Received: 23 October 2018 | Revised: 19 March 2019 | Accepted: 21 May 2019
DOI: 10.1002/nla.2257

RESEARCH ARTICLE

WILEY

Nearness results for real tridiagonal 2-Toeplitz matrices

Natália Bebiano[1,2] | **Susana Furtado**[3,4]

[1]Centro de Matemática, Universidade de Coimbra, Coimbra, Portugal
[2]Departamento de Matemática, Universidade de Coimbra, Coimbra, Portugal
[3]Center for Functional Analysis, Linear Structures and Applications (CEAFEL), Lisboa, Portugal
[4]Faculdade de Economia,Universidade do Porto, Porto, Portugal

Correspondence
Susana Furtado, Center for Functional Analysis, Linear Structures and Applications (CEAFEL), 1049-001 Lisboa, Portugal; or Faculdade de Economia, Universidade do Porto, 4200-464 Porto, Portugal.
Email: sbf@fep.up.pt

Funding information
Centre for Mathematics of the University of Coimbra, Grant/Award Number: (UID/MAT/00324/2019); Portuguese Government through FCT/MEC; European Regional Development Fund through the Partnership Agreement PT2020, Grant/Award Number: UID/MAT/04721/2019

Summary
In this article, we extend the results for Toeplitz matrices obtained by Noschese, Pasquini, and Reichel. We study the distance d, measured in the Frobenius norm, of a real tridiagonal 2-Toeplitz matrix T to the closure $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ of the set formed by the real irreducible tridiagonal normal matrices. The matrices in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$, whose distance to T is d, are characterized, and the location of their eigenvalues is shown to be in a region determined by the field of values of the operator associated with T, when T is in a certain class of matrices that contains the Toeplitz matrices. When T has an odd dimension, the eigenvalues of the closest matrices to T in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ are explicitly described. Finally, a measure of nonnormality of T is studied for a certain class of matrices T. The theoretical results are illustrated with examples. In addition, known results in the literature for the case in which T is a Toeplitz matrix are recovered.

KEYWORDS
eigenvalues, normal matrix, real tridiagonal matrix, structured distance, 2-Toeplitz matrix

AMS CLASSIFICATION
65F15; 65F35; 15B05; 15B57

1 | INTRODUCTION

Tridiagonal matrices have many applications in science. In several situations, these matrices exhibit further special structure such as symmetry, break of reflection or time symmetry, or an r-Toeplitz structure (see, e.g., other works[1–5] and the references therein). For example, several problems in physics and engineering are modeled by using one-dimensional chains[1] such as the Ising model or the structural model for the graphene. Chain models can be represented by tridiagonal r-Toeplitz matrices.

The problems of finding the distance of a matrix to a certain class of matrices have been investigated by several researchers; see other works for example[2–4,6–14] and the references therein. In particular, the distance of a given tridiagonal matrix to normality is of interest because the eigenproblem is well conditioned for normal matrices, whereas for tridiagonal matrices, it is very sensitive to perturbations. A small distance to normality may allow the replacement of the considered matrix by a closest normal matrix and the computation of the eigenvalues of the latter (see the work of Elsner et al.[15] for details on normal matrices).

In this paper, we will focus on $n \times n$ real tridiagonal 2-Toeplitz matrices of the form

$$\begin{bmatrix} \alpha_1 & \beta_1 & & & & & 0 \\ \gamma_1 & \alpha_2 & \beta_2 & & & & \\ & \gamma_2 & \alpha_1 & \beta_1 & & & \\ & & \gamma_1 & \alpha_2 & \beta_2 & & \\ & & & \gamma_2 & \ddots & \ddots & \\ & & & & \ddots & & \\ & & & & & & \\ 0 & & & & & & \end{bmatrix}, \tag{1}$$

which we denote by $T(n;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)$. In view of our results, characterizing the distance of these matrices to a tridiagonal normal matrix is a relevant problem because the spectrum of the closest tridiagonal normal matrix lives in a line segment that can be easily determined.

This paper gives results for 2-Toeplitz matrices that extend results by Noschese et al.,[2–4] where spectral problems and distance to normality of a tridiagonal Toeplitz matrix T are studied. Our results also rely on Gover's work[16] and on our paper,[17] in which the distance to normality of a general tridiagonal complex matrix has been analyzed.

This paper is organized as follows. In Section 2, we characterize the real tridiagonal normal matrices T and describe the location of the eigenvalues when T is 2-Toeplitz. In Section 3, we study the distance d, measured in the Frobenius norm, of T to $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$, the closure of the set formed by the real irreducible tridiagonal normal matrices. The matrices in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$, whose distance to T is d, are characterized, and the location of their eigenvalues is studied for a certain class of matrices T that includes the real tridiagonal Toeplitz matrices, namely, it is shown that these eigenvalues lie in the closure of the field of values, possibly shifted, of the operator associated with T. In Section 4, we study the location of the eigenvalues of T, when its dimension is odd. The eigenvalues of the closest matrix to T in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ are also described. We concentrate on matrices of odd dimension as, in this case, contrarily to the even case, explicit expressions for the eigenvalues of a tridiagonal 2-Toeplitz matrix are known.[16] In Section 5, we study a measure of normality for a certain class of real tridiagonal 2-Toeplitz matrices T of odd size, which includes the real tridiagonal Toeplitz matrices of odd size. In Section 6, the theoretical results are illustrated with some numerical examples. Finally, in Section 7, some concluding remarks are given.

Next, we summarize some notation used throughout the article.

$\mathcal{N}_{\mathcal{T}}$	closure of the subset of $\mathbb{C}^{n\times n}$ formed by the irreducible tridiagonal normal matrices
$\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$	closure of the set formed by the real irreducible tridiagonal normal matrices
$\mathcal{T}_2^{\mathbb{R}}$	set formed by the real tridiagonal 2-Toeplitz matrices

The *structured distance* of $T \in \mathcal{T}_2^{\mathbb{R}}$ to $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$, measured in the Frobenius norm, is denoted and defined as

$$d_F\left(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right) := \min_{T_{\mathcal{N}}\in\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}} \|T - T_{\mathcal{N}}\|_F.$$

We use a similar notation if, instead of $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$, we consider other closed sets of matrices.

2 | PREREQUISITES

We say that a real matrix Z is shifted antisymmetric if $Z+\alpha I$ is antisymmetric for some $\alpha \in \mathbb{R}$. In this case, $i(Z+\alpha I)$ is Hermitian.

In the work of Noschese et al.,[2] the following characterization of a real tridiagonal normal matrix was given.

Theorem 1. *Let T be a real tridiagonal matrix. Then, T is normal if and only if T is a direct sum of symmetric and shifted antisymmetric matrices.*

From Theorem 1, we can give a characterization of the real irreducible tridiagonal normal matrices.

Corollary 1. *Let T be a real irreducible tridiagonal matrix. Then, T is normal if and only if T is symmetric or T is shifted antisymmetric.*

Remark 1. From Corollary 1, we can conclude that the set $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ is formed by the real tridiagonal symmetric and shifted antisymmetric matrices.

Next, we describe the location of the eigenvalues of a real tridiagonal 2-Toeplitz normal matrix. We will show that these eigenvalues lie in the union of two line segments.

Remark 2. If $T \in \mathcal{T}_2^{\mathbb{R}}$ is a reducible normal matrix, then T is a direct sum of normal blocks of size at most 2 and, due to the structure of a 2-Toeplitz matrix, it can be easily verified that either T is symmetric or shifted antisymmetric and, therefore, according to Remark 1, $T \in \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$.

We need to introduce the 2-Toeplitz operator

$$\mathcal{T}(\infty; \beta_1\beta_2, \alpha_1, \alpha_2, \gamma_1\gamma_2) = \begin{bmatrix} \alpha_1 & \beta_1 & 0 & 0 & 0 & \dots \\ \gamma_1 & \alpha_2 & \beta_2 & 0 & 0 & \dots \\ 0 & \gamma_2 & \alpha_1 & \beta_1 & 0 & \dots \\ 0 & 0 & \gamma_1 & \alpha_2 & \beta_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{2}$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2 \in \mathbb{R}$. The symbol of the 2-Toeplitz operator $\mathcal{T}$ is the function $f : [0, 2\pi[\to \mathbb{R}^{2\times2}$, defined by

$$f(\phi) = \begin{bmatrix} \alpha_1 & \beta_1(\phi) \\ \gamma_1(\phi) & \alpha_2 \end{bmatrix}, \tag{3}$$

where

$$\beta_1(\phi) := \beta_1 + \gamma_2 e^{i\phi}, \qquad \gamma_1(\phi) := \gamma_1 + \beta_2 e^{-i\phi}. \tag{4}$$

The eigenvalues of $f(\phi)$ are

$$\frac{1}{2}(\alpha_1 + \alpha_2) \pm \frac{1}{2}\sqrt{(\alpha_1 - \alpha_2)^2 + 4\beta_1(\phi)\gamma_1(\phi)}.$$

The Toeplitz-Hausdorf theorem asserts that the field of values of $\mathcal{T}$, denoted by $W(\mathcal{T})$, is convex. Moreover, its closure, $\overline{W(\mathcal{T})}$, contains the spectrum of $\mathcal{T}$, $\sigma(\mathcal{T})$. If $\mathcal{T}$ is normal, then $\overline{W(\mathcal{T})}$ is the *convex hull* of $\sigma(\mathcal{T})$. For more details, see, for example, the work of Gustafson et al.[18]

Throughout, given $a, b \in \mathbb{C}$ and $S \subset \mathbb{C}$, we denote $a + bS = \{a + bx : x \in S\}$.

