PI4 PROGRESS REPORT: RANDOM MATRIX PENCILS IN GE_n

RUQIAN CHEN, RAEMEON COWAN, ELLIOT KAPLAN, HADRIAN QUAN

Consider the pencil of matrices

$$A + tB,$$

for $t \in \mathbb{C}$, where A, B are random matrices belonging to the Wigner-Ginibre ensemble, GE_n. GE_n is defined to be the ensemble of random matrices whose components are independent, identically distributed random variables in a complex valued Normal distribution.

Explicitly, this means that each component of $A = (a_{r,s})$ [and each component of the matrix $B = (b_{rs})$] is a random variable $a_{r,s}$ whose probability distribution is given by the following probability density function

$$f_{a_{r,s}}(x) := \frac{1}{\sqrt{\pi}} e^{-2x^2} + \frac{i}{\sqrt{\pi}} e^{-2x^2},$$

i.e. the density associated to the distribution $N\left(0, \frac{1}{2}\right) + iN\left(0, \frac{1}{2}\right)$. This is a linear combination of normal distributions with expectation zero, and variance $\frac{1}{2}$ Here $i^2 = -1$.

Recall that the characteristic polynomial of a matrix has roots which are the eigenvalues of the matrix. In particular, we can consider the characteristic polynomial of this matrix pencil for different values of the complex parameter t:

$$det(A + tB + \lambda I) = P(\lambda, t),$$

this defines a bivariate polynomial $P(\lambda, t)$ in the complex parameter t, and the spectral parameter λ.

The following question involves this polynomial $P(\lambda, t)$ and is of interest to both Physicists and Algebraic Geometers. The determinant $det(A + tB + \lambda I)$ reduces the pencil to a degree n polynomial in variable λ, with complex coefficients (the components of A, B are complex). By the fundamental theorem of algebra, we expect that for generic values of t_0, the polynomial $P(\lambda, t_0)$ should have n distinct roots $\lambda_1 \leq \ldots \leq \lambda_n$. *Conversely, for what values of $t \in \mathbb{C}$ do we have* ***fewer than*** n ***distinct eigenvalues***?

In the physics literature, such values are referred to as *level crossings* of the linear operator $H \sim A + tB$ associated to the random matrices. Algebraic geometers call these values *branching points*, in conjunction with the bivariate polynomial $P(\lambda, t)$ becomes multi-valued (has distinct branches). To physicists, the linear operator these matrices represent may encode some physical observable (position, momentum, spin, etc.) and its eigenvalues are 'energy' levels of this observable. Quantum mechanics takes as axiom that these energy levels should be discretely spaced (or quantized), thus the ordering of the eigenvalues $\lambda_1 \leq \ldots \leq \lambda_n$ corresponds to spacings of the 'energy' levels. If two eigenvalues collide for some value of t, and we have fewer than n distinct eigenvalues, than two or more levels have 'crossed'.

To understand when and where these branching points might occur, we'd like to gather some statistical understanding of them. These and related questions include: understanding the Monodromy of these branching points, and determining the Joint Probability Distribution of these branchings points (or barring that, determining the asymptotic behavior of this distribution as the size of the matrix tends to infinity).

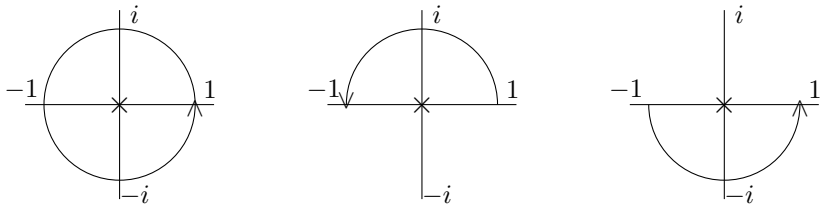
1. Monodromy

Consider the matrix pencil $A + tB$ where $t \in \mathbb{C}P^1$ and A, B are random $n \times n$ complex matrices. The characteristic polynomial, $P(t, \lambda) = \det(A + tB + \lambda I)$ is a polynomial in t and λ. Considered as a polynomial in λ, $P(t, \lambda)$ has a repeated root (that is, $A + tB$ has a repeated eigenvalue) presicely when its discriminant with respect to λ equals zero. This discriminant is a polynomial in t of degree $n(n-1)$ and its roots are precisely the branching points of the equation $P(t, \lambda) = 0$.

