Research Article **Open Access**

Bakytzhan Kurmanbek, Yogi Erlangga, and Yerlan Amanbek*

Inverse properties of a class of seven-diagonal (near) Toeplitz matrices

https://doi.org/10.1515/spma-2021-0148
Received April 20, 2021; accepted September 9, 2021

Abstract: This paper presents the explicit inverse of a class of seven-diagonal (near) Toeplitz matrices, which arises in the numerical solutions of nonlinear fourth-order differential equation with a finite difference method. A non-recurrence explicit inverse formula is derived using the Sherman-Morrison formula. Related to the fixed-point iteration used to solve the differential equation, we show the positivity of the inverse matrix and construct an upper bound for the norms of the inverse matrix, which can be used to predict the convergence of the method.

Keywords: seven-diagonal matrices, Toeplitz, exact inverse, upper bound of norm of inverse

MSC 2020: 15A60, 15B05, 65L10

1 Introduction

Many mathematical problems give rise to a system of equations that involves an inversion of a banded Toeplitz or near Toeplitz matrix. For example, a second-order or fourth-order finite difference approximation to a second-order differential operator results in a tridiagonal and, respectively, pentadiagonal Toeplitz matrix or a near Toeplitz matrix after the inclusion of boundary conditions. Inversions of this class of matrices have been studied extensively, and can be done very efficiently; see, e.g., [1–7]. In addition to the algorithmic development, many authors have contributed to the inverse properties of banded Toeplitz and near Toeplitz matrices, such as exact inverse formulas [8–12], bounds for entries of the inverse matrices, and bounds for the inverse norm [13]. Examining formulas for determinant of such matrices can be also useful to explore the existence and uniqueness of solution related to the ordinary or partial differential problems [14–18].

An improved numerical accuracy can be attained via a higher-order approximation, but at the expense of increased bandwidth of the matrix in the system beyond five diagonals. This increased bandwidth not only increases the computational costs, but also complicates the analysis of the inverse properties. In many cases, the analysis demands for additional conditions such as diagonal dominance or M-matrix [11, 19, 20]. Exact inverse formulas, while can probably still be derived, may not be in an appealing form.

Bakytzhan Kurmanbek: Nazarbayev University, Department of Mathematics, 53 Kabanbay Batyr Ave, Nur-Sultan 010000, Kazakhstan, E-mail: bakytzhan.kurmanbek@nu.edu.kz
Yogi Erlangga: Zayed University, Department of Mathematics, Abu Dhabi Campus, P.O. Box 144534, United Arab Emirates, E-mail: yogi.erlangga@zu.ac.ae
***Corresponding Author: Yerlan Amanbek:** Nazarbayev University, Department of Mathematics, 53 Kabanbay Batyr Ave, Nur-Sultan 010000, Kazakhstan, E-mail: yerlan.amanbek@nu.edu.kz

In this paper, we shall consider the inverse of $n \times n$ seven-diagonal near Toeplitz matrices associated with a fourth-order finite-difference discretization of the fourth-order differential operator d^4/dx^4:

$$A_n = \begin{pmatrix} a_0 & -a_1 & 12 & -1 & 0 & \cdots & \cdots & \cdots & 0 \\ -a_1 & a_2 & -39 & 12 & -1 & \ddots & \ddots & \ddots & \vdots \\ 12 & -39 & 56 & -39 & 12 & -1 & \ddots & \ddots & \vdots \\ -1 & 12 & -39 & 56 & -39 & 12 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & -1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & -39 & 12 \\ \vdots & \ddots & \ddots & \ddots & -1 & 12 & -39 & a_2 & -a_1 \\ 0 & \cdots & \cdots & \cdots & 0 & -1 & 12 & -a_1 & a_0 \end{pmatrix}_{n\times n}, \quad n \geq 7, \tag{1}$$

where $a_0, a_1, a_2 > 0$. The matrix (1) is symmetric, centrosymmetric, nondiagonally dominant, and is not an M-matrix. The perturbation from the Toeplitz structure at the "corner" of the matrix can be caused by the inclusion of boundary conditions in the underlying boundary-value problems.

An instance of application that involves (1) is related to the nonlinear boundary-value problem

$$EI\frac{d^4u}{dx^4} = f(x,u), \quad x \in (0,1) \subset \mathbb{R}, \tag{2}$$

with

$$u(0) = u'(0) = u(1) = u'(1) = 0, \tag{3}$$

Approximating the derivative by the fourth-order finite difference scheme results in the nonlinear system

$$A_n\boldsymbol{u} = h^4 C_{EI}\boldsymbol{f}(\boldsymbol{u}), \quad \boldsymbol{u} \in \mathbb{R}^n, \tag{4}$$

where A_n is a near Toeplitz matrix of the form of (1), h is the meshsize, and C_{EI} is a physical constant. Solution to the nonlinear system (4) can be computed iteratively using a fixed-point method based on the iterands:

$$A_n\boldsymbol{u}^\ell = h^4 C_{EI}\boldsymbol{f}(\boldsymbol{u}^{\ell-1}), \quad \ell = 1, 2, \ldots, \tag{5}$$

for some initial solution vector $\boldsymbol{u}^0 \in \mathbb{R}^n$. For some class of the forcing term f, convergence of this method can be shown to depend on the p-norm of the inverse of A_n, $\|A_n^{-1}\|_p$, where $p \in \{1, 2, \infty\}$.

We approximated the fourth-order derivative by utilizing the following finite difference scheme:

$$\frac{d^4u}{dx^4}(x_i) \approx \frac{1}{6h^4}(-u_{i-3} + 12u_{i-2} - 39u_{i-1} + 56u_i - 39u_{i+1} + 12u_{i+2} - u_{i+3}),$$

where $x_i = ih$ and $u_i \equiv u(x_i)$. We set the boundary condition $u(0) \equiv u_0 = 0$. For ghost points(out of domain), we choose $u_{-1} = u_1$ and $u_{-2} = u_2$ using central difference scheme based on the boundary condition $u'(0) = 0$. The same reasoning applies to cases $i = n - 1$ and n with the boundary conditions $u(1) = u'(1) = 0$. For more details of finite difference method for beam problems, we refer the reader to [21, 22].

In this paper, we first derive an explicit, non-recurrence formula for the inverse of two special cases of the seven-diagonal matrix (1): (i) the Toeplitz case with $a_0 = 56$, $a_1 = 39$ and $a_2 = 56$, and (ii) with $a_0 = 68$, $a_1 = 40$ and $a_2 = 56$, which corresponds to the boundary-value problem (2) and (3). The inverse formulas are used to analyze some properties of the inverse matrices and to construct upper bounds for the norms of the inverses, in terms of the matrix size n (which is linked to the meshsize h). While it is possible to construct a bound which is independent of n, an n-dependent bound is desirable as it can be used to predict more accurately the convergence of the fixed-point method (5) under mesh refinement. We then consider a more general setting associated with (1), where a_0, a_1, a_2 satisfy certain decays in the modulus of the entries of A.

In contrast with the tridiagonal and pentadiagonal cases, there does not exist a large body of results on the inverse of seven diagonal (near) Toeplitz matrices. Literature on inverses of sevendiagonal matrices include [23–25] on algorithm development and [26] on inverse properties. While the class of matrices considered in this paper is quite narrow, our results are new, helpful in analyzing the convergence of the numerical recipe (5), and should contribute to the inverse theories of banded Toeplitz matrices.

The paper is organized as follows. After stating some preliminary results in Section 2, we derive the explicit formula for the Toeplitz matrix and an upper bound for the norms in Section 3. Section 4 is devoted to the formula for and norms of the inverse of the near Toeplitz matrix. Some numerical results are presented in Section 5. Section 6 is devoted to the inverse properties with general parameters in the 2×2 corner block of (1). We finish up the paper with concluding remarks in Section 7.

2 Preliminaries

It is well known that a Toeplitz matrix cannot be decomposed into a product of two Toeplitz matrices. Some classes of Toeplitz matrices however admit a low-rank decomposition of the form

$$A_n = B_n C_n + \sigma U V^T, \quad \sigma \in \mathbb{R}, \tag{6}$$

where B_n and C_n are (near) Toeplitz, and U and V are $n \times m$ matrices, with $m < n$. Furthermore, if A_n is nonsingular, the inverse matrix A_n^{-1} can be computed using the Shermann-Morrison-Woodbury formula [27–29]:

$$A_n^{-1} = D_n^{-1} - \sigma D_n^{-1} U M_m^{-1} V^T D_n^{-1}, \tag{7}$$

where $D_n = B_n C_n$, $M_m = I_m + \sigma V^T D_n^{-1} U \in \mathbb{R}^{m \times m}$, and I_m is the identity matrix of size m.

As we shall see later, for the seven-diagonal matrices considered in this paper, the above low-rank decomposition involves the tridiagonal matrix

$$C_n = \begin{pmatrix} 8 & -1 & 0 & \cdots & 0 \\ -1 & 8 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 8 & -1 \\ 0 & \cdots & 0 & -1 & 8 \end{pmatrix}_{n \times n}. \tag{8}$$

Some properties of C_n are stated the following lemmas.

Lemma 1. *C_n is positive definite, with the inverse $C_n^{-1} = [c_{i,j}^{-1}]$, $i, j = 1, \ldots, n$ given by*

$$c_{i,j}^{-1} = \frac{\gamma_j \gamma_{n+1-i}}{\gamma_{n+1}}$$

for $i \geq j$, and $c_{i,j}^{-1} = c_{j,i}^{-1}$, for $i < j$, where

$$\gamma_k = \left(r_1^k - r_2^k \right) / 2\sqrt{15}, \quad k \in \mathbb{N}, \tag{9}$$

with $r_1 = 4 + \sqrt{15}$ and $r_2 = 4 - \sqrt{15}$.

Proof. The proof can be found in [9]. □

Lemma 2. *Let γ_k be defined as in (9). Then the following holds for any $k \in \mathbb{N}$.*
(i) $\gamma_k + \gamma_{k-2} = 8\gamma_{k-1}$;
(ii) $4 \leq \frac{\gamma_{k+1}}{\gamma_k} \leq 8$.

Proof. Set $r_1^k = (4+\sqrt{15})^k = \alpha_k + \gamma_k\sqrt{15}$ and $r_2^k = (4-\sqrt{15})^k = \alpha_k - \gamma_k\sqrt{15}$. The parameters α_k and γ_k satisfy the recurrence relations

$$\begin{cases} \alpha_k = 4\alpha_{k-1} + 15\gamma_{k-1}, \\ \gamma_k = \alpha_{k-1} + 4\gamma_{k-1}. \end{cases} \tag{10}$$

Solving this linear equation system leads to the statement (i).

