**Electronic Transactions on Numerical Analysis.**
**Volume 65, pp. 479–499, 2026.**

**ISSN 1068–9613.**
**DOI: 10.1553/etna_vol65s479**

**ETNA**
Kent State University **and**
Johann Radon Institute (RICAM)

# COMPUTING THE EXPONENTIAL OF TRIDIAGONAL TOEPLITZ MATRICES WITH APPLICATIONS TO THE NUMERICAL SOLUTION OF THE HEAT EQUATION*

MEHDI TATARI† AND MAJED HAMADI‡

**Abstract.** The computation of the exponential of a tridiagonal matrix and its applications have long attracted considerable interest. One application considered in this work arises from the method of lines for solving the heat equation, where the partial differential equation is transformed into a system of ordinary differential equations (ODEs). The solution of this system depends on the exponential of a tridiagonal Toeplitz matrix. Strang and MacNamara [SIAM Review, 56 (2014), pp. 525–546] proposed an approximate method for computing the exponential of a symmetric tridiagonal Toeplitz matrix arising in the solution of ODEs. Their approach is based on approximating the entries of the matrix exponential by modified Bessel functions of the first kind evaluated at specific values. Consequently, the matrix exponential can be represented as the difference between a Toeplitz matrix and a Hankel matrix. In this paper, we extend this idea to the general class of tridiagonal Toeplitz matrices and improve the stability of the method by approximating the matrix exponential with a banded matrix, thereby making the computational complexity independent of the matrix size. In addition, we present an error analysis for the proposed methods and derive bounds for the entries of the exponential of tridiagonal Toeplitz matrices. As the main contribution of this work, the proposed approach is employed to solve the heat equation, and its uniform stability is established. Furthermore, by means of a splitting technique, the method is extended to two-dimensional problems. Numerical experiments demonstrate the effectiveness and efficiency of the proposed methods and the sharpness of the derived bounds.

**Key words.** tridiagonal Toeplitz matrix, matrix exponential, modified Bessel functions of the first kind, heat equation

**AMS subject classifications.** 65F60, 65M06, 65M12

## 1. Introduction.

The computation of the matrix exponential is an important problem because of its wide range of applications in science and engineering [9, 11, 12, 14, 22]. Indeed, the matrix exponential is perhaps the most extensively studied matrix function. Interest in this topic is largely motivated by its fundamental role in the solution of differential equations [9, 10, 17]. More precisely, the solution of a differential equation at time $t$ is often expressed as the action of $\exp(At)$ on a vector. The specific computational task and the qualitative properties of the underlying problem are important factors influencing the choice of a method for evaluating the matrix exponential.

Various approaches have been proposed for computing the matrix exponential, including methods based on power series expansions, Jordan and Schur decompositions, and Cauchy integral representations [17]. Other widely used techniques include the scaling and squaring method combined with Padé approximants, interpolation methods, and Krylov subspace methods [1, 6, 13, 15, 17, 21, 24].

The exponentials of tridiagonal Toeplitz matrices play a crucial role in various scientific applications and mathematical models [27]. Such matrices arise naturally in the discretization of diffusion operators and are often ill-conditioned. In the numerical solution of differential equations, systems involving tridiagonal Toeplitz matrices frequently need to be solved. Although these systems can be solved efficiently, maintaining the accuracy and preserving the qualitative properties of the resulting solutions remain challenging issues. Methods based on

*Received September 8, 2024. Accepted May 22, 2026. Published online on September 28, 2026. Recommended by Andreas Frommer.
†Corresponding author. Department of Mathematical Sciences, Isfahan University of Technology, Isfahan, 84156-83111, Iran (mtatari@iut.ac.ir).
‡Department of Mathematical Sciences, Sharif University of Technology, Tehran, Iran (majed.hamadi77@sharif.edu).

the matrix exponential provide an alternative approach and have been considered, for example, in [16]. In [18], a splitting method was proposed for computing the exponential of tridiagonal matrices arising in the numerical solution of parabolic partial differential equations (PDEs).

For the heat equation

$$u_t = au_{xx},$$

we know mathematically (and by common experience, if $u$ represents, say, temperature) that $u(x,t)$ is bounded above and below by the extremes attained by the initial data and the values on the boundary up to time $t$. This fact is known as the maximum principle, and an engineering client for our computed results would be rather dismayed if they did not possess this property [25]. In the numerical solution of differential equations, this property is achieved by stringent conditions on the discretization parameters, especially for explicit methods.

Here, based on the matrix exponential method presented in [29], we propose an explicit method for the heat equation using a banded approximation of this method. By using a splitting approach, we generalize the method to the two-dimensional heat equation. We show that both methods preserve the maximum principle unconditionally. We also extend the study of the exponential of tridiagonal Toeplitz matrices to more general cases, including both symmetric and nonsymmetric ones, and present error bounds for each. We show that these methods converge at a rate slightly faster than the geometric order and that the error bounds are sharp in numerical experiments.

To improve the approximation and accelerate the methods, we study bounds for the entries of the matrix exponential of a tridiagonal matrix. Various results have been established in this area, focusing on bounds for the entries of a function of a matrix, particularly for banded matrices [4, 5, 6, 7, 19]. One of the results closely related to the topic of this paper (provided by Iserles in [19]) gives a bound for the exponential of tridiagonal matrices. Additionally, in [8], the exponential of banded matrices was studied, and some bounds for its entries were provided. Here, we provide a bound for the entries of the exponential of a tridiagonal matrix that performs better in numerical experiments compared to the existing bounds. Using these studies, we approximate the exponential of tridiagonal matrices by banded matrices that use Bessel functions. We also use the Fast Fourier Transform (FFT) to compute the product of the approximation of the exponential of tridiagonal matrices by a vector more efficiently and faster, which is helpful in the method we provide for solving heat equations.

The remainder of our work is organized as follows. We end this section with some necessary notation. In Section 2, we consider the problem of computing the exponential of tridiagonal Toeplitz matrices and provide an error analysis of this method. In Section 3, we present an approximation of the exponential of a nonsymmetric tridiagonal matrix and analyze its error. In Section 4, we establish a bound for the entries of the matrix exponential and introduce a banded approximation method. In Section 5, we present a numerical solution and analysis of the heat equation in one and two dimensions. In Section 6, we illustrate our study with numerical examples. Concluding remarks are presented in Section 7.

**Notation.** We define $\mathrm{Toeplitz}\{\mathbf{u},\mathbf{v}\}$ to represent a Toeplitz matrix with $\mathbf{u}$ as its first column and $\mathbf{v}$ as its first row. If $\mathbf{u}=\mathbf{v}$, we simply use the notation $\mathrm{Toeplitz}\{\mathbf{u}\}$. We use $\mathrm{Hankel}\{\mathbf{v}\}$ for a Hankel matrix whose first column is $\mathbf{v}$ and whose last row is the *reverted* $\mathbf{v}$. Finally, for integers $m$ and $n$ with $m<n$, we denote the set $\{m, m+1,\dots,n-1,n\}$ by $\mathtt{m:n}$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)



**2. Exponential of tridiagonal Toeplitz matrices.** In this section, we approximate the exponential of symmetric Toeplitz tridiagonal matrices. Let $\mathbf{T}_n \in \mathbb{C}^{n\times n}$ be the tridiagonal matrix

$$\mathbf{T}_n = \operatorname{tridiag}(a,b,a) = \begin{bmatrix} b & a & & \\ a & b & \ddots & \\ & \ddots & \ddots & a \\ & & a & b \end{bmatrix}, \tag{2.1}$$

where all subdiagonal and superdiagonal elements in $\mathbf{T}_n$ are $a$, the main diagonal elements are $b$, and all other elements in $\mathbf{T}_n$ are zero.