Theorem 2. *Suppose that $T(n; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$ is normal, then its eigenvalues are collinear and live in the line segment*

$$\frac{1}{2}(\alpha_1 + \alpha_2) + [-c, c]$$

where

$$c = \frac{1}{2}\sqrt{(\alpha_1 - \alpha_2)^2 + 4(|\beta_1| + |\beta_2|)^2}, \tag{5}$$

if T is symmetric, and

$$c = i(|\beta_1| + |\beta_2|), \tag{6}$$

if T is shifted antisymmetric.

Proof. By Corollary 1 and Remark 2, T is either symmetric or shifted antisymmetric. Consider the 2-Toeplitz operator $\mathcal{T}(\infty; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2)$. The eigenvalues of the symbol of $\mathcal{T}$ are

$$\lambda_\pm(\phi) = \frac{1}{2}(\alpha_1 + \alpha_2) \pm \frac{1}{2}\sqrt{(\alpha_1 - \alpha_2)^2 + 4\beta_1^2 + 4\beta_2^2 + 8\beta_1\beta_2\cos(\phi)}, \quad 0 \le \phi < 2\pi,$$

if T is symmetric, and

$$\lambda_\pm(\phi) = \alpha \pm i\sqrt{\beta_1^2 + \beta_2^2 - 2\beta_1\beta_2\cos(\phi)}, \quad 0 \le \phi < 2\pi,$$

if T is shifted antisymmetric, where $\alpha_1 = \alpha_2 = \alpha$. Now, the proof follows from Theorem 2.2 in the work of Bebiano et al.[19] □

Remark 3. Considering the notation of Theorem 2, we may conclude, by the proof of this theorem, that, if $T(n; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2)$ is normal, then the closure of the field of values of the normal operator $\mathcal{T}(\infty; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2)$, $\overline{W(\mathcal{T})}$ is the line segment $(\alpha_1 + \alpha_2)/2 + [-c, c]$, which lies in the real axis if T is symmetric and in a shift of the imaginary axis if T is shifted antisymmetric.

3 | DISTANCE TO NORMALITY OF A REAL TRIDIAGONAL 2-TOEPLITZ MATRIX

3.1 | Distance to normality

We will use the following number associated with a given $n \times n$ matrix A:

$$K(A) := tr(A^2) - \frac{tr^2(A)}{n}.$$

Remark 4. Let $T(n; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2), D(n; 0, 0, \alpha_1, \alpha_2, 0, 0) \in \mathcal{T}_2^{\mathbb{R}}$. A calculation shows that

$$K(T-D) = tr(T-D)^2 = \begin{cases} n\beta_1\gamma_1 + (n-2)\beta_2\gamma_2, & \text{if } n \text{ is even} \\ (n-1)(\beta_1\gamma_1 + \beta_2\gamma_2), & \text{if } n \text{ is odd.} \end{cases}$$

Moreover,

$$\begin{aligned} K(T) &= K(T-D) + tr(D^2) - \frac{tr^2(D)}{n} \\ &= \begin{cases} K(T-D) + \frac{n}{4}(\alpha_1 - \alpha_2)^2, & \text{if } n \text{ is even} \\ K(T-D) + \frac{n^2-1}{4n}(\alpha_1 - \alpha_2)^2, & \text{if } n \text{ is odd.} \end{cases} \end{aligned}$$

Remark 5. If $T(n; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$ is such that $\beta_1\gamma_1 \geq 0$ and $\beta_2\gamma_2 \geq 0$, then $K(T) \geq 0$.

Remark 6. From Corollary 12 in the work of Bebiano et al.,[17] if $T(n; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$, there is a real matrix $\hat{T} \in \mathcal{N}_T$ that minimizes $\|T_{\mathcal{N}} - T\|_F$ over $\mathcal{N}_T$. Moreover, $\hat{T} \in \mathcal{T}_2^{\mathbb{R}}$. Because $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}} \subset \mathcal{N}_{\mathcal{T}}$, we have that $d_F(T, \mathcal{N}_{\mathcal{T}}) = d_F(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}})$.

The next result gives the distance of a matrix $T \in \mathcal{T}_2^{\mathbb{R}}$ to the set $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$.

Theorem 3. *Let $T(n; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$. If n is even,*

$$d_F^2\left(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right) = \begin{cases} \frac{n}{4}(\beta_1 - \gamma_1)^2 + \frac{n-2}{4}(\beta_2 - \gamma_2)^2, & \text{if } K(T) \geq 0 \\ \frac{n}{4}(\alpha_1 - \alpha_2)^2 + \frac{n}{4}(\beta_1 + \gamma_1)^2 + \frac{n-2}{4}(\beta_2 + \gamma_2)^2, & \text{if } K(T) < 0, \end{cases}$$

and, if n is odd,

$$d_F^2\left(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right) = \begin{cases} \frac{n-1}{4}\left((\beta_1 - \gamma_1)^2 + (\beta_2 - \gamma_2)^2\right), & \text{if } K(T) \geq 0 \\ \frac{n^2-1}{4n}(\alpha_1 - \alpha_2)^2 + \frac{n-1}{4}\left((\beta_1 + \gamma_1)^2 + (\beta_2 + \gamma_2)^2\right), & \text{if } K(T) < 0. \end{cases}$$

Proof. In Corollary 12 in the work of Bebiano et al.[17] and Remark 6, we have

$$d_F^2\left(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right) = \frac{n}{4}\left(\alpha_1^2 + \alpha_2^2 + \beta_1^2 + \gamma_1^2\right) + \frac{n-2}{4}\left(\beta_2^2 + \gamma_2^2\right) - \frac{n}{8}(\alpha_1 + \alpha_2)^2 - \frac{1}{2}\left|\frac{n}{4}(\alpha_1 - \alpha_2)^2 + n\beta_1\gamma_1 + (n-2)\beta_2\gamma_2\right|$$

if n is even, and

$$d_F^2\left(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right)=\frac{n+1}{4}\alpha_1^2+\frac{n-1}{4}\left(\alpha_2^2+\beta_1^2+\gamma_1^2+\beta_2^2+\gamma_2^2\right)$$
$$-\frac{1}{8n}((n+1)\alpha_1+(n-1)\alpha_2)^2-\frac{n-1}{2}\left|\frac{n+1}{4n}(\alpha_1-\alpha_2)^2+\beta_1\gamma_1+\beta_2\gamma_2\right|$$

if n is odd. Thus, the result follows by some calculations and taking into account Remark 4. □

We observe that in the statement of Theorem 3, the expressions for $d_F^2(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}})$ in both branches ($K(T)\ \geq\ 0$ and $K(T)<0$) take the same value when $K(T)=0$.

Next, we give an upper bound for the normalized distance of a matrix in $\mathcal{T}_2^{\mathbb{R}}$ to $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$.

Corollary 2. *Let $T(n;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)\in\mathcal{T}_2^{\mathbb{R}}$. Then,*

$$\frac{d_F^2\left(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right)}{\|T\|_F^2}\leq\frac{1}{2}.$$

Proof. Case 1. Suppose that n is even and $K(T)\ \geq\ 0$. Then,

$$\|T_0\|_F^2=\frac{n}{2}\left(\beta_1^2+\gamma_1^2+\alpha_1^2+\alpha_2^2\right)+\frac{n-2}{2}\left(\beta_2^2+\gamma_2^2\right).$$

Taking into account Theorem 3 and Remark 4, we have

$$\frac{d_F^2\left(T_0,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right)}{\|T_0\|_F^2}=\frac{1}{2}\frac{n(\beta_1-\gamma_1)^2+(n-2)(\beta_2-\gamma_2)^2}{n\left(\beta_1^2+\gamma_1^2+\alpha_1^2+\alpha_2^2\right)+(n-2)\left(\beta_2^2+\gamma_2^2\right)}\leq\frac{1}{2},$$

where the inequality follows because

$$n\left(\beta_1^2+\gamma_1^2+\alpha_1^2+\alpha_2^2\right)+(n-2)\left(\beta_2^2+\gamma_2^2\right)-\left(n(\beta_1-\gamma_1)^2+(n-2)(\beta_2-\gamma_2)^2\right)$$
$$=n\alpha_1^2+n\alpha_2^2+2(n-2)\beta_2\gamma_2+2n\beta_1\gamma_1=2K(T)+\frac{n}{2}(\alpha_1+\alpha_2)^2\geq 0.$$

The proof is similar if $K(T)<0$.

