

2026 RMT Summer School
Edge Statistics of Sparse Random Graphs (Jiaoyang Huang)
Combined Homework Solutions (HW0–HW4)

Selected participant solutions

Homework 0 | Due: 11:00 pm, Sunday (June 21)

Use the random-matrix sign convention for the Stieltjes transform

$$m_\mu(z) = \int_{\mathbb{R}} \frac{1}{x - z} \mu(dx), \quad z \in \mathbb{C}^+.$$

Problem 1 (Stieltjes Transforms of Measures). Let μ be a finite positive Borel measure on $\mathbb{R}$.

(a) Let $P_\eta(t) = \pi^{-1}\eta/(t^2 + \eta^2)$. Show that

$$\frac{1}{\pi} \operatorname{Im} m_\mu(E + i\eta) = \int_{\mathbb{R}} P_\eta(E - x) \mu(dx),$$

and use this to prove the Stieltjes inversion formula for intervals whose endpoints are continuity points of μ : If $a < b$ are continuity points of μ , then

$$\mu((a, b)) = \lim_{\eta \downarrow 0} \frac{1}{\pi} \int_a^b \operatorname{Im} m_\mu(E + i\eta) dE. \quad (1)$$

(b) Suppose μ_n and μ are probability measures and $m_{\mu_n} \rightarrow m_\mu$ uniformly on compact subsets of $\mathbb{C}^+$. Prove that $\mu_n \Rightarrow \mu$.

(c) Verify that

$$m_{\text{sc}}(z) = \frac{-z + \sqrt{z^2 - 4}}{2}$$

satisfies $1 + z m_{\text{sc}}(z) + m_{\text{sc}}(z)^2 = 0$, where the branch of the square root is chosen so that $\sqrt{z^2 - 4} \sim z$ as $|z| \rightarrow \infty$. Show that $\operatorname{Im} m_{\text{sc}}(z) > 0$ for $z \in \mathbb{C}^+$, and recover the semicircle density

$$\varrho_{\text{sc}}(E) = \frac{\sqrt{4 - E^2}}{2\pi} \mathbf{1}(-2 \leq E \leq 2), \quad (2)$$

from the boundary values of $m_{\text{sc}}(z)$.

Solution

(a) For $z = E + i\eta$,

$$\frac{1}{x - z} = \frac{x - E + i\eta}{(x - E)^2 + \eta^2}.$$

Therefore

$$\frac{1}{\pi} \operatorname{Im} m_\mu(E + i\eta) = \int_{\mathbb{R}} \frac{1}{\pi} \frac{\eta}{(E - x)^2 + \eta^2} \mu(dx) = \int_{\mathbb{R}} P_\eta(E - x) \mu(dx).$$

Consequently, by Tonelli's theorem,

$$\frac{1}{\pi} \int_a^b \operatorname{Im} m_\mu(E + i\eta) dE = \int_{\mathbb{R}} K_\eta(x) \mu(dx),$$

where

$$K_\eta(x) := \int_a^b P_\eta(E - x) dE = \frac{1}{\pi} \left(\arctan \frac{b-x}{\eta} - \arctan \frac{a-x}{\eta} \right).$$

As $\eta \downarrow 0$,

$$K_\eta(x) \longrightarrow \mathbf{1}_{(a,b)}(x) + \frac{1}{2} \mathbf{1}_{\{a,b\}}(x).$$

Moreover, $0 \leq K_\eta \leq 1$. Since a and b are continuity points of μ , $\mu(\{a\}) = \mu(\{b\}) = 0$. Dominated convergence therefore gives

$$\lim_{\eta \downarrow 0} \frac{1}{\pi} \int_a^b \operatorname{Im} m_\mu(E + i\eta) dE = \mu((a, b)).$$

(b) We first prove tightness. For $Y > 0$,

$$Y \operatorname{Im} m_{\mu_n}(iY) = \int_{\mathbb{R}} \frac{Y^2}{x^2 + Y^2} \mu_n(dx).$$

Since μ is a probability measure, $Y \operatorname{Im} m_\mu(iY) \rightarrow 1$ as $Y \rightarrow \infty$. Given $\varepsilon > 0$, choose Y so that

$$1 - Y \operatorname{Im} m_\mu(iY) < \varepsilon.$$

The assumed convergence at iY implies that, for all sufficiently large n ,

$$1 - Y \operatorname{Im} m_{\mu_n}(iY) < 2\varepsilon.$$

For $R \geq Y$,

$$\begin{aligned} \mu_n(\{|x| > R\}) &\leq \frac{R^2 + Y^2}{R^2} \int_{\mathbb{R}} \frac{x^2}{x^2 + Y^2} \mu_n(dx) \\ &= \frac{R^2 + Y^2}{R^2} (1 - Y \operatorname{Im} m_{\mu_n}(iY)) \leq 4\varepsilon. \end{aligned}$$

Increasing R if necessary handles the finitely many remaining indices, so (μ_n) is tight.

Now take any subsequence. By tightness, it has a further subsequence, still denoted (μ_{n_k}) , such that $\mu_{n_k} \Rightarrow \nu$ for some probability measure ν . For each $z \in \mathbb{C}^+$, the function $x \mapsto (x - z)^{-1}$ is bounded and continuous, hence

$$m_\nu(z) = \lim_{k \rightarrow \infty} m_{\mu_{n_k}}(z) = m_\mu(z).$$

Part (a) shows that the Stieltjes transform uniquely determines a finite measure, so $\nu = \mu$. Thus every subsequence has a further subsequence converging weakly to μ , and therefore $\mu_n \Rightarrow \mu$.

- (c) Put $s(z) = \sqrt{z^2 - 4}$ and $m_{\text{sc}}(z) = (s(z) - z)/2$. Since $s(z)^2 = z^2 - 4$, direct substitution gives

$$m_{\text{sc}}(z)^2 + z m_{\text{sc}}(z) + 1 = 0.$$

The chosen branch satisfies

$$s(z) = z \sqrt{1 - \frac{4}{z^2}} = z - \frac{2}{z} + O(|z|^{-3}), \quad m_{\text{sc}}(z) = -\frac{1}{z} + O(|z|^{-3}).$$

In particular, $\text{Im } m_{\text{sc}}(iy) > 0$ for all sufficiently large y . If $\text{Im } m_{\text{sc}}(z) = 0$ at some $z \in \mathbb{C}^+$, then $m_{\text{sc}}(z) \in \mathbb{R} \setminus \{0\}$ and the quadratic equation yields

$$z = -m_{\text{sc}}(z) - \frac{1}{m_{\text{sc}}(z)} \in \mathbb{R},$$

a contradiction. Since $\mathbb{C}^+$ is connected, it follows that $\text{Im } m_{\text{sc}}(z) > 0$ throughout $\mathbb{C}^+$.

The boundary values of the chosen square root are

$$\lim_{\eta \downarrow 0} s(E + i\eta) = \begin{cases} \sqrt{E^2 - 4}, & E > 2, \\ i\sqrt{4 - E^2}, & |E| < 2, \\ -\sqrt{E^2 - 4}, & E < -2. \end{cases}$$

Hence

$$\lim_{\eta \downarrow 0} \frac{1}{\pi} \text{Im } m_{\text{sc}}(E + i\eta) = \frac{\sqrt{4 - E^2}}{2\pi} \mathbf{1}_{\{|E| < 2\}}.$$

The Stieltjes inversion formula therefore recovers

$$\varrho_{\text{sc}}(E) = \frac{\sqrt{4 - E^2}}{2\pi} \mathbf{1}_{[-2, 2]}(E).$$

Problem 2 (Point Processes and Correlation Functions). A point process on $\mathbb{R}$ is a random locally finite Borel measure of the form

$$\xi = \sum_{i \in I} \delta_{X_i},$$

where $\xi(K) < \infty$ almost surely for every compact $K \subset \mathbb{R}$. We will usually assume the process is simple, meaning no point has multiplicity greater than one.

For $k \geq 1$, its k -th factorial moment measure α_k is defined by

$$\int_{\mathbb{R}^k} f(x_1, \dots, x_k) \alpha_k(dx_1 \cdots dx_k) := \mathbb{E} \sum_{i_1, \dots, i_k}^{\neq} f(X_{i_1}, \dots, X_{i_k}),$$

for all nonnegative Borel f , where $\sum^{\neq}$ means the sum is over ordered k -tuples of distinct indices. If α_k is absolutely continuous with respect to Lebesgue measure, its density ρ_k is called the k -point correlation function:

$$\alpha_k(dx_1 \cdots dx_k) = \rho_k(x_1, \dots, x_k) dx_1 \cdots dx_k.$$

Equivalently,

$$\mathbb{E} \sum_{i_1, \dots, i_k}^{\neq} f(X_{i_1}, \dots, X_{i_k}) = \int_{\mathbb{R}^k} f(x_1, \dots, x_k) \rho_k(x_1, \dots, x_k) dx_1 \cdots dx_k.$$

- (a) For bounded Borel sets B_1, B_2 , express

$$\mathbb{E}[\xi(B_1)], \quad \mathbb{E}[\xi(B_1)\xi(B_2)]$$

in terms of α_1, α_2 , and the diagonal contribution.