The (i) part implies $\gamma_k \le 8\gamma_{k-1}$, which is the right inequality of the (ii) part. The left inequality of the (ii) part is proved by induction. For $k = 1$, $\gamma_2/\gamma_1 = 4$. Thus, (ii) holds for $k = 1$. Suppose (ii) also holds for $k = j-1$, i.e., $\gamma_j/\gamma_{j-1} \ge 4$. For $k = j$, by utilizing the (i) part in the process,

$$\frac{\gamma_{j+1}}{\gamma_j} = \frac{\gamma_{j+1}}{\gamma_{j-1}}\frac{\gamma_{j-1}}{\gamma_j} = \left(8\frac{\gamma_j}{\gamma_{j-1}} - 1\right)\frac{\gamma_{j-1}}{\gamma_j} = 8 - \frac{\gamma_{j-1}}{\gamma_j} \ge 4.$$

□

Lemma 3. *For any $i \in \{1, 2, \ldots, n\}$, the following holds*

$$\gamma_{n+1-i}\gamma_{i+1} - \gamma_{n-i}\gamma_i = \gamma_{n+1} \tag{11}$$

Proof. The proof is based on using the definition of γ_i and the fact that $r_1 r_2 = 1$.

$$\begin{aligned} \gamma_{n+1-i}\gamma_{i+1} - \gamma_{n-i}\gamma_i &= \frac{(r_1^{n+1-i} - r_2^{n+1-i})(r_1^{i+1} - r_2^{i+1}) - (r_1^{n-i} - r_2^{n-i})(r_1^i - r_2^i)}{60} = \\ &= \frac{r_1^{n+2} + r_2^{n+2} - r_1^{i+1}r_2^{i+1}(r_1^{n-2i} + r_2^{n-2i}) - r_1^n - r_2^n + r_1^i r_2^i (r_1^{n-2i} + r_2^{n-2i})}{60} = \\ &= \frac{r_1^{n+2} + r_2^{n+2} - r_1^n - r_2^n}{60} = \frac{r_1^{n+2} + r_2^{n+2} - r_1 r_2 (r_1^n + r_2^n)}{60} = \\ &= \frac{(r_1^{n+1} - r_2^{n+1})(r_1 - r_2)}{60} = \gamma_{n+1}\gamma_1 = \gamma_{n+1} \end{aligned}$$

□

Lemma 4. *Let γ_k be defined as in (9). Then, for any $p \in \mathbb{N}$,*

$$\sum_{k=1}^{p} \gamma_k = \frac{1}{6}(\gamma_{p+1} - \gamma_p - 1), \tag{12}$$

$$\sum_{k=1}^{p} k\gamma_k = \frac{1}{6}(p\gamma_{p+1} - (p+1)\gamma_p), \tag{13}$$

$$\sum_{k=1}^{p} k^2\gamma_k = \frac{1}{18}((3p^2+1)\gamma_{p+1} - (3p^2+6p+4)\gamma_p - 1), \tag{14}$$

$$\sum_{k=1}^{p} k^3\gamma_k = \frac{1}{6}((p^3+p)\gamma_{p+1} - (p^3+3p^2+4p+2)\gamma_p). \tag{15}$$

Proof. The proof uses Vieta's formula and Lemma 2. Since the proof for each relation is similar, we shall show the proof only for (12) and (13).

$$\begin{aligned} \sum_{k=1}^{p}\gamma_k &= \sum_{k=1}^{p}\frac{r_1^k - r_2^k}{2\sqrt{15}} = \frac{1}{2\sqrt{15}}\left(\sum_{k=0}^{p} r_1^k - \sum_{k=0}^{p} r_2^k\right) = \frac{1}{2\sqrt{15}}\left(\frac{r_1^{p+1}-1}{r_1-1} - \frac{r_2^{p+1}-1}{r_2-1}\right) \\ &= \frac{1}{12\sqrt{15}}(r_1^{p+1} - r_1^p + r_2^p - r_2^{p+1} + r_2 - r_1) = \frac{1}{6}(\gamma_{p+1} - \gamma_p - 1). \end{aligned}$$

Next,

$$\begin{aligned}\sum_{k=1}^{p} k\gamma_k &= \sum_{k=1}^{p} \frac{k(r_1^k - r_2^k)}{2\sqrt{15}} = \frac{1}{2\sqrt{15}}\left(r_1 \sum_{k=1}^{p} k r_1^{k-1} - r_2 \sum_{k=1}^{p} k r_2^{k-1}\right) \\ &= \frac{1}{2\sqrt{15}}\left(r_1 \left(\sum_{k=1}^{p} r_1^k\right)' - r_2 \left(\sum_{k=1}^{p} r_2^k\right)'\right) \\ &= \frac{1}{2\sqrt{15}}\left(\frac{p r_1^{p+2} - (p+1) r_1^{p+1} + r_1}{(r_1-1)^2} - \frac{p r_2^{p+2} - (p+1) r_2^{p+1} + r_2}{(r_2-1)^2}\right) \\ &= \frac{1}{36}\left(p\gamma_{p+2} - (3p+1)\gamma_{p+1} + (3p+2)\gamma_p - (p+1)\gamma_{p-1}\right).\end{aligned}$$

In the above derivation, we have used the relation (12) to evaluate derivatives of the sum. The relation (13) is obtained from the above equation by applying Lemma 2 multiple times. □

Lemma 5. *Let γ_k be defined as in (9) and $\alpha_k = (4+\sqrt{15})^k - \gamma_k\sqrt{15}$. Then*

$$\alpha_k - \alpha_{k-2} = 30\gamma_{k-1}$$

is true for any $k \in \mathbb{N}$.

Proof. By using the recurrence relations $\gamma_k = \alpha_{k-1} + 4\gamma_{k-1}$ and $\alpha_k = 4\gamma_k - \gamma_{k-1}$ from (10) we obtain

$$\gamma_k - \gamma_{k-2} = 2\alpha_{k-1}$$

Applying this identity to the $(k-2)$ and $(k-1)$ term, we have

$$2(\alpha_k - \alpha_{k-2}) = (\gamma_{k+1} - \gamma_{k-1}) - (\gamma_{k-1} - \gamma_{k-3}) = \gamma_{k+1} - 2\gamma_{k-1} + \gamma_{k-3}.$$

Applying Lemma 2 several times to the above equation leads to the statement in the lemma. □

Lemma 6. *For $p \in \{0, 1, \infty\}$, $\|C_n^{-1}\|_p \le 1/6$.*

Proof. A proof for general diagonally-dominant symmetric tridiagonal matrices $T_n = \mathrm{tril}(-1, b, -1)$ is given in [30]; see also [12] for alternative bounds. For $b > 2$,

$$\|T_n\|_\infty = \frac{1}{b-2} - \frac{2}{r_b^{\frac{n+1}{2}}}, \tag{16}$$

where $r_b = \frac{1}{2}(b + \sqrt{b^2-4})$. For C_n, setting $b = 8$ leads to the bound in the lemma. □

3 The Toeplitz case

In this section, we consider the case where (1) is Toeplitz ($a_0 = 56$, $a_1 = 39$ and $a_2 = 56$). The seven-diagonal matrix A_n can be decomposed into a rank-2 decomposition (6), with $\sigma = 1$,

$$B_n = \begin{pmatrix} 6 & -4 & 1 & 0 & \cdots & \cdots & 0 \\ -4 & 6 & -4 & 1 & \ddots & \ddots & \vdots \\ 1 & -4 & 6 & -4 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & -4 & 6 & -4 & 1 \\ \vdots & \ddots & \ddots & 1 & -4 & 6 & -4 \\ 0 & \cdots & \cdots & 0 & 1 & -4 & 6 \end{pmatrix}_{n\times n}, U = \begin{pmatrix} 4 & 0 \\ -1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & -1 \\ 0 & 4 \end{pmatrix}_{n\times 2}, V = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}_{n\times 2}, \tag{17}$$

and with C_n given in (8). The matrix B_n in the decomposition is nonsingular. The entries of the inverse matrix B_n^{-1} is derived in [8] and given for $i \geq j$ by the formula

$$b_{i,j}^{-1} = -\frac{(n+1-i)(n+2-i)j(j+1)}{6(n+1)(n+2)(n+3)}\left[(i+1)(j-1)(n+3) - i(j+2)(n+1)\right]. \tag{18}$$

Regarding the matrix B_n, we have the following lemma

Lemma 7. *The matrix B_n in* (17) *is positive definite.*

Proof. For any vector $\boldsymbol{x} = (x_1 \ \dots \ x_n)^T \in \mathbb{R}^n$,

$$\begin{aligned}\boldsymbol{x}^T B_n \boldsymbol{x} &= 6(x_1^2 + x_2^2 + \cdots + x_n^2) - 8(x_1x_2 + x_2x_3 + \cdots + x_nx_{n-1}) + 2(x_1x_3 + x_2x_4 + x_3x_5 + \cdots + x_{n-2}x_n)\\ &= x_1^2 + (2x_1 - x_2)^2 + (x_1 - 2x_2 + x_3)^2 + \cdots + (x_{n-2} - 2x_{n-1} + x_n)^2 + (2x_n - x_{n-1})^2 + x_n^2\\ &\geq 0,\end{aligned}$$

with equality holding only when $\boldsymbol{x} = 0$. □

3.1 Exact inverse formula

An explicit formula for A_n^{-1} can be derived by evaluating the right-hand side of the decomposition (7).

Starting with the first term on the right-hand side, let $D_n^{-1} := C_n^{-1}B_n^{-1} = [d_{i,j}^{-1}]$, with

$$\begin{aligned}d_{i,j}^{-1} &= \sum_{k=1}^{n} c_{i,k}^{-1} b_{k,j}^{-1}\\ &= \sum_{k=1}^{j} c_{i,k}^{-1} b_{j,k}^{-1} + \sum_{k=j+1}^{i} c_{i,k}^{-1} b_{k,j}^{-1} + \sum_{k=i+1}^{n} c_{k,i}^{-1} b_{k,j}^{-1}\\ &= \sum_{k=1}^{j} \frac{\gamma_k\gamma_{n+1-i}}{\gamma_{n+1}} b_{j,k}^{-1} + \sum_{k=j+1}^{i} \frac{\gamma_k\gamma_{n+1-i}}{\gamma_{n+1}} b_{k,j}^{-1} + \sum_{k=1}^{n-i} \frac{\gamma_i\gamma_k}{\gamma_{n+1}} b_{n+1-k,j}^{-1},\end{aligned} \tag{19}$$

due to symmetry of B_n and C_n. Furthermore, D_n^{-1} is centrosymmetric, due to the centrosymmetry of B_n^{-1} and C_n^{-1}. By using Lemmas 2–5 and after necessary simplifications, we obtain, for $i \geq j$,

$$d_{i,j}^{-1} = \frac{\gamma_j\gamma_{n+1-i}}{36\gamma_{n+1}} + \eta\left(\zeta_1 + \zeta_2\frac{\gamma_{n+1-i}}{\gamma_{n+1}} + \zeta_3\frac{\gamma_i}{\gamma_{n+1}}\right), \tag{20}$$

where

$$\begin{aligned}\eta &= -\frac{1}{6(n+1)(n+2)(n+3)},\\ \zeta_1 &= \frac{1}{6}j(j+1)\{(j-3i-1)n^3 + (6j+6i^2-12i-4)n^2 + ((-3i^2-3i+10)j - 3i^3 + 18i^2 - 15i - 5)n\\ &\quad + (2i^3 - 3i^2 - 3i + 5)j - 5i^3 + 12i^2 - 6i - 2\},\\ \zeta_2 &= \frac{1}{6}(n+1)j(n+1-j)(n+2-j),\\ \zeta_3 &= \frac{1}{6}(n+1)j(j+1)(n+1-j).\end{aligned}$$

For the second term on the right-hand side of (7), let $M_2 = [m_{i,j}] \in \mathbb{R}^{2\times 2}$. From a direct calculation and the use of Lemmas 2–5 and (20), one can show that

$$m_{11} = m_{22} = 1 + 4d_{1,1}^{-1} - d_{1,2}^{-1} = 1 + \frac{11n^2+5n}{36(n+1)(n+2)} + \frac{1}{36\gamma_{n+1}}\left(\alpha_n - \frac{2+2n\gamma_n}{n+2}\right), \tag{21}$$

$$m_{12} = m_{21} = 4d_{n,1}^{-1} - d_{n,2}^{-1} = \frac{7n+4}{18(n+1)(n+2)} - \frac{1}{18\gamma_{n+1}}\left(2 + \frac{n+\gamma_n}{n+2}\right). \tag{22}$$

Lemma 8. $\det(M) > 0$.

