Exploring Bohemian Matrices

Robert M. Corless
Mark W. Giesbrecht
rcorless@uwaterloo.ca
mwg@uwaterloo.ca
David R. Cheriton School of Computer Science
University of Waterloo
Canada

George Labahn
Leili Rafiee Sevyeri
glabahn@uwaterloo.ca
leili.rafiee.sevyeri@uwaterloo.ca
David R. Cheriton School of Computer Science
University of Waterloo
Canada

Dan Piponi
redacted@epicgames.com
http://blog.sigfpe.com/
USA

ABSTRACT

A Bohemian matrix family is a set of matrices all of whose entries are drawn from a fixed, usually discrete and hence bounded, subset of a field of characteristic zero. Originally these were integers—hence the name, from the acronym BOunded HEight Matrix of Integers (BOHEMI)—but other kinds of entries are also interesting. Some kinds of questions about Bohemian matrices can be answered by numerical computation, but sometimes exact computation is better. In this paper we explore some Bohemian families (symmetric, upper Hessenberg, or Toeplitz) computationally, and discover and prove interesting eigenvalue bounds. We also pose many open problems.

CCS CONCEPTS

• **Computing methodologies → Hybrid symbolic-numeric methods**; **Linear algebra algorithms**; **Symbolic calculus algorithms**.

KEYWORDS

Bohemian matrix, hybrid symbolic-numeric computation, height, characteristic height, upper Hessenberg, Toeplitz, circulant

ACM Reference Format:
Robert M. Corless, Mark W. Giesbrecht, George Labahn, Leili Rafiee Sevyeri, and Dan Piponi. 2022. Exploring Bohemian Matrices. In *Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation (ISSAC '22), July 4–7, 2022, Lille, France.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/xxxx.yyyy

1 HISTORY

The study of matrices with rational integer entries is very old, and the literature too vast to survey coherently here. We instead point to the early survey by Olga Taussky–Todd [36] as an entry point. We are also going to be working with Gaussian integer and algebraic integer entries; see for instance [9] for important work on generalized Hadamard matrices where the entries are roots of unity. The study of *random* matrices where the entries are drawn from discrete distributions is also very advanced; see [34, 35] for instance.

The name "Bohemian" was invented in 2015 at the Fields Institute Thematic Year in Symbolic Computation; the mnemonic is useful because it highlights searching for commonality among features of matrices with discrete populations. This is in essence a *specialization* in the sense of Pólya, and has led now to several workshops, at Manchester in 2018, at ICIAM in 2019, and at SIAM AN in 2021. There have been several publications since, including [14], [20], [19], [13], and the very interesting [4] which explores a connection to the asymptotic spectral theory of Toeplitz matrices [32], which is very much alive today: see e.g. [7] and [2].

There is a significant connection to number theoretic works, as well. Kurt Mahler [28] was interested in the distribution of zeros of polynomials with given *length* (the one-norm of the vector of coefficients) and *height* (the infinity-norm of the vector of coefficients). This is connected with the Littlewood conjecture for polynomials [26] (How large on the unit circle must a polynomial with −1,1 coefficients be?). The numerical visualization of zeros of 0−1 polynomials was apparently first done by Odlyzko and Poonen [29], who proved that the limiting set was connected; later work explained the "holes" [6] and visualizations by Peter Borwein and Loki Jörgenson made several other questions clearer [5]. See also the web pages of John Carlos Baez and of Dan Christensen. Their article *The Beauty of Roots*, published on Baez' website at the previous link, explains quite a few of the visible structures. Then we will see a connection to Kate Stange's work on Schmidt Tessellations https://math.katestange.net/illustration/schmidt-arrangements/. See also [22] and [18] who connect Galois theory and visualization of roots of polynomials (and therefore, although they do not point this out, of eigenvalues of Bohemian matrices).

It is only a small jump from polynomials of bounded height to *matrices* of bounded height; but the questions become (to our minds) even more interesting.

2 THE QUESTIONS WE ARE INTERESTED IN

Note first that because every polynomial written in the monomial basis can be embedded as a Frobenius companion matrix into a matrix of the same *height* (the height of a matrix **A**, as opposed to a polynomial, is the infinity norm of vec**A**; that is, of the matrix reshaped into a vector). Therefore every question about roots of polynomials of bounded height translates directly into a question about eigenvalues of Bohemian matrices. It will become clear as we go that this is one-way; there are indeed questions of Bohemian

ISSAC '22, July 4–7, 2022, Lille, France

ACM ISBN not-an-isbn-number…$15.00
https://doi.org/10.1145/xxxx.yyyy

matrices that do not translate into questions about bounded height polynomials.

The notion of companion matrix has been generalized in several ways, most notably to other polynomial bases; but now there is a new question: for a given integer polynomial, which companion Bohemian matrix has *minimal* height? See [11, 12]. At the present time, we do not have a good algorithm for finding such. It seems clear that lower height companion matrices are likely to be more numerically tractable. As an example, consider the Frobenius companion matrices for the Mandelbrot polynomials versus the Mandelbrot matrices [10]: the Mandelbrot polynomials defined by $p_{n+1}(z) = zp_n^2(z)+1$, $p_0(z) = 0$ have height exponential in the degree d_n, which is itself exponential in n; the Mandelbrot matrices in contrast have height 1, which must be minimal, trivially, because there are no nontrivial integer matrices with lower height. However, the task of *finding* a minimal height companion for a given polynomial is solvable (so far as we know) only by brute force. We do not know the complexity of this problem.

Other questions that could be asked include "Which matrix in the family has the maximum determinant?" This is clearly related to Hadamard matrix questions (a Hadamard matrix is a ± 1 matrix with maximal determinant, achieving the Hadamard bound) and if we allow other roots of unity into the matrix then this is the Butson–Hadamard matrix question [9]. There is a very large literature on this question, but the classification of all Hadamard matrices is still open, so far as we know.

A weaker question is "which matrices in the family have the largest *characteristic height*?" The characteristic height is the height of the characteristic polynomial; as previously noted, the characteristic height might be exponentially larger than the matrix height. This is so for certain upper Hessenberg Toeplitz matrices [14], where a lower bound containing a Fibonacci number is given for the maximum characteristic height in the family studied in that paper.