- (b) Assume that $\mathbb{E} e^{t\xi(K)} < \infty$ for every compact $K \subset \mathbb{R}$ and every $t > 0$. Let $f \in C_c(\mathbb{R})$ with $f \geq 0$, and set $g = e^{-f} - 1$. Starting from

$$\prod_{x \in \xi} (1 + g(x)),$$

derive the expansion of the Laplace functional

$$L_\xi(f) := \mathbb{E} \exp\left(-\int f d\xi\right),$$

in terms of the factorial moment measures α_k .

- (c) Explain why the collection $\{\alpha_k\}_{k \geq 1}$ determines all joint distributions of count vectors

$$(\xi(B_1), \dots, \xi(B_m))$$

for bounded Borel sets $B_1, \dots, B_m$.

Solution

- (a) By definition,

$$\xi(B_1) = \sum_{i \in I} \mathbf{1}_{B_1}(X_i),$$

so

$$\mathbb{E}[\xi(B_1)] = \int_{\mathbb{R}} \mathbf{1}_{B_1}(x) \alpha_1(dx) = \alpha_1(B_1).$$

For the second moment, split the double sum into off-diagonal and diagonal terms:

$$\xi(B_1)\xi(B_2) = \sum_{i \neq j} \mathbf{1}_{B_1}(X_i) \mathbf{1}_{B_2}(X_j) + \sum_i \mathbf{1}_{B_1 \cap B_2}(X_i).$$

Taking expectations gives

$$\mathbb{E}[\xi(B_1)\xi(B_2)] = \alpha_2(B_1 \times B_2) + \alpha_1(B_1 \cap B_2).$$

- (b) Let $K = \text{supp } f$. Since K is compact and ξ is locally finite, $\xi(K) < \infty$ almost surely. Also, $g = 0$ outside K , so the product below is finite for every realization. Since $1 + g = e^{-f}$,

$$\prod_{x \in \xi} (1 + g(x)) = \prod_{x \in \xi} e^{-f(x)} = \exp\left(-\int f d\xi\right).$$

Expanding the finite product gives

$$\prod_{x \in \xi} (1 + g(x)) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{i_1, \dots, i_k}^{\neq} g(X_{i_1}) \cdots g(X_{i_k}).$$

The factor $1/k!$ converts the ordered distinct k -tuples into the unordered subsets arising in the product expansion. Moreover,

$$\sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i_1, \dots, i_k}^{\neq} |g(X_{i_1}) \cdots g(X_{i_k})| \leq \exp(\|g\|_{\infty} \xi(K)),$$

and the right-hand side is integrable by the exponential-moment assumption. We may therefore take expectations term by term. Using the definition of α_k ,

$$L_{\xi}(f) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{R}^k} g(x_1) \cdots g(x_k) \alpha_k(dx_1 \cdots dx_k).$$

Equivalently,

$$L_{\xi}(f) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{R}^k} \prod_{j=1}^k (e^{-f(x_j)} - 1) \alpha_k(dx_1 \cdots dx_k).$$

Solution

- (c) Let $B_1, \dots, B_m$ be bounded Borel sets. Refine them into finitely many disjoint bounded Borel atoms $A_1, \dots, A_r$ so that every B_j is a union of some of the A_i . Each $\xi(B_j)$ is then a fixed sum of the atom counts

$$N_i := \xi(A_i), \quad 1 \leq i \leq r,$$

so it is enough to determine the joint law of $(N_1, \dots, N_r)$.

Choose a compact set K containing $\bigcup_i A_i$. The exponential-moment hypothesis implies

$$\mathbb{E} \exp \left(t \sum_{i=1}^r N_i \right) \leq \mathbb{E} e^{t\xi(K)} < \infty \quad (t > 0).$$

Consequently, the joint probability generating function

$$G(z_1, \dots, z_r) := \mathbb{E} \left[\prod_{i=1}^r z_i^{N_i} \right]$$

is entire on $\mathbb{C}^r$. Put $u_i = z_i - 1$ and define

$$g_u(x) := \sum_{i=1}^r u_i \mathbf{1}_{A_i}(x).$$

Because the A_i are disjoint,

$$G(1 + u_1, \dots, 1 + u_r) = \mathbb{E} \prod_{x \in \xi} (1 + g_u(x)).$$

Applying the expansion from part (b), which is valid for this bounded Borel function by the same domination argument, gives

$$G(1 + u_1, \dots, 1 + u_r) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{R}^k} \prod_{\ell=1}^k g_u(x_\ell) \alpha_k(dx_1 \cdots dx_k).$$

Thus the collection $\{\alpha_k\}_{k \geq 1}$ determines G in a neighborhood of $(1, \dots, 1)$. Since G is entire, it determines G everywhere. Its Taylor coefficients at the origin are precisely the joint probabilities

$$\mathbb{P}(N_1 = n_1, \dots, N_r = n_r),$$

so the joint law of $(N_1, \dots, N_r)$, and hence that of $(\xi(B_1), \dots, \xi(B_m))$, is determined.

Homework 1 | Due: 11:00 pm, Monday (June 22)

Problem 1 (Tridiagonal Model for GOE). Let H be an $N \times N$ GOE matrix. Equivalently,

$$H = \frac{1}{\sqrt{N}}A,$$

where A is real symmetric, $A_{ij} \sim N(0, 1)$ for $i < j$, $A_{ii} \sim N(0, 2)$, and the upper-triangular entries are independent.

- (a) Let $x \in \mathbb{R}^m$ be nonzero. Recall the construction of a Householder reflection $R = I - 2uu^T$ such that $Rx = \|x\|e_1$. Verify that R is orthogonal and symmetric.
- (b) Write

$$A = \begin{pmatrix} a & v^T \\ v & B \end{pmatrix}, \quad a \in \mathbb{R}, \quad v \in \mathbb{R}^{N-1}, \quad B = B^T.$$

Choose an orthogonal matrix $R_1 \in O(N-1)$ with $R_1v = \|v\|e_1$, and set $Q_1 = \text{diag}(1, R_1)$. Show that

$$Q_1 A Q_1^T = \begin{pmatrix} a & \|v\|e_1^T \\ \|v\|e_1 & R_1 B R_1^T \end{pmatrix}.$$

- (c) Use the orthogonal invariance of the GOE to show that, conditionally on v , the lower-right $(N-1) \times (N-1)$ block $R_1 B R_1^T$ has the same distribution as B and is independent of a and $\|v\|$. Identify the distributions of a and $\|v\|$.
- (d) Iterate the preceding step. Prove that A is orthogonally conjugate to a random tridiagonal matrix

$$T_N = \begin{pmatrix} g_1 & \chi_{N-1} & & & \\ \chi_{N-1} & g_2 & \chi_{N-2} & & \\ & \chi_{N-2} & g_3 & \ddots & \\ & & \ddots & \ddots & \chi_1 \\ & & & \chi_1 & g_N \end{pmatrix},$$

where $g_1, \dots, g_N$ are independent $N(0, 2)$ variables, the χ_k 's are independent chi random variables with k degrees of freedom, and all these variables are mutually independent.

Solution

- (a) Let u be a unit vector and $R = I - 2uu^T$. Since uu^T is symmetric,

$$R^T = R.$$

Moreover,

$$R^T R = R^2 = I - 4uu^T + 4u(u^T u)u^T = I,$$

so R is orthogonal.

For the usual Householder construction, when $x \neq \|x\|e_1$, set

$$w := x - \|x\|e_1, \quad u := \frac{w}{\|w\|}.$$

As

$$w^T x = \|x\|^2 - \|x\|x_1 = \frac{1}{2}\|w\|^2,$$

we obtain

$$Rx = x - 2w \frac{w^T x}{\|w\|^2} = x - w = \|x\|e_1.$$

In the degenerate case $x = \|x\|e_1$, one may take $R = I$ (or, when $m \geq 2$, any Householder reflector whose normal is perpendicular to e_1).

(b) Direct block multiplication gives

$$\begin{aligned} Q_1 A Q_1^T &= \begin{pmatrix} 1 & 0 \\ 0 & R_1 \end{pmatrix} \begin{pmatrix} a & v^T \\ v & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & R_1^T \end{pmatrix} \\ &= \begin{pmatrix} a & v^T R_1^T \\ R_1 v & R_1 B R_1^T \end{pmatrix}. \end{aligned}$$

Since $R_1 v = \|v\|e_1$, we also have $v^T R_1^T = (R_1 v)^T = \|v\|e_1^T$. Hence

$$Q_1 A Q_1^T = \begin{pmatrix} a & \|v\|e_1^T \\ \|v\|e_1 & R_1 B R_1^T \end{pmatrix}.$$

(c) Condition on v . Then $R_1 = R_1(v)$ is deterministic, while B remains an independent $(N-1) \times (N-1)$ GOE matrix. Orthogonal invariance therefore yields

$$R_1 B R_1^T \stackrel{d}{=} B \quad \text{conditionally on } v.$$

The conditional law on the left does not depend on v . Thus $R_1 B R_1^T$ is independent of v , and in particular independent of $\|v\|$. It is also independent of a , because a is independent of (v, B) . Consequently, a , $\|v\|$, and $R_1 B R_1^T$ are mutually independent.