Using the eigenvalues of $\mathbf{T}_n$ and its eigenvectors, we approximate the entries of $\exp(\mathbf{T}_n)$. The eigenvalues of $\mathbf{T}_n$ are

$$\lambda_k = b + 2a\cos\left(\frac{k\pi}{n+1}\right), \qquad k = 1,\dots,n,$$

and an eigenvector corresponding to $\lambda_k$ is given by $v^{(k)} = \big(v_1^{(k)},\dots,v_n^{(k)}\big)^T$, where

$$v_j^{(k)} = \sin\left(\frac{j\pi k}{n+1}\right), \qquad j = 1,\dots,n.$$

We can compute $a_{ij}^n$ as the $ij$-th element of $\exp(\mathbf{T}_n)$ as follows:

$$\begin{aligned} a_{ij}^n &= \frac{2\exp(b)}{n+1}\sum_{k=1}^{n}\exp\left(2a\cos\left(\frac{k\pi}{n+1}\right)\right)\sin\left(\frac{k\pi i}{n+1}\right)\sin\left(\frac{k\pi j}{n+1}\right) \\ &= \frac{\exp(b)}{n+1}\sum_{k=1}^{n}\exp\left(2a\cos\left(\frac{k\pi}{n+1}\right)\right)\cos\left(\frac{k\pi(i-j)}{n+1}\right) \\ &\quad - \frac{\exp(b)}{n+1}\sum_{k=1}^{n}\exp\left(2a\cos\left(\frac{k\pi}{n+1}\right)\right)\cos\left(\frac{k\pi(i+j)}{n+1}\right). \end{aligned}$$

We can easily see that for large values of $n$, we can approximate $a_{ij}^n$ by

$$\lim_{n\to+\infty} a_{ij}^n = e^b\big(I_{i-j}(2a) - I_{i+j}(2a)\big), \tag{2.2}$$

where $I_k(\cdot)$ are the modified Bessel functions of the first kind for $k \in \mathbb{Z}$,

$$I_k(x) = \sum_{s=0}^{\infty}\frac{1}{s!(k+s)!}\left(\frac{x}{2}\right)^{k+2s},$$

and we can represent them in integral form (as indicated in [2])

$$I_k(x) = \frac{1}{\pi}\int_0^{\pi} e^{x\cos(\theta)}\cos(k\theta)\mathrm{d}\theta.$$

According to (2.2), we can approximate $\exp(\mathbf{T}_n)$ entirely by

$$\exp(\mathbf{T}_n) \approx \text{Toeplitz}\{\mathbf{u}_n\} - \text{Hankel}\{\mathbf{v}_n\}, \tag{2.3}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)



where $\mathbf{u}_n$ and $\mathbf{v}_n$ are vectors of length $n$,

$$\mathbf{u}_n = \big(e^b I_0(2a), \dots, e^b I_{n-1}(2a)\big), \quad \mathbf{v}_n = \big(e^b I_2(2a), \dots, e^b I_{n+1}(2a)\big).$$

For the remainder of the discussion, we use the notation $\Psi(\mathbf{T}_n) \in \mathbb{C}^{n\times n}$,

$$\Psi(\mathbf{T}_n) = \text{Toeplitz}\{\mathbf{u}_n\} - \text{Hankel}\{\mathbf{v}_n\}. \tag{2.4}$$

In the rest of the paper, the method will be analyzed and extended for more complicated cases that appear in the numerical solution of differential equations.

**2.1. Error analysis.** In this section, we establish an error bound for the approximation presented in the previous section, as given in (2.3). Let $v$ be a real or complex $2\pi$-periodic function on the real line, and define

$$I := \int_0^{2\pi} v(\theta)\mathrm{d}\theta.$$

For any positive integer $N$, the trapezoidal rule for approximating the integral takes the form

$$T_N = \frac{2\pi}{N}\sum_{k=1}^{N} v\left(\frac{2\pi k}{N}\right).$$

The following theorem (proved in [20, p. 211]) provides a sharp error bound for the trapezoidal rule in approximating the integral of $2\pi$-periodic functions. This theorem is also discussed in [30].

THEOREM 2.1 ([20]). *Suppose $v$ is $2\pi$-periodic and analytic and satisfies $|v(\theta)| \le M$ in the strip $-s < Im(\theta) < s$, for some $s > 0$. Then for any $N \ge 1$,*

$$|I - T_N| \le \frac{4\pi M}{\exp(sN) - 1},$$

*and the constant $4\pi$ is the smallest possible.*

Suppose we want to approximate the integral

$$I_{ij} := \frac{1}{2\pi}\int_0^{2\pi} \exp(b + 2a\cos(\theta))\sin(i\theta)\sin(j\theta)\mathrm{d}\theta \tag{2.5}$$

with the trapezoidal rule

$$T_{2n+2}^{(i,j)} := \frac{1}{2n+2}\sum_{k=1}^{2n+2} \exp\left(b + 2a\cos\left(\frac{2k\pi}{2n+2}\right)\right)\sin\left(\frac{2k\pi i}{2n+2}\right)\sin\left(\frac{2k\pi j}{2n+2}\right).$$

In the integral (2.5) the integrand function is

$$v(\theta) = \frac{1}{2\pi}\exp\left(b + 2a\cos(\theta)\right)\sin(i\theta)\sin(j\theta).$$

To apply Theorem 2.1 in this case, the maximum value of $v(\theta)$ in the strip $-s \le \mathrm{Im}(\theta) \le s$ occurs at $\theta = \pm\mathbf{i}s$, with $\mathbf{i}$ the imaginary unit, where an upper bound for the integrand is given by

$$|v(\theta)| \le \frac{1}{2\pi}\exp\left(\mathrm{Re}(b) + 2|a|\cosh(s)\right)\cosh(is)\cosh(js),$$

**ETNA**
Kent State University and
Johann Radon Institute (RICAM)



and by using Theorem 2.1, we obtain

$$\left|I_{ij} - T^{(i,j)}_{2n+2}\right| \leq \frac{2\exp\left(\mathrm{Re}(b) + 2|a|\cosh(s)\right)\cosh(is)\cosh(js)}{\exp(s(2n+2)) - 1}, \tag{2.6}$$

which holds for all values of $s$. Since

$$\left|\frac{2}{\pi}\int_0^\pi \exp(b + 2a\cos(\theta))\sin(i\theta)\sin(j\theta)\mathrm{d}\theta - a^n_{i,j}\right| = 2\left|I_{ij} - T^{(i,j)}_{2n+2}\right|,$$

it follows from (2.6) that

$$\begin{aligned}&\left|\frac{2}{\pi}\int_0^\pi \exp(b + 2a\cos(\theta))\sin(i\theta)\sin(j\theta)\mathrm{d}\theta - a^n_{i,j}\right| \\ &\qquad\leq \frac{4e^{(\mathrm{Re}(b)+2|a|\cosh(s))}\cosh(is)\cosh(js)}{\exp(s(2n+2)) - 1}.\end{aligned} \tag{2.7}$$

The approximation $\Psi(\mathbf{T}_n)$ allows us to replace the elements of the exponential matrix with only $n+2$ modified Bessel functions, which can be computed using a standard software command or by the trapezoidal rule with $m \ll n$ quadrature points, independent of $n$.

