UPPER HESSENBERG AND TOEPLITZ BOHEMIAN MATRIX SEQUENCES: A NOTE ON THEIR ASYMPTOTICAL EIGENVALUES AND SINGULAR VALUES*

MANUEL BOGOYA[†], STEFANO SERRA-CAPIZZANO[†], AND KEN TROTTI[†]

Abstract. In previous works, Bohemian matrices have attracted the attention of several researchers for their rich combinatorial structure, and they have been studied intensively from several points of view, including height, determinants, characteristic polynomials, normality, and stability. Here we consider a selected number of examples of upper Hessenberg and Toeplitz Bohemian matrix sequences whose entries belong to the population $P = \{0, \pm 1\}$, and we propose a connection with the spectral theory of Toeplitz matrix sequences and Generalized Locally Toeplitz (GLT) matrix sequences in order to give results on the localization and asymptotical distribution of their spectra and singular values. Numerical experiments that support the mathematical study are reported. A conclusion section ends the note in order to illustrate the applicability of the proposed tools to more general cases.

Key words. matrix (Bohemian, (upper) Hessenberg, Toeplitz), matrix sequence (Toeplitz, GLT), eigenvalue, singular value, spectral and singular value symbol/distribution

AMS subject classifications. 15B05, 15B36, 15A18, 11C20, 65F08, 65F15

1. Introduction. A set of matrices having entries belonging to a fixed population P is called Bohemian, where the name has neither a geographic nor an anthropological meaning, but it comes from the simple acronym Bounded Height Matrix of Integers, that is, *Bohemi*.

Whereas it is not strictly true, the majority of works (see [8, 27] and the references therein) have focused their attention on the case where P is made up of integers (or subsets of integers) or on the simplest case with $P = \{0, \pm 1\}$; see, e.g., [7, 17]. It is worth mentioning at this point that Olga Taussky-Todd in her instructive work [25] writes the following illuminating incipit:

"This subject is very vast and very old. It includes all of the arithmetic theory of quadratic forms, as well as many other classical subjects, such as Latin squares and matrices with elements +1or -1 which enter into Euler's, Sylvester's or Hadamard's famous conjectures".

Then Gear [16] and again Taussky-Todd [26] considered more general instances with P composed of integers.

A further different point of view has been introduced by Tao and Vu [24], who have considered real, symmetric, random matrices with upper triangular and diagonal entries being independent random variables. In this setting, quite different from the standard one, the authors prove results on the expected spectra.

In this short note, we take the very standard and simple case of upper Hessenberg and Toeplitz Bohemian matrix sequences with the population $P = \{0, \pm 1\}$; see, e.g., [8, 9, 11, 12] and the references therein. In particular in [9, Figure 1 and Figure 2], the plots of the spectra of small-size upper Hessenberg and Toeplitz Bohemian matrices are reported. Furthermore, in the Conclusion Remarks in [9, last lines at Page 17], the authors observe some kind of asymptotics for such spectra, but the authors admit that "many puzzles remain" and that "they do not have any explanation" for the observed asymptotic behavior when the matrix size grows.

For trying to cope with this problem, we put the study in the context of Toeplitz matrix sequences, endowed with a Lebesgue-integrable generating function and in the much more

^{*}Received September 22, 2021. Accepted October 1, 2021. Published online on November 4, 2021. Recommended by Silvia Noschese. This work is partly supported by GNCS-INDAM (Italy).

[†]University of Insubria, via Valleggio, 11 22100 Como, Italy

^{({}johanmanuel.bogoya, s.serracapizzano, kl.trotti}@uninsubria.it).

general setting of the *-algebra of Generalized Locally Toeplitz (GLT) matrix sequences endowed with Lebesgue-measurable symbols (see [5, 14] and the references there reported).

With the help of these tools, we can easily describe the asymptotic behavior of the singular values by identifying the singular value distribution function. However, the considered matrices, also due to their upper Hessenberg nature, are highly non-normal. Consequently, the information on the singular values does not imply any precise result on the eigenvalues in terms of either localization or distribution. For the eigenvalues, the results that we obtain are in fact less precise and a bit more difficult to establish, but in any case we arrive at giving localization results and an indication of the spectral distribution.

We emphasize that only two precise examples of matrix sequences taken from the literature are considered in the paper. However, as it will become clear in the sequel, the techniques are simple and quite general and can be adapted to more general Bohemian matrix sequences with a population of integers.

The present work is organized as follows. In Section 2 we set the notation and the main definitions regarding Bohemian matrices, Hessenberg matrices, and Toeplitz matrices: a separate part concerns the notion of distribution in the eigenvalue and singular value sense, the definition of Toeplitz matrix sequences, and of the *-algebra of GLT matrix sequences with their main properties. In Section 3 we use the machinery of the previous section in the specific cases of upper Hessenberg and Toeplitz Bohemian matrix sequences by indicating how to generalize the results in a wider setting. For supporting the analysis, a selection of numerical experiments is reported with a related critical discussion. A short conclusion section (that is Section 4) ends the present note.

2. Notations, definitions, tools. A family of Bohemian matrices is a set of matrices where the entries are independently sampled from a finite set, usually integers, of bounded height. Such families arise in many applications, e.g., compressed sensing, and, as already emphasized in the introduction, the properties of matrices selected "at random" from such families are of practical and mathematical interest. Studying such matrices leads to many unanswered questions. In the sequel we state the problem of understanding the spectral features of such matrices such as localization and distribution. More in detail, for such matrices we are interested in the eigenvalues and singular values, both in terms of the localization set and of the asymptotical spectral and singular value distribution as the matrix size *N* tends to infinity according to Definition 2.7. Let us now introduce some basic definitions and results.

Let $A \in \mathbb{C}^{N \times N}$. Then we denote by $\Lambda(A)$ the set of eigenvalues λ of the matrix A and by $\Sigma(A)$ the set of singular values σ of A.

DEFINITION 2.1. Let $A \in \mathbb{C}^{N \times N}$ be nonsingular, and consider the uniquely solvable linear system Ax = b. Then the condition number $\mu(A)$ is defined as

$$\mu(A) = \|A\| \cdot \|A^{-1}\|,$$

where $\|\cdot\|$ is the spectral norm.

REMARK 2.2. The condition number $\mu(A)$ of a matrix A gives a bound of how inaccurate the solution x of the linear system Ax = b will be after numerical computations due to roundoff errors. Therefore, when $\mu(A)$ is large, then we say that the matrix A is ill-conditioned.