Finally,

$$a = A_{11} \sim N(0, 2), \quad \|v\| = \left(\sum_{j=2}^N A_{j1}^2 \right)^{1/2} \sim \chi_{N-1},$$

since the entries of v are independent $N(0, 1)$ variables.

Solution

(d) We iterate the preceding decomposition. After $r-1$ steps, the unexposed lower-right block C_{r-1} has size $N-r+1$, has the GOE law, and is independent of all previously exposed variables. Write

$$C_{r-1} = \begin{pmatrix} g_r & w_r^T \\ w_r & B_r \end{pmatrix},$$

where

$$g_r \sim N(0, 2), \quad w_r \sim N(0, I_{N-r}).$$

Choose $R_r \in O(N - r)$ with $R_r w_r = \|w_r\| e_1$, and extend it to the full space by

$$\tilde{Q}_r := \text{diag}(I_r, R_r).$$

Part (c), applied inside C_{r-1} , shows that conjugation by $\tilde{Q}_r$ exposes

$$\|w_r\| \sim \chi_{N-r}$$

as the next off-diagonal entry and leaves an independent GOE block C_r of size $N - r$.

Continuing for $r = 1, \dots, N - 1$ and setting

$$Q := \tilde{Q}_{N-1} \cdots \tilde{Q}_1,$$

we obtain

$$Q A Q^T = \begin{pmatrix} g_1 & \chi_{N-1} & & & \\ \chi_{N-1} & g_2 & \chi_{N-2} & & \\ & \chi_{N-2} & g_3 & \ddots & \\ & & \ddots & \ddots & \chi_1 \\ & & & \chi_1 & g_N \end{pmatrix}.$$

At each stage, the newly exposed pair (g_r, χ_{N-r}) is independent of the remaining GOE block and of all variables exposed earlier. Hence the g_r 's, the χ_k 's, and all cross-families are mutually independent.

Problem 2 (Loop Equations for the Gaussian β -Ensemble). For this problem, fix $\beta > 0$ and work with the Gaussian β -ensemble on $\mathbb{R}^N$, whose density is given by

$$p_{N,\beta}(\lambda_1, \dots, \lambda_N) = \frac{1}{Z_{N,\beta}} \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^\beta \exp\left(-\frac{\beta N}{4} \sum_{i=1}^N \lambda_i^2\right). \quad (3)$$

For $z \in \mathbb{C} \setminus \mathbb{R}$, define the empirical Stieltjes transform

$$m_N(z) := \frac{1}{N} \sum_{i=1}^N \frac{1}{\lambda_i - z}.$$

(a) Show that

$$\partial_{\lambda_i} \log p_{N,\beta}(\lambda) = \beta \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} - \frac{\beta N}{2} \lambda_i.$$

Deduce the integration-by-parts identity

$$\mathbb{E} \sum_{i=1}^N [\partial_{\lambda_i} F_i(\lambda) + F_i(\lambda) \partial_{\lambda_i} \log p_{N,\beta}(\lambda)] = 0$$

for sufficiently regular functions F_i for which the boundary terms vanish.

(b) Apply part (a) with $F_i(\lambda) = (\lambda_i - z)^{-1}$ to derive the first loop equation

$$\mathbb{E} \left[m_N(z)^2 + z m_N(z) + 1 + \frac{1}{N} \left(\frac{2}{\beta} - 1 \right) \partial_z m_N(z) \right] = 0. \quad (4)$$

(c) Let $p \geq 2$, and let $z, z_2, \dots, z_p \in \mathbb{C} \setminus \mathbb{R}$, with $z \neq z_j$ for $j \geq 2$. Prove the rank- p loop equation

$$\begin{aligned} & \mathbb{E} \left[\left(m_N(z)^2 + z m_N(z) + 1 + \frac{1}{N} \left(\frac{2}{\beta} - 1 \right) \partial_z m_N(z) \right) \prod_{j=2}^p m_N(z_j) \right] \\ & + \frac{2}{\beta N^2} \sum_{j=2}^p \mathbb{E} \left[\left(\partial_{z_j} \frac{m_N(z) - m_N(z_j)}{z - z_j} \right) \prod_{\substack{k=2 \\ k \neq j}}^p m_N(z_k) \right] = 0. \end{aligned} \quad (5)$$

If $z = z_j$, interpret the difference quotient by its analytic continuation.

Solution

(a) Taking logarithms in (3),

$$\log p_{N,\beta}(\lambda) = -\log Z_{N,\beta} + \beta \sum_{1 \leq i < j \leq N} \log |\lambda_i - \lambda_j| - \frac{\beta N}{4} \sum_{i=1}^N \lambda_i^2.$$

Therefore

$$\partial_{\lambda_i} \log p_{N,\beta}(\lambda) = \beta \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} - \frac{\beta N}{2} \lambda_i.$$

For admissible F_i , integration of a total derivative gives

$$0 = \int_{\mathbb{R}^N} \sum_{i=1}^N \partial_{\lambda_i} (F_i(\lambda) p_{N,\beta}(\lambda)) d\lambda.$$

Using

$$\partial_{\lambda_i} (F_i p_{N,\beta}) = (\partial_{\lambda_i} F_i + F_i \partial_{\lambda_i} \log p_{N,\beta}) p_{N,\beta}$$

and the assumed vanishing of the boundary terms yields

$$\mathbb{E} \sum_{i=1}^N [\partial_{\lambda_i} F_i(\lambda) + F_i(\lambda) \partial_{\lambda_i} \log p_{N,\beta}(\lambda)] = 0.$$

Solution

(b) Set $r_i = (\lambda_i - z)^{-1}$. Applying part (a) with $F_i = r_i$ gives

$$0 = \mathbb{E} \left[-\sum_i r_i^2 + \beta \sum_{i \neq j} \frac{r_i}{\lambda_i - \lambda_j} - \frac{\beta N}{2} \sum_i \lambda_i r_i \right].$$

Pairing the (i, j) and (j, i) terms,

$$\begin{aligned}\sum_{i \neq j} \frac{r_i}{\lambda_i - \lambda_j} &= \sum_{i < j} \frac{r_i - r_j}{\lambda_i - \lambda_j} \\ &= - \sum_{i < j} r_i r_j \\ &= -\frac{1}{2} \left[\left(\sum_i r_i \right)^2 - \sum_i r_i^2 \right].\end{aligned}$$

Moreover,

$$\sum_i r_i = N m_N(z), \quad \sum_i r_i^2 = N \partial_z m_N(z), \quad \sum_i \lambda_i r_i = N + z N m_N(z).$$

Substitution gives

$$\begin{aligned}0 = \mathbb{E} \left[-N \partial_z m_N(z) - \frac{\beta}{2} (N^2 m_N(z)^2 - N \partial_z m_N(z)) \right. \\ \left. - \frac{\beta N^2}{2} (1 + z m_N(z)) \right].\end{aligned}$$

Factoring $-\beta N^2/2$ proves

$$\mathbb{E} \left[m_N(z)^2 + z m_N(z) + 1 + \frac{1}{N} \left(\frac{2}{\beta} - 1 \right) \partial_z m_N(z) \right] = 0.$$

Solution

(c) Set

$$r_i(w) := \frac{1}{\lambda_i - w}, \quad P := \prod_{j=2}^p m_N(z_j), \quad P_j := \prod_{\substack{k=2 \\ k \neq j}}^p m_N(z_k),$$

and write

$$L(z) := m_N(z)^2 + z m_N(z) + 1 + \frac{1}{N} \left(\frac{2}{\beta} - 1 \right) \partial_z m_N(z).$$

The pointwise algebra in part (b) gives

$$\sum_{i=1}^N [-r_i(z)^2 + r_i(z) \partial_{\lambda_i} \log p_{N,\beta}(\lambda)] = -\frac{\beta N^2}{2} L(z).$$

Applying part (a) with $F_i(\lambda) = r_i(z)P$ therefore yields

$$0 = -\frac{\beta N^2}{2} \mathbb{E}[L(z)P] + \mathbb{E} \sum_{i=1}^N r_i(z) \partial_{\lambda_i} P.$$

Since $\partial_{\lambda_i} m_N(z_j) = -N^{-1} r_i(z_j)^2$,

$$\sum_{i=1}^N r_i(z) \partial_{\lambda_i} P = - \sum_{j=2}^p \left(\frac{1}{N} \sum_{i=1}^N r_i(z) r_i(z_j)^2 \right) P_j.$$

For $w \neq z$,

$$\frac{m_N(z) - m_N(w)}{z - w} = \frac{1}{N} \sum_{i=1}^N r_i(z) r_i(w),$$

and hence

$$\partial_w \frac{m_N(z) - m_N(w)}{z - w} = \frac{1}{N} \sum_{i=1}^N r_i(z) r_i(w)^2.$$

Substituting $w = z_j$ and multiplying the integration-by-parts identity by $-2/(\beta N^2)$ gives exactly (5).

The difference quotient has a removable singularity at $w = z$: its analytic continuation satisfies

$$\left. \frac{m_N(z) - m_N(w)}{z - w} \right|_{w=z} = m'_N(z), \quad \partial_w \left. \frac{m_N(z) - m_N(w)}{z - w} \right|_{w=z} = \frac{1}{2} m''_N(z).$$

This is the required interpretation when $z_j = z$.