Case 2. Suppose that n is odd and $K(T)\ \geq\ 0$. Then,

$$\|T_0\|_F^2=\frac{n-1}{2}\left(\beta_1^2+\beta_2^2+\gamma_1^2+\gamma_2^2+\alpha_2^2\right)+\frac{n+1}{2}\alpha_1^2.$$

Taking into account Theorem 3 and Remark 4, we have

$$\frac{d_F^2\left(T_0,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right)}{\|T_0\|_F^2}=\frac{1}{2}\frac{(n-1)\left((\beta_1-\gamma_1)^2+(\beta_2-\gamma_2)^2\right)}{(n-1)\left(\beta_1^2+\beta_2^2+\gamma_1^2+\gamma_2^2+\alpha_2^2\right)+(n+1)\alpha_1^2}\leq\frac{1}{2},$$

where the inequality follows because

$$(n-1)\left(\beta_1^2+\beta_2^2+\gamma_1^2+\gamma_2^2+\alpha_2^2\right)+(n+1)\alpha_1^2-(n-1)\left((\beta_1-\gamma_1)^2+(\beta_2-\gamma_2)^2\right)$$
$$=\alpha_1^2-\alpha_2^2-2\beta_1\gamma_1-2\beta_2\gamma_2+n\alpha_1^2+n\alpha_2^2+2n\beta_1\gamma_1+2n\beta_2\gamma_2$$
$$=2K(T)+\frac{1}{2n}(\alpha_1-\alpha_2+n\alpha_1+n\alpha_2)^2\geq 0.$$

The proof is similar if $K(T)<0$. □

Remark 7. The distance to $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ of a general real reducible tridiagonal normal matrix T may be nonzero, as T may not belong to $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$. However, if in addition $T \in \mathcal{T}_2^{\mathbb{R}}$, by Remark 2, $T \in \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$, implying that this distance is zero.

3.2 | Closest normal matrix

The next result gives the matrices in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ with minimal distance to a given matrix $T \in \mathcal{T}_2^{\mathbb{R}}$. This result is a simple consequence of Remark 6 and Corollary 12 in the work of Bebiano et al.[17]

Theorem 4. *Let $T(n;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$. Then, $d_F(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}) = \|\hat{T}-T\|_F$, where $\hat{T} = \hat{T}(n;b_1,b_2,a_1,a_2,c_1,c_2)$ with*

$$a_1 = \alpha_1, \qquad a_2 = \alpha_2,$$
$$b_j = c_j = \frac{\beta_j+\gamma_j}{2}, \qquad j = 1,2,$$

if $K(T) \geq 0$, or

$$a_1 = a_2 = \begin{cases} \frac{\alpha_1+\alpha_2}{2}, & \text{if } n \text{ is even} \\ \frac{(n+1)\alpha_1+(n-1)\alpha_2}{2n}, & \text{if } n \text{ is odd,} \end{cases}$$
$$b_j = \frac{\beta_j-\gamma_j}{2}, \qquad c_j = \frac{\gamma_j-\beta_j}{2} \qquad j = 1,2,$$

if $K(T) \leq 0$. Moreover, if $K(T) \neq 0$, then $\hat{T}$ is the unique matrix in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ that minimizes $d_F(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}})$.

Corollary 3. *Let $T(n;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$. Suppose that $K(T) \neq 0$ and let $\hat{T} = \hat{T}(n;b_1,b_2,a_1,a_2,c_1,c_2)$ be the (unique) closest matrix to T in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$. If $K(T) > 0$, the eigenvalues of $\hat{T}$ are real. If $K(T) < 0$, the eigenvalues of $\hat{T}$ lie in a shift of the imaginary axis.*

Proof. The result follows from Theorem 4 taking into account that $\hat{T}$ is real symmetric if $K(T) > 0$, and $i(\hat{T}+\alpha I)$ is Hermitian for some $\alpha \in \mathbb{R}$, if $K(T) < 0$. □

Next, we denote by $\mathcal{S}_{\mathcal{T}}^{\mathbb{R}}$ the subset formed by the real tridiagonal symmetric matrices and by $\mathcal{A}_{\mathcal{T}}^{\mathbb{R}}$ the subset formed by the real tridiagonal shifted antisymmetric matrices. Because, by Remark 1, $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}} = \mathcal{S}_{\mathcal{T}}^{\mathbb{R}} \cup \mathcal{A}_{\mathcal{T}}^{\mathbb{R}}$, we have

$$d_F\left(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right) = d_F\left(T,\mathcal{S}_{\mathcal{T}}^{\mathbb{R}} \cup \mathcal{A}_{\mathcal{T}}^{\mathbb{R}}\right) = \min\left\{d_F\left(T,\mathcal{S}_{\mathcal{T}}^{\mathbb{R}}\right), d_F\left(T,\mathcal{A}_{\mathcal{T}}^{\mathbb{R}}\right)\right\}.$$

Corollary 4. *Let $T(n;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$.*

- *If $K(T) \geq 0$, then $d_F(T,\mathcal{S}_{\mathcal{T}}^{\mathbb{R}}) = d_F(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}})$.*
- *If $K(T) \leq 0$, then $d_F(T,\mathcal{A}_{\mathcal{T}}^{\mathbb{R}}) = d_F(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}})$.*

In particular, if $K(T) = 0$ then $d_F(T,\mathcal{S}_{\mathcal{T}}^{\mathbb{R}}) = d_F(T,\mathcal{A}_{\mathcal{T}}^{\mathbb{R}}) = d_F(T,\mathcal{N}_{\mathcal{T}}^{\mathbb{R}})$.

Proof. From Theorem 4, the unique matrix T that minimizes $\|T_{\mathcal{N}} - T\|_F$ over $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ is symmetric if $K(T) > 0$ and is shifted antisymmetric if $K(T) < 0$. If $K(T) = 0$, there is a symmetric and a shifted antisymmetric matrix that minimizes $\|T_{\mathcal{N}} - T\|_F$ over $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$. Because $\mathcal{S}_{\mathcal{T}}^{\mathbb{R}} \subset \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ and $\mathcal{A}_{\mathcal{T}}^{\mathbb{R}} \subset \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$, the claim follows. □

3.3 | Extension of a result for Toeplitz matrices

In the work of Noschese et al.,[3] the authors showed that the eigenvalues of the normal tridiagonal matrix closest to a given tridiagonal Toeplitz matrix T lie in the closure of the field of values of the operator associated with T. In this section, we consider the real case and extend this result to the class $\mathcal{L}_2^{\mathbb{R}}$ of 2-Toeplitz matrices, which we will introduce next. This class contains the real tridiagonal Toeplitz matrices and the normal matrices in $\mathcal{T}_2^{\mathbb{R}}$.

We denote by $\mathcal{L}_2^{\mathbb{R}}$ the class of matrices $T(n;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$, and their transposes, satisfying one of the following conditions:

- $\gamma_1 = k\beta_1$ and $\beta_2 = k\gamma_2$ for some $k \in \mathbb{R}$;
- $\gamma_1 = k\beta_1$ and $\gamma_2 = k\beta_2$ for some $k \in \mathbb{R}$.

The following observations will be used in the proof of Theorem 5.

Remark 8. The field of values of $T(n;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)$ remains unchanged if β_1 is interchanged with γ_1 or β_2 with γ_2.[20] Moreover, if $T \in \mathcal{L}_2^{\mathbb{R}}$, it can be easily seen from the proof of Theorem 4.2 in the work of Bebiano et al.[19] that the same holds with the field of values of the associated operator. In addition, from Theorem 4, it follows that the matrices in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ closest to $T \in \mathcal{T}_2^{\mathbb{R}}$ are invariant, up to a diagonal unitary similarity, for such interchanges in T.

In Theorem 4.2 in the work of Bebiano et al.,[19] the description of the field of values of an operator $\mathcal{T}(\infty;\beta_1,k\gamma_2,\alpha_1,\alpha_2,k\beta_1,\gamma_2)$ is given. We consider next the case $\mathscr{T}$ is nonnormal, as otherwise the result is given in Remark 3.