In general, the discriminant of $P(t, \lambda)$ has $n(n-1)$ distinct roots, so the equation $P(t, \lambda) = 0$ has $n(n-1)$ distinct branching points. These branching points are *simple* in general, meaning that at each branching point t, the matrix $A + tB$ has $n - 1$ distinct eigenvalues, one of which is repeated twice.

Consider the equation $P(t, \lambda) = 0$ as a function $t = f(\lambda)$. For almost all t, there are n distinct values λ such that $f(\lambda) = t$, but at our branching points there are, in general, only $n - 1$ such values. Monodromy is the phenomenon that occurs when we examine the values of λ for which $f(\lambda) = t$ as we allow t to move in

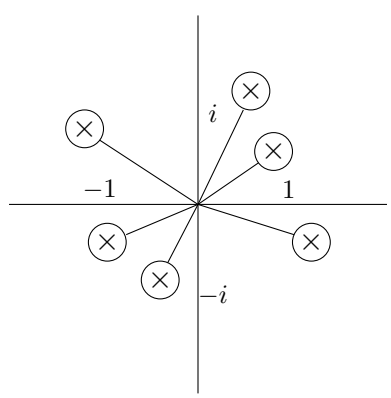
a path around one of these branching points. Consider for example, the function $w = z^2$. The point $w = 1$ has two distinct pre-images: $z = 1$ and $z = -1$. The point $w = 0$ has only one: $z = 0$. If we let w follow a path around this singularity (such as fixing $|w| = 1$ but allowing $\arg(w)$ to go from 0 to 2π, we see that the two preimages of $w = 1$ switch.



The path of w | The path of one of w's preimages | The path of w's other preimage

In order to calculate the monodromy group for a general matrix pencil $A + tB$, we must first choose a starting point for t which is not a branching point. In general, $t = 0$ is not a branching point, and the preimages of 0 are precisely the eigenvalues of A. From this point, we order our branching points and choose paths for t such that each path begins and ends at our starting point, goes around exactly one branching point (in order), and does not overlap any of the other paths. As mentioned, these branching points are all simple in general, so as t makes a path around one branching point and returns to 0, two of the eigenvalues of A will switch. We may think of this as a transposition in the symmetric group on n elements (S_n). We must order the preimages of our starting point (the eigenvalues of A) and note how these preimages are transposed as we follow each path. This gives us a word length $n(n-1)$ of transpositions in S_n. Since our path may be stretched to a circle encompasing all of our branching points, this word must sum to identity.

When A B in GE_n, the arguments of our branching points are uniformly distributed, so we may order our branching points by argument (though the choice of which branching point is first and whether the branching points are ordered clockwise or counterclockwise is arbitrary). The paths we choose will start and end at 0 and go around these branching points in a natural way. An example is shown below:

