hexagon:2611.00927v1

UNCONDITIONAL IRREDUCIBILITY OF CHARACTERISTIC POLYNOMIALS OF SYMMETRIC RADEMACHER MATRICES

TAO HU AND QUANYU TANG

Abstract. We prove that the characteristic polynomial of an $n \times n$ symmetric Rademacher matrix is irreducible over $\mathbb{Q}$ with probability $1 - O(\exp(-cn^{1/4}))$ for some absolute constant $c > 0$. The proof combines the zero-free half-plane for Dedekind zeta functions of Cureton and Thorner with the finite-field rank distribution theorem of Ferber, Jain, Sah, and Sawhney.

1. Introduction

The arithmetic structure of the eigenvalues of an integer matrix is encoded by its characteristic polynomial. Irreducibility over $\mathbb{Q}$ means that the matrix has no nonzero proper invariant subspace defined over $\mathbb{Q}$. For a real symmetric matrix, all eigenvalues are real, but this does not determine whether they belong to a single irreducible polynomial over $\mathbb{Q}$. The question is particularly natural for a symmetric matrix whose independent entries are random signs.

For matrices with all entries independent, the conjecture that the characteristic polynomial is usually irreducible goes back to unpublished work of Babai in the 1970s and was raised again by Vu and P. M. Wood in 2009; see Vu's account [7, Section 11]. A major advance in the study of related random polynomials was the work of Breuillard and Varjú [1]. They connected factorization over $\mathbb{Q}$ with averages of root counts modulo primes, using prime ideals in the fields generated by individual roots. Eberhard [3] adapted this method to characteristic polynomials of matrices with independent entries, under the extended Riemann hypothesis. Here the extended Riemann hypothesis means the Riemann hypothesis for the Dedekind zeta function of every number field. Eberhard also raised the symmetric case [3, Section 1.4].

We consider the following precise question.

Problem 1.1. For each $n \geq 1$, let A_n be a symmetric $n \times n$ matrix whose entries on and above the diagonal are independent and take the values 1 and -1, each with probability $1/2$. Is

$$\lim_{n\to\infty} \mathbb{P}\big(\det(tI_n - A_n) \text{ is irreducible in } \mathbb{Q}[t]\big) = 1?$$

Ferber, Jain, Sah, and Sawhney [4, Theorem 1.1] answered this question under the extended Riemann hypothesis, with an exceptional probability of order $\exp(-cn^{1/4})$. Their main contribution was a uniform theorem for the ranks of $A_n - \lambda I_n$ over finite fields, valid for primes as large as $\exp(\eta n^{1/4})$. We combine that rank theorem with the zero-free theorem of Cureton and Thorner [2] to obtain the same scale of error without the Riemann hypothesis.

Theorem 1.2. *There are absolute constants $c, C > 0$ such that, for every $n \geq 1$,*

$$\mathbb{P}\big(\det(tI_n - A_n) \textit{ is reducible in } \mathbb{Q}[t]\big) \leq C\exp(-cn^{1/4}).$$

Theorem 1.2 answers Problem 1.1 affirmatively. Its analytic input is the following result of Cureton and Thorner [2, Theorem 1.1].

Date: November 23, 2026.

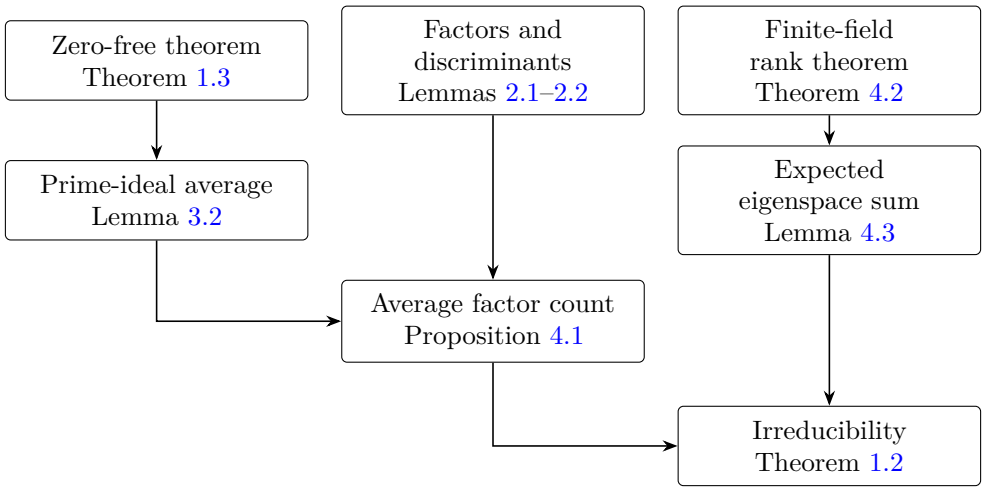


FIGURE 1. The two external theorems and the steps connecting them.