Lemma 1 (See the work of Bebiano et al.[19]). *Consider the nonnormal real operator $\mathcal{T}(\infty;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)$. Suppose that $\gamma_1 = k\beta_1$ and $\beta_2 = k\gamma_2$, for some $k \in \mathbb{R}$. Then, the boundary of $\overline{W(\mathcal{T})}$,$\partial\overline{W(\mathcal{T})}$ is the ellipse*

$$\frac{(x-(\alpha_1+\alpha_2)/2)^2}{A^2} + \frac{y^2}{B^2} = 1,$$

where

$$A^2 = \frac{1}{4}\left((|\beta_1| + |\gamma_2|)^2(k+1)^2 + (\alpha_1-\alpha_2)^2\right),$$

$$B^2 = \frac{1}{4}(|\beta_1| + |\gamma_2|)^2(k-1)^2.$$

We now show that the eigenvalues of a closest normal matrix in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ to a given matrix $T \in \mathcal{L}_2^{\mathbb{R}}$ lie in the field of values (possibly shifted) of the operator $\mathcal{T}$ associated with T. In Theorem 8, we explicitly give the eigenvalues of the matrices in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ closest to a given $T \in \mathcal{T}_2^{\mathbb{R}}$ of odd size, without requiring that $T \in \mathscr{L}_2^{\mathbb{R}}$.

Theorem 5. *Let $T(n;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathcal{L}_2^{\mathbb{R}}$ and consider the operator $\mathcal{T}(\infty;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)$. Then, the eigenvalues of a closest matrix $\widehat{T}$ to T in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ lie*

- *in $\overline{W(\mathcal{T})}$ if $\widehat{T}$ is symmetric or n is even,*
- *in a shift of $\overline{W(\mathcal{T})}$, under the vector $(\frac{\alpha_1-\alpha_2}{2n}, 0)$, if $\widehat{T}$ is shifted antisymmetric and n is odd.*

Proof. If T is normal then, by Corollary 1 and Remarks 1 and 2, $T \in \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$. Then, $\widehat{T} = T$ and the result follows from Theorem 2 and Remark 3.

Now, suppose that T is nonnormal. According to Remark 8, we now may assume that $\gamma_1 = k\beta_1$ and $\beta_2 = k\gamma_2$, where $k \in \mathbb{R}$. The boundary of $\overline{W(\mathcal{T})}$ is the ellipse described in Lemma 1. Suppose that $\widehat{T} \in \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ minimizes the distance of T to $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$. Taking into account Theorem 4, we consider two cases.

Case 1: Suppose that $\widehat{T}$ is symmetric. By Theorems 2 and 4, the spectrum of $\widehat{T}$ lies in $(\alpha_1+\alpha_2)/2 + [-c, c]$, where

$$c^2 = \frac{1}{4}\left((|\beta_1| + |\gamma_2|)^2(k+1)^2 + (\alpha_1-\alpha_2)^2\right),$$

which implies the claim.

Case 2. Suppose that $\widehat{T}$ is shifted antisymmetric.

Case 2.1. Suppose that n is even. By Theorems 2 and 4, the spectrum of $\widehat{T}$ lies in $(\alpha_1+\alpha_2)/2 + i[-c, c]$, where

$$c^2 = \frac{1}{4}(|\beta_1| + |\gamma_2|)^2(k-1)^2, \tag{7}$$

implying the claim.

Case 2.2. Suppose that n is odd. By Theorems 2 and 4, each eigenvalue λ of $\widehat{T}$ lies in $\frac{(n+1)\alpha_1+(n-1)\alpha_2}{2n} + i[-c, c]$, where c satisfies (7). As in the previous case, $\lambda + \frac{\alpha_1+\alpha_2}{2} - \frac{(n+1)\alpha_1+(n-1)\alpha_2}{2n}$ lies in $\overline{W(\mathscr{T})}$, and the claim follows. □

4 | EIGENVALUES OF $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)$

4.1 | Geometrical location of the eigenvalues

In Theorem 2.3 in the work of Gover,[16] the next result was obtained. We use the following notation:

$$s_1 := (\alpha_1-\alpha_2)^2 + 4(\beta_1\gamma_1+\beta_2\gamma_2) \text{ and } s_2 := 8\sqrt{\beta_1\beta_2\gamma_1\gamma_2}. \tag{8}$$

Proposition 1. *The eigenvalues of the matrix $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$ are α_1 and*

$$\lambda_r^{\pm} = \frac{1}{2}(\alpha_1+\alpha_2) \pm \frac{1}{2}\sqrt{s_1 + s_2\cos\frac{r\pi}{m+1}},$$

$r = 1, \ldots, m.$

We observe that an explicit expression for the eigenvalues of a real tridiagonal 2-Toeplitz matrix of even dimension seems difficult to obtain (see Theorem 2.4 in the work of Gover[16]). Therefore, from now on, we will restrict our attention to real tridiagonal 2-Toeplitz matrices $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)$ of odd dimension.

The following observation will be used in Section 5.

Remark 9. Consider the real tridiagonal 2-Toeplitz matrix $T(n;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)$ with n odd.

- If $\beta_1\gamma_1 \ge 0$ and $\beta_2\gamma_2 \ge 0$, then $s_1 - s_2 = (\alpha_1-\alpha_2)^2 + 4(\sqrt{\beta_1\gamma_1} - \sqrt{\beta_2\gamma_2})^2 \ge 0$.
- If $\beta_1\gamma_1 \le 0$ and $\beta_2\gamma_2 \le 0$, we have

$$s_1 \pm s_2 = (\alpha_1-\alpha_2)^2 - 4\left(\sqrt{|\beta_1\gamma_1|} \mp \sqrt{|\beta_2\gamma_2|}\right)^2.$$

We now describe the location of the eigenvalues of a matrix $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathscr{T}_2^{\mathbb{R}}$. We first assume that $\beta_1\beta_2\gamma_1\gamma_2 \ge 0$. In this case, we will use the following notation. For $r, r_1, r_2 \in \{1, \ldots, m\}$, we denote

$$a(r) = \sqrt{\left|s_1 + s_2\cos\frac{r\pi}{m+1}\right|},$$

and

$$S(r_1, r_2) = [a(r_1), a(r_2)].$$

If there is $r \in \{1, \ldots, m\}$ such that

$$s_1 + s_2\cos\frac{r\pi}{m+1} < 0, \tag{9}$$

we denote by m_0 the smallest such r, otherwise let $m_0 = m+1$. Note that (9) holds for any r with $m_0 \le r \le m$.

Theorem 6. *Suppose that $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$ with $\beta_1\beta_2\gamma_1\gamma_2 \ge 0$. Then, the eigenvalues of T distinct from α_1 occur in pairs centered in $\frac{1}{2}(\alpha_1+\alpha_2)$ and satisfy the following.*

- *If $m_0 = m+1$, they lie in the union of the two closed real line segments:*

$$\frac{1}{2}(\alpha_1+\alpha_2) + \left(-\frac{1}{2}S(m,1) \cup \frac{1}{2}S(m,1)\right).$$

- *If $m_0 = 1$, they lie in the union of the two closed line segments on a shift of the imaginary axis:*

$$\frac{1}{2}(\alpha_1+\alpha_2) + \left(-\frac{i}{2}S(1,m) \cup \frac{i}{2}S(1,m)\right).$$

- *If $1 < m_0 \leq m$, they lie in the union of the four closed line segments:*

$$\frac{1}{2}(\alpha_1+\alpha_2)+\left(-\frac{1}{2}S(m_0-1,1)\cup\frac{1}{2}S(m_0-1,1)\cup-\frac{i}{2}S(m_0,m)\cup\frac{i}{2}S(m_0,m)\right).$$

Moreover, the end points of each segment are eigenvalues of T.

Proof. For $r=1,\dots,m_0-1$, we have
$$s_1+s_2\cos\frac{r\pi}{m+1}\geq 0,$$
implying that the eigenvalues $\lambda_r^{\pm}$ of T given in Proposition 1 are real. For $r=m_0,\dots,m$, we have

$$s_1+s_2\cos\frac{r\pi}{m+1}\leq 0.$$

implying that the eigenvalues $\lambda_r^{\pm}$ of T given in Proposition 1 lie in a shift of the imaginary axis. Therefore, the statement concerning the location of the eigenvalues follows taking into account that $|s_1+s_2\cos\frac{r\pi}{m+1}|$ decreases as $r\in\{1,\dots,m_0-1\}$ increases, and increases as $r\in\{m_0,\dots,m\}$ increases. □

We now assume that $\beta_1\gamma_1\beta_2\gamma_2<0$.

