20000\}$	-0.761	-0.790	-0.671	-0.789
$\{252,\dots,20000\}$	-0.735	-0.762	-0.638	-0.761
$\{26,\dots,64\}$	- 1.058	-1.105	-1.065	-1.151
$\{80,\dots,200\}$	-0.854	-0.949	-0.982	-0.968
$\{232,\dots,632\}$	-0.773	-0.840	-0.737	-0.842
$\{796,\dots,2000\}$	- 0.762	-0.805	-0.649	-0.785
$\{2516,\dots,6324\}$	-0.791	-0.741	-0.579	-0.718
$\{7962,\dots,20000\}$	- 0.728	-0.701	-0.584	-0.757

N	$\mathcal{CI}_{N,3}$	$\mathcal{SCI}_{N,3}$	$\mathcal{CB}_{N,3}$	$\mathcal{SCB}_{N,3}$
$\{26,\dots,20000\}$	- 0.713	-0.725	-0.709	-0.729
$\{80,\dots,20000\}$	-0.693	-0.703	**-0.697**	-0.706
$\{252,\dots,20000\}$	- 0.679	-0.691	**-0.688**	-0.696
$\{26,\dots,64\}$	-0.863	-0.918	-0.794	-0.957
$\{80,\dots,200\}$	-0.694	-0.752	- 0.719	-0.750
$\{232,\dots,632\}$	-0.718	-0.716	-0.714	-0.734
$\{796,\dots,2000\}$	-0.602	-0.648	**-0.705**	-0.763
$\{2516,\dots,6324\}$	- 0.614	-0.668	**-0.770**	-0.688
$\{7962,\dots,20000\}$	-0.543	**-0.716**	**-0.671**	- 0.648

Best-fit exponents ($d = 3$) for $\lambda_+(G)$

$$2\sqrt{d-1} \approx \mu_{\mathcal{F}_{N,d}} + \frac{c_{\mu,N,d}}{c_{\sigma,N,d}} N^{m(\mathcal{F}_{N,d})-s(\mathcal{F}_{N,d})} \sigma_{\mathcal{F}_{N,d}}$$

$m(\mathcal{F}_{N,d}) - s(\mathcal{F}_{N,d})$, Bold entries $m(\mathcal{F}) > s(\mathcal{F})$.

N	$\mathcal{CI}_{N,3}$	$\mathcal{SCI}_{N,3}$	$\mathcal{CB}_{N,3}$	$\mathcal{SCB}_{N,3}$
$\{26,\dots,20000\}$	-0.082	-0.103	-0.014	-0.104
$\{80,\dots,20000\}$	-0.068	-0.087	**0.026**	-0.083
$\{252,\dots,20000\}$	- 0.056	-0.071	**0.050**	-0.065
$\{26,\dots,64\}$	-0.195	-0.187	-0.271	-0.194
$\{80,\dots,200\}$	-0.160	-0.197	- 0.263	-0.218
$\{232,\dots,632\}$	-0.055	-0.124	-0.023	-0.108
$\{796,\dots,2000\}$	-0.160	-0.157	**0.056**	-0.022
$\{2516,\dots,6324\}$	-0.177	-0.073	**0.191**	-0.030
$\{7962,\dots,20000\}$	-0.185	**0.015**	**0.087**	-0.109

References

- Y. Alhassid, *The statistical theory of quantum dots*, Rev. Mod. Phys. **72** (2000), no. 4, 895–968.
- N. Alon, *Eigenvalues and expanders*, Combinatorica **6** (1986), no. 2, 83–96.
- N. Alon and V. Milman, λ_1*, isoperimetric inequalities for graphs, and superconcentrators*, J. Combin. Theory Ser. B **38** (1985), no. 1, 73–88.
- J. Baik, P. Deift and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, J. Amer. Math. Soc. **12** (1999), 1119–1178.
- J. Baik and E. M. Rains, *The asymptotics of monotone subsequences of involutions*, Duke Math. J. **109** (2001), 205–281.
- J. Baik and E. M. Rains, *Symmetrized random permutations, in Random Matrix Models and their Applications*, eds. P. Bleher & A. Its, MSRI Publications **40**, Cambridge Univ. Press, 2001, 1–19.

- J. Baik and E. M. Rains, *Limiting distributions for a polynuclear growth model*, J. Stat. Phys. **100** (2000), 523–541.
- Yu. Baryshnikov, *GUEs and queues*, Probab. Th. Rel. Fields **119** (2001), 256–274.
- F. Bien, *Constructions of telephone networks by group representations*, Notices of the AMS **36** (1989), no. 1, 5–22.
- B. Bollobás, *Random Graphs*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2001.
- A. Borodin, A. Okounkov and G. Olshanski, *Asymptotics of Plancherel measures for symmetric groups*, J. Amer. Math. Soc. **13** (2000), 481–515.
- P. Chiu, *Cubic Ramanujan graphs*, Combinatorica **12** (1992), no. 3, 275–285.