Proof. Note that $11n^2 - 9n - 8 > 9(n^2 - n - 1) > 0$ and $36(n+2)\gamma_{n+1} - (2n-2)\gamma_n > (2n-2)(\gamma_{n+1} - \gamma_n) > 0$. Furthermore, $\gamma_{n+1} > \gamma_n$ and $\gamma_{n+1} > n$. Using these inequalities, for $n \geq 2$,

$$\begin{aligned} m_{11} - m_{12} &= 1 + \frac{11n^2 - 9n - 8}{36(n+1)(n+2)} + \frac{1}{18\gamma_{n+1}}\left(\frac{\alpha_n}{2} + 2 - \frac{(n-1)(\gamma_n - 1)}{n+2}\right) \\ &= \frac{11n^2 - 9n - 8}{36(n+1)(n+2)} + \frac{\alpha_n(n+2) + 6n + 6 + 36(n+2)\gamma_{n+1} - (2n-2)\gamma_n}{36(n+2)\gamma_{n+1}} > 0. \end{aligned}$$

Next, for $n \geq 2$, using (22),

$$\begin{aligned} m_{12} &= \frac{(n+1)(\gamma_{n+1} - \gamma_n) + (6n+3)\gamma_{n+1} - 3n^2 - 7n - 4}{18(n+1)(n+2)\gamma_{n+1}} \\ &> \frac{(n+1)(\gamma_{n+1} - \gamma_n) + (n-2)(3n+2)}{18(n+1)(n+2)\gamma_{n+1}} > 0. \end{aligned}$$

Therefore, $m_{11} > m_{12} > 0$, and hence $m_{11}^2 - m_{12}^2 > 0$. □

By combining the above results and simplifying the expressions for $i \geq j$, the explicit formula of the inverse of A_n can be written as

$$a_{i,j}^{-1} = d_{i,j}^{-1} + \frac{1}{m_{11}^2 - m_{12}^2}\left((m_{12}d_{1,j}^{-1} - m_{11}d_{n,j}^{-1})(4d_{i,n}^{-1} - d_{i,n-1}^{-1}) + (m_{12}d_{n,j}^{-1} - m_{11}d_{1,j}^{-1})(4d_{i,1}^{-1} - d_{i,2}^{-1})\right). \tag{23}$$

where the coefficents on the right-hand side are computed using (9), (20), (21) and (22).

Theorem 9. *A_n is positive definite.*

Proof. Since A_n is symmetric, the proof is based on Sylvester's criterion. In this case, we need to show that all upper left $k \times k$ submatrices of A_n have positive determinant, $k = 1, \ldots, n$. For $k = 1, \ldots, 6$, the determinant can be shown to be positive via numerical calculation. For $k \geq 7$, since the submatrices retain the structure of A_n, we only need to show that A_n has positive determinant. Using the generalized matrix determinant lemma, with $D_n = B_nC_n$,

$$\det(A_n) = \det(D_n + UV^T) = \det(I_2 + V^TD^{-1}U)\det(D_n) = \det(M)\det(B_n)\det(C_n).$$

Positive definiteness of B_n and C_n (Lemmas 7 and 1, respectively) implies $\det(B_n) > 0$ and $\det(C_n) > 0$. Together with Lemma 8, we have $\det(A_n) > 0$, which proves the theorem. □

3.2 Bound of norms of inverse

In this section we derive a bound of $\|A_n^{-1}\|_p$, for $p = 1, 2, \infty$. The result is summarized in Theorem 10 below:

Theorem 10. *Let A_n be given as in (1), with $a_0 = 56$ and $a_1 = 39$. Then the following inequality holds for $p \in \{1, 2, \infty\}$:*

$$\|A_n^{-1}\|_p \leq \frac{(n+1)^2(n+3)^2}{2304} + \frac{(n+1)^2}{432} + \frac{n+4}{24}.$$

Proof. By the symmetry of A_n, $\|A^{-1}\|_1 = \|A^{-1}\|_\infty$ and $\|A^{-1}\|_2 \leq \sqrt{\|A^{-1}\|_1\|A^{-1}\|_\infty} = \|A^{-1}\|_\infty$. Therefore, it suffices to prove the result for $p = \infty$.

The positive definiteness of A_n (Theorem 9) implies that $A_n^{-1} > 0$. With

$$\begin{aligned} m_{12}d_{1,j}^{-1} - m_{11}d_{n,j}^{-1} &= m_{12}d_{n,n+1-j}^{-1} - m_{11}d_{n,j}^{-1}, \\ m_{12}d_{n,j}^{-1} - m_{11}d_{1,j}^{-1} &= m_{12}d_{n,j}^{-1} - m_{11}d_{n,n+1-j}^{-1}, \end{aligned}$$

due to centrosymmetry of D_n and M, for the i-th rowsum of A_n^{-1}, we have

$$\begin{aligned}\sum_{j=1}^{n}|a_{i,j}^{-1}| &= \sum_{j=1}^{n} d_{i,j}^{-1} \\ &+ \frac{1}{m_{11}^2 - m_{12}^2}\left((4d_{i,n}^{-1} - d_{i,n-1}^{-1})\sum_{j=1}^{n}(m_{12}d_{1,j}^{-1} - m_{11}d_{n,j}^{-1}) + (4d_{i,1}^{-1} - d_{i,2}^{-1})\sum_{j=1}^{n}(m_{12}d_{n,j}^{-1} - m_{11}d_{1,j}^{-1})\right) \\ &= \sum_{j=1}^{n} d_{i,j}^{-1} - \frac{\sum_{j=1}^{n} d_{n,j}^{-1}}{m_{11}+m_{12}}\left(4(d_{i,n}^{-1} + d_{i,1}^{-1}) - (d_{i,n-1}^{-1} + d_{i,2}^{-1})\right).\end{aligned}$$

Then

$$\|A_n^{-1}\|_\infty = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{i,j}^{-1}| \le \underbrace{\max_{1\le i\le n}\sum_{j=1}^{n} d_{i,j}^{-1}}_{\pi_1} + \frac{1}{m_{11}+m_{12}}\underbrace{\left(\sum_{j=1}^{n} d_{n,j}^{-1}\right)}_{\pi_2}\underbrace{\max_{1\le i\le n} g(i)}_{\pi_3} \tag{24}$$

where

$$g(i) = 4\underbrace{(d_{i,n}^{-1} + d_{i,1}^{-1})}_{\theta_1} - \underbrace{(d_{i,n-1}^{-1} + d_{i,2}^{-1})}_{\theta_2}.$$

First of all, with $d_{i,n}^{-1} = d_{n+1-i,1}^{-1}$,

$$\begin{aligned}\theta_1 &= d_{n+1-i,1}^{-1} + d_{i,1}^{-1} = \frac{\gamma_1(\gamma_i + \gamma_{n+1-i})}{36\gamma_{n+1}} - \frac{1}{36(n+2)}\left(\frac{n(\gamma_i+\gamma_{n+1-i})}{\gamma_{n+1}} + 2(1-3i(n+1-i))\right), \\ \theta_2 &= \frac{\gamma_2(\gamma_i + \gamma_{n+1-i})}{36\gamma_{n+1}} - \frac{1}{36(n+2)}\left(\frac{2(n-1)(\gamma_i+\gamma_{n+1-i})}{\gamma_{n+1}} + 6(n+3-3i(n+1-i))\right).\end{aligned}$$

Hence,

$$g(i) = \frac{(6n+10)(\gamma_{n+1} - \gamma_i - \gamma_{n+1-i})}{36(n+2)\gamma_{n+1}} + \frac{i(n+1-i)}{6(n+2)}.$$

Note that

$$\gamma_i' = \left(\frac{r_1^i - r_2^i}{2\sqrt{15}}\right)' = \frac{r_1^i \ln r_1 - r_2^i \ln r_2}{2\sqrt{15}} = \frac{\alpha_i \ln r_1}{\sqrt{15}}$$

and $\ln r_1 + \ln r_2 = 0$; Thus, $r_1 r_2 = 1$.

Consider $i \in [1, n] \subset \mathbb{R}$. Differentiating $g(i)$ with respect to i and solving the equation, we get

$$g'(i) = \frac{(6n+10)\ln r_1}{36(n+2)\gamma_{n+1}\sqrt{15}}(\alpha_{n+1-i} - \alpha_i) + \frac{n+1-2i}{6(n+2)} = 0.$$

Due to monotonicity of $g'(i)$, there is only solution of the above equation, given by $i = \frac{n+1}{2}$, This critical point is also the maximum point of the $g(i)$. Thus,

$$\pi_3 = \max_{1\le i\le n} g(i) = g\left(\frac{n+1}{2}\right),$$

where

$$g\left(\frac{n+1}{2}\right) = \frac{6n+10}{36(n+2)} - \frac{2(6n+10)\gamma_{\frac{n+1}{2}}}{36(n+2)\gamma_{n+1}} + \frac{(n+1)^2}{24(n+2)} \le \frac{3n^2+18n+23}{72(n+2)} \le \frac{3(n+2)(n+4)}{72(n+2)} = \frac{n+4}{24}.$$

For π_2, we first note that

$$d_{n,j}^{-1} = \frac{c_{n,j}^{-1}}{36} - \frac{1}{36(n+1)(n+2)(n+3)}(f_1 + f_2)$$

where

$$f_1 = \frac{(n+1-j)j(n+1)}{\gamma_{n+1}}(\gamma_{n+1-i}(n+2-j) + \gamma_i(j+1)),$$

$$f_2 = j(j+1)\left(2(j-1)i^3 - 3i(n+1)(i^2+ij+j+(n+2)(n+1-2i))\right.$$
$$\left. + (n+1)\left((n+1)(n+2)(j-1)+j(2n+3)\right)\right).$$

Therefore,

$$\pi_2 = \sum_{j=1}^{n} d_{n,j}^{-1} = \frac{\gamma_1}{36\gamma_{n+1}}\left(\sum_{j=1}^{n}\gamma_j\right) - \frac{1}{36(n+1)(n+2)(n+3)}\left(\sum_{j=1}^{n} f_1 + \sum_{j=1}^{n} f_2\right).$$

Using Lemma 4,

$$\sum_{j=1}^{n}\gamma_j = \frac{1}{6}\left(\gamma_{n+1} - \gamma_n - 1\right).$$

$$\sum_{j=1}^{n} f_1 = \frac{n+1}{\gamma_{n+1}}\left(\gamma_1\sum_{j=1}^{n}(j^3 - 2j^2n - 3j^2 + jn^2 + 3jn + 2j) + \gamma_n\sum_{j=1}^{n}(-j^3 + j^2n + jn + j)\right)$$
$$= \frac{n(n+1)^2(n+2)(n+3)(\gamma_1+\gamma_n)}{12\gamma_{n+1}},$$

$$\sum_{j=1}^{n} f_2 = \sum_{j=1}^{n}(7j^3n + 5j^3 - 7j^2n^2 - 4j^2n + 3j^2 - 7jn^2 - 11jn - 2j) = -\frac{n(n+1)(n+2)(n+3)(7n+1)}{12}.$$

Substitution of all terms gives

$$\pi_2 = \frac{\gamma_1(\gamma_{n+1} - \gamma_n - 1)}{216\gamma_{n+1}} - \frac{1}{36}\left(\frac{n(n+1)(\gamma_1+\gamma_n)}{12\gamma_{n+1}} - \frac{n(7n+1)}{12}\right)$$
$$= \frac{1}{216} + \frac{n(7n+1)}{432} - \frac{(n^2+n+2)(\gamma_n+1)}{432\gamma_{n+1}}.$$