Theorem 1.3. *For every number field K, its Dedekind zeta function satisfies*

$$\zeta_K(s) \neq 0 \qquad \text{if } \Re s > \frac{31}{32} \text{ and } s \neq 1.$$

The proof below uses Theorem 1.3 as an external result. The issue is to extract an estimate that is uniform in the degree and discriminant of K, since the fields arising from the characteristic polynomial vary with both n and the matrix. We give this step explicitly. In fact, any fixed zero-free half-plane $\Re s > 1 - \delta$, common to the relevant fields, is sufficient. The use of weaker information about zeros in irreducibility questions already appears in Breuillard and Varjú [1, Theorems 5 and 6].

Our argument follows the prime-averaging method of Breuillard–Varjú, Eberhard, and Ferber–Jain–Sah–Sawhney [1, 3, 4]. We make one modification to the statistic being averaged. For a fixed symmetric integer matrix A, we use

$$S_A(p) = \sum_{\lambda \in \mathbb{F}_p} \dim_{\mathbb{F}_p} \ker(\overline{A} - \lambda I_n),$$

where $\overline{A}$ is the reduction of A modulo p. Away from explicitly controlled primes, this statistic counts linear factors with their multiplicities. Its prime average therefore approximates the number of rational irreducible factors counted with multiplicity. By contrast, averaging distinct roots detects only distinct irreducible factors. This modification incorporates proper powers directly and removes the separate simple-spectrum argument used in [4, Appendix A].

The finite-field rank estimate is entirely supplied by [4, Theorem 1.2]; the zero-free input is entirely supplied by [2, Theorem 1.1]. The purpose of this note is to give their uniform combination and the multiplicity argument in detail. Figure 1 records the dependencies.

Organization. Section 2 relates finite-field eigenspaces to rational factors and bounds the relevant discriminants. Section 3 proves the uniform prime-ideal estimate. Section 4 combines these statements with the finite-field rank theorem to prove Theorem 1.2.

Notation. The letter p always denotes a rational prime, and $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. We write I_n for the identity matrix, $\mathbf{1}_E$ for the indicator of a set or event E, and $\mathbb{P}$ and $\mathbb{E}$ for probability and expectation. The symbol log denotes the natural logarithm. The symbols $O(\cdot)$ and $\ll$ have absolute implied constants unless a subscript indicates otherwise. Empty products equal 1. All

irreducible factors over $\mathbb{Q}$ are taken to be monic. For a number field K, we write $\mathcal{O}_K$ for its ring of integers and $D_K = |\operatorname{disc}(K)|$ for its absolute discriminant.

Formalization. Theorem 1.2 has been formalized in Lean 4. The formalization is available on GitHub and registered in the Palomar Registry as PALOMAR-2026-11-23-000014. It uses the formalized zero-free theorem of Cureton and Thorner.

Statement on AI usage. The authors acknowledge the use of AI tools for exploratory discussions and language refinement during the preparation of this work. ChatGPT assisted in developing certain technical arguments in the proof of the main theorem. OpenAI's Codex assisted with the Lean 4 formalization of the main theorem.

2. Rational factors and finite-field eigenspaces

Fix a symmetric integer matrix A of size n, with $|A_{ij}| \le 1$. By Gauss's lemma, we may write

$$\chi_A(t) = \det(tI_n - A) = \prod_{i=1}^{r} g_i(t)^{e_i}, \qquad g_i \in \mathbb{Z}[t],$$

where the g_i are distinct monic irreducible polynomials and the e_i are positive integers. Set

$$d_i = \deg g_i, \qquad M(A) = \sum_{i=1}^{r} e_i.$$

Then $\sum_i e_i d_i = n$, and

$$\mathbf{1}_{\{\chi_A \text{ is reducible over } \mathbb{Q}\}} \le M(A) - 1. \tag{2.1}$$

Indeed, $M(A) = 1$ holds exactly when χ_A is irreducible.