We note that the matrix $\mathbf{T}_n$ is a per-symmetric matrix, which implies that both $\exp(\mathbf{T}_n)$ and $\Psi(\mathbf{T}_n)$ are also per-symmetric. Thus, the entries satisfying $i + j > n + 1$ can be mapped to corresponding entries $(i', j')$ such that $i' + j' \leq n + 1$, having the same value in

$$\mathbf{Er}(\mathbf{T}_n) := \exp(\mathbf{T}_n) - \Psi(\mathbf{T}_n). \tag{2.8}$$

This fact allows us to apply the bound obtained for the $(i', j')$-entry in (2.7) to the $(i, j)$-entry as well, since they have the same values in $\mathbf{Er}(\mathbf{T}_n)$.

Let $\|\cdot\|_\infty$ denote the maximum absolute row sum of a matrix.

THEOREM 2.2. *Let $\mathbf{T}_n = \mathrm{tridiag}(a, b, a) \in \mathbb{C}^{n\times n}$, and let $\mathbf{Er}(\mathbf{T}_n)$ be the matrix defined in (2.8). Then, for any $n \geq 1$,*

$$\|\mathbf{Er}(\mathbf{T}_n)\|_\infty \leq \frac{4e^{(\mathrm{Re}(b)+2|a|\cosh(s))}}{e^{s(2n+2)} - 1}\Big(G_n(s) + G_n(-s)\Big), \tag{2.9}$$

*where*

$$G_n(s) = \frac{e^{s(n+2)} - e^{2s}}{e^s - 1}, \qquad \textit{for } s > 0.$$

*Proof.* In (2.7), we have $2\cosh(is)\cosh(js) = \cosh((i + j)s) + \cosh((i - j)s)$. Using the fact that the matrix $\mathbf{Er}(\mathbf{T}_n)$ is both symmetric and per-symmetric, the infinity-norm of this matrix is therefore less than “twice” the sum of the error bounds in (2.7), for $i + j$ and $j - i$ ranging from 2 to $n + 1$. Thus,

$$\begin{aligned}\|\mathbf{Er}(\mathbf{T}_n)\|_\infty &\leq \frac{8e^{(\mathrm{Re}(b)+2|a|\cosh(s))}}{e^{s(2n+2)} - 1}\sum_{k=2}^{n+1}\cosh(ks) \\ &= \frac{4e^{(\mathrm{Re}(b)+2|a|\cosh(s))}}{e^{s(2n+2)} - 1}\left(G_n(s) + G_n(-s)\right). \qquad \square\end{aligned}$$

Theorem 2.2 introduces a family of error bounds for the infinity-norm of the matrix $\mathbf{Er}(\mathbf{T}_n)$, depending on the parameter $s$. One suitable choice for $s$ in the bound (2.9) is the

TABLE 2.1
*Infinity-norm of* $\mathbf{Er}(\mathbf{T}_n)$*, where* $\mathbf{T}_n = \mathrm{tridiag}(1, -2, 1)$*, and the error bound from Corollary 2.3.*

| $n$ | $\lVert\mathbf{Er}(\mathbf{T}_n)\rVert_\infty$ | Error bound |
|---|---|---|
| 1 | $7.99 \cdot 10^{-2}$ | 0.5 |
| 2 | $3.39 \cdot 10^{-2}$ | 0.2 |
| 3 | $9.46 \cdot 10^{-3}$ | $5.77 \cdot 10^{-2}$ |
| 4 | $1.79 \cdot 10^{-3}$ | $1.29 \cdot 10^{-2}$ |
| 5 | $2.85 \cdot 10^{-4}$ | $2.34 \cdot 10^{-3}$ |
| 6 | $3.88 \cdot 10^{-5}$ | $3.60 \cdot 10^{-4}$ |
| 7 | $4.66 \cdot 10^{-6}$ | $4.81 \cdot 10^{-5}$ |
| 8 | $5.02 \cdot 10^{-7}$ | $5.66 \cdot 10^{-6}$ |
| 9 | $4.89 \cdot 10^{-8}$ | $5.96 \cdot 10^{-7}$ |
| 10 | $4.34 \cdot 10^{-9}$ | $5.68 \cdot 10^{-8}$ |

value $s = \ln((n+1)/|a|)$, where $(n+1) > |a|$, which yields sharp estimates in the numerical examples. However, in the next corollary, by using this value of $s$, we simplify the bound from the previous theorem and make it more efficient.

COROLLARY 2.3. *Let* $\mathbf{T}_n = \mathrm{tridiag}(a, b, a) \in \mathbb{C}^{n\times n}$ *be as defined in* (2.1)*, and let* $\mathbf{Er}(\mathbf{T}_n)$ *be the matrix defined in* (2.8)*. Then, for large* $n$*,*

$$\lVert\mathbf{Er}(\mathbf{T}_n)\rVert_\infty \lesssim 2e^{\mathrm{Re}(b)} \left(\frac{|a|e}{n+1}\right)^{n+1}. \tag{2.10}$$

*Proof.* Let $s = \ln((n+1)/|a|)$, where $n+1 > |a|$. Thus, by substituting the chosen value of $s$ into the hyperbolic cosine in (2.7), we obtain $\cosh(ps) \sim 0.5\big((n+1)/|a|\big)^p$, for any $p > 0$, as $n \to \infty$. Therefore,

$$\frac{4e^{\mathrm{Re}(b)+2|a|\cosh(s)}\cosh(is)\cosh(js)}{e^{s(2n+2)}-1} \sim e^{\mathrm{Re}(b)}\left(\frac{|a|^2 e}{(n+1)^2}\right)^{n+1}\left(\frac{n+1}{|a|}\right)^{i+j}, \quad n\to\infty.$$

Similar to the proof of Theorem 2.2 and using the fact that the matrix $\mathbf{Er}(\mathbf{T}_n)$ is per-symmetric, we have

$$\begin{aligned}\lVert\mathbf{Er}(\mathbf{T}_n)\rVert_\infty &\lesssim 2e^{\mathrm{Re}(b)}\left(\frac{|a|^2 e}{(n+1)^2}\right)^{n+1}\sum_{k=2}^{n+1}\left(\frac{n+1}{|a|}\right)^{k} \\ &\sim 2e^{\mathrm{Re}(b)}\left(\frac{|a|^2 e}{(n+1)^2}\right)^{n+1}\left(\frac{n+1}{|a|}\right)^{n+1}, \quad n\to\infty. \qquad \square\end{aligned}$$

According to Corollary 2.3, it is clear that $a$ has a greater effect on the accuracy compared to $b$. Also, since $(n+1)^{n+1}$ appears in the denominator of the error bound, this causes the approximation to converge very quickly as $n$ increases. We show the effectiveness of the bound from Corollary 2.3 with a numerical example.