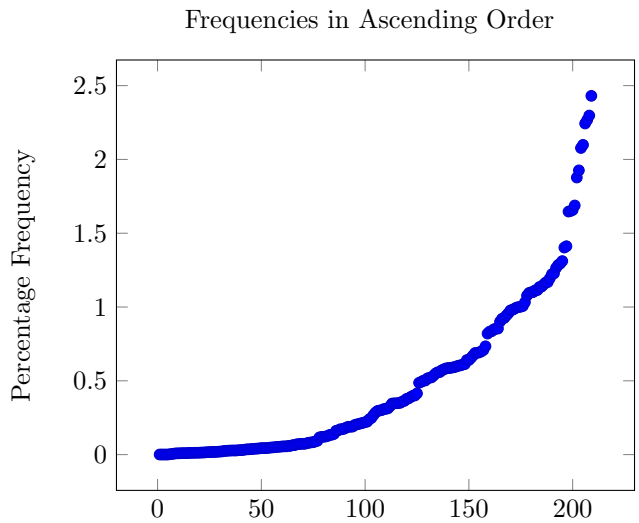


An example of the paths chosen for t for A, $B \in GE_3$

In the case where A, $B \in GE_2$, there is only one possible monodromy group, (1 2)(1 2). When A and B are in GE_3, this number increases to 240 possible cases (technically, 243 possible cases, but words such as (1 2)(1 2)(1 2)(1 2)(1 2)(1 2) appear with probability 0). For A, $B \in GE_4$, there are 181401594 possible cases and for A, $B \in GE_5$ we have 1220910766434156240 possibilities. We were unable to find a way to calculate this number for A, $B \in GE_n$ for a general n. Since there is only one possibility for the size 2 case and so many possibilities for the size 4 case, we figured that it is best to collect data for the size 3 case.

Due to the fact that we only have three preimages in general for each value of t in the size three case, we implemented a code which tracked the three preimages of our starting point as t followed a line to a branching point. The two eigenvalues which collided at this branching point are the two which would switch if we were to trace a path around the branching point. We generated 150000 random matrix pairs in GE_3 and calculated their monodromy groups. Of the 240 possible cases, only 209 were realized and only 204 were realized more

than once. The most common monodromy groups were those of the form (2 3)(1 2)(2 3)(1 2)(2 3)(1 2), which occured 2.43 percent of the time and those of the form (1 2)(1 3)(1 3)(2 3)(2 3)(1 2) which occured 2.29 percent of the time. Monodromy groups in which one permutation occurs four times in a row followed by two occurances of another permutation and the rotations thereof (for example, (1 2)(1 2)(1 2)(1 2)(1 3)(1 3) or (1 2)(2 3)(2 3)(2 3)(2 3)(1 2)) are the most rare, occuring only once or not at all. However the frequencies with which certain monodromy groups occur do not seem to have a discrete distribution, as can be seen in the graph below. This suggests that there is no "grouping" amongst these possibilites. These frequencies seem to be about the same regardless of the distribution chosen for our matrix pair, so long as the distribution is radially symmetric.



One particularly strange and interesting result is that the labeling of the eigenvalues seems to affect the frequencies with which certian monodromy groups appear. In the case of GE_3 matrices, we may relabel the three preimages of $t = 0$ by acting on the labels with members of the symmetric group on three elements. Usually, about half of these six group members change the frequency, generally by either doubling or halving the original frequency. The other half of the group tends to keep the frequency the same, but exactly which members of S_3 do what varies from case to case. We have not been able to find a pattern or an explanation as to why relabeling changes the frequencies in this way.

2. Joint Probability Distributions of Branching Points

Given a prescribed sequence of branching points $(t_1, \ldots, t_{n(n-1)})$, what is the probability that a random choice of matrix pencil in GE_n has those $n(n-1)$ branching points. In other words, we seek the Joint Probability Distribution of the branching points:

$$P[(t_1, \ldots, t_{n(n-1)}) \text{ are the branching points of } A + tB] = ???$$

We have began by trying to tackle this problem for the case of $n = 2$. Using techniques from probability theory, we can determine the joint probability for the branching points by using the joint probability distribution for the eigenvalues [which is known] and multiplying it with the conditional probability for the branching points given those eigenvalues.

For the following simplified pencil,

$$A + tB = \begin{pmatrix} 0 & 0 \\ 0 & \Delta \end{pmatrix} + t \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

where δ is a prescribed complex value, and the components of B are $N(0, 1)$. We shall try to explicitly calculate the formula for the joint probability distribution of the branching points t_1, t_2. First of all, this is equal to the polynomial

$$\begin{aligned} det(A + tB + \lambda I) &= (tb_{11} + \lambda)(\Delta + tb_{22} + \lambda) - t^2 b_{12} b_{21} \\ &= \lambda^2 + \lambda \underbrace{(\Delta + tb_{11} + tb_{22})}_{u(t)} + \underbrace{(\Delta t b_{11} + t^2(b_{11}b_{22} - b_{12}b_{21}))}_{v(t)} =: P(\lambda, t). \end{aligned}$$