Homework 2 | Due: 11:00 pm, Tuesday (June 23)**Setup**

Throughout this homework, $H = H_N$ is a real symmetric Wigner matrix. The entries $(H_{ij})_{1 \leq i < j \leq N}$ are independent, the diagonal entries are independent of the off-diagonal entries, and

$$\mathbb{E}H_{ij} = 0, \quad \mathbb{E}H_{ij}^2 = \frac{1}{N} \ (i < j), \quad \mathbb{E}H_{ii}^2 \leq \frac{C}{N}.$$

Assume the uniform moment condition

$$\mathbb{E}|H_{ij}|^{2p} \leq \frac{C_p}{N^p}, \quad 1 \leq i \leq j \leq N, \ p \in \mathbb{N}.$$

For $z = E + i\eta \in \mathbb{C}^+$, define

$$G(z) := (H - z)^{-1}, \quad m_N(z) := \frac{1}{N} \operatorname{tr} G(z).$$

The semicircle Stieltjes transform is

$$m_{\text{sc}}(z) = \frac{-z + \sqrt{z^2 - 4}}{2}, \quad \operatorname{Im} m_{\text{sc}}(z) > 0,$$

or equivalently

$$1 + zm_{\text{sc}}(z) + m_{\text{sc}}(z)^2 = 0.$$

Set

$$\rho(z) := \operatorname{Im} m_{\text{sc}}(z), \quad \kappa(E) := \left| |E| - 2 \right|.$$

For fixed $\delta > 0$, work on the spectral domain

$$D_\delta := \{E + i\eta : E \in [-3, 3], \ N^{-1+\delta} \leq \eta \leq 1\}.$$

We use stochastic domination: $X \prec Y$ means that for all $\varepsilon, D > 0$,

$$\mathbb{P}(X > N^\varepsilon Y) \leq N^{-D}$$

for all sufficiently large N . All stochastic domination estimates below should be understood uniformly in the relevant indices and spectral parameters.

The only problem in this homework is to prove the following weak local law.

Theorem 1 (Weak local semicircle law). *Fix $\delta > 0$. Uniformly for $z = E + i\eta \in D_\delta$,*

$$|m_N(z) - m_{\text{sc}}(z)| \prec (N\eta)^{-1/4},$$

and

$$\max_i |G_{ii}(z) - m_{\text{sc}}(z)| \prec (N\eta)^{-1/4}.$$

Lemma 1 (Quadratic concentration; use without proof). *Let $x_1, \dots, x_M$ be independent centered real random variables satisfying*

$$\mathbb{E}x_k^2 = N^{-1}, \quad \mathbb{E}|x_k|^{2p} \leq C_p N^{-p} \quad (p \in \mathbb{N}).$$

Let $A = (A_{kl}) \in \mathbb{C}^{M \times M}$ be deterministic, or random but independent of $x = (x_1, \dots, x_M)$. Then

$$\left| \sum_{k,l} x_k A_{kl} x_l - \frac{1}{N} \sum_k A_{kk} \right| \prec \left(\frac{1}{N^2} \operatorname{tr} A A^* \right)^{1/2}.$$

Lemma 2 (Stability of the quadratic equation; use without proof). *Let $z = E + i\eta$, where $E \in [-3, 3]$ and $0 < \eta \leq 1$. Then*

$$z + 2m_{\text{sc}}(z) = \sqrt{z^2 - 4}, \quad |z + 2m_{\text{sc}}(z)| \asymp \sqrt{\kappa(E) + \eta}.$$

Let $u \in \mathbb{C}^+$, assume $|u| \leq C$, and set

$$r := 1 + zu + u^2.$$

If $|u - m_{\text{sc}}(z)| \leq c_0$, where $c_0 > 0$ is a sufficiently small constant, then

$$|u - m_{\text{sc}}(z)| \leq C \frac{|r|}{\sqrt{\kappa(E) + \eta} + \sqrt{|r|}} \leq C \sqrt{|r|}.$$

If $\eta \asymp 1$, the same conclusion holds without the a priori assumption $|u - m_{\text{sc}}(z)| \leq c_0$, provided $|r|$ is sufficiently small and $u \in \mathbb{C}^+$.

Problem 1 (Proof of the Weak Local Law). For a set $T \subset \{1, \dots, N\}$, let $H^{(T)}$ be the minor obtained by removing rows and columns in T , and let

$$G^{(T)}(z) := (H^{(T)} - z)^{-1}.$$

Write $G^{(i)} = G^{(\{i\})}$, and use the normalization

$$m_N^{(i)}(z) := \frac{1}{N} \operatorname{tr} G^{(i)}(z).$$

Sums with a superscript (i) are over indices different from i . For this problem, fix $z \in D_\delta$, and write

$$\Lambda(z) := |m_N(z) - m_{\text{sc}}(z)|.$$

(a) Prove the Ward identity

$$\sum_j |G_{ij}(z)|^2 = \frac{\operatorname{Im} G_{ii}(z)}{\eta}.$$

Prove also the Schur complement identities

$$\frac{1}{G_{ii}} = -z - m_N + \Upsilon_i,$$

where

$$Z_i := \sum_{k,l}^{(i)} H_{ik} G_{kl}^{(i)} H_{li} - \frac{1}{N} \operatorname{tr} G^{(i)}, \quad \Upsilon_i := H_{ii} - Z_i + m_N - m_N^{(i)}.$$

Then average over i to show the random self-consistent equation

$$1 + zm_N + m_N^2 = \frac{1}{N} \sum_i \Upsilon_i G_{ii}.$$

(b) Assume the a priori bounds

$$\Lambda \leq c_0, \quad \max_i |G_{ii}| \leq C_0.$$

Prove the single-site estimate

$$\max_i |\Upsilon_i| \prec \Omega, \quad \Omega := \sqrt{\frac{\rho + \Lambda}{N\eta}} + \frac{1}{N\eta}.$$

In your proof, use the supplied quadratic concentration estimate, Ward's identity for the minor, and the rank-one estimate

$$|m_N - m_N^{(i)}| \leq \frac{C}{N\eta}.$$

(c) Under the same a priori bounds, prove the one-step averaged estimate

$$|m_N(z) - m_{\text{sc}}(z)| \prec \frac{1}{(N\eta)^{1/4}}.$$

Your proof should use the self-consistent equation from part (a), the estimate from part (b), and the stability lemma above.

(d) Under the same a priori bounds, prove the diagonal estimate

$$\max_i |G_{ii}(z) - m_{\text{sc}}(z)| \prec \frac{1}{(N\eta)^{1/4}}.$$

(e) Establish the initial estimate at scale $\eta = 1$. More precisely, show that for $E \in [-3, 3]$,

$$|m_N(E + i) - m_{\text{sc}}(E + i)| \prec N^{-1/4},$$

and

$$\max_i |G_{ii}(E + i) - m_{\text{sc}}(E + i)| \prec N^{-1/4}.$$

Hint: start from the deterministic resolvent bound $\|G(E + i)\| \leq 1$, and repeat the Schur complement argument.

(f) Complete the bootstrap from $\eta = 1$ down to $\eta = N^{-1+\delta}$. Use the lattice

$$\mathcal{L} = \left\{ -3 + aN^{-5} + i(1 - bN^{-5}) : a, b \in \mathbb{Z}_{\geq 0}, 0 \leq a \leq 6N^5, N^{-1+\delta} \leq 1 - bN^{-5} \leq 1 \right\},$$

together with the Lipschitz bounds

$$|\partial_z G_{ij}(z)| \leq \eta^{-2}, \quad |\partial_z m_N(z)| \leq \eta^{-2}, \quad |\partial_z m_{\text{sc}}(z)| \leq C\eta^{-2}.$$

Explain why a union bound over $\mathcal{L}$ and the one-step estimate from parts (b)–(d) prove the weak local law uniformly on D_δ :

$$|m_N(z) - m_{\text{sc}}(z)| \prec (N\eta)^{-1/4}, \quad \max_i |G_{ii}(z) - m_{\text{sc}}(z)| \prec (N\eta)^{-1/4}.$$

Solution

(a) Since H is real symmetric, write its spectral decomposition as

$$H = \sum_{\alpha=1}^N \lambda_{\alpha} u_{\alpha} u_{\alpha}^*, \quad G(z) = \sum_{\alpha=1}^N \frac{u_{\alpha} u_{\alpha}^*}{\lambda_{\alpha} - z}.$$

Hence

$$GG^* = \sum_{\alpha=1}^N \frac{u_{\alpha} u_{\alpha}^*}{|\lambda_{\alpha} - z|^2},$$

and therefore

$$\sum_{j=1}^N |G_{ij}(z)|^2 = (GG^*)_{ii} = \sum_{\alpha=1}^N \frac{|u_{\alpha}(i)|^2}{|\lambda_{\alpha} - z|^2}.$$

On the other hand,

$$\operatorname{Im} G_{ii}(z) = \sum_{\alpha=1}^N |u_{\alpha}(i)|^2 \operatorname{Im} \frac{1}{\lambda_{\alpha} - E - i\eta} = \eta \sum_{\alpha=1}^N \frac{|u_{\alpha}(i)|^2}{|\lambda_{\alpha} - z|^2}.$$