Theorem 7. *Suppose that $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)\in\mathcal{T}_2^{\mathbb{R}}$ with $\beta_1\gamma_1\beta_2\gamma_2<0$. Then, the eigenvalues of T distinct from α_1 lie in*
$$\frac{1}{2}(\alpha_1+\alpha_2)+\left\{x+iy : x^2-y^2=\frac{1}{4}s_1, |s_2|\cos\frac{m\pi}{m+1}\leq 2xy\leq|s_2|\cos\frac{\pi}{m+1}\right\}.$$

Proof. Consider the matrix $S=-\frac{\alpha_1+\alpha_2}{2}I_{2m+1}+T$. The eigenvalues of S are $\frac{\alpha_1-\alpha_2}{2}$ and

$$\lambda_r^{\pm}=\pm\frac{1}{2}\sqrt{s_1+s_2\cos\frac{r\pi}{m+1}},$$

$r=1,\dots,m$. We have
$$\left(\lambda_r^{\pm}\right)^2=\frac{1}{4}\left(s_1+i|s_2|\cos\frac{r\pi}{m+1}\right).$$
Thus, the square of the eigenvalues of S distinct from $\frac{\alpha_1-\alpha_2}{2}$ lie in the line segment

$$\frac{1}{4}s_1+i|s_2|\left[\cos\frac{m\pi}{m+1},\cos\frac{\pi}{m+1}\right],$$

implying that the eigenvalues of S distinct from $\frac{\alpha_1-\alpha_2}{2}$ lie in

$$\left\{x+iy : x^2-y^2=\frac{1}{4}s_1, |s_2|\cos\frac{m\pi}{m+1}\leq 2xy\leq|s_2|\cos\frac{\pi}{m+1}\right\}.$$

The claim concerning the eigenvalues of T follows easily from the one for the eigenvalues of S. □

4.2 | The case of normal matrices

We next give a more precise location of the eigenvalues of a real 2-Toeplitz tridiagonal normal matrix of odd size than the one in Theorem 2, as now the endpoints of the line segments are eigenvalues of the matrix.

Let $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)\in\mathcal{T}_2^{\mathbb{R}}$ be normal. Because of Corollary 1 and Remark 2, T is symmetric or shifted antisymmetric. Thus, $\beta_1\beta_2\gamma_1\gamma_2\geq 0$. Moreover, s_1 and s_2 defined in (8) take the form

$$s_1=(\alpha_1-\alpha_2)^2+4\left(\beta_1^2+\beta_2^2\right),\qquad s_2=8|\beta_1\beta_2|$$

if T is symmetric, and
$$s_1=-4\left(\beta_1^2+\beta_2^2\right),\qquad s_2=8|\beta_1\beta_2|$$
if T is shifted antisymmetric.

Next, we use the notation introduced before Theorem 6.

Remark 10.

- If T is symmetric, then $m_0 = m + 1$, as, for any $r \in \{1, \ldots, m\}$,

$$s_1 + s_2 \cos \frac{r\pi}{m+1} \geq s_1 - s_2 = (\alpha_1 - \alpha_2)^2 + 4(|\beta_1| + |\beta_2|)^2 \geq 0.$$

- If T is shifted antisymmetric (and nonscalar), then $m_0 = 1$, as, for any $r \in \{1, \ldots, m\}$,

$$s_1 + s_2 \cos \frac{r\pi}{m+1} \leq s_1 + s_2 \leq -4(|\beta_1| - |\beta_2|)^2 \leq 0,$$

  where the first inequality is strict if $\beta_1\beta_2 \neq 0$.

From Remark 10 and Theorem 6, we obtain the following result, which improves Theorem 2 for matrices of odd size.

Corollary 5. *Suppose that $T(2m+1; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$ is normal. Then, the eigenvalues of T distinct from α_1 occur in pairs centered in $\frac{1}{2}(\alpha_1 + \alpha_2)$ and satisfy the following:*

- *If T is symmetric, they lie in the union of the two closed real line segments:*

$$\frac{1}{2}(\alpha_1 + \alpha_2) + \left(-\frac{1}{2}S(m,1) \cup \frac{1}{2}S(m,1)\right).$$

- *If T is shifted antisymmetric, they lie in the union of the two closed line segments parallel to the imaginary axis:*

$$\alpha_1 + \left(-\frac{i}{2}S(1,m) \cup \frac{i}{2}S(1,m)\right).$$

  Moreover, the endpoints of each segment are eigenvalues of T.

Remark 11. If $T(2m+1; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$ is normal, the eigenvalues, if T is symmetric, or the imaginary parts of the eigenvalues, if T is shifted antisymmetric, of any principal submatrix of T of order $2m$ interlace those of T. We then obtain accurate inclusion regions for the spectrum of any even size matrix from the precise knowledge of the spectra of two consecutive nested odd size matrices.

As a consequence of Theorem 4 and Proposition 1, we conclude this section by giving explicit expressions for the eigenvalues of the matrices in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ closest to a given $T(2m+1; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2)$.

Theorem 8. *Let $T(2m+1; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$, and $\hat{T} = \hat{T}(2m+1; b_1, b_2, a_1, a_2, c_1, c_2) \in \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ with minimal distance to T. Then, the eigenvalues of $\hat{T}$ are*

- *α_1 and*

$$\frac{1}{2}(\alpha_1 + \alpha_2) \pm \frac{1}{2}\sqrt{s_1 + s_2 \cos \frac{r\pi}{m+1}}, r = 1, \ldots, m,$$

  where

$$s_1 = (\alpha_1 - \alpha_2)^2 + (\beta_1 + \gamma_1)^2 + (\beta_2 + \gamma_2)^2, \qquad s_2 = 2|(\beta_1 + \gamma_1)(\beta_2 + \gamma_2)|,$$

  if $\hat{T}$ is symmetric (in which case $K(T) \geq 0$), and

- *$\frac{(m+1)\alpha_1 + m\alpha_2}{2m+1}$ and*

$$\frac{(m+1)\alpha_1 + m\alpha_2}{2m+1} \pm \frac{1}{2}\sqrt{s_1 + s_2 \cos \frac{r\pi}{m+1}}, r = 1, \ldots, m,$$

  where

$$s_1 = -(\beta_1 - \gamma_1)^2 - (\beta_2 - \gamma_2)^2, \qquad s_2 = 2|(\beta_1 - \gamma_1)(\beta_2 - \gamma_2)|,$$

  if $\hat{T}$ is shifted antisymmetric (in which case $K(T) \leq 0$).

5 | DEPARTURE FROM NORMALITY OF A REAL TRIDIAGONAL 2-TOEPLITZ MATRIX OF ODD DIMENSION

In this section, we will study a measure of nonnormality of a real tridiagonal 2-Toeplitz matrix $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2)$ of odd dimension satisfying $\beta_1\gamma_1\beta_2\gamma_2 \geq 0$ and either $s_1-s_2 \geq 0$ or $s_1+s_2 \leq 0$, where s_1 and s_2 are as in (8). Note that this class of matrices includes

- the matrices $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$ satisfying $\beta_1\gamma_1 \geq 0$ and $\beta_2\gamma_2 \geq 0$, as in this case, by Remark 9, $s_1-s_2 \geq 0$;
- the matrices $T(2m+1,\beta_1,\beta_2,\alpha,\alpha,\gamma_1,\gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$ with constant diagonal and satisfying $\beta_1\gamma_1 \leq 0$ and $\beta_2\gamma_2 \leq 0$, as in this case, by Remark 9, $s_1+s_2 = -4(\sqrt{-\beta_1\gamma_1}-\sqrt{-\beta_2\gamma_2})^2 \leq 0$.

Observe that, in particular, this class contains the real tridiagonal Toeplitz matrices of odd size.

For $A \in \mathbb{C}^{n\times n}$ with eigenvalues $\lambda_1, \dots, \lambda_n$, the departure from normality

$$\Delta_F(A) = \sqrt{\|A\|_F^2 - \Sigma_{h=1}^n |\lambda_h|^2}$$

was introduced by Henrici[9] to measure the nonnormality of a matrix.