DEFINITION 2.3 ([8]). An upper Hessenberg Bohemian matrix $H_N \in \mathbb{C}^{N \times N}$ is a matrix of the form

$$H_N = \begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} & \cdots & h_{1,N} \\ s_1 & h_{2,2} & h_{2,3} & \cdots & h_{2,N} \\ 0 & s_2 & h_{3,3} & \cdots & h_{3,N} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & s_{N-1} & h_{N,N} \end{bmatrix}$$

with $s_k = e^{i\theta_k}$ and usually $s_k \in \{-1, +1\}$ and $h_{i,j} \in P$, where P is a fixed finite set called the population set.

Moreover, the height of the matrix H_N , written height $(H_N) := \|\operatorname{vec}(H_N)\|_{\infty}$, is the largest absolute value of any entry in H_N .

DEFINITION 2.4. Let $f \in L^1([-\pi,\pi])$, and let $\{f_k\}_{k\in\mathbb{Z}}$ be the sequence of its Fourier coefficients defined as

$$f_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) \mathrm{e}^{-\mathrm{i}k\theta} \,\mathrm{d}\theta, \qquad k \in \mathbb{Z}.$$

Then the matrix sequence $\{T_N\}_{N\in\mathbb{N}}$ with $T_N = [f_{i-j}]_{i,j=1}^N$ is called the sequence of Toeplitz matrices generated by f, which in turn is called the generating function of $\{T_N\}_{N\in\mathbb{N}}$, and T_N is denoted by $T_N(f)$.

DEFINITION 2.5. The Wiener class is the set of functions $f(\theta) = \sum_{k=-\infty}^{\infty} f_k e^{ik\theta}$ such that $\sum_{k=-\infty}^{\infty} |f_k| < \infty$.

Note that the Wiener class forms a subalgebra of the continuous and 2π -periodic functions as indicated in the remark below.

REMARK 2.6. Let $\{T_N\}_{N\in\mathbb{N}}$ be a Toeplitz sequence with $T_N = [f_{i-j}]_{i,j=1}^N$. If $\{f_k\}_{k\in\mathbb{Z}}$ is such that $\sum_{k=-\infty}^{\infty} |f_k| < \infty$, then the series $\sum_{k=-\infty}^{\infty} f_k e^{ik\theta}$ converges uniformly in the infinity norm to a continuous and 2π -periodic function f, which belongs to the Wiener class and which is the generating function of $\{T_N\}_{N\in\mathbb{N}}$, i.e., $T_N = T_N(f)$, for all $N \in \mathbb{N}$.

We continue by giving the definition of the spectral distribution in the sense of the eigenvalues and of the singular values together with the associated notion of the symbol.

DEFINITION 2.7. Let $f: [a, b] \to \mathbb{C}$ be a measurable function defined on $[a, b] \subset \mathbb{R}$. Let $C_0(\mathbb{C})$ be the set of continuous functions with compact support over \mathbb{C} , and let $\{A_N\}_N$ be a sequence of matrices of size N with eigenvalues $\lambda_j(A_N)$, j = 1, ..., N. We say that $\{A_N\}_N$ is distributed as the pair (f, [a, b]) in the sense of the eigenvalues, and we write

$$\{A_N\}_N \sim_\lambda (f, [a, b]),$$

if the following limit relation holds for all $F \in C_0(\mathbb{C})$ *:*

(2.1)
$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} F(\lambda_j(A_N)) = \frac{1}{b-a} \int_a^b F(f(t)) \mathrm{d}t.$$

In this case, the function f is called the spectral symbol (or the symbol in the eigenvalue sense) of $\{A_N\}_N$.

In perfect analogy, with $\sigma_j(A_N)$, j = 1, ..., N, denoting the N singular values of A_N , we say that $\{A_N\}_N$ is distributed as the pair (f, [a, b]) in the sense of the singular values, and we write

$$\{A_N\}_N \sim_\sigma (f, [a, b]),$$

if the following limit relation holds for all $F \in C_0(\mathbb{C})$ *:*

(2.2)
$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} F(\sigma_j(A_N)) = \frac{1}{b-a} \int_a^b F(|f(t)|) dt.$$

In this case, the function f is called the symbol in the sense of the singular values of $\{A_N\}_N$.

REMARK 2.8. When f is continuous, an informal interpretation of the limit relation (2.1) is that when the matrix size is sufficiently large, the eigenvalues of A_N can be approximated by a sampling of f on a uniform equispaced grid of the interval [a, b] up to a relatively small number of potential outliers. In the present context, the expression "relatively small" means o(N) with N denoting the matrix size. When |f| is continuous, an informal interpretation of the limit relation (2.2) is that when the matrix size is sufficiently large, the singular values of A_N can be approximated by a sampling of |f| on a uniform equispaced grid of the interval [a, b], again up to a relatively small number of potential outliers.

For Hermitian Toeplitz matrix sequences, the following theorem due to Szegő, Tyrtyshnikov, etc. holds (see, e.g., [14] and the references therein): here we report it in the version due to Tyrtyshnikov and Zamarashkin [29].

THEOREM 2.9. Let $f \in L^1([-\pi,\pi])$ be a real-valued function almost everywhere. Then,

$${T_N(f)}_N \sim_\lambda (f, [-\pi, \pi]).$$

In the general case we have

$$\{T_N(f)\}_N \sim_\sigma (f, [-\pi, \pi]).$$

We end this introductory part by recalling a property of the spectral norm of Toeplitz matrices and stating relevant results on GLT matrix sequences contained, e.g., in [14]. Given a square matrix X of order N, we denote its spectral norm by ||X|| (see also Definition 2.1), that is, its maximal singular value ($||X|| = \max_{i=1,...,N} \sigma_i(X)$), which coincides with the spectral radius in the case of a normal matrix, and we recall that every Hermitian matrix is also normal.

Given a Toeplitz sequence $\{T_N(f)\}_{N \in \mathbb{N}}$ generated by f, it holds that (see [23, Corollary 3.5]):

$$f \in L^{\infty}([-\pi,\pi]) \Rightarrow ||T_N(f)|| \leq ||f||_{\infty}, \quad \forall N \in \mathbb{N}.$$

In short, the GLT class is a matrix sequence algebra obtained as the closure under some algebraic operations among Toeplitz matrix sequences, diagonal sampling matrix sequences (see Definition 2.10), and zero-distributed matrix sequences (see item **GLT4**). Furthermore, the GLT algebra includes any sequence of matrices coming from "reasonable" approximations by local discretization methods (FDs, FVs, FEs, Isogeometric Analysis, etc.) of partial differential equations (see, e.g., [3, 13, 21, 22]). The formal definition is rather technical, difficult, and involves heavy notation: therefore we just give and briefly discuss the notion in one dimension, which is the case of interest in our setting. In the sequel, we report a few properties of the GLT class [14], which are sufficient for studying the spectral features of the matrix sequence $\{A_N\}_N$ we are interested in (see the pioneering work by Tilli [28] and the generalization/modifications provided in [14, 15, 21, 22] for a more detailed discussion).