From Lemma 2(ii), $\gamma_{n+1}/8 \le \gamma_n \le \gamma_n + 1$. Thus,

$$\pi_2 \le \frac{1}{216} + \frac{n(7n+1)}{432} - \frac{(n^2+n+2)}{432\times 8} = \frac{55n^2+7n+14}{3456} \le \frac{56(n+1)^2}{3456} \le \frac{(n+1)^2}{432}.$$

Lastly,

$$\pi_1 = \max_{1\le i\le n}\sum_{j=1}^{n}|d_{i,j}^{-1}| = \|D_n^{-1}\|_\infty \le \|C_n^{-1}\|_\infty\|B_n^{-1}\|_\infty.$$

By using Theorem 4 of [8]

$$\|B_n^{-1}\|_\infty \le \frac{(n+1)^2(n+3)^2}{384}$$

and Lemma 6,

$$\pi_1 = \|D_n^{-1}\|_\infty \le \|C_n^{-1}\|_\infty\|B_n^{-1}\|_\infty \le \frac{(n+1)^2(n+3)^2}{2304}.$$

Summing up π_1, π_2, and π_3 and using the fact that $m_{11} + m_{12} > 1$ leads to the statement in the theorem. □

4 The near Toeplitz case

We now consider the seven-diagonal near Toeplitz matrix (1) with $a_0 = 68$, $a_1 = 40$ and $a_2 = 56$. We shall denote this matrix by $\widetilde{A}_n$ and use "$\widetilde{\cdot}$" to indicate perturbed matrices relevant to $\widetilde{A}_n$. It can be shown that $\widetilde{A}$ admits rank-2 decomposition*

$$\widetilde{A}_n = \widetilde{B}_n C_n + 2UV^T, \tag{25}$$

* In fact, there exist two a rank-2 decomposition of $\widetilde{A}_n$. We choose this version as it shares many same components in the decomposition as in the Toeplitz case.

where C_n, U, and V are given in (8) and (17), and

$$\widetilde{B}_n = \begin{pmatrix} 7 & -4 & 1 & 0 & \cdots & \cdots & 0 \\ -4 & 6 & -4 & 1 & \ddots & \ddots & \vdots \\ 1 & -4 & 6 & -4 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & -4 & 6 & -4 & 1 \\ \vdots & \ddots & \ddots & 1 & -4 & 6 & -4 \\ 0 & \cdots & \cdots & 0 & 1 & -4 & 7 \end{pmatrix}_{n\times n}. \tag{26}$$

The inverse of $\widetilde{B}_n$ is discussed in [13] and is given entry-wise by the explicit formula, for $i \geq j$,

$$\widetilde{b}^{-1}_{i,j} = \beta\left[\epsilon + (j^2-1)(2\delta^2+1)\right], \tag{27}$$

with

$$\begin{aligned} \delta &= n+1-i, \\ \beta &= \frac{\delta j}{6(n+1)(n^2+2n+3)}, \\ \epsilon &= 3[1+\delta(n+1)][1+(i-j)j], \end{aligned}$$

and $\widetilde{b}^{-1}_{j,i} = \widetilde{b}^{-1}_{i,j}$ for $i < j$. Furthermore,

Lemma 11. *$\widetilde{B}_n$ is positive definite.*

Proof. For any nonzero vector $\boldsymbol{x} = (x_1 \dots x_n)^T \in \mathbb{R}^n$, $\boldsymbol{x}^T\widetilde{B}\boldsymbol{x} = \boldsymbol{x}^T B\boldsymbol{x} + x_1^2 + x_n^2 > 0$. □

4.1 Exact inverse formula

Let $\widetilde{D}_n = \widetilde{B}_n C_n = [\widetilde{d}_{i,j}]$. The general inverse formula of $\widetilde{D}_n$ is given by (19), with $\widetilde{d}$ and $\widetilde{b}$ replacing d and b, respectively. If $\widetilde{A}$ is invertible, its inverse can be expressed as

$$\widetilde{A}_n^{-1} = \widetilde{D}_n^{-1} - 2\widetilde{D}_n^{-1}U\widetilde{M}^{-1}V^T\widetilde{D}_n^{-1}, \tag{28}$$

where $\widetilde{M} = [\widetilde{m}_{i,j}] = I_2 + 2V^T\widetilde{D}^{-1}U \in \mathbb{R}^{2\times 2}$ with

$$\widetilde{m}_{11} = 1 + 8\widetilde{d}^{-1}_{1,1} - 2\widetilde{d}^{-1}_{1,2} = 1 - 2\widetilde{d}^{-1}_{n,n-1} + 8\widetilde{d}^{-1}_{n,n} = \widetilde{m}_{22}, \tag{29}$$

$$\widetilde{m}_{12} = -2\widetilde{d}^{-1}_{1,n-1} + 8\widetilde{d}^{-1}_{1,n} = 8\widetilde{d}^{-1}_{n,1} - 2\widetilde{d}^{-1}_{n,2} = \widetilde{m}_{21}. \tag{30}$$

An explicit form for $\widetilde{m}_{11}$ and $\widetilde{m}_{12}$ can be obtained via direct calculations using the formulas (19) and (27) and Lemma 4, and is given by

$$\begin{aligned} \widetilde{m}_{11} &= 1 + \frac{3(n^3+3n^2+n+1)\gamma_{n+1} - 3(n^3+3n^2+4n+2)\gamma_n - 3(n+1)}{9\gamma_{n+1}(n+1)(n^2+2n+3)}, \\ \widetilde{m}_{12} &= \frac{6(2n+1)\gamma_{n+1} - 3(n+1)\gamma_n - 3(n+1)((n+1)^2+1)}{9\gamma_{n+1}(n+1)(n^2+2n+3)}. \end{aligned}$$

Lemma 12. *$\widetilde{M}$ is a positive, diagonally dominant matrix. Moreover,* $\det(\widetilde{M}) > 0$.

Proof. Writing $\widetilde{m}_{11} = 1 + \tau_{11}/9\gamma_{n+1}(n+1)(n^2+2n+3)$ with $\tau_{11} = (3n^3+9n^2+3n+3)\gamma_{n+1} - 3(n^3+3n^2+4n+2)\gamma_n - 3(n+1)$, we have, for $n \geq 1$,

$$\tau_{11} = 3[(n^3+3n^2+n+1)(\gamma_{n+1}-\gamma_n) - (3n+1)\gamma_n - (n+1)]$$

$$
\begin{aligned}
&\geq 3[(n^3+3n^2+n+1)(\gamma_{n+1}-\gamma_n)-3(n+1)(\gamma_n+1)]\\
&= 3\gamma_n[(n^3+3n^2+n+1)(\gamma_{n+1}/\gamma_n-1)-3(n+1)(1+1/\gamma_n)]\\
&\geq 3\gamma_n[(n^3+3n^2+n+1)(4-1)-3(n+1)(1+1/\gamma_n)] \qquad \text{(Lemma 2(ii))}\\
&\geq 3\gamma_n[3(n^3+3n^2+n+1)-6(n+1)]\\
&\geq 9\gamma_n(n^3+3n^2-2n-1)>0.
\end{aligned}
$$

Thus, $\widetilde{m}_{11}>1$.

Let $\tau_{12}=6(2n+1)\gamma_{n+1}-3(n+1)\gamma_n-3(n+1)((n+1)^2+1)$, the numerator term of $\widetilde{m}_{1,2}$. By using Lemma 2(ii), we have $\tau_{12}=6(2n+1)\gamma_{n+1}-3(n+1)\gamma_n-3(n+1)((n+1)^2+1)\geq\frac{45n+21}{4}\gamma_{n+1}-3(n+1)((n+1)^2+1)>3(n+1)(\gamma_{n+1}-(n+1)^2-1)>0$, where the inequality $\gamma_{n+1}-(n+1)^2-1>0$ can be proved by induction. Thus, $\widetilde{m}_{12}>0$, which shows the positivity of $\widetilde{M}$.

Moreover,

$$
\begin{aligned}
\tau_{11}-\tau_{12} &= (3n^3+9n^2-9n-3)\gamma_{n+1}-3(n^3+3n^2+3n+1)\gamma_n+3(n+1)((n+1)^2)\\
&\geq 3(n^3+3n^2-3n-1)\gamma_{n+1}-\frac{3(n+1)^3\gamma_{n+1}}{4}+3(n+1)^3\\
&= \frac{3(3n^3+9n^2-15n-5)\gamma_{n+1}}{4}+3(n+1)^3\\
&\geq 3\gamma_{n+1}(n^2-3n)>0
\end{aligned}
$$

for $n\geq 3$. Hence, $\tau_{11}>\tau_{12}$ and consequently, $\widetilde{m}_{11}>\widetilde{m}_{12}$. □

Theorem 13. *$\widetilde{A}_n$ is positive definite.*

Proof. Let us denote an upper-left $k\times k$ matrix of $\widetilde{A}$ as $\widetilde{A}_{k,k}$. By Silvester's criterion, we need to show $\det\widetilde{A}_{k,k}>0$ for $k\in\{1,2,\ldots,n\}$. For $k\in 1,\ldots,6$, $\det(A_{k,k})>0$ by numerical calculation. For $k=n$, $\det\widetilde{A}_{n,n}=\det\widetilde{A}_n=\det\widetilde{M}\det\widetilde{B}_n\det\widetilde{C}_n>0$, due to Lemmas 6, 11, and 12.

For $k\in\{7,\ldots,n-1\}$, $\widetilde{A}_{k,k}$ is a seven-diagonal nearly Toeplitz matrix (1), with perturbed top-left corner. For any nonzero vector $\boldsymbol{x}\in\mathbb{R}^k$, then

$$
\begin{aligned}
\boldsymbol{x}^T\widetilde{A}_{k,k}\boldsymbol{x} = {}& 68x_1^2+56(x_2^2+\cdots+x_k^2)-80x_1x_2-78(x_2x_3+\cdots+x_{k-1}x_k)\\
&+24(x_1x_3+\cdots+x_{k-2}x_k)-2(x_1x_4+\cdots+x_{k-3}x_k).
\end{aligned}
$$

Consider

$$
S=\sum_{i=1}^{k-3}(ax_i-bx_{i+1}+cx_{i+2}-dx_{i+3})^2
$$

where $a=\sqrt{4-\sqrt{15}}$, $b=(6+\sqrt{15})a$, $c=(9+2\sqrt{15})a$, and $d=(4+\sqrt{15})a$, with properties

$$
\begin{cases}
a^2+b^2+c^2+d^2=56,\\
ab+bc+cd=39,\\
ac+bd=12,\\
ad=1.
\end{cases}
$$

Then

$$
\begin{aligned}
\boldsymbol{x}^T\widetilde{A}_{k,k}\boldsymbol{x} &= S+(68-a^2)x_1^2+(c^2+d^2)x_2^2+d^2x_3^2+a^2x_{k-2}^2+(a^2+b^2)x_{k-1}^2+(a^2+b^2+c^2)x_k^2\\
&\quad-(80-2ab)x_1x_2-2cdx_2x_3-2abx_{k-2}x_{k-1}-(2ab+2bc)x_{k-1}x_k+2bdx_1x_3+2acx_{k-2}x_k\\
&= S+(68-a^2-b^2)x_1^2+d^2x_2^2+(bx_1-cx_2+dx_3)^2-(80-2ab-2bc)x_1x_2\\
&\quad+(ax_{k-2}-bx_{k-1}+cx_k)^2+a^2x_{k-1}^2-2abx_{k-1}x_k+(a^2+b^2)x_k^2\\
&= S+(12+c^2+d^2)x_1^2+d^2x_2^2+(bx_1-cx_2+dx_3)^2-(2+2cd)x_1x_2+(ax_{k-2}-bx_{k-1}+cx_k)^2
\end{aligned}
$$

$$
\begin{aligned}
&+ (ax_{k-1} - bx_k)^2 + a^2 x_k^2 \\
&= S + (\sqrt{12 + c^2 + d^2}x_1 - \frac{1 + cd}{\sqrt{12 + c^2 + d^2}}x_2)^2 + \left(d^2 - \frac{(1 + cd)^2}{12 + c^2 + d^2}\right)x_2^2 + (bx_1 - cx_2 + dx_3)^2 \\
&+ (ax_{k-2} - bx_{k-1} + cx_k)^2 + (ax_{k-1} - bx_k)^2 + a^2 x_k^2 \geq 0,
\end{aligned}
$$

with equality holding only when $\boldsymbol{x} = 0$. (Note that $d^2 - \frac{(1+cd)^2}{12+c^2+d^2} \approx 2.20 > 0$) □

4.2 Bounds of norms of inverse matrix

In this section, we shall derive an upper bound for norms of $\widetilde{A}_n^{-1}$. As we did for A_n^{-1}, the derivation will be given only for $p = \infty$.