We remark that for populations of uniform absolute value (such as roots of unity) then there is a connection between *condition number* and determinant: In [21] we find the following theorem:

Theorem 2.1. *For a row-equilibrated square matrix* $\mathbf{M}$ *of dimension* m*, the condition number* $\kappa\mathbf{M} < 2/|\det \mathbf{M}|$ *and the constant* 2 *is the best possible.*

By *row-equilibrated* is meant that each row has been scaled (by multiplication by a diagonal matrix if necessary) so that the 2–norm of each row is 1. If, for instance, each entry of the matrix has absolute value B, then the 2-norm of each row of the matrix is $\sqrt{m}B$. Put $\mathbf{M} = \mathbf{A}/(B\sqrt{m})$ and so $\kappa M = \|\mathbf{M}\|\|\mathbf{M}^{-1}\| = \|\mathbf{A}\|\|\mathbf{A}^{-1}\| < 2/|\det \mathbf{M}| = 2B\sqrt{m}/|\det \mathbf{A}|$.

Thus, in contrast to the usual case where the size of the determinant is quite decoupled from the conditioning, we see that for some classes of Bohemian matrices the size of the determinant actually tells us something. [The connection can be weak; this upper bound can be quite pessimistic. Nonetheless it is a useful bound.]

The characteristic height is connected to the numerical conditioning of the characteristic polynomial; indeed one may take the Lebesgue constant for the polynomial [17, ch. 8] on the interval $-1 \le x \le 1$ to be the characteristic height.

One can ask about the distribution of condition numbers in a family: how likely is a matrix picked "at random" from the family to be ill-conditioned? One can ask about the distribution of *eigenvalue* condition numbers [1]. See also [31]. How many matrices in the family are singular? How many have multiple eigenvalues? How many different eigenvalues are there? How many matrices have nontrivial Jordan structure? How many matrices have nontrivial integer Smith form? How many different characteristic polynomials are there in the family? How many different minimal polynomials are there? How many non-normal matrices are there? What is the typical departure from normality? How many orthogonal matrices are there? How many pairs of commuting matrices are there? What is the distribution of eigenvalues? We will typically be interested in complex eigenvalues, but the distribution of real eigenvalues is a classical topic in random matrix theory.

There are an infinite number of further questions that can be asked. We can then ask them again about block matrices; about matrix polynomials; and about commuting matrix families for multivariate polynomial equations.

Typically, one asks these questions in an asymptotic form; one wants answers valid in the limit as the dimension m goes to infinity. Examples of this include work by Tao and Vu, who show that for general unstructured matrices the distribution of eigenvalues divided by $\sqrt{m}$ is asymptotically uniform on the unit disk [34, 35]. This is *not* true for structured matrices; see e.g. [15, 16] where skew-symmetric tridiagonal matrices (independently of dimension, so no scaling is needed) are seen to be confined to a diamond shape. We explain that mystery in this paper, using a century-old theorem, which deserves to be better-known.

We largely back off from the asymptotic aspect, here, and try to answer questions about these families for "small" to "moderate" dimensions. For some populations, computation can go quite far, if one is careful to account for symmetries and be efficient [38]; here we concentrate on using "off-the-shelf" technology to explore questions as far as we can.

3 THE EFFECT OF MATRIX STRUCTURE

The first results were about general (dense, full, unstructured) matrices, including [34]. Or, rather, the first results were on real symmetric matrices or Hermitian matrices [37]. We look at other structures here, in order to get another view. We begin with complex symmetric matrices.

3.1 Complex Symmetric Matrices

A complex symmetric matrix $\boldsymbol{A}$ satisfies $\boldsymbol{A}^T = \boldsymbol{A}$, where ^{T} is the real transpose operation. These occur, for instance in Bézout matrices for polynomials with complex coefficients. Unlike Hermitian matrices, the eigenvalues of complex symmetric matrices need not be real. Indeed, any matrix may be brought by similarity transformation to a complex symmetric matrix [25, Thm 4.4.9]. In many cases this can be done by unitary similarity; see for instance the characterizations of when this can be done, in [27].

Here let us examine a specific complex symmetric family, with population $-1 \pm i$ where $i = (0, 1)$ is the square root of -1. At dimension m, such a matrix has $m(m+1)/2$ free entries, each of which can be either of $-1 \pm i$. This gives $2^{m(m+1)/2}$ such matrices; this growth is (much) faster than exponential. Still, examining eigenvalues of small dimension examples can tell us much. If we take $m = 6$, then

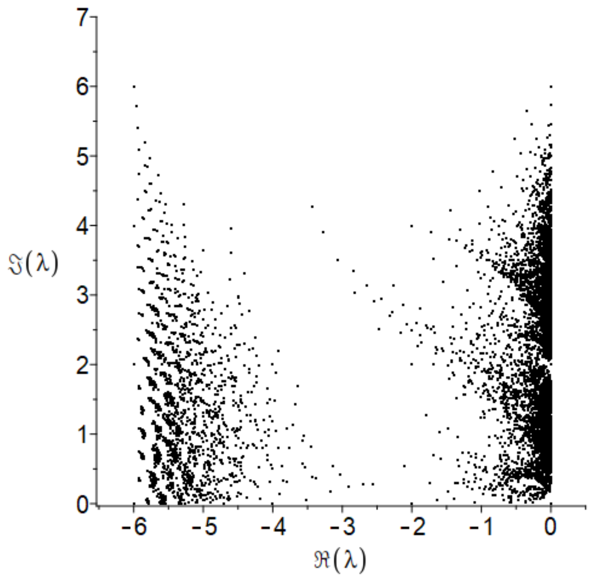

Figure 1: All eigenvalues with $\Im\lambda \geq 0$ of symmetric dimension $m = 6$ matrices with entries $-1 \pm i$. The set is symmetric about the real axis. We see the eigenvalues apparently confined to the strip $-6 \leq \Re\lambda \leq 0$, and bounded below and above by $-6 \leq \Im\lambda \leq 6$. There are several other unexplained features of this eigenvalue distribution.

the number of such matrices is only 2^{21}, slightly more than 2 million. The eigenvalues of all these matrices can be computed in a reasonable time, and plotted. As depicted in Figure 1, they seem confined to a strip in the left-half plane.

We now prove that this will always be true at any dimension.