Let us start by recalling the definition of diagonal sampling matrices.

DEFINITION 2.10. Given a Riemann-integrable function a defined over [0, 1]. By a diagonal sampling matrix of order N we mean

$$D_N(a) = \operatorname{diag}_{j=1,\dots,N} a\left(\frac{j}{N}\right).$$

Throughout, we use the following notation

$$\{A_N\}_{N\in\mathbb{N}} \sim_{\mathrm{GLT}} \psi(x,\theta), \qquad (x,\theta) \in [0,1] \times [-\pi,\pi],$$

to say that the sequence $\{A_N\}_{N \in \mathbb{N}}$ is a GLT sequence with GLT symbol $\psi(x, \theta)$. Here we report five main features of the GLT class (see [14, 21, 22]).

- **GLT1** Let $\{A_N\}_{N\in\mathbb{N}} \sim_{\text{GLT}} \psi(x,\theta)$ with $\psi: \Omega \to \mathbb{C}, \Omega = [0,1] \times [-\pi,\pi]$. Then it holds that $\{A_N\}_{N\in\mathbb{N}} \sim_{\sigma} (\psi,\Omega)$. In the case where the matrices A_N are Hermitian, $\{A_N\}_{N\in\mathbb{N}} \sim_{\lambda} (\psi,\Omega)$.
- **GLT2** The set of GLT sequences form a *-algebra, i.e., it is closed under linear combinations, products, inversion (whenever the symbol vanishes, at most, in a set of zero Lebesgue measure), transposed conjugation. Hence, the sequence obtained via algebraic operations on a finite set of input GLT sequences is still a GLT sequence, and its symbol is obtained by following the same algebraic manipulations for the corresponding symbols of the input GLT sequences.
- **GLT3** Every Toeplitz sequence $\{T_N(f)\}_{N\in\mathbb{N}}$ generated by an $L^1([-\pi,\pi])$ -function $f(\theta)$ is such that $\{T_N(f)\}_{N\in\mathbb{N}} \sim_{\text{GLT}} f(\theta)$, with the specifications reported in item **GLT1**. Every diagonal sampling sequence $\{D_N(a)\}_{N\in\mathbb{N}}$, where a(x) is a Riemann-integrable function, is such that $\{D_N(a)\}_{N\in\mathbb{N}} \sim_{\text{GLT}} a(x)$.
- **GLT4** Every sequence which is distributed as the constant zero in the singular value sense is a GLT sequence with symbol zero and vice versa. In formulae, $\{A_N\}_{N\in\mathbb{N}} \sim_{\sigma} (0, \Omega)$, $\Omega = [0, 1] \times [-\pi, \pi]$, if and only if $\{A_N\}_{N\in\mathbb{N}} \sim_{\text{GLT}} 0$. These matrix sequences are called zero-distributed matrix sequences.

3. Main results and numerical experiments. Consider the Bohemian matrices A_N , $B_N \in \mathbb{R}^{N \times N}$, with $P = \{-1, 1\}$ and $s_k = 1, k = 1, \dots, N$, of the form

$$A_N = \begin{bmatrix} -1 & -1 & -1 & \cdots & -1 \\ 1 & -1 & -1 & \cdots & -1 \\ 0 & 1 & -1 & \cdots & -1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -1 \end{bmatrix}, \quad B_N = \begin{bmatrix} 1 & (-1)^1 & (-1)^2 & \cdots & (-1)^{N-1} \\ 1 & 1 & (-1)^1 & \cdots & (-1)^{N-2} \\ 0 & 1 & 1 & \cdots & (-1)^{N-3} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 1 \end{bmatrix}.$$

We note that A_N and B_N are Toeplitz matrices, but retrieving the symbol of the matrix sequences $\{A_N\}_N, \{B_N\}_N$ is not trivial. Indeed, the respective coefficients are constant in modulus, and therefore the generating function in the sense of the Toeplitz terminology (see Definition 2.4) cannot exist due to the Riemann-Lebesgue Lemma. As a consequence, we have to try the use of the GLT tools for finding a (Lebesgue-measurable) symbol in the sense of Definition 2.7. More precisely we write

$$A_N = I_N - R_N, \qquad B_N = I_N + S_N,$$

with

$$\tilde{I}_N = \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}, \qquad R_N = \begin{bmatrix} 1 & \cdots & \cdots & 1 \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & & \ddots & \vdots \\ & & & & 1 \end{bmatrix},$$

ETNA Kent State University and Johann Radon Institute (RICAM)

UPPER HESSENBERG AND TOEPLITZ BOHEMIAN MATRIX SEQUENCES

$$S_N = \begin{bmatrix} (-1)^0 & (-1)^1 & \cdots & (-1)^{N-1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & (-1)^1 \\ & & & (-1)^0 \end{bmatrix},$$

and we note that

$$R_N^{-1} = \begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{bmatrix}, \qquad S_N^{-1} = \begin{bmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & & 1 \end{bmatrix},$$

Now, since R_N^{-1} , S_N^{-1} are banded Toeplitz matrices, from Remark 2.6 and GLT3 it holds that

(3.1)
$$\{ R_N^{-1} \}_N \sim_{\text{GLT}} 1 - e^{-i\theta}, \\ \{ S_N^{-1} \}_N \sim_{\text{GLT}} 1 + e^{-i\theta},$$

 $\theta \in [-\pi, \pi]$, since $R_N^{-1} = T_N(1 - e^{-i\theta})$, $S_N^{-1} = T_N(1 + e^{-i\theta})$ in the sense of Definition 2.4. Through property **GLT2**, equation (3.1) can be written as

$$\{R_N\}_N \sim_{\text{GLT}} \frac{1}{1 - e^{-i\theta}}, \\ \{S_N\}_N \sim_{\text{GLT}} \frac{1}{1 + e^{-i\theta}},$$

and hence, by item GLT2, we deduce

$$\{A_N\}_N \sim_{\text{GLT}} a(\theta) = e^{i\theta} - \frac{1}{1 - e^{-i\theta}} = \frac{e^{i\theta} - 2}{1 - e^{-i\theta}}$$
$$\{B_N\}_N \sim_{\text{GLT}} b(\theta) = e^{i\theta} + \frac{1}{1 + e^{-i\theta}} = \frac{2 + e^{i\theta}}{1 + e^{-i\theta}}$$

with $a(\theta), b(\theta) \colon [-\pi, \pi] \to \mathbb{C}$.