This proves the Ward identity

$$\sum_j |G_{ij}(z)|^2 = \frac{\operatorname{Im} G_{ii}(z)}{\eta}.$$

To obtain the Schur complement identity, isolate the i th row and column of $H - z$. The Schur complement formula gives

$$\begin{aligned} \frac{1}{G_{ii}} &= H_{ii} - z - \sum_{k,l}^{(i)} H_{ik} G_{kl}^{(i)} H_{li} \\ &= H_{ii} - z - m_N^{(i)} - Z_i \\ &= -z - m_N + \Upsilon_i. \end{aligned}$$

Multiplying by G_{ii} yields

$$1 + zG_{ii} + m_N G_{ii} = \Upsilon_i G_{ii}.$$

Averaging over i now gives

$$1 + zm_N + m_N^2 = \frac{1}{N} \sum_i \Upsilon_i G_{ii}.$$

Solution

(b) From

$$\Upsilon_i = H_{ii} - Z_i + m_N - m_N^{(i)},$$

it suffices to estimate the three terms on the right. The moment assumptions and a union bound imply

$$\max_i |H_{ii}| \prec N^{-1/2}.$$

Moreover, uniformly for $E \in [-3, 3]$ and $0 < \eta \leq 1$,

$$\rho(z) = \int_{-2}^2 \frac{\eta \varrho_{\text{sc}}(x)}{(x - E)^2 + \eta^2} dx \geq c\eta,$$

so $N^{-1/2} \leq C\sqrt{\rho/(N\eta)} \leq C\Omega$.

Conditioning on $H^{(i)}$ and applying quadratic concentration gives

$$|Z_i| \prec \left(\frac{1}{N^2} \text{tr} G^{(i)} (G^{(i)})^* \right)^{1/2}.$$

The Ward identity for the minor implies

$$\text{tr} G^{(i)} (G^{(i)})^* = \frac{1}{\eta} \text{Im tr} G^{(i)} = \frac{N}{\eta} \text{Im} m_N^{(i)}.$$

By the rank-one estimate,

$$\text{Im} m_N^{(i)} \leq \text{Im} m_N + |m_N^{(i)} - m_N| \leq \rho + \Lambda + \frac{C}{N\eta}.$$

Consequently,

$$\begin{aligned} |Z_i| &\prec \left(\frac{\rho + \Lambda}{N\eta} + \frac{C}{N^2\eta^2} \right)^{1/2} \\ &\prec \sqrt{\frac{\rho + \Lambda}{N\eta}} + \frac{1}{N\eta} = \Omega. \end{aligned}$$

Finally,

$$|m_N - m_N^{(i)}| \leq \frac{C}{N\eta} \leq C\Omega.$$

Combining these estimates, uniformly in i ,

$$\max_i |\Upsilon_i| \prec \Omega.$$

(c) Set

$$r := 1 + zm_N + m_N^2.$$

By part (a), part (b), and the a priori bound on the diagonal resolvent entries,

$$|r| \leq \frac{1}{N} \sum_i |\Upsilon_i G_{ii}| \prec \Omega.$$

Since $\Lambda \leq c_0$, $\rho \leq C$, and $N\eta \geq N^\delta$, we have

$$\Omega \prec (N\eta)^{-1/2} + (N\eta)^{-1} \prec (N\eta)^{-1/2}.$$

The stability lemma, applied with $u = m_N$, therefore yields

$$|m_N(z) - m_{\text{sc}}(z)| \leq C\sqrt{|r|} \prec (N\eta)^{-1/4}.$$

Solution

(d) The Schur complement identity and the semicircle equation give

$$\frac{1}{G_{ii}} = -z - m_N + \Upsilon_i, \quad \frac{1}{m_{\text{sc}}} = -z - m_{\text{sc}}.$$

Subtracting,

$$\frac{1}{G_{ii}} - \frac{1}{m_{\text{sc}}} = -(m_N - m_{\text{sc}}) + \Upsilon_i.$$

Equivalently,

$$G_{ii} - m_{\text{sc}} = G_{ii} m_{\text{sc}} ((m_N - m_{\text{sc}}) - \Upsilon_i).$$

Since $\max_i |G_{ii}| \leq C_0$ and $|m_{\text{sc}}(z)| \leq C$ on the domain, parts (b) and (c) imply

$$\begin{aligned} \max_i |G_{ii} - m_{\text{sc}}| &\leq C \left(|m_N - m_{\text{sc}}| + \max_i |\Upsilon_i| \right) \\ &\prec (N\eta)^{-1/4} + \Omega \prec (N\eta)^{-1/4}. \end{aligned}$$

Solution

(e) Fix $z = E + i$. The deterministic resolvent bound gives

$$\|G(E + i)\| \leq 1, \quad \max_i |G_{ii}(E + i)| \leq 1, \quad |m_N(E + i)| \leq 1.$$

We repeat the estimates in part (b), now without assuming a priori closeness to m_{sc} . The diagonal moment bound gives $\max_i |H_{ii}| \prec N^{-1/2}$, and the rank-one estimate gives

$$\max_i |m_N - m_N^{(i)}| \leq \frac{C}{N}.$$

For the quadratic term, quadratic concentration and Ward's identity for the minor yield

$$|Z_i| \prec \left(\frac{1}{N^2} \text{tr} G^{(i)} (G^{(i)})^* \right)^{1/2} = \left(\frac{\text{Im} m_N^{(i)}}{N} \right)^{1/2}.$$

Since $\|G^{(i)}(E + i)\| \leq 1$, we have $0 \leq \text{Im} m_N^{(i)} \leq 1$, and hence

$$\max_i |Z_i| \prec N^{-1/2}.$$

Thus

$$\max_i |\Upsilon_i| \prec N^{-1/2}.$$

Using the self-consistent equation and $\max_i |G_{ii}| \leq 1$,

$$|1 + z m_N + m_N^2| \leq \frac{1}{N} \sum_i |\Upsilon_i G_{ii}| \prec N^{-1/2}.$$

Here $m_N \in \mathbb{C}^+$ and $|m_N| \leq 1$. Since $\eta = 1$, the second part of the stability lemma applies without an a priori bound on $|m_N - m_{\text{sc}}|$, and gives

$$|m_N(E + i) - m_{\text{sc}}(E + i)| \prec N^{-1/4}.$$

Finally, the identity from part (d), together with $|G_{ii}| \leq 1$, $|m_{\text{sc}}| \leq C$, and $\max_i |\Upsilon_i| \prec N^{-1/2}$, gives

$$\max_i |G_{ii}(E+i) - m_{\text{sc}}(E+i)| \prec N^{-1/4}.$$

Solution

(f) Write

$$z_{a,b} := -3 + aN^{-5} + i\eta_b, \quad \eta_b := 1 - bN^{-5}.$$

The lattice has polynomial size, $|\mathcal{L}| = O(N^{10})$, so all pointwise stochastic-domination estimates below may be made simultaneous on $\mathcal{L}$ by choosing the failure exponent larger and applying a union bound.

At $b = 0$, part (e), followed by a union bound over the $O(N^5)$ choices of a , gives the desired estimates along the top row of the lattice. We now proceed downwards by induction on b . Suppose the weak local law holds on row b , and let

$$z = z_{a,b}, \quad z' = z_{a,b+1}.$$

The Lipschitz bounds give

$$\begin{aligned} |m_N(z') - m_{\text{sc}}(z')| &\leq |m_N(z) - m_{\text{sc}}(z)| + |m_N(z') - m_N(z)| + |m_{\text{sc}}(z') - m_{\text{sc}}(z)| \\ &\leq |m_N(z) - m_{\text{sc}}(z)| + CN^{-5}\eta_{b+1}^{-2}, \end{aligned}$$

and similarly

$$\max_i |G_{ii}(z') - m_{\text{sc}}(z')| \leq \max_i |G_{ii}(z) - m_{\text{sc}}(z)| + CN^{-5}\eta_{b+1}^{-2}.$$

Choose the small exponent in the definition of stochastic domination below $\delta/4$. Since

$$(N\eta_b)^{-1/4} \leq N^{-\delta/4}, \quad N^{-5}\eta_{b+1}^{-2} \leq N^{-3-2\delta},$$

the induction hypothesis implies, for all sufficiently large N , the a priori bounds

$$|m_N(z') - m_{\text{sc}}(z')| \leq c_0, \quad \max_i |G_{ii}(z')| \leq C_0.$$

Parts (b)–(d) may therefore be applied at z' , and they reset the estimates to

$$|m_N(z') - m_{\text{sc}}(z')| \prec (N\eta_{b+1})^{-1/4}, \quad \max_i |G_{ii}(z') - m_{\text{sc}}(z')| \prec (N\eta_{b+1})^{-1/4}.$$

This completes the induction. Because there are only $O(N^{10})$ lattice points and stochastic domination allows an arbitrarily large failure exponent, a union bound proves the two estimates uniformly on $\mathcal{L}$.