When does this polynomial $P(\lambda, t)$ have repeated roots in λ? Via the quadratic formula, we know this is equivalent to specifying when the discriminant $Discr(t) = u(t)^2 - 4v(t)$ vanishes. In other words, for what $t \in \mathbb{C}$ do we have

$$u^2(t) - 4v(t) = t^2((b_{22} - b_{11})^2 - 2b_{11}b_{22}) + t(2\Delta(b_{22} - b_{11})) + \Delta^2 = 0 \quad ?$$

This is quite exciting! The roots of this polynomial in t will be the branching points of $A + tB$. Further, the roots of this discriminant now depends on its linear and second-order terms. Which are algebraic formulas in our random variables b_{rs}. And, the second order term depends to some extent on the linear term, so we can compute both with (relative?) ease.

The explicit formula for these roots is given by

$$\tau_{\pm} = \frac{\Delta[-(b_{22} - b_{11}) \pm i2\sqrt{b_{12}b_{21}})]}{(b_{22} - b_{11}^2 + 4b_{12}b_{21})}.$$

Considering one of the roots, and noting that we can write the denominator as a difference of squares, we have

$$\tau_{+} = \frac{\Delta}{(b_{22} - b_{11})^2 - i2\sqrt{b_{12}b_{21}}} = \frac{1}{X - iY} = \alpha + i\beta$$

where $\Delta X = (b_{22} - b_{11})^2$ and $\Delta Y = 2\sqrt{b_{12}b_{21}}$. Now we seek to determine the density functions associated to the random variables X and Y. X is the square of normally distributed random variables, and Y is the square root of a product of normally distributed random variables. These can both be evaluated via the method of transformations, allowing us to write X and Y (and thus τ) in terms of α and β.

The Joint Probability Distribution for the Eigenvalues for a 2×2 real matrix may be given explicitly. Since a 2×2 real matrix may have two real eigenvalues (λ_1, λ_2) or a complex conjugate eigenvalue pair $(z, \overline{z})$, the distribution may be is given by

$$P(\lambda_1, \lambda_2) = |(\lambda_2 - \lambda_1)||\lambda_1|^2|\lambda_2|^2 \frac{e^{-(|\lambda_1|^2 + |\lambda_2|^2)}}{4} + |z|^4 e^{2\mathrm{Re}(z^2)} \mathrm{erfc}(\sqrt{2}\mathrm{Im}(z))\mathrm{erf}(\sqrt{2}\mathrm{Im}(z))$$

where erf, and erfc denote the error function and complementary error function respectively.

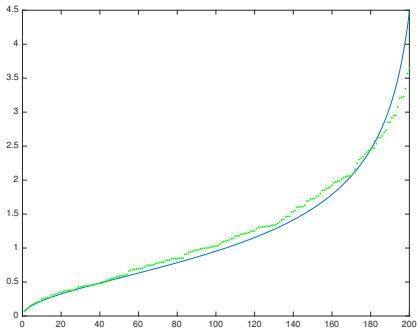
After several more calculations, we can find that the Joint Probability Distribution on the eigenvalues is given by

$$\int_0^\infty \frac{\exp\left[\frac{-\Delta\left|\frac{\alpha}{\alpha^2+\beta^2}\right|}{4}\right]\Delta^3 K_0\left(\frac{\Delta^2\beta^2}{16(\alpha^2+\beta^2)^2}\right)}{(\alpha^2+\beta^2)^{3/2}128\pi^{3/2}\sqrt{\Delta\alpha}} \int_{-\infty}^{\infty} |\Delta + \lambda_1|^2|\lambda_1|^2 e^{-|\lambda_1|^2} e^{-|\Delta+\lambda_1|^2} d\lambda_1 d\Delta$$
$$+ \int_{\mathbb{C}} \frac{\exp\left[\frac{-\mathrm{Im}(z)\left|\frac{\alpha}{\alpha^2+\beta^2}\right|}{2}\right]\mathrm{Im}(z)^2 K_0\left(\frac{\mathrm{Im}(z)^2\beta^2}{4(\alpha^2+\beta^2)^2}\right)}{(\alpha^2+\beta^2)^{3/2}8\pi^{3/2}\sqrt{2\mathrm{Im}(z)\alpha}} |z|^4 e^{2\mathrm{Re}(z^2)} \mathrm{erfc}(\sqrt{2}\mathrm{Im}(z))\mathrm{erf}(\sqrt{2}\mathrm{Im}(z)) dz.$$