REMARK 3.1. From **GLT1**, it holds that $\{A_N\}_N \sim_{\sigma} a(\theta)$ and $\{B_N\}_N \sim_{\sigma} b(\theta)$. Moreover, when taking the absolute value of the symbols $a(\theta), b(\theta)$, we find

$$|a(\theta)| = \frac{|e^{i\theta} - 2|}{|1 - e^{-i\theta}|} = \frac{\sqrt{5 - 4\cos(\theta)}}{\sqrt{2 - 2\cos(\theta)}} = \sqrt{\frac{1}{2 - 2\cos(\theta)} + 2},$$
$$|b(\theta)| = \frac{|2 + e^{i\theta}|}{|1 + e^{-i\theta}|} = \sqrt{\frac{5 + 4\cos(\theta)}{2 + 2\cos(\theta)}} = \sqrt{\frac{1}{2 + 2\cos(\theta)} + 2}.$$

It is now straightforward to see that $|a(\theta)|$, $|b(\theta)|$ are both symmetric with respect to $\theta = 0$, and, as a consequence of Remark 2.8, a better approximation of the singular values is obtained by restricting the definition interval of the functions $|a(\theta)|$, $|b(\theta)|$ from $[-\pi, \pi]$ to $[0, \pi]$.

If we look at Figure 3.7, we can fully appreciate the agreement of the singular values of the two matrix sequences with respect to their symbols, in accordance with Definition 2.7 and Remark 2.8: notice that we have plotted a monotonically increasing rearrangement of the symbols since the singular values are in non-decreasing order.

Now we continue with the analysis of the eigenvalues.

REMARK 3.2. Let $D_N = \underset{j=1,...,N}{\text{diag}} ((-1)^j)$ be the diagonal matrix whose *j*-th diagonal

entry is $(-1)^j$. Then B_N can be obtained from A_N through the transformation

$$B_N = -D_N A_N D_N.$$

Since $D_N = D_N^{-1}$, it follows that $D_N A_N D_N$ has the same spectrum as A_N , and therefore it holds that $\Lambda(B_N) = -\Lambda(A_N)$. Let $E_N = \underset{j=1,\ldots,N}{\text{diag}} (\epsilon^{-j})$ be the diagonal matrix whose j-th

diagonal element is ϵ^{-j} . Then, following the Beam-Warming "replacement trick" [2], which consists in multiplying on the left by E_N and on the right by its inverse E_N^{-1} , we obtain the matrices

$$\tilde{A}_{N} = \begin{bmatrix} -1 & -\epsilon^{1} & -\epsilon^{2} & \cdots & -\epsilon^{N-1} \\ \frac{1}{\epsilon} & -1 & -\epsilon^{1} & \cdots & -\epsilon^{N-2} \\ 0 & \frac{1}{\epsilon} & -1 & \cdots & -\epsilon^{N-3} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{1}{\epsilon} & -1 \end{bmatrix}, \quad \tilde{B}_{N} = \begin{bmatrix} 1 & (-\epsilon)^{1} & (-\epsilon)^{2} & \cdots & (-\epsilon)^{N-1} \\ \frac{1}{\epsilon} & 1 & (-\epsilon)^{1} & \cdots & (-\epsilon)^{N-2} \\ 0 & \frac{1}{\epsilon} & 1 & \cdots & (-\epsilon)^{N-3} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{1}{\epsilon} & 1 \end{bmatrix},$$

which are respectively similar to A_N and B_N .

According to the Definition 2.4 of a Toeplitz matrix generated by a function, the matrices \tilde{A}_N, \tilde{B}_N are respectively generated by the functions

(3.2)

$$\tilde{a}_{\epsilon}(\theta) = \frac{1}{\epsilon} e^{i\theta} - \sum_{k=0}^{\infty} \epsilon^{k} e^{-ik\theta},$$

$$\tilde{b}_{\epsilon}(\theta) = \frac{1}{\epsilon} e^{i\theta} + \sum_{k=0}^{\infty} (-\epsilon)^{k} e^{-ik\theta}$$

In both cases, the series of the absolute values of the Fourier coefficients satisfies

$$\left|\frac{1}{\epsilon}\right| + \sum_{k=0}^{\infty} \left|\epsilon^k\right| = \frac{1}{\epsilon} + \frac{1}{1-\epsilon} < \infty, \qquad 0 < \epsilon < 1.$$

Therefore, according to Remark 2.6, the functions $\tilde{a}_{\epsilon}(\theta)$, $\tilde{b}_{\epsilon}(\theta)$ belong to the Wiener class. Hence, by using the GLT machinery we have

$$\{\tilde{A}_N\}_N \sim_{\text{GLT}} \tilde{a}_{\epsilon}(\theta) = \frac{1}{\epsilon} e^{i\theta} - \sum_{k=0}^{\infty} \epsilon^k e^{-ik\theta} = \frac{1}{\epsilon} e^{i\theta} - \frac{1}{1 - \epsilon e^{-i\theta}},$$
$$\{\tilde{B}_N\}_N \sim_{\text{GLT}} \tilde{b}_{\epsilon}(\theta) = \frac{1}{\epsilon} e^{i\theta} + \sum_{k=0}^{\infty} (-\epsilon)^k e^{-ik\theta} = \frac{1}{\epsilon} e^{i\theta} + \frac{1}{1 + \epsilon e^{-i\theta}},$$

and more precisely, $\tilde{A}_N = T_N(\tilde{a}_{\epsilon}), \tilde{B}_N = T_N(\tilde{b}_{\epsilon}).$

To find the optimal ϵ we introduce the Gershgorin Theorem (see [30] for further details and insights).

THEOREM 3.3 (Gershgorin Theorem [30]). Let $A = [a_{i,j}] \in \mathbb{C}^{N \times N}$, and set

$$\Gamma_i(A) = \{ z \in \mathbb{C} \colon |z - a_{i,i}| \leq r_i(A) \}, \qquad i = 1, \dots, N,$$

$$\Gamma(A) = \bigcup_{i=1}^N \Gamma_i(A),$$

where $\Gamma_i(A)$ is called the *i*-th Gershgorin circle of A, $\Gamma(A)$ is the Gershgorin set of A, and

$$r_i(A) = \sum_{\substack{j=1,\ldots,N\\ j\neq i}} |a_{i,j}|, \qquad i = 1,\ldots,N,$$

is the radius of the *i*-th Gershgorin circle of A. Then it holds that

$$\lambda \in \Gamma(A), \qquad \forall \lambda \in \Lambda(A).$$

The optimal ϵ for \tilde{A}_N or \tilde{B}_N in the Gershgorin sense is the one that minimizes the Gershgorin set of \tilde{A}_N or \tilde{B}_N , respectively.