- G. Davidoff, P. Sarnak, and A. Valette, *Elementary Number Theory, Group Theory, and Ramanujan Graphs*, London Mathematical Society, Student Texts **55**, Cambridge University Press, 2003.
- J. Dodziuk, *Difference equations, isoperimetric inequality and transience of certain random walks*, Trans. Amer. Math. Soc. **284** (1984), no. 2, 787–794.
- J. Friedman, *A proof of Alon's second eigenvalue conjecture*, Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, 720–724 (electronic), ACM, New York, 2003.
- M. Gaudin, *Sur la loi limite de l'espacement des valeurs propres d'une matrice aléatoire*, Nucl. Phys. **25** (1961) 447–458.
- J. Gravner, C. A. Tracy and H. Widom, *Limit theorems for height fluctuations in a class of discrete space and time growth models*, J. Stat. Phys. **102** (2001), 1085–1132.

- O. Goldreich, R. Impagliazzo, L. Levin, R. Venkatesan, and D. Zuckerman, *Security preserving amplification of hardness*. In 31st Annual Symposium on Foundations of Computer Science, Vol. I, II (St. Louis, MO, 1990), 318–326, IEEE Comput. Soc. Press, Los Alamitos, CA, 1990.
- D. Jakobson, S. D. Miller, I. Rivin, and Z. Rudnick, *Eigenvalue spacings for regular graphs*. Pages 317–327 in *Emerging Applications of Number Theory (Minneapolis, 1996)*, The IMA Volumes in Mathematics and its Applications, Vol. 109, Springer, New York, 1999.
- K. Johansson, *Discrete orthogonal polynomial ensembles and the Plancherel measure*, Ann. Math. **153** (2001), 259–296.
- K. Johansson, *Non-intersecting paths, random tilings and random matrices*, Probab. Th. Rel. Fields **123** (2002), 225–280.
- K. Johansson, *Toeplitz determinants, random growth and determinantal processes*, ICM Vol. III (2002), 53–62.

- A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277.
- G. A. Margulis, *Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators (Russian)*, Problemy Peredachi Informatsii **24** (1988), no. 1, 51–60; translation in Problems Inform. Transmission **24** (1988), no. 1, 39–46.
- B. McKay, *The expected eigenvalue distribution of a large regular graph*, Linear Algebra Appl. **40** (1981), 203–216.
- M. Mehta, *On the statistical properties of level spacings in nuclear spectra*, Nucl. Phys. **18** (1960), 395–419.
- M. Mehta, *Random Matrices*, 2nd edition, Academic Press, Boston, 1991.
- S. J. Miller, T. Novikoff and A. Sabelli, *The distribution of the second largest eigenvalue in families of random regular graphs*, to appear in Experimental Mathematics.

- M. Morgenstern, *Existence and explicit constructions of $q+1$ regular Ramanujan graphs for every prime power q*, J. Combin. Theory Ser. B **62** (1994), no. 1, 44–62.
- M. Ram Murty, *Ramanujan graphs*, J. Ramanujan Math. Soc. **18** (2003), no. 1, 33–52.
- A. Odlyzko, *On the distribution of spacings between zeros of the zeta function*, Math. Comp. **48** (1987), no. 177, 273–308.
- A. Odlyzko, *The 10^{22}-nd zero of the Riemann zeta function*, Proc. Conference on Dynamical, Spectral and Arithmetic Zeta-Functions, M. van Frankenhuysen and M. L. Lapidus, eds., Amer. Math. Soc., Contemporary Math. series, 2001, http://www.research.att.com/∼amo/doc/zeta.html
- Pippenger, *Super concentrators*, SIAM Journal Comp. **6** (1977), 298–304.
- M. Prähofer and H. Spohn, *Statistical self-similarity of one-dimensional growth processes*, Physica A **279** (2000), 342–352.

- M. Prähofer and H. Spohn, *Universal distributions for growth processes in* $1+1$ *dimensions and random matrices*, Phys. Rev. Letts. **84** (2000), 4882–4885.
- P. Sarnak *Some applications of modular forms*, Cambridge Trusts in Mathemetics, Vol. 99, Cambridge University Press, Cambridge, 1990.
- A. Soshnikov, *A note on universality of the distribution of the largest eigenvalue in certain classes of sample covariance matrices*, preprint (arXiv: math.PR/0104113).
- M. Sipser and D. A. Spielman, *Expander codes*, IEEE Trans. Inform. Theory **42** (1996), no. 6, part 1, 1710–1722.
- C. A. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Commun. Math. Phys. **159** (1994), 151–174.
- C. Tracy and H. Widom, *On Orthogonal and Symplectic Matrix Ensembles*, Communications in Mathematical Physics **177** (1996), 727–754.

- C. Tracy and H. Widom, *Distribution functions for largest eigenvalues and their applications*, ICM Vol. I (2002), 587–596.
- M. G. Vavilov, P. W. Brouwer, V. Ambegaokar and C. W. J. Beenakker, *Universal gap fluctuations in the superconductor proximity effect*, Phys. Rev. Letts. **86** (2001), 874–877.
- E. Wigner, *Statistical Properties of real symmetric matrices*. Pages 174–184 in *Canadian Mathematical Congress Proceedings*, University of Toronto Press, Toronto, 1957.