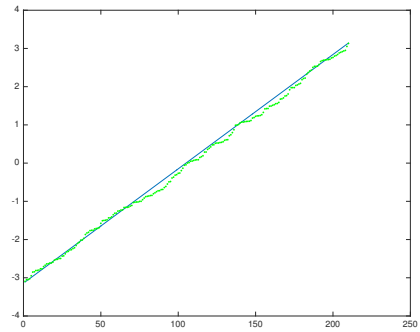
Where K_0 is Bessel's function of the second kind. We have little hope of determining the analogous distribution for the case of $n = 3$.

3. GRAPHICS

Here we consider three ensembles - GE, GUE and GOE. In all cases, we computed one trial of 15-by-15 matrices, twenty trials of 10-by-10 matrices and fifty trials of 5-by-5 matrices. We attached the diagrams of the empirical distribution of branching points below. For each case, we provide graphs of the radius and the angle distribution. The data points (colored green) are sorted from the smallest to the largest. The theoretical curves (colored blue) are obtained by inverting the distribution functions of the uniform distribution.

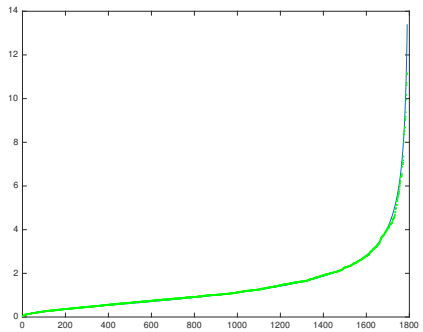


(A) radius

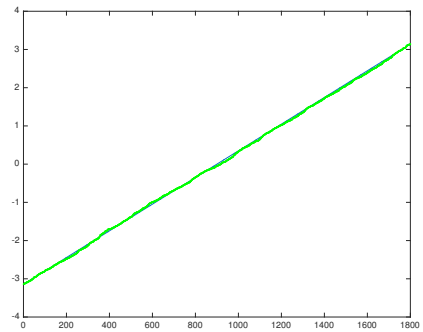


(B) angle