PROPOSITION 3.4. The optimal ϵ in the Gershgorin sense is $\epsilon \approx \frac{1}{2}$ for both matrices \tilde{A}_N, \tilde{B}_N , at least for N large enough.

Proof. We note that, since the main diagonals of \tilde{A}_N , \tilde{B}_N are constant, the center of the Gershgorin circles for these matrices does not depend on the row index. Therefore, we only have to minimize the maximum radius. Let $r_i(\tilde{A}_N), r_i(\tilde{B}_N)$ be the radius of the *i*-th Gershgorin circle for the matrices \tilde{A}_N , \tilde{B}_N , respectively. We note that, due to the presence of the absolute value, for all i = 1, ..., N, it holds that

$$r_i := r_i(\tilde{A}_N) = r_i(\tilde{B}_N) \leqslant \frac{1}{\epsilon} + \sum_{j=0}^{i-1} \epsilon^j \leqslant \frac{1}{\epsilon} + \frac{\epsilon}{1-\epsilon}.$$

The thesis follows by considering the previous bound for r_i as a function of ϵ and by setting its derivative to zero.

Let us fix $\epsilon = 0.5$. Then Proposition 3.4 tells us that $r_i \leq 3$, and in this case, Theorem 3.3 tells us that the Gershgorin set $\Gamma(\tilde{A}_N)$ is a circle centered at z = -1 with radius ≤ 3 . Similarly, $\Gamma(\tilde{B}_N)$ is a circle centered at z = 1 with radius ≤ 3 . By looking at Figures 3.2 and 3.4 we can see that the optimal radius is $\sqrt{5}$.

Figure 3.1 displays the real and imaginary parts of $\Lambda(A_N)$, $\Lambda(\tilde{A}_N)$ and Figure 3.2 displays $\Lambda(A_N)$, $\Lambda(\tilde{A}_N)$ in the complex plane for N = 200, 1600. We note that, even with N = 200, the difference between the two spectra is visible, and the difference becomes more marked as N increases. Indeed, the condition number $\mu(E_N) = \epsilon^{-N}$ of the matrix E_N increases exponentially when increasing N, and, in accordance with Remark 2.2, the matrix \tilde{A}_N is severely ill-conditioned. Therefore, the computation of its eigenvalues is affected by severe numerical instabilities.

Figures 3.3 and 3.4 display $\Lambda(B_N)$, $\Lambda(\tilde{B_N})$ for N = 200, 1600. We note that since the matrix E_N is also involved in the computation of \tilde{B}_N , the computation of $\Lambda(\tilde{B}_N)$ is affected by numerical instability.

Figures 3.1 to 3.4 show that the theoretically optimal ϵ does not yield a good estimate of the spectrum of the matrices A_N, B_N through \tilde{A}_N, \tilde{B}_N . Therefore, we vary $\epsilon \in [0.1, 1]$ and compute the error $e_{\epsilon}^{A_N}$ between A_N and \tilde{A}_N to see its behavior.

We define $e_{\epsilon}^{A_N}$ as

$$e_{\epsilon}^{A_N} = \left\| \operatorname{sort} \Lambda(A_N) - \operatorname{sort} \Lambda(\tilde{A}_N) \right\|,$$

where sort $\Lambda(X)$ denotes the complex vector containing the eigenvalues of X ordered with increasing real part, and, in case of two equal real parts, we order with respect to increasing imaginary part. $e_{\epsilon}^{B_N}$ is defined similarly with B_N in place of A_N .

Figure 3.5 displays $e_{\epsilon}^{A_N}$ and $e_{\epsilon}^{B_N}$ when varying $\epsilon \in [0.1, 1]$ for different values of N. A comparison between the images in Figure 3.5 shows that $e_{\epsilon}^{A_N}$ behaves similarly to $e_{\epsilon}^{B_N}$. Both

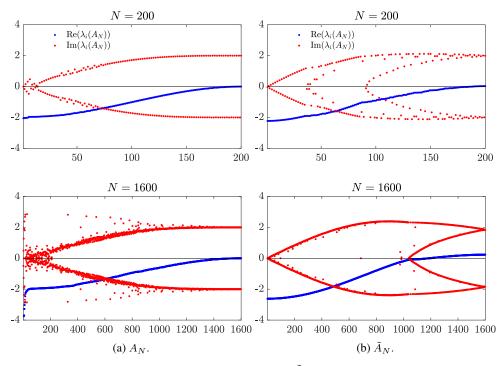


FIG. 3.1. Real and imaginary parts of the matrices A_N , \tilde{A}_N with $\epsilon = 0.5$ and N = 200, 1600.

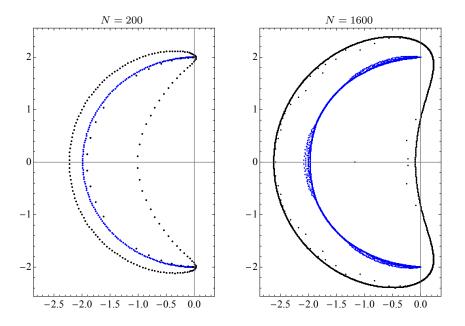


FIG. 3.2. Spectrum in the complex plane of the matrices A_N (blue dots) and \tilde{A}_N (black dots) with $\epsilon = 0.5$ and N = 200, 1600.

ETNA Kent State University and Johann Radon Institute (RICAM)

85

UPPER HESSENBERG AND TOEPLITZ BOHEMIAN MATRIX SEQUENCES

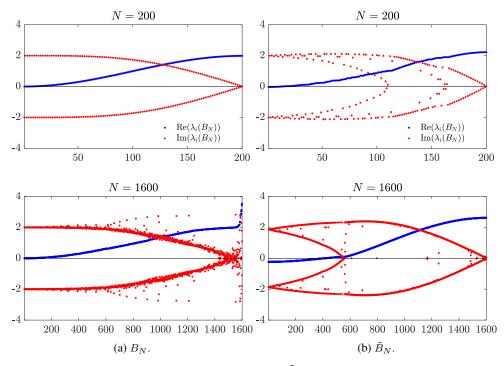


FIG. 3.3. Real and imaginary parts of matrices B_N , \tilde{B}_N with $\epsilon = 0.5$ and N = 200, 1600.

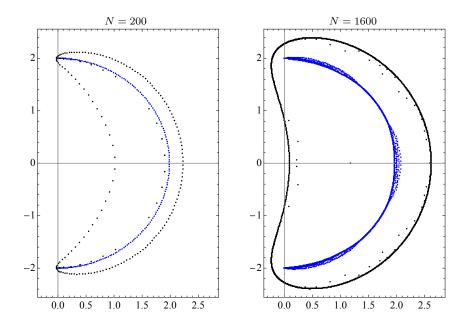


FIG. 3.4. Spectrum in the complex plane of the matrices B_N (blue dots) and \tilde{B}_N (black dots) with $\epsilon = 0.5$ and N = 200, 1600.