Positive definiteness of $\widetilde{A}_n$ implies that $\widetilde{A}_n^{-1}$ is positive. Consequently,

$$
\sum_{j=1}^{n} |\widetilde{a}_{i,j}^{-1}| = \sum_{j=1}^{n} \widetilde{a}_{i,j}^{-1} = \sum_{j=1}^{n} \widetilde{d}_{i,j}^{-1} - \frac{2}{\widetilde{m}_{11} + \widetilde{m}_{12}} \left(\sum_{j=1}^{n} \widetilde{d}_{n,j}^{-1}\right) \left(4(\widetilde{d}_{i,n} + \widetilde{d}_{i,1}^{-1}) - (\widetilde{d}_{i,n-1}^{-1} + \widetilde{d}_{i,2}^{-1})\right).
$$

The following inequality can be derived using the above expression:

$$
\begin{aligned}
\|\widetilde{A}^{-1}\|_\infty &= \max_i \sum_{j=1}^{n} |\widetilde{a}_{i,j}^{-1}| = \max_i \left\{ \sum_{j=1}^{n} \widetilde{d}_{i,j}^{-1} - \frac{2}{\widetilde{m}_{11} + \widetilde{m}_{12}} \left(\sum_{j=1}^{n} \widetilde{d}_{n,j}^{-1}\right) \left(4(\widetilde{d}_{i,n} + \widetilde{d}_{i,1}^{-1}) - (\widetilde{d}_{i,n-1}^{-1} + \widetilde{d}_{i,2}^{-1})\right) \right\} \\
&\leq \underbrace{\max_i \sum_{j=1}^{n} \widetilde{d}_{i,j}^{-1}}_{\widetilde{\pi}_1} + \frac{2}{\widetilde{m}_{11} + \widetilde{m}_{12}} \underbrace{\left(\sum_{j=1}^{n} \widetilde{d}_{n,j}^{-1}\right)}_{\widetilde{\pi}_2} \underbrace{\max_i \widetilde{g}(i)}_{\widetilde{\pi}_3},
\end{aligned} \tag{31}
$$

where

$$
\widetilde{g}(i) = 4\underbrace{(\widetilde{d}_{i,n}^{-1} + \widetilde{d}_{i,1}^{-1})}_{\widetilde{\theta}_1} - \underbrace{(\widetilde{d}_{i,n-1}^{-1} + \widetilde{d}_{i,2}^{-1})}_{\widetilde{\theta}_2}. \tag{32}
$$

With $\|B_n^{-1}\|_\infty \leq (n+1)^2((n+1)^2 + 8)/384$ (see [9]) and Lemma 6, we have

$$
\widetilde{\pi}_1 = \max_i \sum_{j=1}^{n} \widetilde{d}_{i,j}^{-1} = \|\widetilde{D}^{-1}\|_\infty \leq \|\widetilde{B}^{-1}\|_\infty \|C^{-1}\|_\infty \leq \frac{(n+1)^2((n+1)^2 + 8)}{2304}.
$$

Next,

$$
\widetilde{d}_{n,j}^{-1} = \frac{\gamma_1}{\gamma_{n+1}} \left(\sum_{k=1}^{j} \gamma_k \widetilde{b}_{k,j}^{-1} + \sum_{k=j+1}^{n} \gamma_k \widetilde{b}_{k,j}^{-1}\right) = \frac{\gamma_1}{\gamma_{n+1}} \left(\sum_{k=1}^{j} \gamma_k \widetilde{b}_{j,k}^{-1} + \sum_{k=j+1}^{n} \gamma_k \widetilde{b}_{k,j}^{-1}\right).
$$

By using the explicit formula for $\widetilde{b}_{i,j}^{-1}$, $i \geq j$ and Lemma 4, after tedious calculation we get

$$
\widetilde{d}_{n,j}^{-1} = \mu(\nu_3 j^3 + \nu_2 j^2 + \nu_1 j + \nu_0 \gamma_j), \tag{33}
$$

where

$$
\begin{aligned}
\mu &= \frac{\gamma_1}{36\gamma_{n+1}(n+1)(n^2 + 2n + 3)}, \\
\nu_0 &= n^3 + 3n^2 + 5n + 3, \\
\nu_1 &= 2(2n+1)\gamma_{n+1} - (n+1)\gamma_n - (n^3 + 3n^2 + 4n + 2), \\
\nu_2 &= (4n^2 + 5n - 3)\gamma_{n+1} - n(n+2)\gamma_n + 2n^2 + 4n + 3,
\end{aligned}
$$

$$\nu_3 = -2(2n+1)\gamma_{n+1} + (n+1)\gamma_n - (n+1).$$

Therefore,

$$\begin{aligned}
\widetilde{\pi}_2 &= \sum_{j=1}^{n} \widetilde{d}_{n,j}^{-1} = \kappa \left(\nu_3 \sum_{j=1}^{n} j^3 + \nu_2 \sum_{j=1}^{n} j^2 + \nu_1 \sum_{j=1}^{n} j + \nu_0 \sum_{j=1}^{n} \gamma_j \right), \\
&= \kappa \left(\nu_3 \frac{n^2(n+1)^2}{4} + \nu_2 \frac{n(n+1)(2n+1)}{6} + \nu_1 \frac{n(n+1)}{2} + \nu_0 \frac{\gamma_{n+1} - \gamma_n - 1}{6} \right) \\
&= \frac{\gamma_1}{36\gamma_{n+1}} \left(\frac{1}{6}(2n^2+n+1)(\gamma_{n+1}-1) + \frac{n^2}{2(n^2+2n+3)} - \frac{1}{12}(n^2+2n+2)\gamma_n \right) \\
&\le \frac{\gamma_1}{36\gamma_{n+1}} \left(\frac{1}{6}(2n^2+n+1)\gamma_{n+1} + \frac{n^2}{2(n^2+2n+3)} - \frac{1}{12}(n^2+2n+2)\gamma_n \right).
\end{aligned}$$

With Lemma 2(ii) and $\gamma_{n+1} \ge 4\gamma_n \ge 4^2\gamma_{n-1} \ge 4^{n+1}$, for $n \ge 2$,

$$\begin{aligned}
\widetilde{\pi}_2 &\le \frac{1}{36} \left(\frac{1}{6}(2n^2+n+1) + \frac{n^2}{2\gamma_{n+1}(n^2+2n+3)} - \frac{1}{96}(n^2+2n+2) \right) \\
&= \frac{1}{36} \left(\frac{31n^2+14n+14}{96} + \frac{n^2}{2\gamma_{n+1}(n^2+2n+3)} \right) \\
&\le \frac{1}{36} \left(\frac{31n^2+14n+14}{96} + \frac{1}{2\cdot 4^3} \right) \\
&\le \frac{1}{3456}(31n^2+14n+15).
\end{aligned}$$

We now construct an estimate for $\widetilde{\pi}_3$. Using (19), we have

$$\begin{aligned}
\widetilde{d}_{i,1}^{-1} &= \frac{\gamma_{n+1-i}}{\gamma_{n+1}} \sum_{k=1}^{i} \gamma_k \widetilde{b}_{k,1}^{-1} + \frac{\gamma_i}{\gamma_{n+1}} \sum_{k=1}^{n-i} \gamma_k \widetilde{b}_{n+1-k,1}^{-1}, \\
\widetilde{d}_{i,n}^{-1} &= \frac{\gamma_i}{\gamma_{n+1}} \sum_{k=1}^{n+1-i} \gamma_k \widetilde{b}_{k,1}^{-1} + \frac{\gamma_{n+1-i}}{\gamma_{n+1}} \sum_{k=1}^{i-1} \gamma_k \widetilde{b}_{n+1-k,1}^{-1}, \\
\widetilde{d}_{i,2}^{-1} &= \begin{cases} \dfrac{\gamma_1}{\gamma_{n+1}} \displaystyle\sum_{k=1}^{n} \gamma_k \widetilde{b}_{k,n-1}^{-1}, & i = 1, \\ \dfrac{\gamma_{n+1-i}}{\gamma_{n+1}} \displaystyle\sum_{k=1}^{i} \gamma_k \widetilde{b}_{k,2}^{-1} + \dfrac{\gamma_i}{\gamma_{n+1}} \displaystyle\sum_{k=1}^{n-i} \gamma_k \widetilde{b}_{n+1-k,2}^{-1}, & 2 \le i \le n, \end{cases} \\
\widetilde{d}_{i,n-1}^{-1} &= \begin{cases} \dfrac{\gamma_i}{\gamma_{n+1}} \displaystyle\sum_{k=1}^{n+1-i} \gamma_k \widetilde{b}_{k,2}^{-1} + \dfrac{\gamma_{n+1-i}}{\gamma_{n+1}} \displaystyle\sum_{k=1}^{i-1} \gamma_k \widetilde{b}_{n+1-k,2}^{-1}, & 1 \le i \le n-1, \\ \dfrac{\gamma_1}{\gamma_{n+1}} \displaystyle\sum_{k=1}^{n} \gamma_k \widetilde{b}_{k,n-1}^{-1}, & i = n. \end{cases}
\end{aligned}$$

Hence

$$\begin{aligned}
\widetilde{\theta}_1 &= \frac{\gamma_i}{\gamma_{n+1}} \sum_{k=1}^{n-i} \gamma_k \mathcal{B}_1(k) + \frac{\gamma_i \gamma_{n+1-i}}{\gamma_{n+1}} (\widetilde{b}_{i,1}^{-1} + \widetilde{b}_{n+1-i,1}^{-1}) + \frac{\gamma_{n+1-i}}{\gamma_{n+1}} \sum_{k=1}^{i-1} \left\{ \gamma_k \mathcal{B}_1(k) \right\}, \\
\widetilde{\theta}_2 &= \frac{\gamma_i}{\gamma_{n+1}} \sum_{k=1}^{n-i} \left\{ \gamma_k \mathcal{B}_2(k) \right\} + \frac{\gamma_i \gamma_{n+1-i}}{\gamma_{n+1}} (\widetilde{b}_{i,2}^{-1} + \widetilde{b}_{n+1-i,2}^{-1}) + \frac{\gamma_{n+1-i}}{\gamma_{n+1}} \sum_{k=1}^{i-1} \left\{ \gamma_k \mathcal{B}_2(k), \right\}.
\end{aligned}$$

where, for $k = 1, \ldots, n$,

$$\mathcal{B}_1(k) = \widetilde{b}_{k,1}^{-1} + \widetilde{b}_{n+1-k,1}^{-1} = \frac{-k^2 + (n+1)k}{2(n+1)},$$

$$\mathcal{B}_2(k) = \widetilde{b}_{k,2}^{-1} + \widetilde{b}_{n+1-k,2}^{-1} = \frac{-2k^2 + 2(n+1)k - (n+1)}{n+1},$$

after direct calculations of each $\widetilde{b}$ using (27). Furthermore, $4\mathcal{B}_1(k) - \mathcal{B}_2(k) = 1$.