FIGURE 1. GE, $n = 15$, $t = 1$

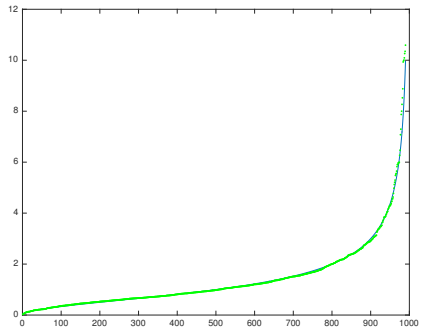


(A) radius

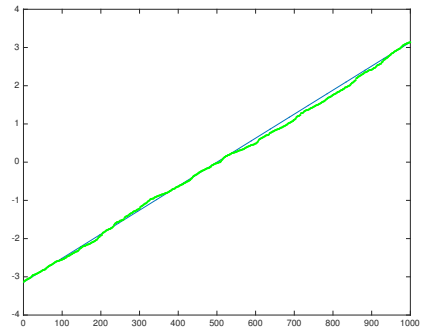


(B) angle

FIGURE 2. GE, $n = 10$, $t = 20$



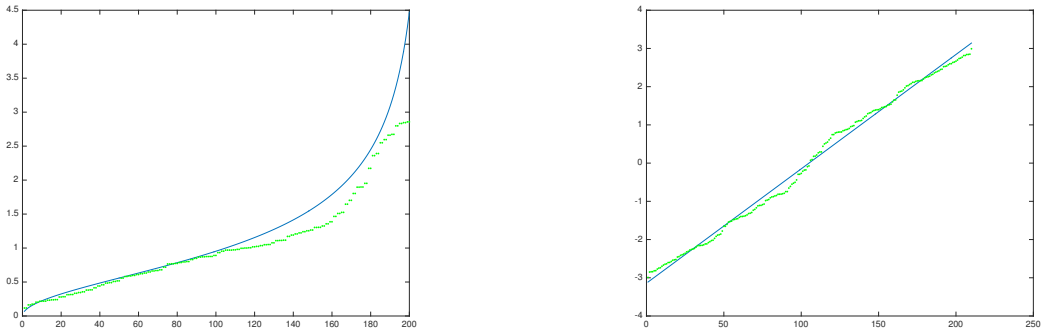
(A) radius



(B) angle

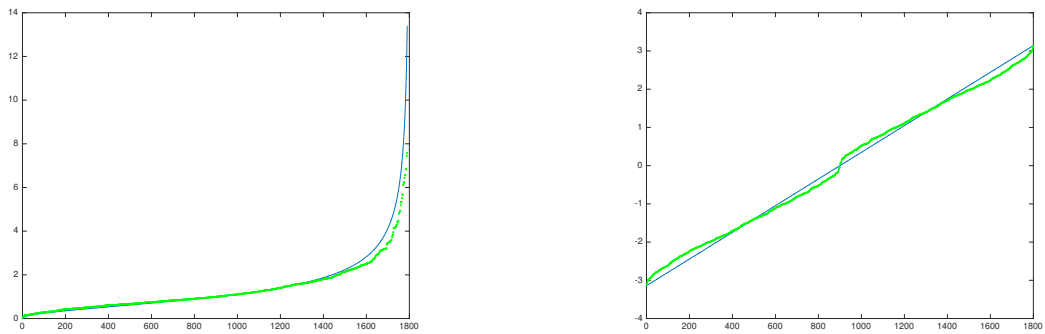
FIGURE 3. GE, $n = 5$, $t = 50$

In the GE case, we have shown that for each integer n, GE ensemble of size n has the uniform distribution on the Riemann Sphere as the expected value of the empirical distribution. Simulation-wise, this can also be seen from the diagrams of twenty trials of 10-by-10 matrices and fifty trials of 5-by-5 matrices. The empirical curves fit the theoretical curve closely. However, the one trial of 15-by-15 matrices show that the empirical distribution of one matrix of small size doesn't have to be the uniform distribution.