Since A is diagonalizable over $\mathbb{R}$, its minimal polynomial is the squarefree polynomial

$$m_A(t) = \prod_{i=1}^{r} g_i(t).$$

Put $\Delta_A = \operatorname{disc}(m_A) \neq 0$. Choose a root θ_i of g_i and set $K_i = \mathbb{Q}(\theta_i)$ and $D_i = D_{K_i}$. Every conjugate of θ_i is real, so each K_i is totally real.

For every prime p, define

$$S_A(p) = \sum_{\lambda \in \mathbb{F}_p} \dim_{\mathbb{F}_p} \ker(\overline{A} - \lambda I_n).$$

Distinct eigenspaces form a direct sum, and hence

$$0 \le S_A(p) \le n. \tag{2.2}$$

For any number field K, let $P_K(p)$ be the number of prime ideals $\mathfrak{p}$ of $\mathcal{O}_K$ with $\#(\mathcal{O}_K/\mathfrak{p}) = p$. In particular, $0 \le P_K(p) \le [K : \mathbb{Q}]$, including at ramified primes.

Lemma 2.1. *For every prime $p \nmid \Delta_A$,*

$$S_A(p) = \sum_{i=1}^{r} e_i P_{K_i}(p).$$

Proof. Reduction modulo p preserves the identity $m_A(A) = 0$. Since $p \nmid \Delta_A$, the reduction of m_A is squarefree. Thus $\overline{A}$ is diagonalizable over an algebraic closure of $\mathbb{F}_p$. The reductions of the g_i are squarefree and pairwise coprime. Each root in $\mathbb{F}_p$ of the reduction of g_i therefore has algebraic and geometric multiplicity e_i. If $R_{g_i}(p)$ denotes the number of these roots, then

$$S_A(p) = \sum_{i} e_i R_{g_i}(p).$$

The discriminant product formula shows that $\operatorname{disc}(g_i)$ divides Δ_A. The discriminant-index identity gives

$$|\operatorname{disc}(g_i)| = D_i[\mathcal{O}_{K_i} : \mathbb{Z}[\theta_i]]^2.$$

Consequently p does not divide the index on the right. Dedekind's factorization theorem [6, Chapter 3, Theorem 27] identifies the irreducible factors of $g_i \bmod p$ with the prime ideals above p, with factor degree equal to residue degree. Its degree-one case yields $R_{g_i}(p) = P_{K_i}(p)$, proving the identity. □

Lemma 2.2. *The discriminants above satisfy*

$$\log|\Delta_A| \le n(n-1)\log(2n), \qquad \sum_{i=1}^{r} e_i \log D_i \le n(n-1)\log(2n).$$

Proof. Every eigenvalue of A belongs to $[-n, n]$, by the bound on its row sums. If h is a monic squarefree polynomial of degree d all of whose roots are eigenvalues of A, the discriminant product formula gives

$$|\operatorname{disc}(h)| \le (2n)^{d(d-1)}.$$

Applying this to m_A, whose degree is at most n, proves the first estimate. Applying it to g_i, and using $D_i \le |\operatorname{disc}(g_i)|$, gives

$$\sum_i e_i \log D_i \le \log(2n) \sum_i e_i d_i(d_i - 1) \le (n-1)\log(2n) \sum_i e_i d_i = n(n-1)\log(2n).$$

□

3. A uniform average over prime ideals

Fix the nonnegative weight

$$w(u) = 48\left(u - \frac{1}{2}\right)(1-u)\mathbf{1}_{[1/2,1]}(u).$$

It satisfies $\int_0^\infty w(u)\,du = 1$ and $0 \le w(u) \le 3$. Its Mellin transform

$$W(s) = \int_0^\infty w(u) u^{s-1}\,du$$

is entire and has $W(1) = 1$. Moreover,

$$|W(\sigma + it)| \ll (1 + |t|)^{-2}, \qquad \frac{3}{4} \le \sigma \le 2. \tag{3.1}$$

To verify this, substitute $u = e^v$ and integrate twice by parts on $[-\log 2, 0]$. The first boundary terms vanish because w vanishes at both endpoints; the second boundary terms are bounded uniformly in the indicated interval of σ.