EXAMPLE 2.4. Let us consider $\mathbf{T}_n = \mathrm{tridiag}(1, -2, 1)$, a type of matrix that appears in the numerical solution of the heat equation. In Table 2.1 we compare the error bound obtained in Corollary 2.3 with the "infinity-norm" of the approximation error, $\lVert\mathbf{Er}(\mathbf{T}_n)\rVert_\infty$, where "`expm`" is used for computing $\exp(\mathbf{T}_n)$. The results in the table show how effectively the bound in (2.10) estimates $\lVert\mathbf{Er}(\mathbf{T}_n)\rVert_\infty$. For example, for $n = 10$, this estimate is $5.68 \cdot 10^{-8}$, which is not far from the actual error $4.34 \cdot 10^{-9}$.

In the next section, we find an approximation for nonsymmetric Toeplitz tridiagonal matrices.

**3. Nonsymmetric Toeplitz tridiagonal matrices.** In this section, we consider the case of matrices of the form $\mathbf{A}_n = \operatorname{tridiag}(a, b, c)$, which denotes a tridiagonal Toeplitz matrix with $a$ on the subdiagonal, $b$ on the main diagonal, and $c$ on the superdiagonal, while all other entries are zero.

The technique discussed in the previous section can be extended to $\mathbf{A}_n$ through a simple similarity transformation, as stated in the following theorem:

THEOREM 3.1. *For nonzero $a$ and $c$, the tridiagonal matrix $\mathbf{A}_n = \operatorname{tridiag}(a, b, c)$ satisfies the similarity relation*

$$\mathbf{A}_n = \mathbf{D}_n \mathbf{T}_n \mathbf{D}_n^{-1},$$

*where*

$$\mathbf{D}_n = \operatorname{diag}(\rho_1, \ldots, \rho_n), \quad \rho_k = \left(\sqrt{\frac{a}{c}}\right)^{k-1}, \quad \text{and} \quad \mathbf{T}_n = \operatorname{tridiag}\left(c\sqrt{\frac{a}{c}}, b, c\sqrt{\frac{a}{c}}\right).$$

*Proof.* The proof is easily obtained using matrix multiplication. $\square$

Let $\mathbf{A}$ and $\mathbf{B}$ be two $n \times n$ matrices. The Hadamard product of $\mathbf{A}$ and $\mathbf{B}$ is defined by $(\mathbf{A} \circ \mathbf{B})_{ij} = (\mathbf{A})_{ij}(\mathbf{B})_{ij}$, for all $i, j \in \{1, \ldots, n\}$. In the next theorem, we use the Hadamard product to find an approximation for the elements of $\exp(\mathbf{A}_n)$ by utilizing the results from the previous section for the symmetric case.

THEOREM 3.2. *Let $\mathbf{A}_n = \operatorname{tridiag}(a, b, c) \in \mathbb{C}^{n\times n}$. The matrix $\exp(\mathbf{A}_n)$ can be approximated by $\mathbf{F}_n \circ \Psi(\mathbf{T}_n)$ with the error bound*

$$\| \exp(\mathbf{A}_n) - \mathbf{F}_n \circ \Psi(\mathbf{T}_n) \|_\infty \leq 2\Delta e^{\operatorname{Re}(b)} \left(\frac{|z|e}{n+1}\right)^{n+1},$$

*where $\mathbf{T}_n = \operatorname{tridiag}(z, b, z)$, $z = c\sqrt{\frac{a}{c}}$, $\Psi(\mathbf{T}_n)$ as defined in (2.3), and $\mathbf{F}_n = \operatorname{Toeplitz}\{\mathbf{u}, \mathbf{v}\}$, where $\mathbf{u} = (a_0, a_1, \ldots, a_{n-1})$, $\mathbf{v} = (a_0, a_{-1}, \ldots, a_{-(n-1)})$, with $a_k = \left(\sqrt{\frac{a}{c}}\right)^k$, and $\Delta = \max_{i,j} |(\mathbf{F}_n)_{ij}|$.*

*Proof.* According to (2.3) and Theorem 3.1, we have

$$\exp(\mathbf{A}_n) \simeq \mathbf{D}_n \Psi(\mathbf{T}_n) \mathbf{D}_n^{-1},$$

and the $ij$-th element of the right-hand side of the last approximation can be rewritten as

$$\frac{\rho_i}{\rho_j} (\Psi(\mathbf{T}_n))_{ij} = \left(\sqrt{\frac{a}{c}}\right)^{i-j} (\Psi(\mathbf{T}_n))_{ij}.$$

Therefore, $\exp(\mathbf{A}_n) \simeq \mathbf{F}_n \circ \Psi(\mathbf{T}_n)$. For the error bound, we use Corollary 2.3 and the fact that for matrices $C$ and $D$, we have $\|C \circ D\|_\infty \leq \max_{i,j} |(C)_{ij}| \|D\|_\infty$. $\square$

Although the previous theorem provides a method for computing the exponential of non-symmetric Toeplitz matrices, there are challenges in its practical implementation. Regardless of the values of $a$, $b$, and $c$, some entries of $\mathbf{F}_n$ exceed 1, which, in computations for large values of $n$, may lead to numerical issues such as multiplying very large numbers by very small ones in $\mathbf{F}_n \circ \Psi(\mathbf{T}_n)$. On the other hand, computing each element of the exponential matrix is not necessary. There are well-known results in this area, called the *decay bound of matrix function elements*, which have been studied in [4, 5, 6, 7]. These results find uniform exponential decay bounds for elements of matrix functions. For our case of $\exp(\mathbf{A}_n)$, there exist constants $C > 0$ and $0 < q < 1$, such that each $ij$-th element of this matrix is bounded by $Cq^{|j-i|}$. In the next section, inspired by the decay bound results, we present a new strategy for addressing these issues and approximating $\exp(\mathbf{A}_n)$ by a banded matrix.

**ETNA**
**Kent State University and**
**Johann Radon Institute (RICAM)**



**4. Decay estimates for the exponential of tridiagonal matrices.** Here, we show that the exponential of a tridiagonal matrix can be approximated by a banded matrix, which leads to stabilization and a reduction in the computational complexity of the method. To achieve this, we first establish a bound for the entries of $\exp(\mathbf{T}_n)$ and then return to the main problem.

**4.1. Decay results.** Based on the results obtained in the previous sections, we derive a bound for the entries of the exponential of tridiagonal matrices.