We shall now use the above intermediate results and Lemmas 4 to derive an expression for $\widetilde{g}$:

$$\begin{aligned}
\widetilde{g}(i) &= 4\widetilde{\theta}_1(i) - \widetilde{\theta}_2(i) \\
&= \frac{\gamma_i}{\gamma_{n+1}} \sum_{k=1}^{n-i} \{\gamma_k(4\mathcal{B}_1(k) - \mathcal{B}_2(k))\} + \frac{\gamma_i\gamma_{n+1-i}}{\gamma_{n+1}}(4\mathcal{B}_1(k) - \mathcal{B}_2(k)) + \frac{\gamma_{n+1-i}}{\gamma_{n+1}} \sum_{k=1}^{i-1} \{\gamma_k(4\mathcal{B}_1(k) - \mathcal{B}_2(k))\} \\
&= \frac{\gamma_i}{\gamma_{n+1}} \sum_{k=1}^{n-i} \gamma_k + \frac{\gamma_i\gamma_{n+1-i}}{\gamma_{n+1}} + \frac{\gamma_{n+1-i}}{\gamma_{n+1}} \sum_{k=1}^{i-1} \gamma_k \\
&= \frac{4}{3}\frac{\gamma_i\gamma_{n+1-i}}{\gamma_{n+1}} - \frac{1}{6}\frac{\gamma_i\gamma_{n-i}}{\gamma_{n+1}} - \frac{1}{6}\frac{\gamma_{i-1}\gamma_{n+1-i}}{\gamma_{n+1}} - \frac{\gamma_i + \gamma_{n+i-1}}{6\gamma_{n+1}}.
\end{aligned} \tag{34}$$

Considering $i \in [1, n] \subset \mathbb{R}$ and with

$$\begin{aligned}
(\gamma_i\gamma_{n+1-i})' &= \frac{1}{2\sqrt{15}}\gamma_{n+1-2i} \ln\left(\frac{r_1}{r_2}\right), \\
(\gamma_i\gamma_{n-i})' &= \frac{1}{2\sqrt{15}}\gamma_{n-2i} \ln\left(\frac{r_1}{r_2}\right), \\
(\gamma_{i-1}\gamma_{n+1-i})' &= \frac{1}{2\sqrt{15}}\gamma_{n-2i+2} \ln\left(\frac{r_1}{r_2}\right),
\end{aligned}$$

we have

$$\begin{aligned}
\widetilde{g}'(i) &= \frac{4}{3}\left(\frac{\gamma_i\gamma_{n+1-i}}{\gamma_{n+1}}\right)' - \frac{1}{6}\left(\frac{\gamma_i\gamma_{n-i}}{\gamma_{n+1}}\right)' - \frac{1}{6}\left(\frac{\gamma_{i-1}\gamma_{n+1-i}}{\gamma_{n+1}}\right)' - \left(\frac{\gamma_i + \gamma_{i+1}}{6\gamma_{n+1}}\right)' \\
&= -\frac{r_1^i(1 - r_1^{n+1-2i})\ln r_1 - r_2^i(1 - r_2^{n+1-2i})\ln r_2}{12\sqrt{15}\gamma_{n+1}} \\
&= \frac{\ln r_1(\alpha_{n+1-i} - \alpha_i)}{6\sqrt{15}\gamma_{n+1}}.
\end{aligned}$$

The critical point is $i = (n+1)/2$, which is also the maximum of $\widetilde{g}(i)$. Therefore,

$$\begin{aligned}
\widetilde{\pi}_3 \le \widetilde{K}((n+1)/2) &= \frac{4}{3}\frac{\gamma_{\frac{n+1}{2}}^2}{\gamma_{n+1}} - \frac{1}{3}\frac{\gamma_{\frac{n+1}{2}}\gamma_{\frac{n-1}{2}}}{\gamma_{n+1}} - \frac{1}{3}\frac{\gamma_{\frac{n+1}{2}}}{\gamma_{n+1}} \\
&\le \frac{4}{3}\frac{\gamma_{\frac{n+1}{2}}^2}{\gamma_{n+1}} - \frac{1}{24}\frac{\gamma_{\frac{n+1}{2}}^2}{\gamma_{n+1}} = \frac{31\gamma_{\frac{n+1}{2}}^2}{24\gamma_{n+1}} = \frac{31}{48\sqrt{15}}\frac{(r_1^{\frac{n+1}{2}} - r_2^{\frac{n+1}{2}})^2}{r_1^{n+1} - r_2^{n+1}} \\
&= \frac{31}{48\sqrt{15}}\frac{1 - 1/r_1^{n+1}}{1 + 1/r^{n+1}} \le \frac{31}{48\sqrt{15}}.
\end{aligned}$$

Theorem 14. *For the matrix* (1)*, with* $a_0 = 68$ *and* $a_1 = 40$*, the following inequality holds for* $p \in \{1, 2, \infty\}$*:*

$$\|\widetilde{A}_n^{-1}\|_p \le \frac{(n+1)^2((n+1)^2 + 14)}{2304}.$$

Proof. Notice that

$$\begin{aligned}
\widetilde{m}_{11} + \widetilde{m}_{12} &= 1 + \frac{3(n^3 + 3n^2 + 15n + 9)\gamma_{n+1} - 3(n^3 + 3n^2 + 5n + 3)\gamma_n + 3(n+1)(n^2 + 2n + 3)}{9\gamma_{n+1}(n+1)(n^2 + 2n + 3)} \\
&= 1 + \frac{1}{3} + \frac{10n + 6}{3(n+1)(n^2 + 2n + 3)} - \frac{1}{3}\frac{\gamma_n}{\gamma_{n+1}} + \frac{1}{3\gamma_{n+1}} \\
&\ge \frac{4}{3} - \frac{1}{3}\frac{\gamma_n}{\gamma_{n+1}} \ge \frac{4}{3} - \frac{1}{3} \times \frac{1}{4} = \frac{5}{4},
\end{aligned}$$

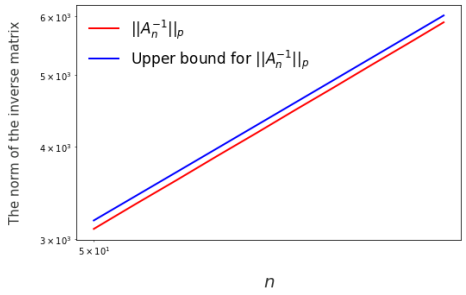
after applying Lemma 2(ii). By using $\widetilde{\pi}_1, \widetilde{\pi}_2, \widetilde{\pi}_3$ as given above, the bound (31) reads, for $n \geq 1$,

$$\begin{aligned}\|\widetilde{A}^{-1}\|_\infty &\leq \frac{1}{2304}((n+1)^2((n+1)^2+8) + \frac{31(31n^2+14n+15)}{103680\sqrt{15}} \\ &\leq \frac{(n+1)^2((n+1)^2+8)}{2304} + \frac{6(n+1)^2}{2304} = \frac{(n+1)^2((n+1)^2+14)}{2304}.\end{aligned}$$

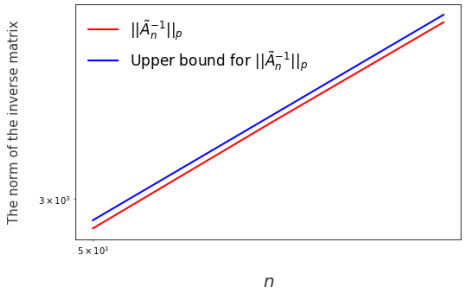
□

5 Numerical examples

We have computed the norms of exact inverses of A_n and $\widetilde{A}$ for various size n and the proposed upper bounds from Theorem 10 and 14. The computational results are presented in Figure 1 various matrix size n, which suggests a good estimate provided by the theorem.



(a) A_n^{-1} (Toeplitz case) norm and the upper bound computations



(b) $\widetilde{A}_n^{-1}$ (nearly Toeplitz case) norm and upper bound computations

Figure 1: Evaluation of norm of inverse of matrices and bound in the log scale.

By utilizing the fixed point iteration (5) and assuming $f(\boldsymbol{u}) = e^{-\boldsymbol{u}}$ we have computed the convergence rate, $Lp := h^4 C_{EI} \|A^{-1}\|_p$; see for more details in [13]. In our setting the iteration convergence was achieved when $\|\boldsymbol{u}^{\ell+1} - \boldsymbol{u}^{\ell}\|_p < 10^{-6}$ for $p \in \{1, 2, \infty\}$. For $n = 50$ and $C_{EI} = 1$ we obtained $L1 = 4.69 \times 10^{-4}$ and $L2 = 3.60 \times 10^{-4}$ the upper bound from Theorem 14 is 4.72×10^{-4}. The solution of (4) is presented in Figure 2.

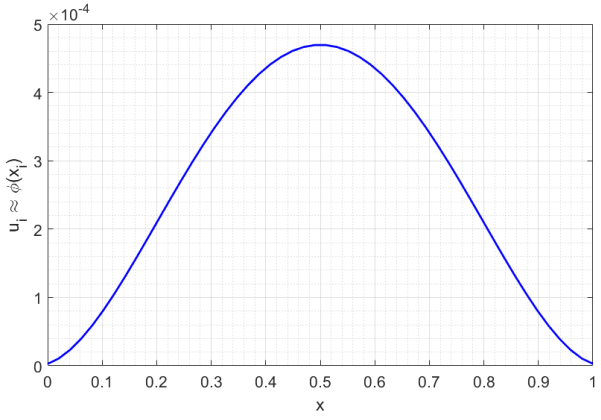


Figure 2: The solution of (4) is presented in blue for $n = 50$ and $C_{EI} = 1$.

6 Generalization to perturbed 2 × 2 corner blocks

In this section, we analyze exact inverse properties of A_n given by (1), with $n \geq 7$. To proceed with the analysis, we assume that $a_0, a_1, a_2 > 0$ and are chosen such that A_n^{-1} exists, and consider the rank-2 decomposition of A_n as follows:

$$A_n = B_n C_n + UV^T, \tag{35}$$

where B_n and C_n are given in (17) and (8), respectively, with

$$U^T = \begin{bmatrix} u_1 & u_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & u_2 & u_1 \end{bmatrix}_{n\times 2}, \quad \text{and } V^T = \begin{bmatrix} v_1 & v_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & v_2 & v_1 \end{bmatrix}_{n\times 2}.$$

The inverse of A_n is given by the Sherman-Morrison formula

$$A_n^{-1} = D_n^{-1} - D_n^{-1} U M_2^{-1} V^T D_n^{-1},$$

where $D_n = B_n C_n$ and $M_2 = I_2 + V^T D_n^{-1} U \in \mathbb{R}^{2\times 2}$. The inverse of B_n, C_n, and D_n are given entry-wise by the formula (18), Lemma 1, and (20), respectively. From (35), relations between $a_0, a_1, a_2, u_1, u_2, v_1$, and v_2 can be established: $u_1 v_1 = a_0 - 52$, $u_1 v_2 = 39 - a_1$, $u_2 v_1 = 38 - a_1$, and $u_2 v_2 = a_2 - 56$.