Finally, take any $z = E + i\eta \in D_\delta$ and choose $z_0 \in \mathcal{L}$ with $|z - z_0| \leq 2N^{-5}$. The imaginary parts of z and z_0 are comparable, and the Lipschitz bounds imply

$$|m_N(z) - m_N(z_0)| + |m_{\text{sc}}(z) - m_{\text{sc}}(z_0)| \leq CN^{-5}\eta^{-2} \leq CN^{-3-2\delta}.$$

The same estimate holds for each G_{ii} . Since

$$N^{-3-2\delta} = o((N\eta)^{-1/4})$$

uniformly on D_δ , the lattice estimates extend to the whole domain:

$$|m_N(z) - m_{\text{sc}}(z)| \prec (N\eta)^{-1/4}, \quad \max_i |G_{ii}(z) - m_{\text{sc}}(z)| \prec (N\eta)^{-1/4}.$$

Homework 3 | Due: 11:00 pm, Wednesday (June 24)**Setup**

Fix $\beta > 0$. The Gaussian β -ensemble is the probability measure on $\mathbb{R}^N$ with density

$$\frac{1}{Z_{N,\beta}} \prod_{i < j} |\lambda_i - \lambda_j|^\beta \exp\left(-\frac{\beta N}{4} \sum_{i=1}^N \lambda_i^2\right).$$

For $z = E + i\eta \in \mathbb{C}^+$, set

$$s(z) := m_N(z) := \frac{1}{N} \sum_{i=1}^N \frac{1}{\lambda_i - z}.$$

Let

$$m(z) := m_{\text{sc}}(z) = \frac{-z + \sqrt{z^2 - 4}}{2}, \quad m(z)^2 + zm(z) + 1 = 0, \quad \text{Im } m(z) > 0.$$

Define the self-consistent error

$$P(z) := s(z)^2 + zs(z) + 1.$$

Throughout the homework, fix $\kappa > 0$ and work in the bulk domain

$$|E| \leq 2 - \kappa, \quad 0 < \eta \leq 1.$$

All constants may depend on β and κ , but not on N , η , z , or the even moment q .

We shall start from the rank- p loop equation derived in Homework 1: for $z, z_2, \dots, z_p \in \mathbb{C} \setminus \mathbb{R}$,

$$\begin{aligned} & \mathbb{E} \left[\left(s(z)^2 + zs(z) + 1 + \frac{1}{N} \left(\frac{2}{\beta} - 1 \right) \partial_z s(z) \right) \prod_{j=2}^p s(z_j) \right] \\ & + \frac{2}{\beta N^2} \sum_{j=2}^p \mathbb{E} \left[\left(\partial_{z_j} \frac{s(z) - s(z_j)}{z - z_j} \right) \prod_{\substack{k=2 \\ k \neq j}}^p s(z_k) \right] = 0. \end{aligned}$$

Problem 1 (High-Moment Local Law for the Gaussian β -Ensemble). Let $q \geq 2$ be an even integer.

(a) Derive the recursive high-moment identity for $P(z)$. Define

$$f(z, w) := \begin{cases} \partial_w \frac{s(z) - s(w)}{z - w}, & w \neq z, \\ \frac{1}{2} s''(z), & w = z. \end{cases}$$

Prove that for $q \geq 4$,

$$\begin{aligned} \mathbb{E}|P(z)|^q &= -\frac{1}{N} \left(\frac{2}{\beta} - 1 \right) \mathbb{E} \left[s'(z) P(z)^{q/2-1} \overline{P(z)}^{q/2} \right] \\ &\quad - \frac{2(q/2-1)}{\beta N^2} \mathbb{E} \left[f(z, z) (2s(z) + z) P(z)^{q/2-2} \overline{P(z)}^{q/2} \right] \\ &\quad - \frac{q}{\beta N^2} \mathbb{E} \left[f(z, \bar{z}) (2s(\bar{z}) + \bar{z}) |P(z)|^{q-2} \right]. \end{aligned}$$

For $q = 2$, interpret the identity with the middle line omitted.

(b) Derive the moment bound

$$\mathbb{E}|P(z)|^q \leq \frac{(Cq)^{q/2}}{(N\eta)^q} \mathbb{E}(\operatorname{Im} s(z) + |2s(z) + z|)^q. \quad (6)$$

(c) Prove the bulk stability estimate. Since the real part E is bounded away from the edges ± 2 , prove the linear stability bound

$$|s(z) - m(z)| \leq C_\kappa |P(z)|, \quad (7)$$

and the square-root stability bound

$$|s(z) - m(z)| \leq C_\kappa |P(z)|^{1/2}. \quad (8)$$

Finally, deduce that

$$\operatorname{Im} s(z) + |2s(z) + z| \leq C_\kappa \left(1 + |P(z)|^{1/2}\right).$$

(d) Derive the moment recursion

$$\mathbb{E}|P(z)|^q \leq \frac{(Cq)^{q/2}}{(N\eta)^q} \left(1 + \mathbb{E}|P(z)|^{q/2}\right). \quad (9)$$

Then use Cauchy–Schwarz to show that, with

$$A := \frac{(Cq)^{q/2}}{(N\eta)^q},$$

one has

$$\mathbb{E}|P(z)|^q \leq C(A + A^2).$$

(e) Finish the high-moment local law:

$$\mathbb{E}|s(z) - m(z)|^q \leq \frac{(Cq)^{q/2}}{(N\eta)^q}. \quad (10)$$

Hint: use both the linear and the square-root stability bounds.

Solution

(a) Set $r = q/2$ and $c_\beta = 2/\beta - 1$. By expanding a polynomial $F(x, y)$ into monomials and applying the rank- p loop equation with repeated spectral parameters z and w , we obtain

$$\begin{aligned} \mathbb{E}[P(z)F(s(z), s(w))] &= -\frac{c_\beta}{N} \mathbb{E}[s'(z)F(s(z), s(w))] \\ &\quad - \frac{2}{\beta N^2} \mathbb{E}[f(z, z) \partial_1 F(s(z), s(w)) + f(z, w) \partial_2 F(s(z), s(w))]. \end{aligned}$$

Indeed, when $F(x, y) = x^a y^b$, the sum in the loop equation consists of a identical terms with $z_j = z$ and b identical terms with $z_j = w$, which gives precisely the two derivatives above.

Now take $w = \bar{z}$ and

$$F(x, y) = (x^2 + zx + 1)^{r-1} (y^2 + \bar{z}y + 1)^r.$$

Since $s(\bar{z}) = \overline{s(z)}$, we have

$$F(s(z), s(\bar{z})) = P(z)^{r-1} \overline{P(z)}^r,$$

while

$$\partial_1 F(s(z), s(\bar{z})) = (r-1)(2s(z) + z)P(z)^{r-2} \overline{P(z)}^r,$$

$$\partial_2 F(s(z), s(\bar{z})) = r(2s(\bar{z}) + \bar{z})|P(z)|^{q-2}.$$

Consequently,

$$\begin{aligned} \mathbb{E}|P(z)|^q &= -\frac{1}{N} \left(\frac{2}{\beta} - 1 \right) \mathbb{E} \left[s'(z) P(z)^{r-1} \overline{P(z)}^r \right] \\ &\quad - \frac{2(r-1)}{\beta N^2} \mathbb{E} \left[f(z, z) (2s(z) + z) P(z)^{r-2} \overline{P(z)}^r \right] \\ &\quad - \frac{2r}{\beta N^2} \mathbb{E} \left[f(z, \bar{z}) (2s(\bar{z}) + \bar{z}) |P(z)|^{q-2} \right]. \end{aligned}$$

Since $2r = q$, this is the stated recursion (with the conjugate term in the last line). For $q = 2$, one has $r - 1 = 0$, so the middle term is absent.

Solution

(b) We first record the elementary Stieltjes-transform derivative estimates. Since

$$s'(z) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(\lambda_i - z)^2}, \quad f(z, z) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(\lambda_i - z)^3},$$

and

$$f(z, \bar{z}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(\lambda_i - z)(\lambda_i - \bar{z})^2},$$

we have

$$|s'(z)| \leq \frac{\operatorname{Im} s(z)}{\eta}, \quad |f(z, z)| + |f(z, \bar{z})| \leq \frac{2 \operatorname{Im} s(z)}{\eta^2}.$$

Let

$$R := \operatorname{Im} s(z) + |2s(z) + z|.$$

Taking absolute values in part (a), using $|2s(\bar{z}) + \bar{z}| = |2s(z) + z|$, gives

$$\mathbb{E}|P|^q \leq \frac{C}{N\eta} \mathbb{E}[R|P|^{q-1}] + \frac{Cq}{(N\eta)^2} \mathbb{E}[R^2|P|^{q-2}].$$

For the first term, Young's inequality yields

$$\frac{C}{N\eta} R|P|^{q-1} \leq \frac{1}{4}|P|^q + \frac{C^q}{(N\eta)^q} R^q.$$

For the second term, when $q \geq 4$, Young's inequality with conjugate exponents $q/(q-2)$ and $q/2$ gives

$$\frac{Cq}{(N\eta)^2} R^2 |P|^{q-2} \leq \frac{1}{4} |P|^q + \frac{(Cq)^{q/2}}{(N\eta)^q} R^q.$$

For $q = 2$, the same estimate follows directly because $|P|^{q-2} = 1$. Taking expectations, absorbing the two $\frac{1}{4} \mathbb{E}|P|^q$ terms, and enlarging C gives

$$\mathbb{E}|P(z)|^q \leq \frac{(Cq)^{q/2}}{(N\eta)^q} \mathbb{E}(\operatorname{Im} s(z) + |2s(z) + z|)^q.$$

Solution

(c) Directly from the explicit semicircle transform (or by compactness after including $\eta = 0$),

$$\inf_{\substack{|E| \leq 2-\kappa \\ 0 < \eta \leq 1}} \operatorname{Im} m(E + i\eta) \geq c_\kappa > 0.$$

Moreover,

$$\begin{aligned} P(z) &= s(z)^2 + zs(z) + 1 - (m(z)^2 + zm(z) + 1) \\ &= (s(z) - m(z))(s(z) + m(z) + z). \end{aligned}$$

Since $s(z), m(z), z \in \mathbb{C}^+$,

$$|s(z) + m(z) + z| \geq \operatorname{Im}(s(z) + m(z) + z) \geq \operatorname{Im} m(z) \geq c_\kappa.$$

Therefore

$$|s(z) - m(z)| \leq C_\kappa |P(z)|,$$

which proves the linear stability estimate.