THEOREM 4.1. *For a symmetric matrix* $\mathbf{T}_n = \mathrm{tridiag}(a, b, a)$ *and* $i, j \in \{1, \dots, n\}$,

$$\left|\big(\exp(\mathbf{T}_n)\big)_{ij}\right| \le 2e^{\mathrm{Re}(b)} \left(\frac{|a|e}{n+1}\right)^{n+1} + e^{\mathrm{Re}(b)} \left(I_{|i-j|}(2|a|) + I_{i+j}(2|a|)\right).$$

*Proof.* The proof is easily obtained using

$$\left|\big(\exp(\mathbf{T}_n)\big)_{ij}\right| \le \|\exp(\mathbf{T}_n) - \Psi(\mathbf{T}_n)\|_\infty + |(\Psi(\mathbf{T}_n))_{ij}|,$$

by applying Corollary 2.3, and the definition of $\Psi(\mathbf{T}_n)$ in (2.4). ∎

We focus on finding a more natural bound that does not use Bessel functions, a result that is actually studied through the research conducted on the Bessel bounds. We need the following lemmas:

LEMMA 4.2 ([23]). *For* $x \in \mathbb{R}$ *with* $x > 0$ *and* $k > -\frac{1}{2}$, *the Bessel functions* $I_k(x)$ *of the first kind satisfy the inequality*

$$I_n(x) < \frac{x^k}{2^k k!} \exp(x).$$

Clearly, for any nonzero $x \in \mathbb{C}$ and $k \ge 0$, we have

$$|I_k(x)| \le I_k(|x|) < \frac{|x|^k}{2^k k!} \exp(|x|).$$

This means that for a fixed $x \in \mathbb{C}$, as $k$ tends to infinity, $I_k(x)$ tends to 0. This implies that for sufficiently large values of $n$ (the size of the matrix), the Bessel functions used in $\Psi(\mathbf{T}_n)$, which then have very large indices, will be significantly small.

LEMMA 4.3 ([26]). *For* $x \in \mathbb{R}$ *with* $x > 0$ *and* $k > -\frac{1}{2}$, *the Bessel functions* $I_k(x)$ *of the first kind satisfy the inequality*

$$I_{k+1}(x) < I_k(x).$$

Using Lemma 4.3, for all $x \in \mathbb{C}$ and $m, k \in \mathbb{N}$, if $m > k \ge 0$, we have

$$|I_m(x)| < I_k(|x|).$$

Now, we can easily state a bound for the $ij$-th element of $\Psi(\mathbf{T}_n)$, as defined in (2.4),

$$|(\Psi(\mathbf{T}_n))_{ij}| \le \exp(\mathrm{Re}(b) + 2|a|) \left(\frac{|a|^{|i-j|}}{|i-j|!} + \frac{|a|^{|i+j|}}{|i+j|!}\right). \tag{4.1}$$

THEOREM 4.4. *For a symmetric matrix* $\mathbf{T}_n = \mathrm{tridiag}(a, b, a)$ *and* $i, j \in \{1, \dots, n\}$,

$$\left|\big(\exp(\mathbf{T}_n)\big)_{ij}\right| \le 2e^{\mathrm{Re}(b)} \left(\frac{|a|e}{n+1}\right)^{n+1} + e^{(\mathrm{Re}(b)+2|a|)} \left(\frac{|a|^{|i-j|}}{|i-j|!} + \frac{|a|^{|i+j|}}{|i+j|!}\right), \tag{4.2}$$

**ETNA**
**Kent State University and**
**Johann Radon Institute (RICAM)**



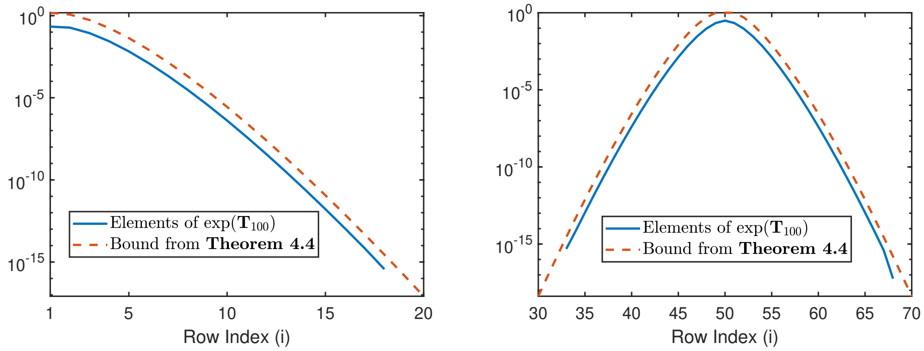



FIG. 4.1. *The elements* $(\exp(\mathbf{T}_n))_{ij}$ *and their bound in Theorem 4.4 for* $n = 100$*, with* $i = 1 : 20$ *and* $j = 1$ *(left) and* $i = 30 : 70$ *and* $j = 50$ *(right).*

*and for a nonsymmetric tridiagonal matrix* $\mathbf{A}_n = \mathrm{tridiag}(a, b, c)$ *and* $i, j \in \{1, \ldots, n\}$,

$$\begin{aligned} \left|\left(\exp(\mathbf{A}_n)\right)_{ij}\right| \le 2\left|\sqrt{\frac{a}{c}}\right|^{i-j} e^{\mathrm{Re}(b)}\left(\frac{|z|e}{n+1}\right)^{n+1} \\ + e^{(\mathrm{Re}(b)+2|z|)}\left|\sqrt{\frac{a}{c}}\right|^{i-j}\left(\frac{|z|^{|i-j|}}{|i-j|!} + \frac{|z|^{|i+j|}}{|i+j|!}\right), \end{aligned} \tag{4.3}$$

*where* $z = c\sqrt{\frac{a}{c}}$.

*Proof.* Using Theorem 4.1 and the bound obtained in (4.1), the result is at hand. To establish the inequality (4.3), we use

$$|\left(\mathbf{F}_n \circ \exp(\mathbf{T}_n)\right)_{ij}| = \left|\sqrt{a/c}\right|^{i-j} |\left(\exp(\mathbf{T}_n)\right)_{ij}|,$$

where $\mathbf{T}_n = \mathrm{tridiag}(z, b, z)$; see Theorem 3.1 and Theorem 3.2. Using the inequality (4.2) for $\mathbf{T}_n$ completes the proof. ∎

To clarify our last results, we provide an example for the bound of the elements found in Theorem 4.4.

EXAMPLE 4.5. Let us consider the matrix from Example 2.4, $\mathbf{T}_n = \mathrm{tridiag}(1, -2, 1)$, with $n = 100$ as the size of $\mathbf{T}_n$. In Figure 4.1, we illustrate $(\exp(\mathbf{T}_n))_{ij}$ for some $i$ and $j$, and the bound from Theorem 4.4, inequality (4.2). On the left-hand side of Figure 4.1, we display the output for $i = 1 : 20$ and $j = 1$, and on the right-hand side, we display it for $i = 30 : 70$ and $j = 50$. The results show that the bound from Theorem 4.4 effectively estimates the exact values of the entries of the matrix $\exp(\mathbf{T}_n)$. It is worth mentioning that if instead of the bound in Theorem 4.4, the bound in Theorem 4.1 is used to estimate the elements of the exponential of tridiagonal matrices, then the accuracy of the estimation will be significantly higher; see Section 6.2.