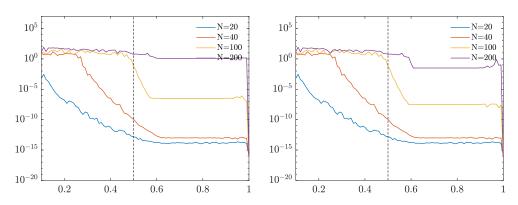


FIG. 3.5. Errors $e_{\epsilon}^{A_N}$ (left) and $e_{\epsilon}^{B_N}$ (right) for varying $\epsilon \in [0.1, 1]$, with N = 20, 40, 100, 200.

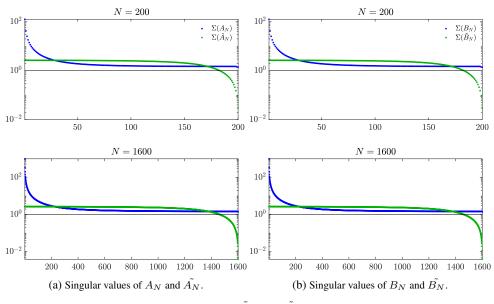


FIG. 3.6. Singular values of matrices A_N , $\tilde{A_N}$, B_N , $\tilde{B_N}$ with $\epsilon = 0.5$ and N = 200, 1600.

errors decrease with increasing ϵ up to $\epsilon \approx 0.6$, and then the error stagnates and collapses to machine precision when $\epsilon = 1$, i.e., when \tilde{A}_N , \tilde{B}_N coincide with A_N , B_N , respectively. We also note that $e_{\epsilon}^{A_N}$, $e_{\epsilon}^{B_N}$ increase significantly with increasing N to $e_{\epsilon}^{A_N} \approx e_{\epsilon}^{B_N} \approx 1$ for N = 200.

The reason for such a large error is again the high condition number of the transformation matrix E_N , which is involved in the computation of the matrices \tilde{A}_N, \tilde{B}_N . From a computational point of view, the matrices \tilde{A}_N, \tilde{B}_N are so ill-conditioned that even with $\epsilon \approx 0.9$, the computation of the spectrum is affected by strong numerical instabilities. Here we want to show that, even in the case of the singular values, the computation is affected by numerical instability, and the results are not reliable. Figure 3.6 provides the singular values of the matrices $A_N, \tilde{A}_N, B_N, \tilde{B}_N$. Again, we note that the strong ill-conditioning of the matrices \tilde{A}_N, \tilde{B}_N leads to large errors in the computation of $\Sigma(\tilde{A}_N), \Sigma(\tilde{B}_N)$, which are visibly different from $\Sigma(A_N), \Sigma(B_N)$.

ETNA Kent State University and Johann Radon Institute (RICAM)

87

UPPER HESSENBERG AND TOEPLITZ BOHEMIAN MATRIX SEQUENCES

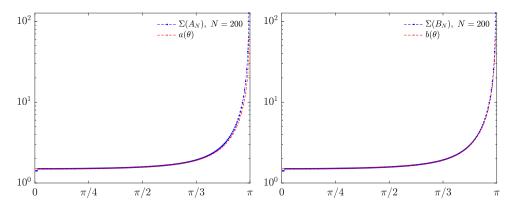


FIG. 3.7. Comparison between the singular values of A_N (left), B_N (right) and their monotonically increasing respective symbols.

On the other hand, Figure 3.7 shows that, according to Remark 2.8, the uniform sampling of the functions |a(x)|, |b(x)| overlaps the plot of $\Sigma(A_N)$, $\Sigma(B_N)$, ordered from the smallest to the largest. This shows that the theoretical derivations are much more informative than the numerical calculations. In other words, the numerical experiments do not give the right information due to numerical instabilities: the theoretical derivations are much more informative.

We now use the limiting set concept to provide a visual understanding of the spectrum of A_N and B_N . Let \mathbb{T} be the unit complex circumference, and write $t = e^{i\theta}$. For a function $h: \mathbb{T} \to \mathbb{C}$ in $L^1(\mathbb{T})$, the *limiting set* of the Toeplitz matrix sequence $\{T_N(h)\}_N$ is given by

$$LS(h) := \liminf_{N \to \infty} \Lambda(T_N(h)).$$

According to the work of Schmidt and Spitzer [20], when h is a Laurent polynomial, the limiting set LS(h) turns out to be a finite union of analytic arcs, and we can expect that the eigenvalues of $T_N(h)$ approximate it (in the Hausdorff metric) as $N \to \infty$. In such a case, LS(h) can be obtained by

(3.3)
$$\operatorname{LS}(h) = \bigcap_{\varrho \in (0,\infty)} \Lambda(T(h_{\varrho})).$$

where $h_{\varrho}(t) := h(\varrho t)$ and $T(h_{\varrho})$ is the well-known Toeplitz operator with symbol h_{ϱ} . Thanks to the classical Coburn lemma (see, e.g., [6, Theorem 1.10, p. 20]),

$$\Lambda(T(h_{\rho})) = \mathcal{R}(h_{\rho}) \cup \{\lambda \in \mathbb{C} \setminus \mathcal{R}(h_{\rho}) : \text{ wind}(h_{\rho}, \lambda) \neq 0\},\$$

where wind (h_{ϱ}, λ) stands for the winding number of h_{ϱ} around the point λ . In plane words, the spectrum of $T(h_{\varrho})$ is the range of h_{ϱ} together with all its enclosed points in \mathbb{C} . The result (3.3) was extended by Day in [10] to rational functions. Since we are interested in the spectra of A_N and B_N , for $N \in \mathbb{N}$, we can take a look at the respective limiting sets.