Write $\Delta = s(z) - m(z)$. Then

$$P(z) = \Delta^2 + (2m(z) + z)\Delta.$$

The quantity $2m(z) + z$ is uniformly bounded on the bulk domain, so the linear estimate gives

$$|\Delta|^2 \leq |P(z)| + |2m(z) + z| |\Delta| \leq C_\kappa |P(z)|.$$

Thus

$$|s(z) - m(z)| \leq C_\kappa |P(z)|^{1/2}.$$

Finally, $m(z)$ and $2m(z) + z$ are uniformly bounded on the same domain, and hence

$$\begin{aligned} \operatorname{Im} s(z) + |2s(z) + z| &\leq \operatorname{Im} m(z) + |s(z) - m(z)| + |2m(z) + z| + 2|s(z) - m(z)| \\ &\leq C_\kappa \left(1 + |P(z)|^{1/2}\right). \end{aligned}$$

Solution

(d) Combining the moment bound in part (b) with the final estimate in part (c),

$$\begin{aligned}\mathbb{E}|P(z)|^q &\leq \frac{(Cq)^{q/2}}{(N\eta)^q} \mathbb{E}\left[C_\kappa \left(1 + |P(z)|^{1/2}\right)\right]^q \\ &\leq \frac{(Cq)^{q/2}}{(N\eta)^q} \left(1 + \mathbb{E}|P(z)|^{q/2}\right),\end{aligned}$$

after enlarging C to absorb the factors depending on κ . This is the claimed moment recursion.

Set

$$A := \frac{(Cq)^{q/2}}{(N\eta)^q}, \quad B := \mathbb{E}|P(z)|^q.$$

By Cauchy–Schwarz, $\mathbb{E}|P(z)|^{q/2} \leq B^{1/2}$, and therefore

$$B \leq A(1 + B^{1/2}).$$

Equivalently, with $X = B^{1/2}$,

$$X^2 - AX - A \leq 0,$$

so

$$X \leq \frac{A + \sqrt{A^2 + 4A}}{2}.$$

It follows that

$$\mathbb{E}|P(z)|^q = B \leq C(A + A^2).$$

Solution

(e) Let

$$A := \frac{(Cq)^{q/2}}{(N\eta)^q}.$$

By part (d),

$$\mathbb{E}|P(z)|^q \leq C(A + A^2).$$

The linear stability bound yields

$$\mathbb{E}|s(z) - m(z)|^q \leq C_\kappa^q \mathbb{E}|P(z)|^q \leq C_\kappa^q C(A + A^2),$$

whereas the square-root stability bound and Cauchy–Schwarz give

$$\begin{aligned}\mathbb{E}|s(z) - m(z)|^q &\leq C_\kappa^q \mathbb{E}|P(z)|^{q/2} \\ &\leq C_\kappa^q (\mathbb{E}|P(z)|^q)^{1/2} \leq C_\kappa^q C(A + A^2)^{1/2}.\end{aligned}$$

If $A \leq 1$, the first estimate is bounded by $C_\kappa^q CA$; if $A \geq 1$, the second is bounded by $C_\kappa^q CA$. Enlarging the constant inside $(Cq)^{q/2}$ to absorb the κ -dependent factors, we conclude that

$$\mathbb{E}|s(z) - m(z)|^q \leq \frac{(Cq)^{q/2}}{(N\eta)^q}.$$

Homework 4 | Due: 11:00 pm, Thursday (June 25)

Setup. Throughout this homework, $d \geq 3$ is fixed.

Problem 1 (Local Tree-Likeness in the Configuration Model). Let $G_{N,d}^{\text{conf}}$ be the configuration model. Recall its construction: for each vertex $v \in [N]$, create d half-edges attached to v . Then choose a uniformly random perfect matching of all dN half-edges. Each matched pair gives an edge between the two vertices to which the paired half-edges are attached. The result is a random d -regular multigraph; loops and multiple edges are allowed. For a vertex o , write $B_R(o, G_{N,d}^{\text{conf}})$ for the radius- R neighborhood of o in this multigraph. A loop or a multiple edge should be counted as creating a cycle in the explored neighborhood.

Prove the following proposition.

Proposition 1 (Local tree-likeness). *Fix a small constant $c > 0$ and set*

$$R = \frac{c}{4} \log_{d-1} N.$$

For the random configuration model $G_{N,d}^{\text{conf}}$, with probability $1 - o(1)$, the following hold.

- (i) *All but at most N^c vertices of $G_{N,d}^{\text{conf}}$ have a tree neighborhood of radius R .*
- (ii) *Every radius- R neighborhood in $G_{N,d}^{\text{conf}}$ contains at most one cycle.*

The following outline may be helpful.

- (a) Show that a radius- R ball in a d -regular tree contains at most $C(d-1)^R$ vertices and $C(d-1)^R$ half-edges.
- (b) Recall the breadth-first exploration of $B_R(o, G_{N,d}^{\text{conf}})$. Initially, only the root o is discovered, and the first frontier is $F_0 = \{o\}$. For $r = 1, \dots, R$, we look at all unmatched half-edges incident to the vertices in the current frontier F_{r-1} , that is, to the vertices at distance $r-1$ from o . We then expose their partners in the random matching. If the partner of such a half-edge is incident to a vertex that has not been discovered before, we add this vertex to the next frontier F_r . We call the step a *collision* if the partner is incident to an already discovered vertex, or if two active half-edges are matched to half-edges incident to the same previously undiscovered vertex. Equivalently, a collision is any event in which the exploration fails to grow like a tree.

Show that, for a fixed root o , the probability of seeing at least one collision before the radius- R exploration finishes is at most

$$\frac{C(d-1)^{2R}}{N}.$$

Deduce that the expected number of vertices whose radius- R neighborhood is not a tree is at most $CN^{c/2}$, and prove part (i).

- (c) Show that, for a fixed root o , the probability that $B_R(o, G_{N,d}^{\text{conf}})$ contains at least two cycles is at most

$$\frac{C(d-1)^{4R}}{N^2}.$$

Use a union bound over $o \in [N]$ to prove part (ii).

Solution

Throughout, C denotes a constant depending only on the fixed degree d , which may change from line to line. We freely use that $(d-1)^R = N^{c/4} = o(N)$.

(a) Size of a tree ball. In the infinite d -regular tree the root has d neighbours and every non-root vertex has $d-1$ children, so the number of vertices at distance exactly $r \geq 1$ is $d(d-1)^{r-1}$. Hence

$$|B_R(o, \mathbb{T}_d)| = 1 + \sum_{r=1}^R d(d-1)^{r-1} = 1 + d \frac{(d-1)^R - 1}{d-2} \leq 1 + \frac{d}{d-2} (d-1)^R \leq C(d-1)^R,$$

using $d \geq 3$ (so $\frac{d}{d-2} \leq 3$). Each vertex carries exactly d half-edges, so the ball meets at most $d|B_R| \leq C(d-1)^R$ half-edges. Any graph of maximum degree d has its radius- R ball dominated by that of the tree, so the bounds

$$|B_R(o, G_{N,d}^{\text{conf}})| \leq C(d-1)^R, \quad \#\{\text{half-edges met}\} \leq C(d-1)^R$$

hold *deterministically* for the explored neighbourhood, irrespective of collisions (collisions only slow the growth).

(b) Collision probability and part (i). Reveal the matching one pair at a time along the breadth-first exploration: at each step take an unmatched active half-edge h and expose its partner $\pi(h)$, which, conditionally on the history $\mathcal{F}_{s-1}$ before step s , is uniform over the currently unmatched half-edges. By (a), throughout the exploration:

- the number of reveal steps is at most the number of half-edges met, $m \leq C(d-1)^R$;
- a step is a collision exactly when $\pi(h)$ is one of the (still unmatched) half-edges incident to an already-discovered vertex; there are at most $d|B_R| \leq C(d-1)^R$ of these;
- at least $dN - C(d-1)^R \geq dN/2$ half-edges remain to be chosen from (for N large, since $(d-1)^R = o(N)$).