Theorem 3.1. *If the symmetric matrix* $\boldsymbol{A}$ *has dimension m, entries drawn from* $-1 \pm i$*, and eigenvalue* λ*, then* $-m \leq \Re\lambda \leq 0$ *and* $-m \leq \Im\lambda \leq m$*.*

Proof. Write $\boldsymbol{A} = -\boldsymbol{E} + i\boldsymbol{M}$ where $\boldsymbol{E} = \mathbf{e}\mathbf{e}^T$ is the rank-one matrix that has all 1s, and $\boldsymbol{M}$ is symmetric and has entries only ± 1. We use the Bendixon–Bromwich–Hirsch theorem [24, Fact 5, p. 16-2] (original references [3, 8, 23]) as follows. Note that $(\boldsymbol{A} + \boldsymbol{A}^*)/2 = -\boldsymbol{E}$ is the Hermitian part of $\boldsymbol{A}$, while $(\boldsymbol{A} - \boldsymbol{A}^*)/(2i) = \boldsymbol{M}$ is the skew-Hermitian part. The Bendixon–Bromwich–Hirsch theorem says that if the eigenvalues of $-\boldsymbol{E}$ are written $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_m$ (they are all real because $\boldsymbol{E}$ is Hermitian) and the eigenvalues of $\boldsymbol{M}$ are written $\nu_1 \geq \nu_2 \geq \ldots \geq \nu_m$, then $\mu_m \leq \Re\lambda \leq \mu_1$ and $\nu_m \leq \Im\lambda \leq \nu_1$. But a short computation shows that the eigenvalues of $-\boldsymbol{E}$ are $-m$ and 0 with multiplicity $m-1$, and the Gerschgorin disk theorem shows that the eigenvalues of $\boldsymbol{M}$ lie in the union of circles centred at 1 of radius $m-1$ and centred at -1 of radius $m-1$. This establishes the theorem. □

We now use this method to explain why skew-symmetric tridiagonal matrices with population 1 and i (the population considered in [16] and [15]) are confined to a diamond shape.

3.1.1 Squares, Diamonds, and Lozenges. Take skew-symmetric tridiagonal matrices with population $-1 \pm i$. Then a matrix $\boldsymbol{A}$ from this family can be written as $\boldsymbol{A} = \boldsymbol{S} + i\boldsymbol{T}$ where the superdiagonal of $\boldsymbol{S}$ is -1 and the superdiagonal of $\boldsymbol{T}$ is $\pm$. Both are skew-symmetric: $\boldsymbol{S}^T = -\boldsymbol{S}$ and $\boldsymbol{T}^T = -\boldsymbol{T}$. The Hermitian part of $\boldsymbol{A}$ is $(\boldsymbol{A} + \boldsymbol{A}^H)/2 = i\boldsymbol{T}$ and the skew-Hermitian part is $(\boldsymbol{A} - \boldsymbol{A}^H)/(2i) = -i\boldsymbol{S}$. [Both are Hermitian, and have real eigenvalues.] Application of the Gerschgorin circle theorem to each of these shows that the eigenvalues of either matrix are confined to the interval $-2 \leq \mu \leq 2$. By the theorem of Bendixon–Bromwich–Hirsch cited earlier, the eigenvalues of $\boldsymbol{A}$ are confined to the *square* $-2 \leq \Re\lambda \leq 2$, $-2 \leq \Im\lambda \leq 2$.

To show that skew-symmetric tridiagonal matrices with population 1 and i are confined to a diamond $|\Re\lambda| + |\Im\lambda| \leq \sqrt{2}$, multiply the matrix by $-1 + i$, which rotates its eigenvalues by $\pi/4$ and stretches them by $\sqrt{2}$; but now the matrix population is $-1 \pm i$ and it's still skew-symmetric, and hence confined to the square as described above. Rotate the square back by $\pi/4$ and shrink by $\sqrt{2}$, and the result follows. This was the curious fact, remarked on in [15], which needed to be explained.

3.1.2 Characteristic compression. Another interesting fact that arises in these computations is that there is significant *compression* for this family in computing characteristic polynomials: at dimension 2 there are 8 different matrices possible, but only 6 different characteristic polynomials; at dimension 3 there are 20 different characteristic polynomials shared amongst 64 different matrices; at dimension 4 there are 90 different characteristic polynomials shared amongst 1024 different matrices; at dimension 5 there are 538 different characteristic polynomials shared amongst 32,768 matrices. The largest exhaustive computation we did with this family was for $m = 6$, where there were only 4,970 characteristic polynomials shared among 2,097,152 matrices. Clearly the growth rate of the number of polynomials is significantly slower than the growth rate of the number of matrices. We do not have any method for computing all the characteristic polynomials of symmetric matrices without first computing all the different matrices. Such would be of clear value for this family.

3.2 Upper Hessenberg Matrices

An upper Hessenberg matrix is a matrix of the form

$$\begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} & \cdots & h_{1,m} \\ h_{2,1} & h_{2,2} & \ddots & & \vdots \\ 0 & h_{3,2} & h_{3,3} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{m,m-1} & u_{m,m} \end{bmatrix}. \tag{3.1}$$

That is, it is zero below the first subdiagonal. If any entry of the first subdiagonal is zero, then the matrix is said to be reducible, because the matrix then separates into blocks containing distinct eigenvalues; we restrict our attention to irreducible matrices and indeed we specify that all the subdiagonal entries $h_{j+1,j} = -1$. If the subdiagonal entries all have $|h_{j,j+1}| = 1$ we say that the matrix is *unit* upper Hessenberg. If all the diagonal entries are zero, we say that it is Zero Diagonal.

Remark 1. *Suppose that all $h_{i,j}$ are roots of unity. Consider the 5 by 5 case, for definiteness. Take four roots of unity, as yet unspecified: call them s_2, s_3, s_4, s_5. Form the diagonal matrix $\boldsymbol{D} = \mathrm{diag}(1, s_2, s_3, s_4, s_5)$ and perform the similarity transform $\boldsymbol{DHD}^{-1}$. The result is*

$$\begin{bmatrix} h_{1,1} & \frac{h_{1,2}}{s_2} & \frac{h_{1,3}}{s_3} & \frac{h_{1,4}}{s_4} & \frac{h_{1,5}}{s_5} \\ s_2 h_{2,1} & h_{2,2} & \frac{s_2 h_{2,3}}{s_3} & \frac{s_2 h_{2,4}}{s_4} & \frac{s_2 h_{2,5}}{s_5} \\ 0 & \frac{s_3 h_{3,2}}{s_2} & h_{3,3} & \frac{s_3 h_{3,4}}{s_4} & \frac{s_3 h_{3,5}}{s_5} \\ 0 & 0 & \frac{s_4 h_{4,3}}{s_3} & h_{4,4} & \frac{s_4 h_{4,5}}{s_5} \\ 0 & 0 & 0 & \frac{s_5 h_{5,4}}{s_4} & h_{5,5} \end{bmatrix} . \tag{3.2}$$

Now choose $s_2 = \overline{h_{2,1}}$, $s_3 = s_2\overline{h_{3,2}}$, $s_4 = s_3\overline{h_{4,3}}$, and $s_5 = s_4\overline{h_{5,4}}$. This forces the subdiagonal entries to be 1, leaves the diagonal as it was before, and shuffles the upper triangle to be possibly different roots of unity to what they were before. Since the elements of the upper triangle were independent, the resulting formulae will range over the entire set of possibilities as we make the $h_{i,j}$ range over the entire set of possibilities. It is for this specific population that we started studying unit upper Hessenberg matrices.