Clearly, we have

$$\frac{\Delta_F^2(A)}{\|A\|_F^2} \leq 1.$$

Theorem 9. *Consider the matrix $T(2m+1;\beta_1,\beta_2,\alpha_1,\alpha_2,\gamma_1,\gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$. Suppose that $\beta_1\beta_2\gamma_1\gamma_2 \geq 0$. Let s_1 and s_2 be as in (8).*

- *If $s_1-s_2 \geq 0$, then $\Delta_F(T) = \sqrt{m((\beta_1-\gamma_1)^2+(\beta_2-\gamma_2)^2)}$.*
- *If $s_1+s_2 \leq 0$, then $\Delta_F(T) = \sqrt{m((\alpha_1-\alpha_2)^2+(\beta_1+\gamma_1)^2+(\beta_2+\gamma_2)^2)}$.*

Proof. We have

$$\|T\|_F^2 = (m+1)\alpha_1^2 + m\left(\alpha_2^2+\beta_1^2+\beta_2^2+\gamma_1^2+\gamma_2^2\right).$$

In addition,

$$\Sigma_{h=1}^{2m+1}|\lambda_r|^2 = \alpha_1^2 + \Sigma_{r=1}^m\left(|\lambda_r^+|^2+|\lambda_r^-|^2\right),$$

where

$$\lambda_r^{\pm} = \frac{1}{2}(\alpha_1+\alpha_2) \pm \frac{1}{2}\sqrt{s_1+s_2\cos\frac{r\pi}{m+1}},$$

$r = 1, \dots, m$. Suppose that $\beta_1\beta_2\gamma_1\gamma_2 \geq 0$.

Recall that

$$\sum_{r=1}^m \cos\frac{r\pi}{m+1} = 0.$$

Suppose that $s_1+s_2 \leq 0$. Then, $s_1+s_2\cos\frac{r\pi}{m+1} \leq 0$ for any $r = 1, \dots, m$, implying that

$$\begin{aligned}|\lambda_r^+|^2 = |\lambda_r^-|^2 &= \frac{1}{4}(\alpha_1+\alpha_2)^2 - \frac{1}{4}\left(s_1+s_2\cos\frac{r\pi}{m+1}\right)\\ &= \alpha_1\alpha_2 - (\beta_1\gamma_1+\beta_2\gamma_2) - 2\sqrt{\beta_1\beta_2\gamma_1\gamma_2}\cos\frac{r\pi}{m+1}.\end{aligned}$$

Hence,

$$\begin{aligned}
\Delta_F^2(T) &= \|T\|_F^2 - \left(\alpha_1^2 + \sum_{r=1}^{m}\left(|\lambda_r^+|^2 + |\lambda_r^-|^2\right)\right) \\
&= (m+1)\alpha_1^2 + m\left(\alpha_2^2 + \beta_1^2 + \beta_2^2 + \gamma_1^2 + \gamma_2^2\right) \\
&\quad - \left(\alpha_1^2 + 2m(\alpha_1\alpha_2 - \beta_1\gamma_1 - \beta_2\gamma_2) - 4\sqrt{\beta_1\beta_2\gamma_1\gamma_2}\sum_{r=1}^{m}\cos\frac{r\pi}{m+1}\right) \\
&= m(\alpha_1-\alpha_2)^2 + m\left(\beta_1^2 + \beta_2^2 + \gamma_1^2 + \gamma_2^2\right) + 2m(\beta_1\gamma_1 + \beta_2\gamma_2) \\
&= m(\alpha_1-\alpha_2)^2 + m(\beta_1+\gamma_1)^2 + m(\beta_2+\gamma_2)^2.
\end{aligned}$$

Suppose that $s_1 - s_2 \geq 0$. Then, $s_1 + s_2\cos\frac{r\pi}{m+1} \geq 0$ for any $r = 1, \ldots, m$, implying that

$$|\lambda_r^{\pm}|^2 = \left(\frac{1}{2}(\alpha_1+\alpha_2) \pm \frac{1}{2}\sqrt{s_1 + s_2\cos\frac{r\pi}{m+1}}\right)^2,$$

and

$$\begin{aligned}
|\lambda_r^+|^2 + |\lambda_r^-|^2 &= \frac{1}{2}(\alpha_1+\alpha_2)^2 + \frac{1}{2}\left(s_1 + s_2\cos\frac{r\pi}{m+1}\right) \\
&= \alpha_1^2 + \alpha_2^2 + 2(\beta_1\gamma_1 + \beta_2\gamma_2) + 4\sqrt{\beta_1\beta_2\gamma_1\gamma_2}\cos\frac{r\pi}{m+1}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\Delta_F^2(T) &= \|T\|_F^2 - \left(\alpha_1^2 + \sum_{r=1}^{m}\left(|\lambda_r^+|^2 + |\lambda_r^-|^2\right)\right) \\
&= \|T\|_F^2 - \left(\alpha_1^2 + m\left(\alpha_1^2 + \alpha_2^2 + 2(\beta_1\gamma_1 + \beta_2\gamma_2)\right) + 4\sqrt{\beta_1\beta_2\gamma_1\gamma_2}\sum_{r=1}^{m}\cos\frac{r\pi}{m+1}\right) \\
&= (m+1)\alpha_1^2 + m\left(\alpha_2^2 + \beta_1^2 + \beta_2^2 + \gamma_1^2 + \gamma_2^2\right) - \left(\alpha_1^2 + m\left(\alpha_1^2 + \alpha_2^2 + 2(\beta_1\gamma_1 + \beta_2\gamma_2)\right)\right) \\
&= m\left(\beta_1^2 + \beta_2^2 + \gamma_1^2 + \gamma_2^2\right) - 2m(\beta_1\gamma_1 + \beta_2\gamma_2) \\
&= m(\beta_1-\gamma_1)^2 + m(\beta_2-\gamma_2)^2.
\end{aligned}$$

□

From Remark 9 and Theorem 9, we obtain the next result.

Corollary 6. *Consider the matrix $T(2m+1; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$. Suppose that $\beta_1\gamma_1 \geq 0$ and $\beta_2\gamma_2 \geq 0$. Then,*

$$\Delta_F(T) = \sqrt{m(\beta_1-\gamma_1)^2 + m(\beta_2-\gamma_2)^2}.$$

Corollary 7. *Consider the matrix $T(2m+1; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2) \in \mathcal{T}_2^{\mathbb{R}}$. Suppose that $\beta_1\gamma_1 \geq 0$ and $\beta_2\gamma_2 \geq 0$. Then,*

$$\Delta_F(T) = \sqrt{2}d_F\left(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right).$$

Proof. From Remark 5, we have $K(T) \geq 0$, where $K(T)$ is defined in Section 3.1. Thus, taking into account Theorem 3 and Corollary 6, the claim follows. □

When T is a real tridiagonal Toeplitz matrix, we obtain the following consequence of Theorem 9, presented in the work of Noschese et al.[3]

Corollary 8. *Consider the real tridiagonal Toeplitz matrix $T(2m+1; \beta, \beta, \alpha, \alpha, \gamma, \gamma)$. Then, $\Delta_F(T) = \sqrt{2m}\,||\beta| - |\gamma||$.*

Proof. The result follows from Theorem 9 taking into account that, by Remark 9, $s_1 - s_2 \geq 0$ if $\beta\gamma \geq 0$ and $s_1 + s_2 \leq 0$ otherwise, where $s_1 = 8\beta\gamma$ and $s_2 = 8|\beta\gamma|$. □

6 | NUMERICAL EXPERIMENTS

We next give some examples that illustrate the theory discussed above.

Example 1. Consider the tridiagonal 2-Toeplitz matrix $T(2m+1; 1, 2, a, -a, -2, -1)$, where $a \in \mathbb{R} \setminus \{0\}$. In order to study the location of the eigenvalues of T and of the closest matrices to T in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$, consider the 2-Toeplitz operator $\mathcal{T}(\infty; 1, 2, a, -a, -2, -1)$. By Lemma 1, we find that $\partial\overline{W(\mathcal{T})}$ is the ellipse

$$\frac{x^2}{1+a^2} + \frac{y^2}{9} = 1. \tag{10}$$

According to Proposition 1, the eigenvalues of T are a and $\pm\sqrt{a^2 - 4 + 4\cos\frac{r\pi}{m+1}}$, $r = 1, \dots, m$. By Remark 4, we have $K(T) = 4m(\frac{m+1}{2m+1}a^2 - 2)$. From Theorem 4,

$$d_F\left(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right) = \|\widehat{T}_1 - T\|_F$$

if $K(T) \geq 0$, and

$$d_F\left(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}\right) = \|\widehat{T}_2 - T\|_F$$

if $K(T) \leq 0$, where

$$\widehat{T}_1 = \widehat{T}_1\left(2m+1; -\frac{1}{2}, \frac{1}{2}, a, -a, -\frac{1}{2}, \frac{1}{2}\right) \quad \text{and} \quad \widehat{T}_2 = \widehat{T}_2\left(2m+1; \frac{3}{2}, \frac{3}{2}, \frac{a}{2m+1}, \frac{a}{2m+1}, -\frac{3}{2}, -\frac{3}{2}\right).$$

According to Theorem 8, the eigenvalues of $\widehat{T}_1$ are a and $\pm\frac{1}{2}\sqrt{4a^2 + 2 + 2\cos\frac{r\pi}{m+1}}$, $r = 1, \dots, m$, and the eigenvalues of $\widehat{T}_2$ are $\frac{a}{2m+1}$ and $\frac{a}{2m+1} \pm \frac{3}{2}i\sqrt{2 - 2\cos\frac{r\pi}{m+1}}$, $r = 1, \dots, m$.