We first record the uniform logarithmic-derivative estimate needed for the contour argument. Throughout this section K has degree $d = [K : \mathbb{Q}]$ and absolute discriminant $D = D_K$. We define its logarithmic-derivative coefficients by

$$-\frac{\zeta_K'}{\zeta_K}(s) = \sum_{m \ge 1} \frac{a_K(m)}{m^s}, \qquad \Re s > 1.$$

Lemma 3.1. *Let $0 < \delta \le 1/2$, and suppose that every nontrivial zero ρ of ζ_K satisfies $\Re\rho \le 1 - \delta$. With $\sigma_0 = 1 - \delta/2$, one has, for every $t \in \mathbb{R}$,*

$$\left|\frac{\zeta_K'}{\zeta_K}(\sigma_0 + it)\right| \ll_\delta \log D + d\log(|t| + 3). \tag{3.2}$$

Proof. The Euler product gives

$$0 \le a_K(m) \le d\Lambda(m), \qquad a_K(p) = P_K(p)\log p, \tag{3.3}$$

where $\Lambda(p^j) = \log p$ for $j \ge 1$ and $\Lambda(m) = 0$ otherwise. Indeed, for $m = p^j$, the coefficient divided by $\log p$ is the sum of the residue degrees $f_{\mathfrak{p}}$ over prime ideals $\mathfrak{p} \mid p$ for which $f_{\mathfrak{p}} \mid j$. This sum is at most d. It follows that

$$\left|\frac{\zeta_K'}{\zeta_K}(2+it)\right| \le d\sum_{m\ge1}\frac{\Lambda(m)}{m^2} \ll d.$$

Let r_1 and r_2 be the numbers of real embeddings and pairs of complex embeddings of K, respectively, so that $r_1 + 2r_2 = d$. Put

$$G_K(s) = \left(\pi^{-s/2}\Gamma(s/2)\right)^{r_1}\left((2\pi)^{-s}\Gamma(s)\right)^{r_2},$$

where Γ is the gamma function. The Hadamard product for the completed Dedekind zeta function gives

$$\frac{\zeta_K'}{\zeta_K}(s) = B_K + \sum_\rho\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right) - \frac{1}{2}\log D$$
$$-\frac{1}{s} - \frac{1}{s-1} - \frac{G_K'}{G_K}(s),$$

where zeros are counted with multiplicity and $B_K = -\sum_\rho \Re(1/\rho)$. These identities are unconditional; see [5, Section 2, equations (2.5)–(2.6)]. We do not use the GRH-dependent estimates in that paper.

Take real parts at $z = 2 + it$. Since $0 < \Re\rho < 1$, the zero terms have positive real parts and are comparable to $(1 + (t - \Im\rho)^2)^{-1}$. The preceding Euler-product bound and Stirling's formula therefore give

$$\sum_\rho \frac{1}{1+(t-\Im\rho)^2} \ll \log D + d\log(|t|+3). \tag{3.4}$$

Now subtract the Hadamard identities at $s = \sigma_0 + it$ and $z = 2 + it$. Both B_K and the $1/\rho$ terms cancel. The zero-free assumption implies $\sigma_0 - \Re\rho \ge \delta/2$, and hence

$$\left|\frac{1}{s-\rho} - \frac{1}{z-\rho}\right| = \frac{2-\sigma_0}{|s-\rho|\,|z-\rho|} \ll_\delta \frac{1}{1+(t-\Im\rho)^2}.$$

The difference series is thus absolutely convergent and is bounded by (3.4). The gamma-factor terms are $O(d\log(|t|+3))$, and the rational terms are $O_\delta(1)$. This proves (3.2). The same argument is uniform for $\sigma_0 \le \Re s \le 2$ when $|\Im s| \ge 1$. □

Lemma 3.2. *Under the assumptions of Lemma 3.1, uniformly for $x \ge 4$,*

$$\frac{1}{x}\sum_p P_K(p)\log p\, w(p/x) = 1 + O_\delta\left((\log D + d)x^{-\delta/2} + dx^{-1/2}\log x\right).$$

The implied constant is independent of K.

Proof. Mellin inversion, with absolute convergence on $\Re s = 2$, gives

$$\sum_{m\ge1} a_K(m)w(m/x) = -\frac{1}{2\pi i}\int_{\Re s=2}\frac{\zeta_K'}{\zeta_K}(s)x^s W(s)\,ds.$$

Shift the contour to $\Re s = \sigma_0 = 1 - \delta/2$. No zero is crossed, and the simple pole of ζ_K at 1 contributes $xW(1) = x$. The uniform bound at large height established in Lemma 3.1, together with (3.1), makes the horizontal integrals tend to zero. On the new vertical line, (3.2) yields

$$\sum_{m\ge1} a_K(m)w(m/x) = x + O_\delta\big((\log D + d)x^{1-\delta/2}\big).$$

By (3.3), the terms with $m = p$ are precisely the prime sum in the statement. All remaining nonzero terms have $m = p^j$ with $j \geq 2$, so their total contribution is at most

$$3d \sum_{p \leq \sqrt{x}} \log p \left\lfloor \frac{\log x}{\log p} \right\rfloor \leq 3d\sqrt{x} \log x.$$

Removing them and dividing by x proves the result. □

4. From finite-field ranks to irreducibility

For the matrix A considered in Section 2, set

$$Z_A(x) = \frac{1}{x} \sum_p S_A(p) \log p \, w(p/x).$$

The following proposition makes the dependence on the matrix uniform.