Many populations have zero as their mean value (e.g. $\{-1, 1\}$ or indeed roots of unity; or $\{-1, 0, 1\}$). In that case, the eigenvalues are typically symmetric about zero and a simplified picture is obtained simply by setting all the diagonal entries to zero, in which case the Gerschgorin circles are all centred at 0. Sometimes we will have to transform back to the nonzero diagonal case, but a surprising amount of information is retained even with the simplification of insisting on a zero diagonal.

Upper Hessenberg matrices occur in the numerical solution of eigenvalue problems via QR iteration as an intermediate step: every matrix $\mathbf{A}$ is unitarily similar to an upper Hessenberg matrix, and this conversion can occur in $O(n^3)$ operations. Subsequent QR iterations are then much cheaper. These matrices also occur in other contexts (and indeed have been known to be reinvented). For us, they become an object of study in and of themselves; admittedly, we remain chiefly interested in eigenvalues.

We will use the following lemma repeatedly, to ensure that we see all the eigenvalues of the family in the window of the plot.

Lemma 3.2. *Suppose that every entry of a unit upper Hessenberg zero diagonal matrix $\mathbf{H}$ has magnitude at most B: that is, $|h_{i,j}| \le B$. Then every eigenvalue λ of $\mathbf{H}$ has*

$$|\lambda| \le 1 + 2\sqrt{B} \, . \tag{3.3}$$

This bound is independent of the dimension. The proof uses an idea already present in [32], and doubtless in other places, namely to choose a diagonal matrix $\boldsymbol{D} = \mathrm{diag}(1, r, r^2, \ldots, r^{m-1})$ with a free parameter $r > 1$ and consider the similar matrix $\boldsymbol{DHD}^{-1}$ which has the same eigenvalues as $\boldsymbol{H}$. Suppose without loss of generality that the subdiagonal entries of the zero diagonal unit upper Hesssenberg $\boldsymbol{H}$ are all -1. Then

$$\boldsymbol{DHD}^{-1} = \begin{bmatrix} 0 & h_{1,2}/r & h_{1,3}/r^2 & \cdots & h_{1,m}/r^{m-1} \\ -r & 0 & h_{2,3}/r & & h_{2,m}/r^{m-2} \\ 0 & -r & 0 & h_{3,4}/r & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -r & 0 \end{bmatrix} . \tag{3.4}$$

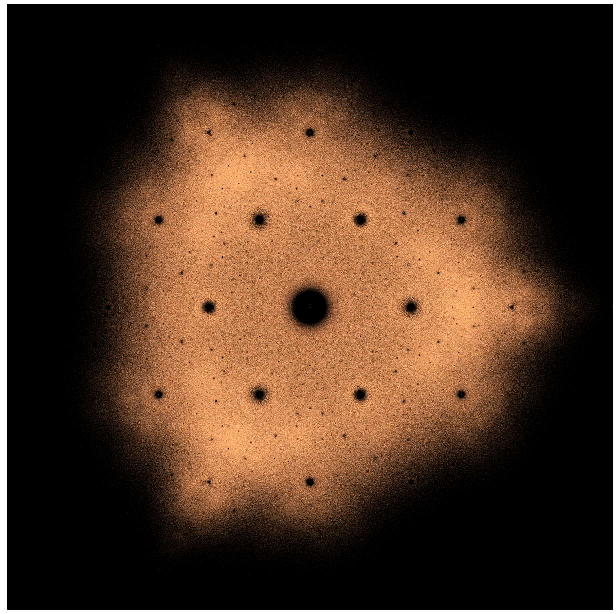

Figure 2: Eigenvalues of a sample of 5 million upper Hessenberg matrices of dimension $m = 5$ matrices with population cube roots of unity. The image is visually indistinguishable from the density plot of all 14, 348, 907 unit upper Hessenberg zero diagonal matrices of dimension $m = 6$.

The Gerschgorin disk for the first row has radius at most $B/(r-1)$ by comparison with a geometric series; the second and all subsequent rows have the bound $r + B/(r-1)$ for the radius which is larger because $r > 1$. To minimize this bound, we write it as $1 + r - 1 + B/(r-1)$ and use the AGM inequality to say that this is minimized when $r - 1 = B/(r-1)$ or $r = 1 + \sqrt{B}$; this gives the value of the Gershgorin radius as $1 + 2\sqrt{B}$, as desired.

Remark 2. *This is the only Gerschgorin-like theorem that we are aware of that gives a bound on eigenvalues that is independent of the dimension and depends on the square root of the bound for the entries in the matrix instead of the more usual linear power of the bound. Of course, if one multiplies a matrix $\boldsymbol{A}$ by a constant factor, then the eigenvalues must also be multiplied by that factor; but we cannot perform such a multiplication here and remain in the class of unit upper Hessenberg matrices.*

3.3 Upper Hessenberg Toeplitz Matrices

A Toeplitz matrix $\boldsymbol{T}$ has constant elements on every diagonal, $t_{i,j} = t_{0,j-i}$. The authors of [14] found that the Toeplitz subset of unit upper Hessenberg matrices maximized the *characteristic height*, and so decided to study that subset directly; it contains only exponentially many elements ($\#P^{m-1}$ entries, rather than $\#P^{O(m^2)}$) and has several other interesting features. For large dimension, the spectral theory connects to the well-known asymptotic spectral theory for Toeplitz matrices, although because the population does not decay the connection is not as straightforward as one would like: see [18] for a careful treatment.

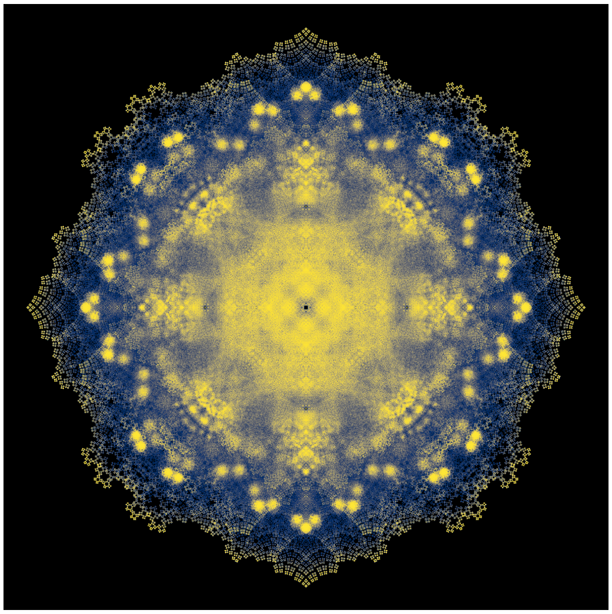

Figure 3: Density plot of eigenvalues of all $1,048,576$ **upper Hessenberg Toeplitz matrices of dimension** 11 **with zero diagona,** -1 **subdiagonal, and population** ± 1, $\pm i$ **(fourth roots of unity) otherwise. Brighter colours correspond to higher density.**