**4.2. Banded approximation.** Theorem 4.4 not only provides a bound for the entries of the matrix exponential of a tridiagonal matrix but also clearly states that for sufficiently large matrices, the entries of the matrix $\exp(\mathbf{T}_n)$ become significantly small as we move away from the main diagonal. This decay depends on the part of the theorem's bound that involves $|i - j|$. Therefore, in practice, we can ignore computing entries for which the bounds from Theorem 4.4 are smaller than a given machine epsilon $\varepsilon > 0$. This means that we can approximate $\exp(\mathbf{T}_n)$ with a $d$-banded matrix for some $d > 0$, a matrix in which the elements

ETNA
Kent State University and
Johann Radon Institute (RICAM)



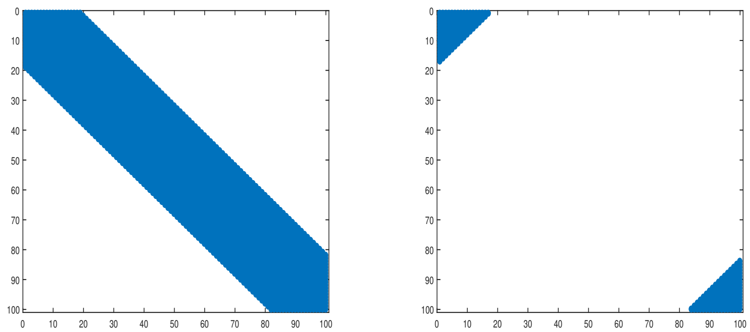


FIG. 4.2. *Sparsity pattern of the 100-by-100 Toeplitz (left) and Hankel (right) matrices with elements having magnitudes larger than $\varepsilon = 10^{-16}$, where these matrices are used in $\Psi(\mathbf{T}_n)$ for $\mathbf{T}_n = \operatorname{tridiag}(1,-2,1)$.*

corresponding to indices $(i,j)$ with $|i-j| > d$ are zero. This fact helps us make our approach more efficient.

For example, set $\varepsilon = 10^{-16}$ and $\mathbf{T}_n = \operatorname{tridiag}(1,-2,1)$, with $n = 100$. The entries of the corresponding Toeplitz and Hankel matrices in $\Psi(\mathbf{T}_n)$ with magnitudes larger than $\varepsilon$ are illustrated in Figure 4.2. This figure shows that we can approximate $\exp(\mathbf{T}_n)$ with a 19-banded matrix. Here, we approximate the matrix $\exp(\mathbf{T}_n)$ with a $d$-banded matrix, which we denote as $\Psi^{[d]}(\mathbf{T}_n)$.

For a given $d > 0$, we define $\Psi^{[d]}(\mathbf{T}_n)$, where $\mathbf{T}_n = \operatorname{tridiag}(a,b,a)$. Using only modified Bessel functions of the first kind of indices 0 to $d+2$, i.e., the modified Bessel functions $I_0(2a), I_1(2a), \ldots, I_d(2a), I_{d+1}(2a), I_{d+2}(2a)$, we define vectors of length $n$,

$$\begin{aligned}\mathbf{u}^{[d]} &= (e^b I_0(2a), e^b I_1(2a), \ldots, e^b I_{d-1}(2a), e^b I_d(2a), 0, \ldots, 0),\\ \mathbf{v}^{[d]} &= (e^b I_2(2a), \ldots, e^b I_{d+1}(2a), e^b I_{d+2}(2a), 0, \ldots, 0),\end{aligned}$$

such that $\Psi^{[d]}(\mathbf{T}_n)$ is given by the difference of a Toeplitz and a Hankel matrix,

$$\Psi^{[d]}(\mathbf{T}_n) = \operatorname{Toeplitz}\left\{\mathbf{u}^{[d]}\right\} - \operatorname{Hankel}\left\{\mathbf{v}^{[d]}\right\}. \tag{4.4}$$

Here, $\Psi^{[d]}(\mathbf{T}_n)$ is a $d$-banded matrix approximation of $\exp(\mathbf{T}_n)$. We approximate $\exp(\mathbf{A}_n)$, where $\mathbf{A}_n = \operatorname{tridiag}(a,b,c)$, using Theorem 3.1. We define $\mathbf{F}_n^{[d]} = \operatorname{Toeplitz}\left\{\mathbf{x}^{[d]}, \mathbf{y}^{[d]}\right\}$, where

$$\mathbf{x}^{[d]} = (a_0, a_1, \ldots, a_d, \underbrace{0, \ldots, 0}_{n-d-1 \text{ times}}), \quad \mathbf{y}^{[d]} = (a_0, a_{-1}, \ldots, a_{-d}, \underbrace{0, \ldots, 0}_{n-d-1 \text{ times}}),$$

with $a_k = \left(\sqrt{\frac{a}{c}}\right)^k$. We approximate the exponential of nonsymmetric tridiagonal matrices as

$$\exp(\mathbf{A}_n) \approx \mathbf{F}_n^{[d]} \circ \Psi^{[d]}(\mathbf{T}_n), \tag{4.5}$$

where the appropriate matrix $\mathbf{T}_n$ is defined in Theorem 3.2.

Note that in $\Psi^{[d]}(\mathbf{T}_n)$, we only need to compute $d+3$ modified Bessel functions at $2a$ using the trapezoidal rule with $m$ quadrature points. Therefore, the complexity of the method is $\mathcal{O}(\max\{m,d\}\cdot d)$. To obtain the error bound for this approximation, we estimate it by

$$\begin{aligned}\|\exp(\mathbf{T}_n) - \Psi^{[d]}(\mathbf{T}_n)\|_\infty &\le \|\exp(\mathbf{T}_n) - \Psi(\mathbf{T}_n)\|_\infty + \|\Psi(\mathbf{T}_n) - \Psi^{[d]}(\mathbf{T}_n)\|_\infty\\ &\le 2e^{\operatorname{Re}(b)}\left(\frac{|a|e}{n+1}\right)^{n+1} + \frac{4|a|^{d+1}}{(d+1)!}\exp(\operatorname{Re}(b) + 3|a|).\end{aligned} \tag{4.6}$$

The bound (4.6) is easily obtained by utilizing Corollary 2.3, Lemma 4.2, and Lemma 4.3. In the nonsymmetric case $\mathbf{A}_n = \mathrm{tridiag}(a, b, c)$, only $\Delta$ as used in Theorem 3.2 is multiplied in the above bound, and $a$ is replaced by $z = c\sqrt{\frac{a}{c}}$. In Section 6.3 we examine the effectiveness of the bound in (4.6) and test it for different matrices.

**4.3. Fast matrix-vector product.** The matrix-vector product with a Toeplitz matrix can be computed using the Fast Fourier Transform (FFT) in $\mathcal{O}(n\log(n))$ operations [3]. We note that $\Psi(\mathbf{T}_n)$ is a Toeplitz-plus-Hankel matrix. Therefore, for a vector $\mathbf{v}$, the computation of $\Psi(\mathbf{T}_n)\mathbf{v}$ can be performed efficiently. The Toeplitz part can be handled using the FFT in $\mathcal{O}(n\log(n))$ operations. Similarly, the Hankel part can also be treated efficiently by decomposing it into the product of a Toeplitz matrix and a backward identity matrix. Using this decomposition, the multiplication of a vector by a Hankel matrix can also be carried out in $\mathcal{O}(n\log(n))$ operations. Thus, computing $\Psi(\mathbf{T}_n)\mathbf{v}$ requires only $\mathcal{O}(2n\log(n))$ operations.

**5. Approximation of the exponential of a tridiagonal matrix and the heat equation.** In this section, using the studies conducted in the previous sections on tridiagonal matrices, we present an approximation of the solution of the one-dimensional and two-dimensional heat equation that leads to a stable method.