Hence the conditional collision probability at each step satisfies

$$\mathbb{P}[\text{collision at step } s \mid \mathcal{F}_{s-1}] \leq \frac{C(d-1)^R}{dN/2} \leq \frac{C(d-1)^R}{N},$$

and a union bound over the $m \leq C(d-1)^R$ steps gives

$$\mathbb{P}[\text{some collision before radius } R] \leq m \cdot \frac{C(d-1)^R}{N} \leq \frac{C(d-1)^{2R}}{N},$$

which is the claimed bound. Now $B_R(o)$ is a tree iff no collision occurs, and $(d-1)^{2R} = N^{c/2}$, so for fixed o ,

$$\mathbb{P}[B_R(o) \text{ not a tree}] \leq \frac{CN^{c/2}}{N} = CN^{c/2-1}.$$

Let $X = \#\{o \in [N] : B_R(o) \text{ is not a tree}\}$. Then $\mathbb{E}[X] = \sum_o \mathbb{P}[B_R(o) \text{ not a tree}] \leq N \cdot CN^{c/2-1} = CN^{c/2}$, and Markov's inequality yields

$$\mathbb{P}[X \geq N^c] \leq \frac{\mathbb{E}[X]}{N^c} \leq CN^{-c/2} = o(1).$$

Thus with probability $1 - o(1)$ at most N^c vertices fail to have a tree radius- R neighbourhood, proving (i).

(c) Two cycles and part (ii). Each collision creates at most one cycle in the explored neighbourhood, so “ $B_R(o)$ has ≥ 2 cycles” implies “ ≥ 2 collision steps occur.” Let A_s be the event that step s is a collision; from (b), $\mathbb{P}[A_s \mid \mathcal{F}_{s-1}] \leq p := C(d-1)^R/N$, with $m \leq C(d-1)^R$ steps. For $s < s'$,

$$\mathbb{P}[A_s \cap A_{s'}] = \mathbb{E}[\mathbf{1}_{A_s} \mathbb{P}[A_{s'} \mid \mathcal{F}_{s'-1}]] \leq p \mathbb{P}[A_s] \leq p^2,$$

so

$$\mathbb{P}[\geq 2 \text{ collisions}] \leq \sum_{s < s'} \mathbb{P}[A_s \cap A_{s'}] \leq \binom{m}{2} p^2 \leq \frac{1}{2} (C(d-1)^R)^2 \left(\frac{C(d-1)^R}{N} \right)^2 = \frac{C(d-1)^{4R}}{N^2}.$$

Since $(d-1)^{4R} = N^c$, this is $\leq CN^{c-2}$ for each fixed o , and a union bound over $o \in [N]$ gives, using $c < 1$,

$$\mathbb{P}[\exists o : B_R(o) \text{ has } \geq 2 \text{ cycles}] \leq N \cdot CN^{c-2} = CN^{c-1} = o(1).$$

Hence with probability $1 - o(1)$ every radius- R neighbourhood has at most one cycle, proving (ii). Parts (i) and (ii) hold simultaneously with probability $1 - o(1)$, establishing the Proposition. $\square$

Problem 2 (Cheeger’s Constant). Let $G = (V, E)$ be a connected d -regular graph with $|V| = n$. Define

$$h_G := \min_{\emptyset \neq S \subsetneq V} \frac{|\partial S|}{\min\{|S|, |S^c|\}},$$

where ∂S denotes the set of edges crossing from S to S^c . After labeling the vertices by $V = \{1, \dots, n\}$, write $f_i = f(i)$. The goal is to prove that, for any function $f : V \rightarrow \mathbb{R}$,

$$h_G \sum_{1 \leq i < j \leq n} |f_i - f_j| \leq n \sum_{\{i,j\} \in E} |f_i - f_j|. \quad (11)$$

(a) Relabel the vertices so that

$$f_1 \geq f_2 \geq \dots \geq f_n,$$

and set $S_k := \{1, \dots, k\}$ for $1 \leq k \leq n-1$. Prove the two identities

$$\sum_{1 \leq i < j \leq n} |f_i - f_j| = \sum_{k=1}^{n-1} (f_k - f_{k+1}) k(n-k)$$

and

$$\sum_{\{i,j\} \in E} |f_i - f_j| = \sum_{k=1}^{n-1} (f_k - f_{k+1}) |\partial S_k|.$$

(b) Use the definition of h_G and the two identities above to prove (11).

(c) Suppose $f : V \rightarrow \mathbb{R}$ is 1-Lipschitz on edges, meaning that $|f(u) - f(v)| \leq 1$ for every $\{u, v\} \in E$. Let m be a median of f , i.e. a number such that

$$|\{v \in V : f(v) \leq m\}| \geq \frac{n}{2} \quad \text{and} \quad |\{v \in V : f(v) \geq m\}| \geq \frac{n}{2}.$$

Use (11) to prove that

$$\frac{1}{n} \sum_{v \in V} |f(v) - m| \leq \frac{d}{h_G}.$$

Hint: first compare $\sum_{i < j} |f_i - f_j|$ with $\frac{n}{2} \sum_v |f(v) - m|$.

Solution

(a) The two identities. Relabel the vertices so that $f_1 \geq \dots \geq f_n$. Then, for $i < j$,

$$|f_i - f_j| = f_i - f_j = \sum_{k=i}^{j-1} (f_k - f_{k+1}).$$

Therefore

$$\begin{aligned} \sum_{1 \leq i < j \leq n} |f_i - f_j| &= \sum_{1 \leq i < j \leq n} \sum_{k=1}^{n-1} (f_k - f_{k+1}) \mathbf{1}_{\{i \leq k < j\}} \\ &= \sum_{k=1}^{n-1} (f_k - f_{k+1}) \#\{(i, j) : i \leq k < j\} \\ &= \sum_{k=1}^{n-1} (f_k - f_{k+1}) k(n-k). \end{aligned}$$

For an undirected edge $\{i, j\}$, write it with $i < j$. Applying the same telescoping expansion to the edge sum gives

$$\begin{aligned} \sum_{\{i, j\} \in E} |f_i - f_j| &= \sum_{\{i, j\} \in E} \sum_{k=1}^{n-1} (f_k - f_{k+1}) \mathbf{1}_{\{i \leq k < j\}} \\ &= \sum_{k=1}^{n-1} (f_k - f_{k+1}) \#\{\{i, j\} \in E : i \leq k < j\}. \end{aligned}$$

The edges counted in the last display are precisely the edges crossing from $S_k = \{1, \dots, k\}$ to S_k^c . Hence

$$\sum_{\{i, j\} \in E} |f_i - f_j| = \sum_{k=1}^{n-1} (f_k - f_{k+1}) |\partial S_k|.$$

(b) Cheeger's inequality in L^1 . By the definition of h_G , for every $1 \leq k \leq n-1$,

$$|\partial S_k| \geq h_G \min\{|S_k|, |S_k^c|\} = h_G \min\{k, n-k\}.$$

Moreover

$$\min\{k, n-k\} \geq \frac{k(n-k)}{n},$$

since, if $k \leq n-k$, then $k \geq k(n-k)/n$, and the other case is symmetric. Thus

$$n|\partial S_k| \geq h_G k(n-k).$$

Since $f_k - f_{k+1} \geq 0$, multiplying by $f_k - f_{k+1}$, summing in k , and using part (a) gives

$$\begin{aligned} h_G \sum_{1 \leq i < j \leq n} |f_i - f_j| &= h_G \sum_{k=1}^{n-1} (f_k - f_{k+1}) k(n-k) \\ &\leq n \sum_{k=1}^{n-1} (f_k - f_{k+1}) |\partial S_k| \\ &= n \sum_{\{i,j\} \in E} |f_i - f_j|. \end{aligned}$$

This is (11).

(c) Median bound for Lipschitz functions. A median m minimizes the convex function

$$x \mapsto \sum_{i=1}^n |f_i - x|.$$

Consequently, for each j ,

$$\sum_{i=1}^n |f_i - m| \leq \sum_{i=1}^n |f_i - f_j|.$$

Summing this inequality over j gives

$$n \sum_{i=1}^n |f_i - m| \leq \sum_{j=1}^n \sum_{i=1}^n |f_i - f_j| = 2 \sum_{1 \leq i < j \leq n} |f_i - f_j|,$$

and hence

$$\frac{1}{n} \sum_{i=1}^n |f_i - m| \leq \frac{2}{n^2} \sum_{1 \leq i < j \leq n} |f_i - f_j|.$$

Using (11) and the assumption that f is 1-Lipschitz on edges,

$$\sum_{1 \leq i < j \leq n} |f_i - f_j| \leq \frac{n}{h_G} \sum_{\{i,j\} \in E} |f_i - f_j| \leq \frac{n}{h_G} |E|.$$

Since G is d -regular, $|E| = dn/2$. Combining the two preceding displays yields

$$\frac{1}{n} \sum_{v \in V} |f(v) - m| \leq \frac{2}{n^2} \cdot \frac{n}{h_G} \cdot \frac{dn}{2} = \frac{d}{h_G},$$

as required.