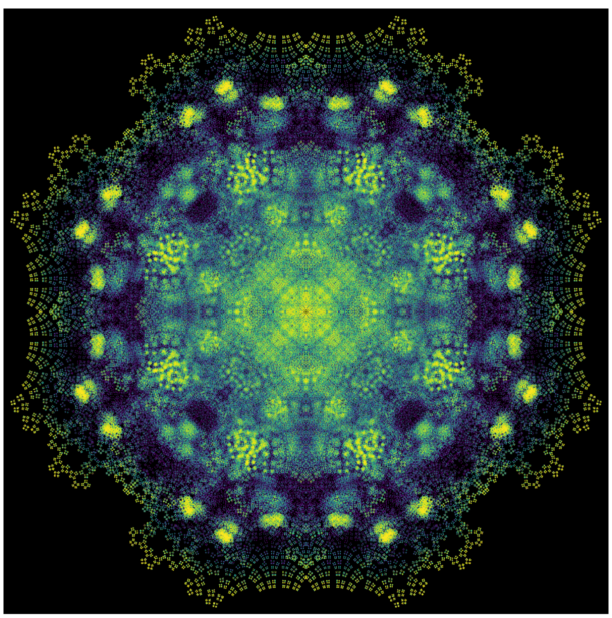

Figure 4: Density plot of eigenvalues of all $262,144$ **upper Hessenberg Toeplitz matrices of dimension** 10 **with zero diagonal,** -1 **subdiagonal, and population** $\pm 1 \pm i$ **(four corners of a square) otherwise. Brighter colours correspond to higher density.**

4 THE FRACTAL EDGES

At the edges of Figures 3 and 4 we see clear indication of fractal gasket-like structures. When the population is third roots of unity, we see Sierpinski gaskets; when the population only has two elements, we see pairs, and pairs of pairs, recursively; with five-element populations we see recursive pentagonal structures. The phenomenon seems universal for upper Hessenberg Toeplitz matrices. We now give an explanation for this behaviour in this section.

The key to understanding this is to consider the characteristic polynomials. For this structure, there is a recurrence relation for the characteristic polynomial, say $Q_n(z; t_1, t_2, \ldots, t_{n-1})$, of a unit upper Hessenberg zero diagonal matrix, namely

$$Q_{n+1}(z; t_1, t_2, \ldots, t_n) = zQ_n(z; t_1, t_2, \ldots, t_{n-1}) - \sum_{k=1}^{n} (-1)^k t_k Q_{n-k}(z; t_1, t_2, \ldots, t_{n-k-1}) \,. \tag{4.5}$$

The final term is $\pm t_n Q_0(z)$ and $Q_0(z) = 1$. There is a somewhat more complicated recurrence relation for a general upper Hessenberg matrix; see [14].

We make an observation: for all of the populations where we have seen this gasket-like behaviour, the arithmetic mean of the population is zero: $\sum_{p \in P} t_i = 0$. For definiteness, let us fix the population to be third roots of unity; the argument will be similar for all such zero-mean populations.

For each of the three choices of t_n, the resulting characteristic polynomial $Q_{n+1}(z)$ can be written as a *fixed* polynomial $zQ_n(z) - \sum_{k=1}^{n-1} (-1)^k t_k Q_{n-k}$ plus t_n. This final term perturbs that fixed polynomial in one of three ways. If we use a homotopy argument, replacing t_n by st_n where $0 \le s \le 1$, we see the roots of Q_{n+1} arising by paths emanating in three directions from the roots of that fixed polynomial.

That fixed polynomial can be viewed as a branching perturbation from z^m, where first t_1 is varied from 0 to its ultimate choice; then t_2, and so on.

This recursive construction is not a linear one: the structures resembling Sierpinski gaskets seen in the close-up in Figure 5 are clearly not rigid triangles, but rather have been distorted into curved shapes. Nonetheless we believe the above explanation is one way of understanding why this structure arises.

5 RAYLEIGH QUOTIENTS

Recall the Rayleigh Quotient:

$$r = \frac{\mathbf{y}^T \boldsymbol{A} \mathbf{x}}{\mathbf{y}^T \mathbf{x}} \,. \tag{5.6}$$

If $\mathbf{x}$ is an approximate eigenvector, and $\mathbf{y}^T$ an approximate left eigenvector, then this quotient is a least-squares approximation to an eigenvalue of $\boldsymbol{A}$. If we replace $\boldsymbol{A}$ by $\boldsymbol{A}^{-1}$ above, then this is an approximation to an eigenvalue of $\boldsymbol{A}^{-1}$, and typically the largest one; this of course is the reciprocal of the smallest eigenvalue of $\boldsymbol{A}$. We will consider this not as an eigenvalue approximation, but as a process in its own right, and plot the results of a single iteration of this on a Bohemian family, with both $\mathbf{y}$ and $\mathbf{x}$ taken to be the first elementary vectors; so the result is the top left corner of the inverse

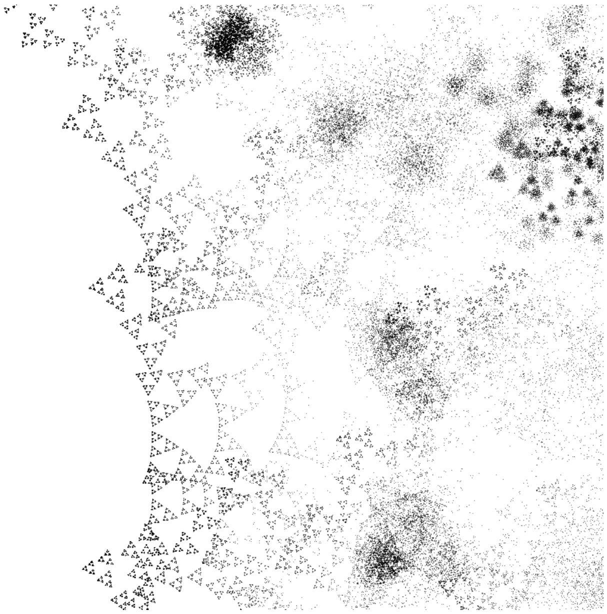

Figure 5: A close-up (window $-2.5 \le \Re\lambda \le -1.5$, $0 \le \Im\lambda \le 1$) of an 800 by 800 density plot of the eigenvalues of all $531,441$ upper Hessenberg zero diagonal matrices with population cube roots of unity. The resemblance to a Sierpinski gasket is striking.