Proposition 4.1. *Let $0 < \delta \leq 1/2$. Suppose that, for every i, the function ζ_{K_i} has no zero in $\Re s > 1 - \delta$. Uniformly for $x \geq 4$ and all the matrices A under consideration,*

$$|Z_A(x) - M(A)| \ll_\delta n^2 \log(2n) x^{-\delta/2} + nx^{-1/2} \log x + n^3 \log(2n) x^{-1}. \tag{4.1}$$

Proof. Lemma 2.1 identifies $S_A(p)$ with $\sum_i e_i P_{K_i}(p)$ when $p \nmid \Delta_A$. At every prime both quantities lie in $[0, n]$, by (2.2) and $\sum_i e_i d_i = n$. Replacing one by the other in $Z_A(x)$ therefore changes its value by at most

$$\frac{3n}{x} \sum_{p \mid \Delta_A} \log p \leq \frac{3n \log |\Delta_A|}{x}.$$

Apply Lemma 3.2 to each K_i and multiply by e_i. The main terms sum to $M(A)$. The total error is

$$\ll_\delta \left(\sum_i e_i \log D_i + n \right) x^{-\delta/2} + nx^{-1/2} \log x + \frac{n \log |\Delta_A|}{x}.$$

Lemma 2.2 proves (4.1). □

We next state the part of the rank-distribution theorem of Ferber, Jain, Sah, and Sawhney [4, Theorem 1.2] that is needed here.

Theorem 4.2. *There is an absolute constant $\eta > 0$ such that, for every prime $2 < p \leq \exp(\eta n^{1/4})$, every $\lambda \in \mathbb{F}_p$, and $k \in \{0, 1\}$,*

$$\mathbb{P}\left(\dim_{\mathbb{F}_p} \ker(\overline{A}_n - \lambda I_n) = k \right) = \frac{\alpha_p}{\prod_{j=1}^{k} (p^j - 1)} + O\left(\exp\left(-\frac{\eta n}{\log p} \right) \right).$$

Here $\alpha_p = \prod_{j=0}^{\infty} (1 - p^{-(2j+1)})$. The implied constant is absolute and the error is uniform in all the displayed parameters.

Lemma 4.3. *For every prime in the range of Theorem 4.2,*

$$\mathbb{E} S_{A_n}(p) \leq 1 + C \left(\frac{n}{p^2} + np \exp\left(-\frac{\eta n}{\log p} \right) \right), \tag{4.2}$$

where $C > 0$ is absolute.

Proof. Fix $\lambda \in \mathbb{F}_p$, and abbreviate

$$\kappa = \dim_{\mathbb{F}_p} \ker(\overline{A}_n - \lambda I_n), \qquad \varepsilon = \exp\left(-\frac{\eta n}{\log p} \right).$$

Write

$$\beta_p = \frac{\alpha_p}{1-p^{-1}} = \prod_{j=1}^{\infty}(1-p^{-(2j+1)}).$$

Since $p \geq 3$,

$$0 \leq 1-\beta_p \leq \sum_{j=1}^{\infty} p^{-(2j+1)} = \frac{p^{-3}}{1-p^{-2}} \leq \frac{9}{8}p^{-3}.$$

The two cases of Theorem 4.2 now give

$$\mathbb{P}(\kappa \geq 1) = 1-\alpha_p + O(\varepsilon) \leq p^{-1} + Cp^{-3} + C\varepsilon,$$
$$\mathbb{P}(\kappa \geq 2) = 1-\alpha_p - \frac{\alpha_p}{p-1} + O(\varepsilon) = 1-\beta_p + O(\varepsilon) \leq Cp^{-3} + C\varepsilon.$$

Use the elementary inequality

$$\kappa \leq \mathbf{1}_{\{\kappa \geq 1\}} + (n-1)\mathbf{1}_{\{\kappa \geq 2\}}$$

to obtain

$$\mathbb{E}\kappa \leq p^{-1} + Cnp^{-3} + Cn\varepsilon.$$

Summing over $\lambda \in \mathbb{F}_p$ proves (4.2). □

We now prove Theorem 1.2.