From (3.2) we obtain

$$\tilde{a}_{\epsilon}(\theta) = \left(\frac{1}{\epsilon}\right)t - \frac{\left(\frac{1}{\epsilon}\right)t}{\left(\frac{1}{\epsilon}\right)t - 1}, \quad \tilde{b}_{\epsilon}(\theta) = \left(\frac{1}{\epsilon}\right)t + \frac{\left(\frac{1}{\epsilon}\right)t}{\left(\frac{1}{\epsilon}\right)t + 1},$$

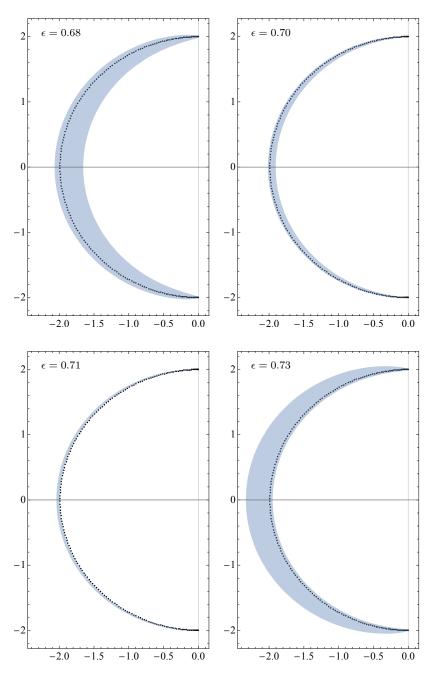


FIG. 3.8. The spectrum of A_N for N = 200 (black dots), and the spectrum of $T(\tilde{a}_{\epsilon})$ for different values of ϵ (shaded regions).

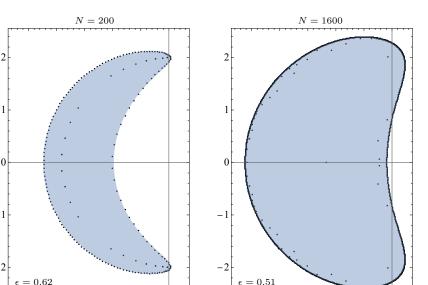


FIG. 3.9. The eigenvalues of \tilde{A}_N (black dots) and the spectrum of $T(\tilde{a}_{\epsilon})$ (shaded regions).

-2.5

-2.0

-1.5

-1.0 -0.5

0.0

and hence \tilde{a}_{ϵ} and \tilde{b}_{ϵ} are rational functions in the variable t. Using the relations

0.0

$$[\tilde{a}_{\epsilon}]_{\varrho}(\theta) = \left(\frac{\varrho}{\epsilon}\right)t - \frac{\left(\frac{\varrho}{\epsilon}\right)t}{\left(\frac{\varrho}{\epsilon}\right)t - 1} = \tilde{a}_{\epsilon/\varrho}(\theta), \quad [\tilde{b}_{\epsilon}]_{\varrho}(\theta) = \left(\frac{\varrho}{\epsilon}\right)t + \frac{\left(\frac{\varrho}{\epsilon}\right)t}{\left(\frac{\varrho}{\epsilon}\right)t + 1} = \tilde{b}_{\epsilon/\varrho}(\theta),$$

together with (3.3) we obtain

-2.5

-2.0 -1.5 -1.0 -0.5

$$\begin{split} \mathrm{LS}(\tilde{a}_{\epsilon}) &= \ \bigcap_{\varrho \in (0,\infty)} \Lambda(T([\tilde{a}_{\epsilon}]_{\varrho})) = \bigcap_{\epsilon \in (0,\infty)} \Lambda(T(\tilde{a}_{\epsilon})), \\ \mathrm{LS}(\tilde{b}_{\epsilon}) &= \ \bigcap_{\varrho \in (0,\infty)} \Lambda(T([\tilde{b}_{\epsilon}]_{\varrho})) = \bigcap_{\epsilon \in (0,\infty)} \Lambda(T(\tilde{b}_{\epsilon})). \end{split}$$

The previous equations suggest a strategy for approximating a limiting set: plot the spectrum $\Lambda(T(\tilde{a}_{\epsilon}))$ for different values of ϵ , trying to get the "thinner" possible region. Figure 3.8 illustrates a sequence of four images plotting $\Lambda(A_N)$, for N = 200, together with $\Lambda(T(\tilde{a}_{\epsilon}))$ for carefully selected values of ϵ . From these plots, it is easy to deduce that the set $\mathrm{LS}(\tilde{a}_{\epsilon}) = \{z \in \mathbb{C} : |z| = 2, \operatorname{Re}(z) \leq 0\}$ and hence the eigenvalues of A_N must be arbitrarily close to this set. A similar analysis can be applied to \tilde{b}_{ϵ} producing the equality $\mathrm{LS}(\tilde{b}_{\epsilon}) = \{z \in \mathbb{C} : |z| = 2, \operatorname{Re}(z) \geq 0\}.$

We now turn our attention to the spectra of \hat{A}_N and \hat{B}_N . A classical theorem (see, e.g., [4, Corollary 2.18, p. 28]) asserts that in our case the eigenvalues of \tilde{A}_N are contained in $\Lambda(T(\tilde{a}_{\epsilon}))$ with the possible exception of o(1) outliers for every sufficiently large N. We can use this result to explain the behavior of $\Lambda(\tilde{A}_N)$ and $\Lambda(\tilde{B}_N)$ shown in Figures 3.2 and 3.4. Figure 3.9 shows that most of the eigenvalues of \tilde{A}_N approach the boundary of $\Lambda(T(\tilde{a}_{\epsilon}))$ (i.e. $\mathcal{R}(\tilde{a}_{\epsilon}))$ for selected values of ϵ . Thus, because of the numerical restrictions induced by the severe ill-conditioning of the considered matrices, the eigenvalues of \tilde{A}_N are evolving like the dynamic process sketched in Figure 3.8.

89

ETNA Kent State University and Johann Radon Institute (RICAM)

4. Concluding remarks and open problems. In the previous literature, Bohemian matrices have attracted the attention of several researchers for their rich combinatorial structure, and they have been studied intensively from several points of view, including height, determinants, characteristic polynomials, normality, and stability.

Here we considered a selected number of examples of upper Hessenberg and Toeplitz Bohemian matrix sequences whose entries belong to the population $P = \{0, \pm 1\}$. In this setting, we proposed a connection with the spectral theory of Toeplitz matrix sequences and Generalized Locally Toeplitz (GLT) matrix sequences in order to give localization and asymptotical distribution results on their spectra and singular values. Numerical experiments support the mathematical study.

For appreciating our study, we refer to the last 7 lines of the concluding remarks in [9] regarding the spectral properties of such matrices: indeed we can affirm that our Toeplitz and GLT machinery sheds some light on the spectral and singular value structure of upper Hessenberg and Toeplitz Bohemian matrix sequences with the population $P = \{0, \pm 1\}$. Our mathematical tools do not depend on the population P, and hence, more cases can be considered as long as we remain in the GLT setting.