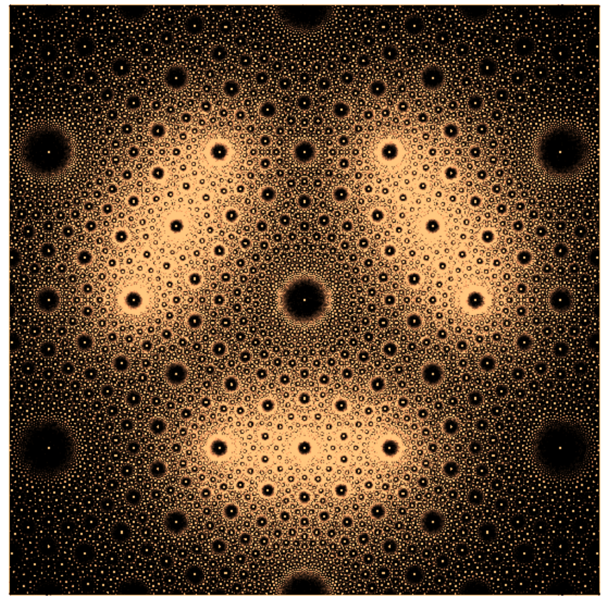

Figure 6: Density plot of the upper left corner element of A^{-1} where A is sampled randomly from the 8 by 8 upper Hessenberg Bohemian (*not* unit or zero diagonal) family with population cube roots of unity. There are $3^{43} > 3.2 \cdot 10^{20}$ such matrices; we sampled only $5 \cdot 10^5$ of these and gave a density plot on a 2048 by 2048 grid, enhanced by an adjoint-based filter.

of our Bohemian matrix. See figure 6. This can also be computed by the recurrence relation (4.5), as the ratio of two determinants, from Cramer's rule:

$$r = \frac{Q_m(0; t_1, t_2, \ldots, t_{m-1})}{Q_{m-1}(0; t_1, t_2, \ldots, t_{m-2})} \tag{5.7}$$

Perhaps surprisingly to a numerical analyst, this offers an effective way to perform this computation when the population consists of small Gaussian integers, which can be represented as complex "flints" and are not subject to rounding error when ring arithmetic is carried out in floats; the only division occurs at the end, and so rounding error is trivial. Even with roots of unity, the rounding errors are generally not of serious consequence.

What we are computing here is representable in (if cube roots of unity are used) $\mathbb{Q}(\sqrt{-3})$, and moreover the rational numbers involved will not have overly large values. This suggests that the appearance in Figure 6 can be explained as Schmidt arrangements, as done by Katherine Stange. See her blog at https://math.katestange.net/, but note in particular [33] and her papers previously cited.

6 ON VISUALIZATION

This raises the question of the visualization techniques being used. Most of the figures in this present paper use only the simplest techniques: Figure 1 is a simple plot. Figure 5 is a greyscale density plot on an 800 by 800 grid. Figures 3 and 4 use a *colourized* density plot, where the colouring scheme was chosen by using a cumulative frequency count in order to attempt to equalize the *density* of eigenvalues using colour; this verges on true computer imaging techniques but is actually very crude. The technique has some value because it is relatively faithful to the underlying mathematics: brighter colours correspond to higher eigenvalue density, and when the cividis or viridis colour palette is used, the colours are relatively even perceptually. The copper palette of Figure 2 retains the correlation of brightness to density, but has a smaller colour range.

The appearance of Figure 6, however, depends on some rather more professional techniques, as described in [30]. The basic idea is to estimate spatial derivative information (using TensorFlow Gradient) and use that to enhance the figure, making the density visible even with relatively sparse data (for this figure, only 500, 000 matrices were used, and the computational cost was substantially lower than for the other figures). The picture remains faithful to the underlying mathematics, however.

7 CONCLUDING REMARKS

The notion of a *Bohemian matrix* seems to be a remarkably productive one, with substantial connections to very active areas of research, including visualizations in number theory, combinatorial design, random matrices in physics, numerical analysis, and computer algebra. In an earlier section, we asked several questions (most of which are unanswered in general) and claimed that there were infinitely many more. We will talk here about which of these we think is worth tackling next.

First, many of the combinatorial questions, such as "how many different characteristic polynomials are there" for a given Bohemian family (say unit upper Hessenberg with population $(-1, 0, 1)$ for concreteness), need to be answered by theoretical advances or inspired guessing; brute algebraic computing might lead to answers, but insight seems needed. One is forced to look at polynomials (and thus computer algebra, however implemented) because of the *multiple-eigenvalue* problem: in those circumstances, numerical computation of eigenvalues is ill-conditioned and one cannot really count things by "clustering" nearby eigenvalues. One is often tempted, when thinking of random matrices, to say that "multiple eigenvalues never happen" but of course this is not true. Supplying constraints (either on the population or the matrix structure) significantly enhances the probability that multiplicity will be encountered.

One topic some of us looked at briefly before was the concept of *stable* matrices. Which Bohemian matrices have all their eigenvalues strictly in the left half-plane? For those matrices $\boldsymbol{A}$, and those matrices only, the solutions to the linear differential equation $\dot{y} = \boldsymbol{A}y$ will ultimately decay to zero. If the dimensions are large, then one may have to consider pseudospectra (and thus matrix non-normality) as well. A similar question asks which Bohemian matrices have all their eigenvalues inside the unit circle? For those matrices, the solution to the linear difference equation $y_{n+1} = \boldsymbol{A}y_n$ will ultimately decay. If the matrix is non-normal, again pseudospectra will play a role.

It can be unsatisfactory to compute the eigenvalues of a matrix $\boldsymbol{A}$ and check to see if they are all in the left half plane; rounding errors may drift some of them into the right half plane. Computation of the characteristic polynomial, and subsequent use of the Routh–Hurwitz criterion, seems in order. One would like to take advantage of the compression seen earlier: instead of computing eigenvalues of several million matrices, instead compute the roots of the (equivalent) several thousand characteristic polynomials; or, better yet, apply the Routh–Hurwitz criterion, which is a rational criterion, to make the decision in an arena uncontaminated by rounding errors.