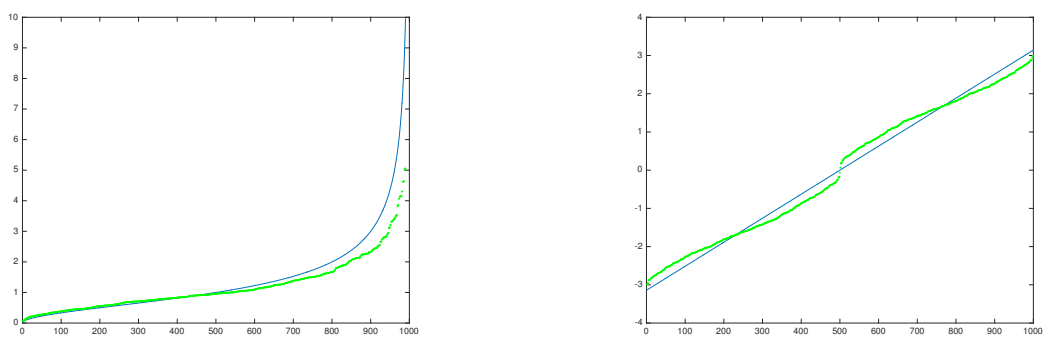
(A) radius (B) angle

FIGURE 4. GUE, $n = 15$, $t = 1$



(A) radius (B) angle

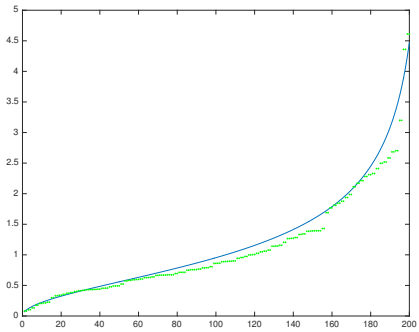
FIGURE 5. GUE, $n = 10$, $t = 20$



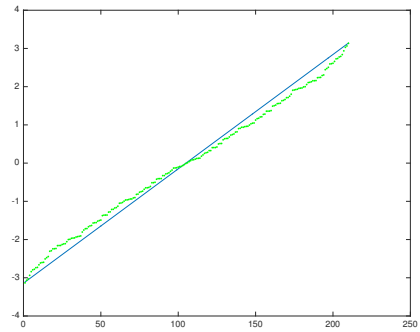
(A) radius (B) angle

FIGURE 6. GUE, $n = 5$, $t = 50$

In the GUE case, the matrices we tested do not satisfy the uniform distribution. We observe that the graphs for the 10-by-10 case are closer to the theoretical curves than the graphs for the 5-by-5 case. It is possible that in the limiting case as n goes to infinity, the empirical distribution is the uniform distribution.

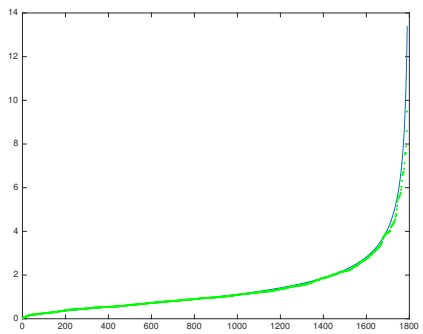


(A) radius

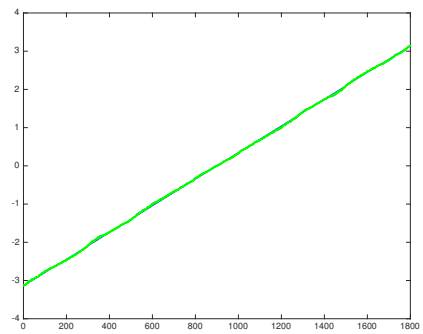


(B) angle

FIGURE 7. GOE, $n = 15$, $t = 1$

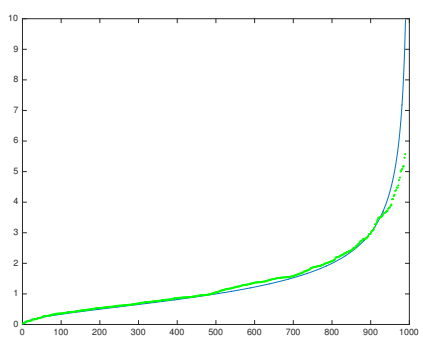


(A) radius

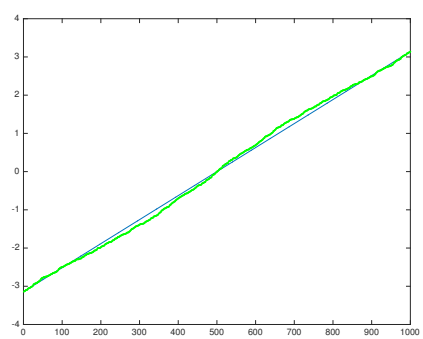


(B) angle

FIGURE 8. GOE, $n = 10$, $t = 20$



(A) radius



(B) angle

FIGURE 9. GOE, $n = 5$, $t = 50$

In the GOE case, we see that as matrix size grows from 5 to 10, the graphs approach the theoretical curves. This suggests that as n grows, the expectation of the empirical distribution could be the uniform distribution. However, the one trial of 15-by-15 matrices show that the empirical distribution of one matrix of small size doesn't have to be the uniform distribution.