**5.1. One-dimensional heat equation.** Let us consider the one-dimensional heat equation,

$$u_t = au_{xx}, \qquad x \in [b, c], \qquad t > 0, \qquad a > 0, \tag{5.1}$$

with initial condition $u(x, 0) = u_0(x)$, for $x \in [b, c]$, and homogeneous Dirichlet boundary conditions $u(b, t) = u(c, t) = 0$, for $t > 0$. Consider points $x_j$ in $[b, c]$ with $x_j = b + j\Delta x$, for $j = 1, \dots, n$, where $\Delta x = \frac{c-b}{n+1}$ and $n \in \mathbb{N}$. By applying the central difference discretization of the second-order spatial derivative for $u_{xx}$ in (5.1), the equation can be written as

$$\frac{\mathrm{dU}_j}{\mathrm{dt}} = a\frac{\mathrm{U}_{j+1} - 2\mathrm{U}_j + \mathrm{U}_{j-1}}{(\Delta x)^2}, \quad j = 1, \dots, n,$$

where

$$\mathrm{U}_j(t) \approx u(x_j, t)\,, \quad j = 1, \dots, n. \tag{5.2}$$

The heat equation (5.1) is transformed into a system of ordinary differential equations,

$$\frac{\mathrm{d}\mathbf{U}}{\mathrm{dt}} = \mathbf{K}_n\mathbf{U}, \tag{5.3}$$

where $\mathbf{U}(t) = (\mathrm{U}_1(t), \dots, \mathrm{U}_n(t))^T$ with $\mathrm{U}_j(t)$ in (5.2), and

$$\mathbf{K}_n = \mathrm{tridiag}\left(\frac{a}{(\Delta x)^2}, \frac{-2a}{(\Delta x)^2}, \frac{a}{(\Delta x)^2}\right) \in \mathbb{R}^{n\times n}.$$

The solution of the system in (5.3) is

$$\mathbf{U} = \exp(\mathbf{K}_n\mathrm{t})\mathbf{U}_0,$$

where $\mathbf{U}_0 = (u_0(x_1), \dots, u_0(x_n))^T$. Moreover, by choosing a time step size $\Delta t$ and defining $\mathbf{U}^{(k)} := \mathbf{U}(k\Delta t)$, the method can be written as

$$\mathbf{U}^{(k)} = \exp(\mathbf{T}_n)\mathbf{U}^{(k-1)}, \tag{5.4}$$

**ETNA**
**Kent State University and**
**Johann Radon Institute (RICAM)**



where $\mathbf{T}_n = \text{tridiag}(\mu, -2\mu, \mu)$ and $\mu = \frac{a\Delta t}{(\Delta x)^2}$. As discussed in the previous sections, we can approximate $\exp(\mathbf{T}_n)$ by $\Psi(\mathbf{T}_n)$ and use the FFT to compute $\Psi(\mathbf{T}_n)\mathbf{U}^{(k-1)}$ at each step with $\mathcal{O}(n\log(n))$ complexity as detailed in Section 4.3.

To reduce computational complexity, we can use the approximation $\Psi^{[d]}(\mathbf{T}_n)$ from Section 4 as in (4.4). The method corresponding to (5.4) for a suitable value of $d > 0$ is

$$\mathbf{U}^{(k,d)} = \Psi^{[d]}(\mathbf{T}_n)\mathbf{U}^{(k-1,d)}, \tag{5.5}$$

where $\Psi^{[d]}(\mathbf{T}_n)$ is used instead of $\exp(\mathbf{T}_n)$. The method presented in (5.5) provides an approximation of the solution of the heat equation in (5.1) through its output, the vector $\mathbf{U}^{(k,d)}$. One advantage of the method (5.5) is that it unconditionally preserves the maximum principle, which we prove in the following.

Let $\|\cdot\|_\infty$ denote the maximum absolute value of elements in a vector. The following theorem establishes the stability of the numerical method (5.5).

THEOREM 5.1. *Let $\mathbf{U}^{(k,d)}$ be the vector obtained from method (5.5) at the $k$-th step, and let $\mathbf{U}_0$ be the initial vector. They satisfy the inequality*

$$\|\mathbf{U}^{(k,d)}\|_\infty < \|\mathbf{U}_0\|_\infty.$$

*In other words, the method in (5.5) preserves the maximum principle unconditionally.*

*Proof.* First, from (5.5), we have

$$\mathbf{U}^{(k,d)} = \left(\Psi^{[d]}(\mathbf{T}_n)\right)^k \mathbf{U}_0.$$

Since $\|\Psi^{[d]}(\mathbf{T}_n)\|_\infty \leq \|\Psi(\mathbf{T}_n)\|_\infty$, we proceed to find a bound for $\|\Psi(\mathbf{T}_n)\|_\infty$. Since the Bessel functions satisfy $I_m(\cdot) = I_{-m}(\cdot)$ for integer values of $m$ and since we have that $(\Psi(\mathbf{T}_n))_{ij} = \exp(-2\mu)\left(I_{i-j}(2\mu) - I_{i+j}(2\mu)\right)$, using the fact that the Bessel functions are decreasing with respect to their index for positive arguments, i.e., $I_{m+1}(2\mu) \leq I_m(2\mu)$, we can conclude that $(\Psi(\mathbf{T}_n))_{ij} < \exp(-2\mu)I_{|i-j|}(2\mu)$. In this case, since for each row of $\Psi(\mathbf{T}_n)$ the term $\exp(-2\mu)I_{|i-j|}(2\mu)$ includes exactly one instance where the Bessel function has index 0 and at most two instances with indices ranging from 1 to $n-1$, we obtain that

$$\|\Psi(\mathbf{T}_n)\|_\infty \leq \exp(-2\mu)\left(I_0(2\mu) + 2\sum_{p=1}^{n-1} I_p(2\mu)\right).$$

Since the modified Bessel functions of the first kind satisfy (see [2])

$$\exp(2\mu) = I_0(2\mu) + 2\sum_{p=1}^{\infty} I_p(2\mu),$$

we get $\|\Psi(\mathbf{T}_n)\|_\infty < 1$. $\square$

**5.2. Two-dimensional heat equation.** Consider the two-dimensional heat equation

$$u_t = a\nabla^2 u, \quad \mathbf{x} = (x,y) \in \Omega \subset \mathbb{R}^2,\ t > 0,\ a > 0, \tag{5.6}$$

with boundary and initial conditions

$$\begin{aligned} u(\mathbf{x},t) &= 0, && \mathbf{x} \in \partial\Omega,\ t > 0, \\ u(\mathbf{x},0) &= u_0(\mathbf{x}), && \mathbf{x} \in \Omega. \end{aligned}$$