6.0.0.1 Explicit inverse formula.

Let $M_2 = [m_{i,j}]_{i,j=1,2}$. The entries of M_2 are determined explicitly by the formula

$$\begin{aligned}
m_{1,1} &= 1 + (a_0 - 52)d_{1,1}^{-1} + (38 - a_1)d_{1,2}^{-1} + (39 - a_1)d_{2,1}^{-1} + (a_2 - 56)d_{2,2}^{-1};\\
m_{1,2} &= (38 - a_1)d_{1,n-1}^{-1} + (a_0 - 52)d_{1,n}^{-1} + (a_2 - 56)d_{2,n-1}^{-1} + (39 - a_1)d_{2,n}^{-1};\\
m_{2,1} &= (39 - a_1)d_{n-1,1}^{-1} + (a_2 - 56)d_{n-1,2}^{-1} + (a_0 - 52)d_{n,1}^{-1} + (38 - a_1)d_{n,2}^{-1};\\
m_{2,2} &= 1 + (a_2 - 56)d_{n-1,n-1}^{-1} + (39 - a_1)d_{n-1,n}^{-1} + (38 - a_1)d_{n,n-1}^{-1} + (a_0 - 52)d_{n,n}^{-1}.
\end{aligned} \tag{36}$$

Due to centrosymmetry of D_n, we have that $m_{1,1} = m_{2,2}$ and $m_{1,2} = m_{2,1}$. Thus, M_2 is symmetric. The d^{-1}'s in the above formulas for $m_{i,j}^{-1}$ are given as follows:

$$\begin{aligned}
d_{1,1}^{-1} &= \frac{\gamma_1\gamma_n}{36\gamma_{n+1}} + \eta\left(\frac{n(n+1)^2}{6}\frac{\gamma_n}{\gamma_{n+1}} + \frac{n(n+1)}{3}\frac{\gamma_1}{\gamma_{n+1}} - \frac{n(n+1)(3n+1)}{3}\right),\\
d_{1,2}^{-1} &= d_{n,n-1}^{-1} = \frac{\gamma_1\gamma_{n-1}}{36\gamma_{n+1}} + \eta\left((n+1)(n-1)\frac{\gamma_1}{\gamma_{n+1}} + \frac{n(n-1)(n+1)}{3}\frac{\gamma_n}{\gamma_{n+1}} - \frac{n(n-1)(13n+7)}{6}\right),\\
d_{2,1}^{-1} &= \frac{\gamma_1\gamma_{n-1}}{36\gamma_{n+1}} + \eta\left(\frac{n(n+1)^2}{6}\frac{\gamma_{n-1}}{\gamma_{n+1}} + \frac{n(n+1)}{3}\frac{\gamma_2}{\gamma_{n+1}} - \frac{(n+1)(6n^2-8n+3)}{3}\right),\\
d_{2,2}^{-1} &= \frac{\gamma_2\gamma_{n-1}}{36\gamma_{n+1}} + \eta\left(\frac{n(n-1)(n+1)}{3}\frac{\gamma_{n-1}}{\gamma_{n+1}} + (n+1)(n-1)\frac{\gamma_2}{\gamma_{n+1}} - n(n-1)(5n-3)\right),\\
d_{1,n-1}^{-1} &= d_{n,2}^{-1} = \frac{\gamma_2\gamma_1}{36\gamma_{n+1}} + \eta\left(\frac{n(n-1)(n+1)}{3}\frac{\gamma_1}{\gamma_{n+1}} + (n+1)(n-1)\frac{\gamma_n}{\gamma_{n+1}} - 7n^2 + 3n + 8\right),\\
d_{1,n}^{-1} &= d_{n,1}^{-1} = \frac{\gamma_1\gamma_1}{36\gamma_{n+1}} + \eta\left(\frac{n(n+1)^2}{6}\frac{\gamma_1}{\gamma_{n+1}} + \frac{n(n+1)}{3}\frac{\gamma_n}{\gamma_{n+1}} - \frac{(n+1)(7n-3)}{3}\right),\\
d_{2,n-1}^{-1} &= d_{n-1,2}^{-1} = \frac{\gamma_2\gamma_2}{36\gamma_{n+1}} + \eta\left(\frac{n(n-1)(n+1)}{3}\frac{\gamma_2}{\gamma_{n+1}} + (n+1)(n-1)\frac{\gamma_{n-1}}{\gamma_{n+1}} - 19n^2 + 24n + 27\right),\\
d_{2,n}^{-1} &= d_{n-1,1}^{-1} = \frac{\gamma_1\gamma_2}{36\gamma_{n+1}} + \eta\left(\frac{n(n+1)^2}{6}\frac{\gamma_2}{\gamma_{n+1}} + \frac{n(n+1)}{3}\frac{\gamma_{n-1}}{\gamma_{n+1}} - \frac{(n+1)(19n-24)}{3}\right).
\end{aligned} \tag{37}$$

Via direct calculations, $A_n^{-1} = [a_{i,j}^{-1}]$ can be determined entry-wise by the following formula:

$$a_{i,j}^{-1} = d_{i,j}^{-1}$$

$$
\begin{aligned}
&- d_{i,1}^{-1}[m_{1,1}^{-1}((a_0-52)d_{1,j}^{-1}+(39-a_1)d_{2,j}^{-1})+m_{1,2}^{-1}((39-a_1)d_{n-1,j}^{-1}+(a_0-52)d_{n,j}^{-1})]\\
&- d_{i,2}^{-1}[m_{1,1}^{-1}((38-a_1)d_{1,j}^{-1}+(a_2-56)d_{2,j}^{-1})+m_{1,2}^{-1}((a_2-56)d_{n-1,j}^{-1}+(38-a_1)d_{n,j}^{-1})]\\
&- d_{i,n-1}^{-1}[m_{1,2}^{-1}((38-a_1)d_{1,j}^{-1}+(a_2-56)d_{2,j}^{-1})+m_{1,1}^{-1}((a_2-56)d_{n-1,j}^{-1}+(38-a_1)d_{n,j}^{-1})]\\
&- d_{i,n}^{-1}[m_{1,2}^{-1}((a_0-52)d_{1,j}^{-1}+(39-a_1)d_{2,j}^{-1})+m_{1,1}^{-1}((39-a_1)d_{n-1,j}^{-1}+(a_0-52)d_{n,j}^{-1})].
\end{aligned}
$$

We note here that the above exact inverse formula holds for any choice of a_0, a_1, and a_2, provided that they lead to an invertible A_n.

6.0.0.2 Bound of norms of the inverse.

To derive a bound for norms of A_n^{-1}, we start by writing the inverse as $A_n^{-1} = D_n^{-1}(I_2 - UM_2^{-1}V^TD_n^{-1})$. Then,

$$
\begin{aligned}
\|A_n^{-1}\|_p &= \|D_n^{-1}(I_2 - UM_2^{-1}V^TD_n^{-1})\|_p\\
&\le \|D_n^{-1}\|_p\|I_2 - UM_2^{-1}V^TD_n^{-1}\|_p\\
&\le \|D_n^{-1}\|_p(1+\|UM_2^{-1}V^T\|_p\|D_n^{-1}\|_p)\\
&\le \|B_n^{-1}\|_p\|C_n^{-1}\|_p(1+\|UM_2^{-1}V^T\|_p\|B_n^{-1}\|_p\|C_n^{-1}\|_p),
\end{aligned}
$$

for $p \in \{1, 2, \infty\}$. From Theorem 4 of [8], $\|B_n^{-1}\|_p \le (n+1)^2(n+3)^3/384$, and from Lemma 6, $\|C_n^{-1}\|_p \le 1/6$. So, we need to construct a bound for $\|UM_2^{-1}V^T\|_p$. First, notice that, using the symmetry of M_2,

$$
UM^{-1}V^T = \begin{bmatrix}
(a_0-52)m_{1,1}^{-1} & (39-a_1)m_{1,1}^{-1} & 0 & \cdots & 0 & (a_0-52)m_{1,2}^{-1} & (39-a_1)m_{1,2}^{-1}\\
(38-a_1)m_{1,1}^{-1} & (a_2-56)m_{1,1}^{-1} & 0 & \cdots & 0 & (38-a_1)m_{1,2}^{-1} & (a_2-56)m_{1,2}^{-1}\\
0 & 0 & 0 & \cdots & 0 & 0 & 0\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots\\
0 & 0 & 0 & \cdots & 0 & 0 & 0\\
(38-a_1)m_{1,2}^{-1} & (a_2-56)m_{1,2}^{-1} & 0 & \cdots & 0 & (38-a_1)m_{1,1}^{-1} & (a_2-56)m_{1,1}^{-1}\\
(a_0-52)m_{1,2}^{-1} & (39-a_1)m_{1,2}^{-1} & 0 & \cdots & 0 & (a_0-52)m_{1,1}^{-1} & (39-a_1)m_{1,1}^{-1}
\end{bmatrix}.
$$

For $p = 1$,

$$
\begin{aligned}
\|UM^{-1}V^T\|_1 &= \max\{colsum_1, colsum_2, colsum_{n-1}, colsum_n\} = \max\{colsum_1, colsum_2\}\\
&= \left(\left|m_{1,1}^{-1}\right| + \left|m_{1,2}^{-1}\right|\right)\max\{|a_0-52|+|a_1-38|, |a_1-39|+|a_2-56|\}. \qquad (38)
\end{aligned}
$$

Note that $\left|m_{1,1}^{-1}\right| + \left|m_{1,2}^{-1}\right| = \|M_2^{-1}\|_1$. Similarly, for $p = \infty$, one can show that

$$
\|UM^{-1}V^T\|_\infty = \|M_2^{-1}\|_\infty \max\{|a_0-52|+|a_1-39|, |a_1-38|+|a_2-56|\}, \qquad (39)
$$

where $\|M_2^{-1}\|_\infty = \|M_2^{-1}\|_1$.

For the remaining part of this section, we shall focus on the derivation of bound for a specific choice of a_0, a_1, and a_2, as set in Theorem 15; The cases presented in Sections 3 and 4 are two specific cases under this choice. We remark, however, that it is possible to derive a bound for wider options of a_0, a_1, a_2. Due its lengthy details, this will be reported separately in a follow-up report.

Theorem 15. *For the matrix* (1)*, with* $a_0 \ge a_2 \ge 56$, $a_1 \ge 39$ *and* $6a_0 - 35a_1 + 28a_2 > 494$*, the following inequality holds for* $p \in \{1, 2, \infty\}$*:*

$$
\|A_n^{-1}\|_p \le \frac{(n+1)^2(n+3)^2}{2304}\left(1 + \frac{42(a_0+a_1-90)}{6a_0-35a_1+28a_2-494}\frac{(n+1)^2(n+3)^2}{2304}\right).
$$

Proof. With $\Delta = m_{1,1}^2 - m_{1,2}^2$, the determinant of M_2, we have

$$
\left|m_{1,1}^{-1}\right| + \left|m_{1,2}^{-1}\right| = \frac{|m_{1,1}| + |m_{1,2}|}{|\Delta|} = \frac{1}{|m_{1,1} \pm m_{1,2}|}. \qquad (40)
$$

From (36),

$$m_{1,1} \pm m_{1,2} = 1 + (a_0 - 52)(d_{1,1}^{-1} \pm d_{1,n}^{-1}) + (38 - a_1)(d_{1,2}^{-1} \pm d_{1,n-1}^{-1}) + (39 - a_1)(d_{2,1}^{-1} \pm d_{2,n}^{-1}) + (a_2 - 56)(d_{2,2}^{-1} \pm d_{2,n-1}^{-1}).$$