As an example, consider the symmetric matrices with population $-1 \pm i$ of dimension $m = 6$. We already know from Theorem 3.1 that all eigenvalues lie in $\Re\lambda \le 0$. But are there any of the $4,970$ characteristic polynomials of these 2^{21} matrices which have *all* of their roots in the left half plane? Yes. By applying Maple's Hurwitz tool to the characteristic polynomials (which were actually computed using Python and exported in a JSON container to Maple) we identified 1328 of these polynomials, all of whose roots were strictly in $\Re\lambda < 0$. Indeed, the maximum real part was approximately $-1.03 \cdot 10^{-5}$. Corresponding to these 1328 polynomials were $966,240$ matrices, or about 46% of the total.

For upper Hessenberg matrices with population $(-1 - i, -1, -1 + i)$ and dimension $m = 4$, out of $1,594,323$ matrices we find $365,307$ distinct characteristic polynomials. Of these, only $14,604$ (associated with $66,782$ matrices, about 4.2% of the total number of matrices) have all their roots in the left half plane. The maximum real part is about $-1/14,000$. This is enough to give a respectable picture even without enhancements. See Figure 7. For comparison, we plot the eigenvalues of the whole familiy in Figure **??**. One sees immediately that the majority of matrices in this collection have eigenvalues in the right half-plane.

We now leave you with more questions. How many *unimodular* matrices are in any given family? How many matrices have inverses

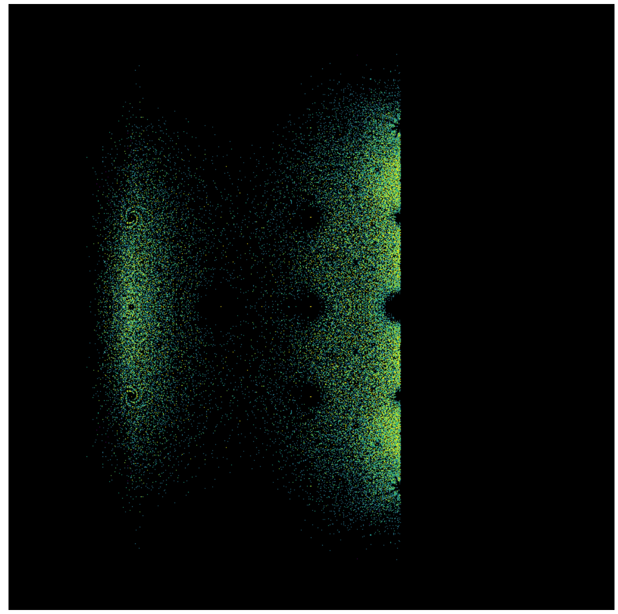

Figure 7: An 800 by 800 grid density plot of the roots of all $14,604$ stable characteristic polynomials, weighted by the number of matrices out of the possible $66,782$ with that polynomial. The Bohemian family is upper Hessenberg, population $-1 - i$, -1, $-1 + i$, and dimension $m = 4$. The plot is on $-L - 1 \le \Re\lambda < L - 1$, $-L \le \Im\lambda \le L$, where $L = 1 + 2 \cdot 2^{1/4}$.

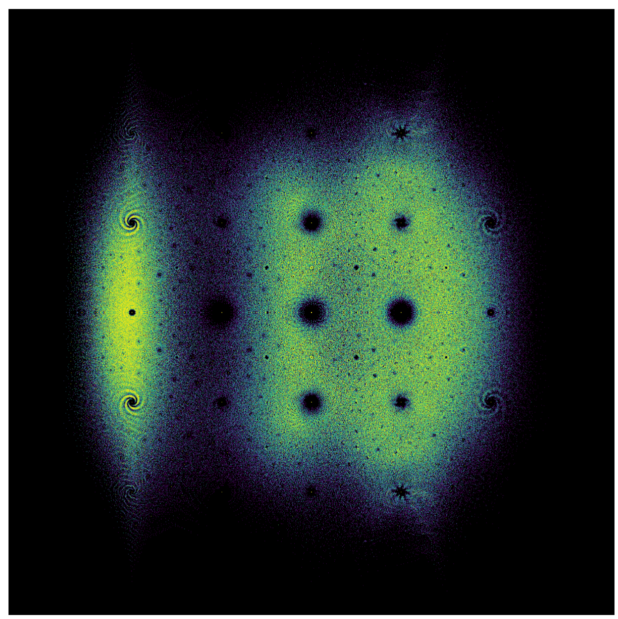

Figure 8: A 1200 by 1200 grid density plot of the eigenvalues of the whole family, to compare with the plot of Figure 7. The Bohemian family is upper Hessenberg, population $-1 - i$, -1, $-1 + i$, and dimension $m = 4$. The plot is on $-L - 1 \le \Re\lambda < L - 1$, $-L \le \Im\lambda \le L$, where $L = 1 + 2 \cdot 2^{1/4}$. The spirals are completely unexplained.

that are also Bohemian (in the same family?) [In that case, we say that the matrix has *rhapsody*. Since number theory is the Queen of mathematics, this joke will never not be funny.] Can we solve the *inverse* eigenvalue problem over a given Bohemian family? That is, given an eigenvalue, can we decide if there is a Bohemian matrix with that eigenvalue? This would help us to compute *minimal height companion matrices*, for instance. What questions would you like to ask, of a Bohemian family?

ACKNOWLEDGMENTS

The Python code we used for many of the experiments was written largely by Eunice Y. S. Chan (with some help by Rob Corless) for another project; we thank her for her work on that code. We also thank her for many discussions on Bohemian matrices and colouring algorithms. We thank Neil J. Calkin for valuable discussions about finding good questions to direct research. We thank Ilias Kotsireas for pointing out several references relevant to the Hadamard matrix literature, and for helpful discussions. We thank Nick Higham for an independent proof of Theorem 3.1, and for searching out the original references to Hirsch, Bromwich, and Bendixon. We also thank Owen Maresh (@graveolens) for pointing out the connection to Kate Stange's work.

REFERENCES

[1] Diego Armentano and Carlos Beltrán. The polynomial eigenvalue problem is well conditioned for random inputs. *SIAM Journal on Matrix Analysis and Applications*, 40(1):175–193, 2019.

[2] Mauricio Barrera, Albrecht Böttcher, Sergei M Grudsky, and Egor A Maximenko. Eigenvalues of even very nice Toeplitz matrices can be unexpectedly erratic. In *The diversity and beauty of applied operator theory*, pages 51–77. Springer, 2018.

[3] Ivar Bendixson. Sur les racines d'une équation fondamentale. *Acta Mathematica*, 25(0):359–365, 1902.

[4] Manuel Bogoya, Stefano Serra-Capizzano, and Ken Trotti. Upper Hessenberg and Toeplitz Bohemian matrix sequences: a note on their asymptotical eigenvalues and singular values. *Electronic Transactions on Numerical Analysis*, 55:76–91, 2022.