If $K(T) \geq 0$, by Theorem 5, the eigenvalues of $\widehat{T}_1$ lie in the region $\overline{W(\mathcal{T})}$ bounded by the ellipse. If $K(T) \leq 0$, that is, $a^2 \leq 2(2 - \frac{1}{m+1})$, the eigenvalues of $\widehat{T}_2$ lie in a line segment that approaches the vertical axis of the ellipse as $m \to \infty$.

For $m = 4$, Figures 1 and 2 display the spectra in the complex plane of the matrix T (green crosses) and of the closest normal matrices $\widehat{T}_1$ and $\widehat{T}_2$ (blue circles) when $a = 4$ and $a = 1$, respectively. In each case, the ellipse is the boundary of $\overline{W(\mathcal{T})}$. The horizontal and vertical axes are the real and imaginary axes. The eigenvalues are computed with the MATLAB (version R2016a) function eig. Note that, for $a = 1$, we have $K(T) < 0$ whereas for $a = 4$, we have $K(T) > 0$.

In Figure 3, we study how the deviation of T from normality depends on a. We use two different measures for this deviation, the squared normalized departure $\frac{\Delta_F^2(T)}{\|T\|_F^2}$ and

$$\frac{N}{\|T\|_F^2} = \frac{\|TT^T - T^TT\|_F}{\|T\|_F^2},$$

which does not depend on the calculation of the eigenvalues of T. The horizontal axis represents the values of the parameter a. The thick lines represent the distances to the symmetric matrix $\widehat{T}_1$ and to the shifted antisymmetric matrix $\widehat{T}_2$. Note that the value $a = \sqrt{18/5}$ for which the two lines intersect is such that $K(T) = 0$. For $a < \sqrt{18/5}$, we have $K(T) < 0$, and the distance of T to the shifted antisymmetric matrix $\widehat{T}_2$ is smaller than to the symmetric matrix $\widehat{T}_1$, whereas for $a > \sqrt{18/5}$ we have $K(T) > 0$ and the distance of T to the shifted antisymmetric matrix $\widehat{T}_2$ is greater than to the symmetric matrix $\widehat{T}_1$ (see Theorem 4). The parts of the thick lines below the crossing gives $d_F^2(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}})/\|T\|_F^2$ (see Theorem 3). The dashing line represents $N/|T\|_F^2$ and the thin line represents $\Delta_F^2(T)/|T\|_F^2$. Note that, for $a^2 \geq 8$, $\Delta_F^2(T)/\|T\|_F^2$ can be obtained from Theorem 9, as in this case we have $s_1 - s_2 = 4a^2 - 32 \geq 0$.

We next include an example for a real tridiagonal 2-Toeplitz matrix $T(2m+1; \beta_1, \beta_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2)$ with $\beta_1\beta_2\gamma_1\gamma_2 < 0$.

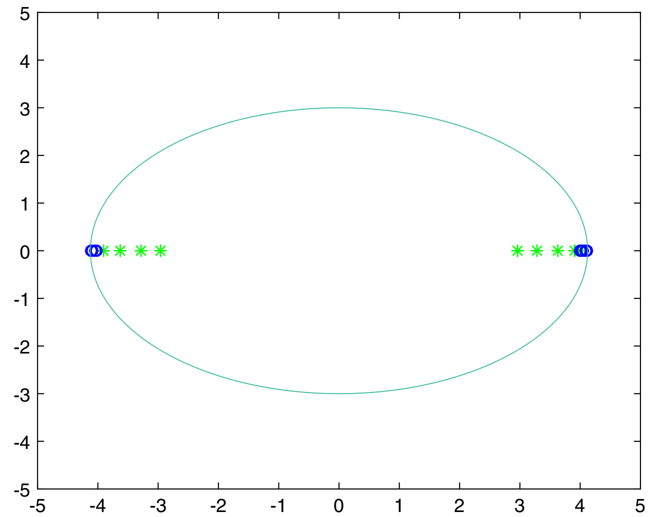

FIGURE 1 $a = 4$ in Example 1

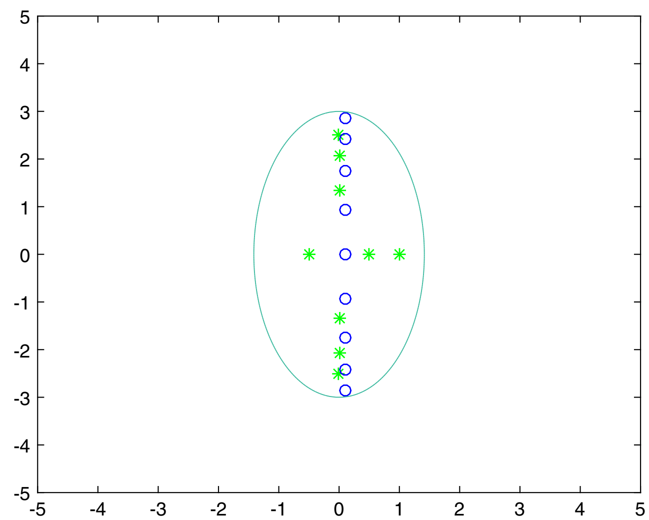

FIGURE 2 $a = 1$ in Example 1

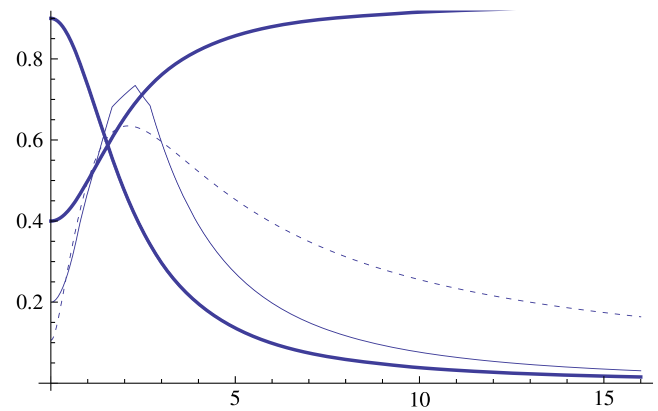

FIGURE 3 Deviation of T from normality in Example 1

Example 2. Consider the tridiagonal 2-Toeplitz matrix $T(2m+1; 1, 1, a, -a, -1, 1)$, where $a \in \mathbb{R}$, and the associated 2-Toeplitz operator $\mathcal{T}(\infty; 1, 1, a, -a, -1, 1)$. By Theorem 4.1 in the work of Bebiano et al.,[19] $\overline{W(\mathcal{T})}$ is the convex hull of the ellipses,

$$x^2 + (1 + a^2)y^2 \pm 2xy = a^2.$$

When $a = 0$, the above ellipses degenerate into line segments, namely, the diagonals of the square region with vertices $1 \pm i$ and $-1 \pm i$, which represents $\partial \overline{W(\mathcal{T})}$.