Direct calculations using (37) result in

$$d_{1,1}^{-1} + d_{1,n}^{-1} = \frac{\gamma_n + \gamma_1 + \gamma_{n+1}(3n-1)}{18\gamma_{n+1}(n+2)},$$
$$d_{1,2}^{-1} + d_{1,n-1}^{-1} = \frac{(\gamma_{n-1} + \gamma_2)(n+2) - 2(\gamma_n + \gamma_1)(n-1) + \gamma_{n+1}(13n-16)}{36\gamma_{n+1}(n+2)},$$
$$d_{2,1}^{-1} + d_{2,n}^{-1} = \frac{\gamma_{n-1} + \gamma_2 + \gamma_{n+1}(6n-7)}{18\gamma_{n+1}(n+2)},$$
$$d_{2,2}^{-1} + d_{2,n-1}^{-1} = \frac{(\gamma_{n-1} + \gamma_2)(n+3) + \gamma_{n+1}(5n-9)}{6\gamma_{n+1}(n+2)},$$

and

$$d_{1,1}^{-1} - d_{1,n}^{-1} = \frac{(\gamma_n - \gamma_1)(n+1) + \gamma_{n+1}(n-1)^2}{6\gamma_{n+1}(n+2)(n+3)},$$
$$d_{1,2}^{-1} - d_{1,n-1}^{-1} = \frac{\gamma_{n-1} - \gamma_2}{36\gamma_{n+1}} + \frac{(n-3)(13n^2 - 9n - 16)}{36(n+1)(n+2)(n+3)} - \frac{(\gamma_n - \gamma_1)(n-3)(n-1)}{18\gamma_{n+1}(n+2)(n+3)},$$
$$d_{2,1}^{-1} - d_{2,n}^{-1} = \frac{(\gamma_{n-1} - \gamma_2)(n+1) + \gamma_{n+1}(n-3)(2n-3)}{6\gamma_{n+1}(n+2)(n+3)},$$
$$d_{2,2}^{-1} - d_{2,n-1}^{-1} = \frac{(\gamma_{n-1} - \gamma_2)(n+7)(n+1)^2 + \gamma_{n+1}(5n+3)(n-3)^2}{6\gamma_{n+1}(n+1)(n+2)(n+3)}.$$

Using Lemma 2(ii) and with $n \geq 7$, the above sums can be bounded as follows:

$$\frac{1}{7} < d_{1,1}^{-1} \pm d_{1,n}^{-1} < \frac{1}{6},$$
$$\frac{1}{3} < d_{1,2}^{-1} \pm d_{1,n-1}^{-1} < \frac{1}{2},$$
$$\frac{1}{6} < d_{2,1}^{-1} \pm d_{2,n}^{-1} < \frac{1}{3},$$
$$\frac{2}{3} < d_{2,2}^{-1} \pm d_{2,n-1}^{-1} < 1.$$

The above bounds, together with $a_0 \geq a_2 \geq 56$ and $a_1 \geq 39$ leads to the inequality

$$m_{1,1} \pm m_{1,2} > 1 + \frac{a_0 - 52}{7} + \frac{38 - a_1}{2} + \frac{39 - a_1}{3} + \frac{2(a_2 - 56)}{3} = \frac{6a_0 - 35a_1 + 28a_2 - 494}{42}.$$

When $6a_0 - 35a_1 + 28a_2 > 494$, we have $0 < m_{1,1} \pm m_{1,2} = |m_{1,1} \pm m_{1,2}|$. Thus,

$$|m_{1,1} \pm m_{1,2}| > \frac{6a_0 - 35a_1 + 28a_2 - 494}{42},$$

or

$$\left|m_{1,1}^{-1}\right| + \left|m_{1,2}^{-1}\right| < \frac{42}{6a_0 - 35a_1 + 28a_2 - 494}.$$

Next, we use the fact that the maximum of any two positive real numbers can be written as

$$\max\{x, y\} = \frac{x + y + |x - y|}{2}.$$

Thus, for $p = 1$,

$$\max\{|a_0 - 52| + |a_1 - 38|, |a_1 - 39| + |a_2 - 56|\} = \frac{|a_0 - 52| + |a_1 - 38| + |a_1 - 39| + |a_2 - 56| + ||a_0 - 52| + |a_1 - 38| - |a_1 - 39| - |a_2 - 56||}{2}.$$

With $a_0 \geq a_2 \geq 56$ and $a_1 \geq 39$, we obtain

$$\max\{|a_0 - 52| + |a_1 - 38|, |a_1 - 39| + |a_2 - 56|\} = a_0 + a_1 - 90.$$

Similarly, for $p = \infty$, we get

$$\max\{|a_0 - 52| + |a_1 - 39|, |a_1 - 38| + |a_2 - 56|\} = a_0 + a_1 - 91.$$

Since $a_0 + a_1 - 90 > a_0 + a_1 - 91$, for (38) and (39), we have

$$\|UM^{-1}V^T\|_\infty < \|UM^{-1}V^T\|_1 \leq \frac{42(a_0 + a_1 - 90)}{6a_0 - 35a_1 + 28a_2 - 494}.$$

Finally, substitution of Lemma (6) and Theorem 4 of [8] in

$$\begin{aligned}\|\bar{A}^{-1}\|_p &\leq \|B^{-1}\|_p\|C^{-1}\|_p(1 + \|UM^{-1}V^T\|_p\|B^{-1}\|_p\|C^{-1}\|_p) \\ &\leq \frac{(n+1)^2(n+3)^2}{2304}\left(1 + \frac{42(a_0 + a_1 - 90)}{6a_0 - 35a_1 + 28a_2 - 494}\frac{(n+1)^2(n+3)^2}{2304}\right).\end{aligned}$$

□

As the bound given by Theorem 15 is applicable for the two special cases considered in Sections 3 and 4, this general result is expected to be no better than the bound derived specifically for the two special cases (Theorems 10 and 14). One can verify this easily by substituting the appropriate values of a_0, a_1, a_2. While Theorem 15 is important due to its generality, application to our specific numerical problem will lead to a rather pessimistic convergence of the fixed-point method.

7 Conclusions

In this paper, we derived the explicit formula of the inverse of seven-diagonal matrices associated with a fourth-order nonlinear boundary-value problem and give upper bounds for its norms in terms of n. The analytical bounds were compared numerically, whose quality suggesting a great potential for other applications such as numerical analysis involving the fourth-order differential operator. A generalization to a class of near-Toeplitz matrices with perturbed 2×2 block corners was also presented. Results for more general classes are possible to derive and will be considered in the future's work.

Acknowledgement: YA wishes to acknowledge the research grant, No AP08052762, from the Ministry of Education and Science of the Republic of Kazakhstan and the Nazarbayev University Faculty Development Competitive Research Grant (NUFDCRG), Grant No 110119FD4502.

Data Availability Statement: Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] W. F. Trench. Inversion of Toeplitz band matrices. *Mathematics of Computation*, 28:1089–1095, 1974.

[2] M. El-Mikkawy and E. D. Rahmo. A new recursive algorithm for inverting general periodic pentadiagonal and anti-pentadiagonal matrices. *Applied Mathematics and Computation*, 207:164–170, 2009.

[3] A. D. A. Hadj and M. Elouafi. A fast numerical algorithm for the inverse of a tridiagonal and pentadiagonal matrix. *Applied Mathematics and Computation*, 202:441–445, 2008.

[4] M. E. Kanal, N. A. Baykara, and M. Demiral. Theory and algorithm of the inversion method for pentadiagonal matrices. *Journal of Mathematical Chemistry*, 50:289–299, 2012.

[5] M. El-Mikkawy and F. Atlan. A new recursive algorithm for inverting general k-tridiagonal matrices. *Applied Mathematics Letters*, 44:34–39, 2015.
[6] A. Tănăsescu and P. G. Popescu. A fast singular value decomposition algorithm of general k-tridiagonal matrices. *Journal of Computational Science*, 31:1–5, 2019.
[7] M. A. El-Shehawey, G. A. El-Shreef, and A. Sh. Al-Henawy. Analytical inversion of general periodic tridiagonal matrices. *Journal of Mathematical Analysis and Applications*, 345(1):123–134, 2008.
[8] W. D. Hoskins and P. J. Ponzo. Some properties of a class of band matrices. *Mathematics of Computation*, 26(118):393–400, 1972.
[9] M. Dow. Explicit inverses of Toeplitz and associated matrices. *ANZIAM Journal*, 44:E185–E215, 2002.
[10] R. Peluso and T. Politi. Some improvements for two-sided bounds on the inverse of diagonally dominant tridiagonal matrices. *Linear Algebra and its Applications*, 330:1–14, 2001.
[11] D. A. Lavis and B. W. Southern. The inverse of a symmetric banded matrix. *Reports on Mathematical Physics*, 37:137–146, 1997.
[12] L. S. L. Tan. Explicit inverse of tridiagonal matrix with applications in autoregressive modelling. *IMA Journal of Applied Mathematics*, 84:679–695, 2019.
[13] B. Kurmanbek, Y. Erlangga, and Y. Amanbek. Explicit inverse of near Toeplitz pentadiagonal matrices related to higher order difference operators. *Results in Applied Mathematics*, 11:100164, 2021.
[14] Y. Amanbek, Z. Du, Y. Erlangga, C. M. da Fonseca, B. Kurmanbek, and A. Pereira. Explicit determinantal formula for a class of banded matrices. *Open Mathematics*, 18(1):1227–1229, 2020.
[15] Z. Cinkir. An elementary algorithm for computing the determinant of pentadiagonal Toeplitz matrices. *Journal of Computational and Applied Mathematics*, 236(9):2298–2305, 2012.
[16] M. Anđelić and C.M. da Fonseca. Some determinantal considerations for pentadiagonal matrices. *Linear and Multilinear Algebra*, DOI: 10.1080/03081087.2019.1708845.
[17] J. T. Jia. On a structure-preserving matrix factorization for the determinants of cyclic pentadiagonal Toeplitz matrices. *Journal of Mathematical Chemistry*, 57(8):2007–2017, 2019.
[18] B. Kurmanbek, Y. Amanbek, and Y. Erlangga. A proof of Anđelić-Fonseca conjectures on the determinant of some toeplitz matrices and their generalization. *Linear and Multilinear Algebra*, DOI: 10.1080/03081087.2020.1765959.
[19] D. S. Meek. The inverse of Toeplitz band matrices. *Linear Algebra and its Applications*, 49:117–129, 1983.
[20] V. Eijkhout and B. Polman. Decay rates of inverses of banded M-matrices that are near to Toeplitz matrices. *Linear Algebra and its Applications*, 109:247–277, 1988.
[21] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method*. Englewood Cliffs, N. J., Prentice-Hall, Inc., 1973.
[22] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathematics*, volume 37. Springer Science & Business Media, 2010.
[23] J. T. Jia and X. L. Lin. A new computational algorithm for inverting general periodic seven-diagonal matrices. *Pure and Applied Mathematics*, 26:1040–1046, 2010.
[24] X. L. Lin, P. P. Huo, and J. T. Jia. A new recursive algorithm for inverting general periodic sevendiagonal and anti-sevendiagonal matrices. *Far East Journal on Applied Mathematics*, 86:41–55, 2014.
[25] Y. Lin and X. Lin. A computational algorithm for the inverse of a sevendiagonal matrix. *Advances in Computer Science Research*, 58:298–302, 2016.
[26] Z. Huang and T. Z. Huang. Lower and upper bounds for inverse elements of strictly diagonally dominant seventh-diagonal matrices. *Journal of Applied Mathematics, Statistics and Informatics*, 27:943–953, 2009.
[27] J. Sherman and W. J. Morrison. Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics*, 21(1):124–127, 1950.
[28] M. A. Woodbury. *Inverting Modified Mmatrices*. Statistical Research Group, 1950.
[29] E. Bodewig. *Matrix Calculus*. Elsevier, 2014.
[30] Y. Amanbek, Y. Erlangga, and B. Kurmanbek. Bounds of inverse of tridiagonal (near) Toeplitz matrices. *Manuscript*, 2021.