For giving a concrete example of the potential of our spectral tools, consider an example of upper Hessenberg Toeplitz Bohemian matrix with increasing entries belonging to the integers, that is,

$$X_N = \begin{bmatrix} -1 & -2 & -3 & \cdots & -N+1 \\ 1 & -1 & -2 & \cdots & -N+2 \\ 0 & 1 & -1 & \cdots & -N+3 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -1 \end{bmatrix}.$$

We have

$$X_N = T_N(e^{i\theta}) - R_N^{-2} = T_N(e^{i\theta}) - T_N^{-2}(1 - e^{-i\theta}).$$

Therefore,

$$\{X_N\}_N \sim_{\text{GLT}} e^{i\theta} - \frac{1}{(1 - e^{-i\theta})^2}$$

and consequently all the Toeplitz and GLT machinery of our note can be used successfully again, including the Beam-Warming "replacement trick" [2].

A further direction which deserves to be explored is the case where the population is constituted by a selection of small blocks of fixed dimension instead of scalars. In that case the Toeplitz and GLT machinery can be used again but in the block setting, which has been developed in detail in [1]; see also the references therein.

Finally, we emphasize that the high numerical instability in computing the eigenvalues of the proposed large matrices gives an indication that we are far from a normal setting. The latter is also confirmed by the fact that the singular values and their asymptotic symbol differ substantially from the moduli of the eigenvalues, as indicated theoretically by the use of the Beam-Warming "replacement trick". In this direction a connection with the theory of ϵ -pseudospectra (see [18, 19] and the references there reported) should exist, and also this direction deserves to be the subject of future investigations.

REFERENCES

- [1] G. BARBARINO, C. GARONI, AND S. SERRA-CAPIZZANO, Block generalized locally Toeplitz sequences: theory and applications in the unidimensional case, Electron. Trans. Numer. Anal., 53 (2020), pp. 28–112. http://etna.ricam.oeaw.ac.at/vol.53.2020/pp28-112.dir/pp28-112.pdf
- R. M. BEAM AND R. F. WARMING, The asymptotic spectra of banded Toeplitz and quasi-Toeplitz matrices, SIAM J. Sci. Comput., 14 (1993), pp. 971–1006.
- B. BECKERMANN AND S. SERRA-CAPIZZANO, On the asymptotic spectrum of finite element matrix sequences, SIAM J. Numer. Anal., 45 (2007), pp. 746–769.
- [4] A. BÖTTCHER AND S. M. GRUDSKY, Toeplitz Matrices, Asymptotic Linear Algebra, and Functional Analysis, Birkhäuser, Basel, 2000.
- [5] ——, Spectral Properties of Banded Toeplitz Matrices, SIAM, Philadelphia, 2005.
- [6] A. BÖTTCHER AND B. SILBERMANN, Introduction to Large Truncated Toeplitz Matrices, Springer, New York, 1999.
- [7] C. BRIAT, Sign properties of Metzler matrices with applications, Linear Algebra Appl., 515 (2017), pp. 53-86.
- [8] E. Y. S. CHAN, R. M. CORLESS, L. GONZALEZ-VEGA, J. R. SENDRA, J. SENDRA, AND S. E. THORNTON, Upper Hessenberg and Toeplitz Bohemians, Linear Algebra Appl., 601 (2020), pp. 72–100.
- [9] ——, Upper Hessenberg and Toeplitz Bohemians, Preprint on arXiv, 2018. https://arxiv.org/abs/1809.10664
- [10] K. DAY, Toeplitz matrices generated by the Laurent series expansion of an arbitrary rational function, Trans. Amer. Math. Soc., 206 (1975), pp. 224–245.
- [11] Z. DU, C. M. DA FONSECA, Y. XU, AND J. YE, Disproving a conjecture of Thornton on Bohemian matrices, Open Math., 19 (2021), pp. 505–514.
- [12] M. FASI AND G. M. NEGRI PORZIO, Determinants of normalized bohemian upper Hessenberg matrices, Electronic J. Linear Algebra, 36 (2020), pp. 352–366.
- [13] C. GARONI, C. MANNI, F. PELOSI, S. SERRA-CAPIZZANO, AND H. SPELEERS, On the spectrum of stiffness matrices arising from isogeometric analysis, Numer. Math., 127 (2014), pp. 751–799.
- [14] C. GARONI AND S. SERRA-CAPIZZANO, Generalized Locally Toeplitz Sequences: Theory and Applications. Vol. I, Springer, Cham, 2017.
- [15] ——, Generalized Locally Toeplitz Sequences: Theory and Applications. Vol. II, Springer, Cham, 2018.
- [16] C. W. GEAR, A simple set of test matrices for eigenvalue programs, Math. Comp., 23 (1969), pp. 119–125.
- [17] E. KILIC AND D. TASCI, On the generalized Fibonacci and Pell sequences by Hessenberg matrices, Ars Combin., 94 (2010), pp. 161–174.
- [18] L. REICHEL AND L. N. TREFETHEN, Eigenvalues and pseudo-eigenvalues of Toeplitz matrices, Linear Algebra Appl., 162/164 (1992), pp. 153–185.
- [19] L. N. TREFETHEN AND M. EMBREE, Spectra and Pseudospectra. The Behavior of Nonnormal Matrices and Operators, Princeton University Press, Princeton, 2005.
- [20] P. SCHMIDT AND F. SPITZER, The Toeplitz matrices of an arbitrary Laurent polynomial, Math. Scand., 8 (1960), pp. 15–38.
- [21] S. SERRA-CAPIZZANO, Generalized locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations, Linear Algebra Appl., 366 (2003), pp. 371–402.
- [22] ——, The GLT class as a generalized Fourier analysis and applications, Linear Algebra Appl., 419 (2006), pp. 180–233.
- [23] S. SERRA-CAPIZZANO AND P. TILLI, On unitarily invariant norms of matrix-valued linear positive operators, J. Inequal. Appl., 7 (2002), pp. 309–330.
- [24] T. TAO AND V. VU, Random matrices have simple spectrum, Combinatorica, 37 (2017), pp. 539–553.
- [25] O. TAUSSKY, Matrices of rational integers, Bull. Amer. Math. Soc., 66 (1960), pp. 327–345.
- [26] _____, Some computational problems involving integral matrices, J. Res. Nat. Bur. Standards Sect. B, 65B (1961), pp. 15–17.
- [27] S. THORNTON, *The characteristic polynomial database*, available at http://bohemianmatrices.com/cpdb, Sept. 2018.
- [28] P. TILLI, Locally Toeplitz sequences: spectral properties and applications, Linear Algebra Appl., 278 (1998), pp. 91–120.
- [29] E. E. TYRTYSHNIKOV AND N. L. ZAMARASHKIN, Spectra of multilevel Toeplitz matrices: advanced theory via simple matrix relationships, Linear Algebra Appl., 270 (1998), pp. 15–27.
- [30] R. S. VARGA, Geršgorin and Jis Circles, Springer, Berlin, 2004.