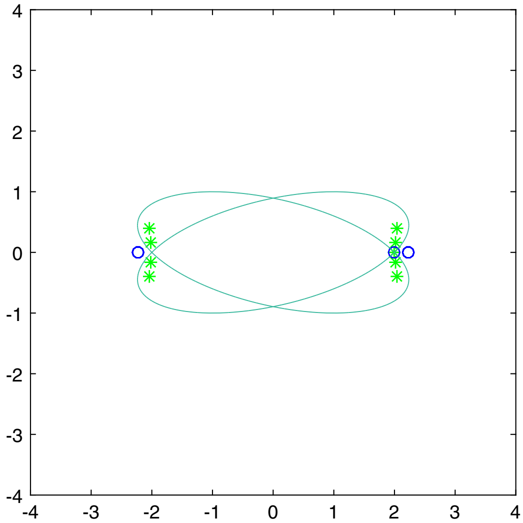

FIGURE 4 $a = 2$ in Example 2

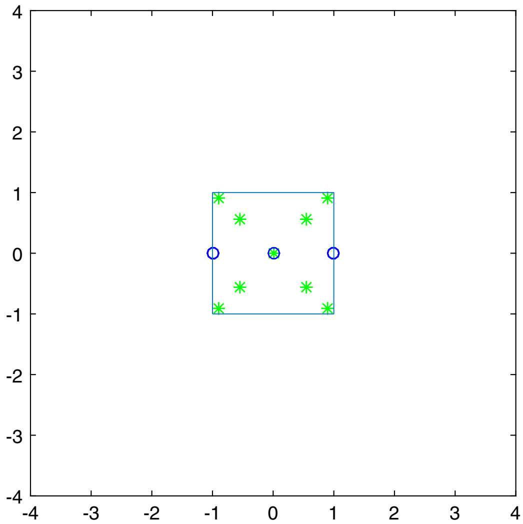

FIGURE 5 $a = 0$ in Example 2

According to Proposition 1, the eigenvalues of T are a and

$$\pm\sqrt{a^2 + 2i\cos\frac{r\pi}{m+1}}, \qquad r = 1, \ldots, m.$$

By Remark 4, we have $K(T) = 4a^2\frac{m^2+m}{2m+1} \geq 0$. From Theorem 4, $d_F(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}) = \|\widehat{T}_1 - T\|_F$, and, if $a = 0$, $d_F(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}}) = \|\widehat{T}_1 - T\|_F = \|\widehat{T}_2 - T\|_F$, where $\widehat{T}_1 = \widehat{T}_1(2m+1, 0, 1, a, -a, 0, 1)$ and $\widehat{T}_2 = \widehat{T}_2(2m+1, 1, 0, 0, 0, -1, 0)$. The spectrum of $\widehat{T}_2$ is $\sigma(\widehat{T}_2) = \{i, -i, 0\}$, where the eigenvalues i, $-i$ and 0 have multiplicities m, m and 1, respectively, and the spectrum of $\widehat{T}_1$ is

$$\sigma(\widehat{T}_1) = \left\{-\sqrt{a^2+1}, \sqrt{a^2+1}, a\right\},$$

where $\sqrt{a^2+1}$, $-\sqrt{a^2+1}$ and a have multiplicities m, m and 1, respectively. The eigenvalues of $\widehat{T}_1$ lie in $\overline{W(\mathcal{T})}$.

This is illustrated in Figure 4 for $m = 4$ and $a = 2$. The figure displays the spectra in the complex plane of the matrix T (green crosses) and of the closest normal matrix $\widehat{T}_1$ (blue circles). The convex hull of the ellipses is $\overline{W(\mathcal{T})}$. The horizontal and vertical axes are the real and imaginary axes. The eigenvalues are computed with the MATLAB (version R2016a) function eig. In Figures 5 and 6, we consider $a = 0$ and display the spectra in the complex plane of the matrix T (green crosses) and of the closest normal matrices $\widehat{T}_1$ and $\widehat{T}_2$ (blue circles), respectively. In each case, the square region is $\overline{W(\mathcal{T})}$.

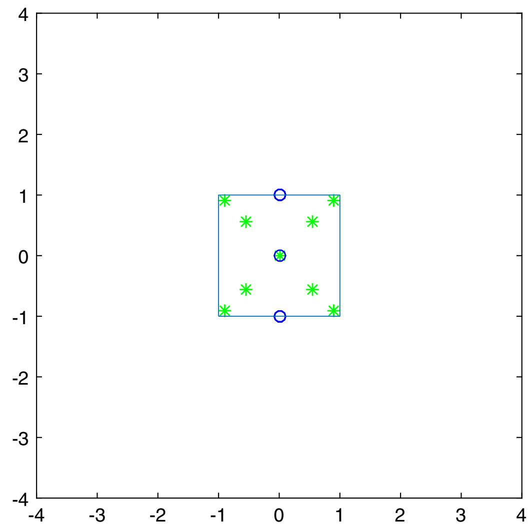

FIGURE 6 $a = 0$ in Example 2

TABLE 1 Deviation of T from normality

a	$N/\|\| T \|\|_F^2$	$\Delta_F^2(T)/\|\| T \|\|_F^2$	$d_F^2(T, \mathcal{N}_{\mathcal{T}}^{\mathbb{R}})/\|\| T \|\|_F^2$
1	0.5426	0.4676	0.32
2	0.4583	0.2798	0.1538
4	0.2867	0.0977	0.05
8	0.1534	0.0269	0.0135
16	0.07809	0.0069	0.0034

Finally, in Table 1, we give the deviation of T from normality for several values of a. We also include in the table the squared normalized distance of T to $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$, according to Theorem 3.

7 | CONCLUSIONS

In this paper, known results in the literature for real tridiagonal Toeplitz matrices have been generalized. We focused on the study of real tridiagonal 2-Toeplitz matrices T, described the distance of such matrices to the closure $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ of the set formed by the real irreducible tridiagonal normal matrices, and characterized the matrices in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ whose distance to T is d. We verified that the closest matrices to T in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ are also 2-Toeplitz matrices. For a certain class of matrices T, including the real tridiagonal Toeplitz matrices, we showed that the eigenvalues of the closest normal matrix in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$ lie in the closure of the field of values, possibly shifted, of the operator associated with T. When T has an odd dimension, we can find in the work of Gover[16] closed formulae for the eigenvalues of T. Based on this result, we gave a more precise location of the eigenvalues of T and of the closest matrix to T in $\mathcal{N}_{\mathcal{T}}^{\mathbb{R}}$. Finally, a measure of nonnormality of T is analyzed. Examples illustrating the theoretical results obtained were provided.

ACKNOWLEDGEMENTS

The authors thank the referees for the valuable comments that helped improve this paper. There are no conflicts of interest to this work. This work was partially supported by the Centre for Mathematics of the University of Coimbra (UID/MAT/00324/2019), funded by the Portuguese Government through FCT/MEC and cofunded by the European Regional Development Fund through the Partnership Agreement PT2020. This work was partially supported by project UID/MAT/04721/2019.

ORCID

Susana Furtado https://orcid.org/0000-0003-0395-5972

REFERENCES

1. Hariprasad M, Murugesan V. Chain models and the spectra of tridiagonal k-Toeplitz matrices. arXiv:1506.05317.
2. Noschese S, Pasquini L, Reichel L. The structured distance to normality of an irreducible real tridiagonal matrix. Electron Trans Numer Anal. 2007;28:65–77.
3. Noschese S, Pasquini L, Reichel L. Tridiagonal Toeplitz matrices: properties and novel applications. Numer Linear Algebra Appl. 2013;20:302–326.
4. Noschese S, Reichel L. The structured distance to normality of banded Toeplitz matrices. BIT Numer Math. 2009;49:629–640.
5. Reichel L, Trefethen LN. Eigenvalues and pseudo-eigenvalues of Toeplitz matrices. Linear Algebra Appl. 1992;162-164:153–185.
6. Demmel JW. On condition numbers and the distance to the nearest ill-posed problem. Numerische Mathematik. 1987;51:251–289.
7. Demmel JW. Nearest defective matrices and the geometry of ill-conditioning. In: Cox MG, Hammarling S, editors. Reliable numerical computation. Oxford, UK: Clarendon Press, 1990; p. 35–55.
8. Elsner L, Paardekooper MHC. On measures of nonnormality of matrices. Linear Algebra Appl. 1987;92:107–123.
9. Henrici P. Bounds for iterates, inverses, spectral variation and field of values of non-normal matrices. Numerische Mathematik. 1962;4:24–40.
10. Higham NJ. Matrix nearness problems and applications. In: Gover MJC, Barnett S, editors. Applications of matrix theory. Oxford, UK: Oxford University Press, 1989; p. 1–27.
11. Lee SL. A practical upper bound for departure from normality. SIAM J Matrix Anal Appl. 1995;16:462–468.
12. Lee SL. Best available bounds for departure from normality. SIAM J Matrix Anal Appl. 1996;17:984–991.
13. Ruhe A. On the closeness of eigenvalues and singular values for almost normal matrices. Linear Algebra Appl. 1975;11:87–93.
14. Ruhe A. Closest normal matrix finally found! BIT Numer Math. 1987;27:585–598.
15. Elsner L, Ikramov KD. Normal matrices: an update. Linear Algebra Appl. 1998;285:291–303.
16. Gover MJC. The eigenproblem of a tridiagonal 2-Toeplitz matrix. Linear Algebra Appl. 1994;197:63–78.
17. Bebiano N, Furtado S. Structured distance to normality of tridiagonal matrices. Linear Algebra Appl. 2018;552:239–255.
18. Gustafson KE, Rao DKM. Numerical range: The field of values of linear operators and matrices. New York, NY: Springer; 1997.
19. Bebiano N, da Providência J, Nata A. The numerical range of banded biperiodic Toeplitz operators. J Math Anal Appl. 2013;398:189–197.
20. Brown ES, Spitkovsky IM. On matrices with elliptical numerical ranges. Linear Multilinear Algebra. 2004;52:177–193.

How to cite this article: Bebiano N, Furtado S. Nearness results for real tridiagonal 2-Toeplitz matrices. *Numer Linear Algebra Appl*. 2019;26:e2257. https://doi.org/10.1002/nla.2257