[5] Peter Borwein and Loki Jörgenson. Visible structures in number theory. *The American Mathematical Monthly*, 108(10):897–910, 2001.

[6] Peter Borwein and Christopher Pinner. Polynomials with $\{0,+\ 1,\text{-}1\}$ coefficients and a root close to a given point. *Canadian Journal of Mathematics*, 49(5):887–915, 1997.

[7] Albrecht Böttcher and Bernd Silbermann. *Introduction to large truncated Toeplitz matrices*. Springer Science & Business Media, 2012.

[8] T. J. I'a Bromwich. On the roots of the characteristic equation of a linear substitution. *Acta Mathematica*, 30(0):297–304, 1906.

[9] AT Butson. Generalized Hadamard matrices. *Proceedings of the American Mathematical Society*, 13(6):894–898, 1962.

[10] Neil Calkin, Eunice Chan, and Robert Corless. Some facts and conjectures about Mandelbrot polynomials. *Maple Transactions*, 1(1):13, October 2021.

[11] Eunice Chan and Robert Corless. A new kind of companion matrix. *The Electronic Journal of Linear Algebra*, 32:335–342, February 2017.

[12] Eunice Y. S. Chan and Robert M. Corless. Minimal height companion matrices for Euclid polynomials. *Mathematics in Computer Science*, 13(1-2):41–56, July 2018.

[13] Eunice Y. S. Chan, Robert M. Corless, Laureano González-Vega, J. Rafael Sendra, and Juana Sendra. Inner Bohemian inverses. *Applied Mathematics and Computation*, accepted January 11, 2022, 2022.

[14] Eunice Y.S. Chan, Robert M. Corless, Laureano Gonzalez-Vega, J. Rafael Sendra, Juana Sendra, and Steven E. Thornton. Upper Hessenberg and Toeplitz Bohemians. *Linear Algebra and its Applications*, 601:72–100, September 2020.

[15] Robert Corless. Skew-symmetric tridiagonal Bohemian matrices. *Maple Transactions*, 1(2), October 2021.

[16] Robert M. Corless. What can we learn from Bohemian matrices? *Maple Transactions*, 1:31, 7 2021.

[17] Robert M. Corless and Nicolas Fillion. *A Graduate Introduction to Numerical Methods*. Springer, 2013.

[18] Gabriel Dorfsman-Hopkins and Candy Xu. Searching for rigidity in algebraic starscapes. *arXiv preprint arXiv:2107.06328*, 2021.

[19] Zhibin Du, Carlos M. da Fonseca, Yingqiu Xu, and Jiahao Ye. Disproving a conjecture of Thornton on Bohemian matrices. *Open Mathematics*, 19(1):505–514, 2021.

[20] Massimiliano Fasi and Gian Maria Negri Porzio. Determinants of normalized Bohemian upper Hessenberg matrices. *The Electronic Journal of Linear Algebra*, 36:352–366, 2020.

[21] Heinrich W Guggenheimer, Alan S Edelman, and Charles R Johnson. A simple estimate of the condition number of a linear system. *The College Mathematics Journal*, 26(1):2–5, 1995.

[22] Edmund Harriss, Katherine E Stange, and Steve Trettel. Algebraic number starscapes. *arXiv preprint arXiv:2008.07655*, 2020.

[23] M. A. Hirsch. Sur les racines d'une équation fondamentale. *Acta Mathematica*, 25(0):367–370, 1902.

[24] Leslie Hogben. *Handbook of linear algebra*. CRC Press, 2nd edition, 2013.

[25] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[26] J. E. Littlewood. *Some problems in real and complex analysis*. 1968.

[27] Xuhua Liu, Brice M. Nguelifack, and Tin-Yau Tam. Unitary similarity to a complex symmetric matrix and its extension to orthogonal symmetric lie algebras. *Linear Algebra and its Applications*, 438(10):3789–3796, May 2013.

[28] Kurt Mahler. On two extremum properties of polynomials. *Illinois Journal of Mathematics*, 7(4):681–701, 1963.

[29] Andrew Odlyzko. Zeros of polynomials with 0–1 coefficients. In Bruno Salvy, editor, *Algorithms Seminar 1992–1993*. INRIA, 1992. Summary by Xavier Gourdon.

[30] Dan Piponi. Two tricks for the price of one: Linear filters and their transposes. *Journal of Graphics, GPU, and Game Tools*, 14(1):63–72, January 2009.

[31] T. Ratnarajah, R. Vaillancourt, and M. Alvo. Eigenvalues and condition numbers of complex random matrices. *SIAM Journal on Matrix Analysis and Applications*, 26(2):441–456, January 2004.

[32] Palle Schmidt and Frank Spitzer. The Toeplitz matrices of an arbitrary Laurent polynomial. *Mathematica Scandinavica*, 8(1):15–38, 1960.

[33] Katherine E Stange. Visualizing the arithmetic of imaginary quadratic fields. *International Mathematics Research Notices*, 2018(12):3908–3938, February 2017.

[34] Terence Tao and Van Vu. On random$\pm$1 matrices: singularity and determinant. *Random Structures & Algorithms*, 28(1):1–23, 2006.

[35] Terence Tao and Van Vu. Random matrices have simple spectrum. *Combinatorica*, 37(3):539–553, 2017.

[36] Olga Taussky. Matrices of rational integers. *Bulletin of the American Mathematical Society*, 66(5):327–345, 1960.

[37] Craig A. Tracy and Harold Widom. Level-spacing distributions and the Airy kernel. *Physics Letters B*, 305(1-2):115–118, May 1993.

[38] Miodrag Živković. Classification of small (0, 1) matrices. *Linear algebra and its applications*, 414(1):310–346, 2006.

A SOFTWARE USED

The code base used here is a combination of Maple (including the Image and ColorTools packages) and Python. They are interoperable to an extent, in that data built in one can be saved to a JSON container and read into the other. Figure 1 and Figures 7–8 were produced using polynomials computed in Python and transferred to Maple for image processing. All others were produced in Python. For some things the Python code is faster, but for others (including polynomial manipulation) Maple is faster. An early version of the Maple code is available at [16]. The Python code will be released soon. Older software for Bohemian matrices can be found at https://github.com/BohemianMatrices.

We have not yet systematically taken advantage of parallelism, even though many of the computations here are trivially parallel. One example (not shown) of computing the eigenvalues of 2^{31} matrices on a 32 core machine was successful; and the plot in Figure 6 used GPU computing on Google's COLAB through TensorFlow at https://colab.research.google.com/drive/1dpk1aM_6e9k6Ep0sf3iMFdMKvZW8zphP?usp=sharing. It is clear that there is significant scope for parallel investigations.