Proof. Fix $\delta = 1/32$. Theorem 1.3 verifies the zero-free hypotheses in Proposition 4.1 for every realization of A_n. For sufficiently large n, put

$$x = \exp\left(\frac{\eta}{2}n^{1/4}\right). \tag{4.3}$$

Every prime contributing to $Z_{A_n}(x)$ then satisfies

$$2 < p, \qquad x/2 \leq p \leq x < \exp(\eta n^{1/4}).$$

By (4.2),

$$\mathbb{E}S_{A_n}(p) \leq 1 + O\left(nx^{-2} + nx\exp\left(-\frac{\eta n}{\log x}\right)\right)$$

uniformly on this support. Lemma 3.2 for $K = \mathbb{Q}$ gives

$$\frac{1}{x}\sum_{p} \log p\, w(p/x) = 1 + O_\delta\left(x^{-\delta/2} + x^{-1/2}\log x\right).$$

All the weights are nonnegative. Taking expectations in (4.1) and using the last two estimates yields

$$\begin{aligned} 0 \leq \mathbb{E}M(A_n) - 1 \ll_\delta\ & n^2\log(2n)x^{-\delta/2} + nx^{-1/2}\log x \\ & + n^3\log(2n)x^{-1} + nx^{-2} + nx\exp\left(-\frac{\eta n}{\log x}\right). \end{aligned} \tag{4.4}$$

For the choice (4.3), the first term on the right of (4.4) is

$$n^2\log(2n)\exp\left(-\frac{\delta\eta}{4}n^{1/4}\right).$$

It is bounded by $C\exp(-cn^{1/4})$ after decreasing $c > 0$. The next three terms satisfy bounds of the same form. The last term equals

$$n\exp\left(\frac{\eta}{2}n^{1/4} - 2n^{3/4}\right)$$

and is smaller still. Finally, (2.1) gives

$$\mathbb{P}(\chi_{A_n} \text{ is reducible over } \mathbb{Q}) \leq \mathbb{E}M(A_n) - 1 \leq C\exp(-cn^{1/4}).$$

Increasing C covers the finitely many remaining values of n. □

Remark. Only totally real fields occur in this argument. More generally, if one assumes a common zero-free half-plane $\Re s > 1 - \delta$, for a fixed $0 < \delta \leq 1/2$, only for Dedekind zeta functions of totally real number fields, the proof gives the same conclusion with constants depending on δ. No independence between different primes or different values of λ is needed: the probabilistic step uses only linearity of expectation.

References

[1] E. Breuillard and P. P. Varjú, *Irreducibility of random polynomials of large degree*, Acta Math. **223** (2019), no. 2, 195–249. https://doi.org/10.4310/ACTA.2019.v223.n2.a1.

[2] D. Cureton and J. Thorner, *A quasi-Riemann hypothesis for Dedekind zeta functions*, Hexagon preprint 2610.00200v2 (2026), 36 pp. https://hexagonmath.org/2610.00200v2.

[3] S. Eberhard, *The characteristic polynomial of a random matrix*, Combinatorica **42** (2022), no. 4, 491–527. https://doi.org/10.1007/s00493-020-4657-0.

[4] A. Ferber, V. Jain, A. Sah, and M. Sawhney, *Random symmetric matrices: rank distribution and irreducibility of the characteristic polynomial*, Math. Proc. Cambridge Philos. Soc. **174** (2023), no. 2, 233–246. https://doi.org/10.1017/S0305004122000226.

[5] L. Grenié and G. Molteni, *Explicit versions of the prime ideal theorem for Dedekind zeta functions under GRH*, Math. Comp. **85** (2016), no. 298, 889–906. https://doi.org/10.1090/mcom3031.

[6] D. A. Marcus, *Number Fields*, second edition, Universitext, Springer, Cham, 2018. https://doi.org/10.1007/978-3-319-90233-3.

[7] V. H. Vu, *Recent progress in combinatorial random matrix theory*, Probab. Surv. **18** (2021), 179–200. https://doi.org/10.1214/20-PS346.

Department of Information, Risk, and Operations Management, McCombs School of Business, The University of Texas at Austin, Austin, TX 78712, USA

Email address: tao_hu@utexas.edu

School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, P. R. China

Email address: tangquanyu827@gmail.com