**ETNA**
Kent State University and
Johann Radon Institute (RICAM)



Consider a uniform rectangular grid of points $(x_i, y_j)$ in the region $\Omega = (b, c) \times (d, e)$, where $x_i = b+i\Delta x$, for $i = 1, \ldots, n_x$, and $y_j = d+j\Delta y$, for $j = 1, \ldots, n_y$. Here, the discretization parameters in the $x$- and $y$-directions are given by $\Delta x = \frac{c-b}{n_x+1}$ and $\Delta y = \frac{e-d}{n_y+1}$, with $n_x, n_y \in \mathbb{N}$. Spatial discretization of (5.6) using central second-order formulas is performed for $i = 1, \ldots, n_x$ and $j = 1, \ldots, n_y$,

$$\frac{\mathrm{dU}_{i,j}}{\mathrm{dt}} = a\frac{\mathrm{U}_{i+1,j} - 2\mathrm{U}_{i,j} + \mathrm{U}_{i-1,j}}{(\Delta x)^2} + a\frac{\mathrm{U}_{i,j+1} - 2\mathrm{U}_{i,j} + \mathrm{U}_{i,j-1}}{(\Delta y)^2}.$$

where the approximate solution is denoted by $\mathrm{U}_{i,j}(t)$,

$$\mathrm{U}_{i,j}(t) \approx u(x_i, y_j, t)\,, \quad i = 1, \ldots, n_x, \quad j = 1, \ldots, n_y,$$

i.e., time remains a continuous variable. We now transform the heat equation into the form (see [28])

$$\frac{\mathrm{d}\mathbf{U}}{\mathrm{dt}} = \mathbf{A}_n\mathbf{U}, \tag{5.7}$$

where $n = n_x n_y$ and $\mathbf{A}_n = \mathrm{tridiag}(\mathbf{D}_{n_x}, \mathbf{S}_{n_x}, \mathbf{D}_{n_x}) \in \mathbb{R}^{n\times n}$ is block tridiagonal. The matrix $\mathbf{D}_{n_x} \in \mathbb{R}^{n_x\times n_x}$ is diagonal, with diagonal entries $\frac{a}{(\Delta y)^2}$, while $\mathbf{S}_{n_x} \in \mathbb{R}^{n_x\times n_x}$ is the following symmetric tridiagonal matrix:

$$\mathbf{S}_{n_x} = \mathrm{tridiag}\left(\frac{a}{(\Delta x)^2}, -\frac{2a}{(\Delta x)^2} - \frac{2a}{(\Delta y)^2}, \frac{a}{(\Delta x)^2}\right).$$

Defining the tridiagonal matrix $\mathbf{H}_{n_y} = \mathrm{tridiag}(1, 0, 1) \in \mathbb{R}^{n_y\times n_y}$, $\mathbf{A}_n$ can be written as

$$\mathbf{A}_n = \frac{a}{(\Delta y)^2}\mathbf{H}_{n_y} \otimes \mathbf{I}_{n_x} + \mathbf{I}_{n_y} \otimes \mathbf{S}_{n_x},$$

where $\mathbf{I}_{n_i}$ is the identity matrix and $\otimes$ is the Kronecker product. By definition of the Kronecker sum, we have

$$\mathbf{A}_n = \frac{a}{(\Delta y)^2}\mathbf{H}_{n_y} \oplus \mathbf{S}_{n_x}.$$

Therefore (see [8]),

$$\exp\left(\mathbf{A}_n\mathrm{t}\right) = \exp\left(\frac{a}{(\Delta y)^2}\mathbf{H}_{n_y}\mathrm{t}\right) \otimes \exp\left(\mathbf{S}_{n_x}\mathrm{t}\right).$$

Thus, the solution of the system (5.7) is

$$\mathbf{U} = \left(\exp\left(\frac{a}{(\Delta y)^2}\mathbf{H}_{n_y}\mathrm{t}\right) \otimes \exp\left(\mathbf{S}_{n_x}\mathrm{t}\right)\right)\mathbf{U}_0.$$

Let $\mu_x = \frac{a\Delta t}{(\Delta x)^2}$, $\mu_y = \frac{a\Delta t}{(\Delta y)^2}$, $\mathbf{T}_{n_x} = \mathrm{tridiag}(\mu_x, -2\mu_x-2\mu_y, \mu_x)$, $\mathbf{K}_{n_y} = \mathrm{tridiag}(\mu_y, 0, \mu_y)$, and by choosing a time step size $\Delta t$, the numerical method is given by

$$\mathbf{U}^{(k+1)} = \left(\exp(\mathbf{K}_{n_y}) \otimes \exp(\mathbf{T}_{n_x})\right)\mathbf{U}^{(k)},$$

where $\mathbf{U}^{(k)} := \mathbf{U}(k\Delta t)$. More precisely, a splitting based on the Kronecker product is implemented for solving the two-dimensional heat equation. Using the approximation method

**ETNA**
Kent State University and
Johann Radon Institute (RICAM)



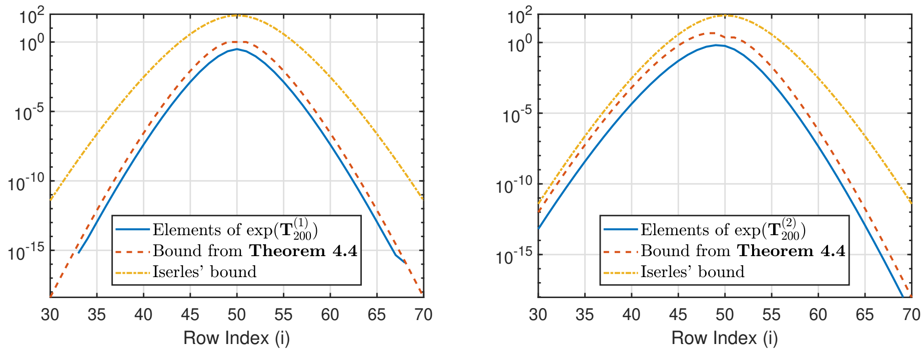



FIG. 6.1. *The elements* $|(\exp(\mathbf{T}_n))_{30:70,50}|$ *and their Iserles' bounds, along with bounds from Theorem 4.4 for* $\mathbf{T}_{200}^{(1)} = \mathrm{tridiag}(1, -2, 1)$ *(left) and for* $\mathbf{T}_{200}^{(2)} = \mathrm{tridiag}(1, -2, 2)$ *(right).*

presented in Section 2, for suitable values of $d_x$ and $d_y$, the following numerical method can be considered:

$$\mathbf{U}^{(k+1,d_x,d_y)} = \left(\Psi^{[d_y]}(\mathbf{K}_{n_y}) \otimes \Psi^{[d_x]}(\mathbf{T}_{n_x})\right) \mathbf{U}^{(k,d_x,d_y)}. \tag{5.8}$$

Similar to Theorem 5.1, we can prove that

$$\left\| \Psi^{[d_y]}(\mathbf{K}_{n_y}) \otimes \Psi^{[d_x]}(\mathbf{T}_{n_x}) \right\|_\infty < 1,$$

which shows the stability of the method in (5.8) and that the method unconditionally preserves the maximum principle.

**6. Numerical examples.** In this section, we present various numerical experiments to investigate the quality of our methods, approximations, and error bounds. All experiments were conducted in MATLAB R2021a. We use the command "`besseli`" to compute the values of the modified Bessel functions of the first kind. We compute $\exp(A)$ with entries $(\exp(A))_{ij}$ using the MATLAB command "`expm`" to obtain the error of our approximations to machine precision.

**6.1. Iserles' bound for the entries of the exponential of tridiagonal matrices.** In [19, Theorem 2.2], Iserles established a bound for the entries of the exponential of tridiagonal matrices. To state this bound, consider a tridiagonal matrix $A$, and define $\rho := \max_{i,j} |(A)_{ij}|$. By using modified Bessel functions of the first kind, $I_k(\cdot)$, he proved the following bound:

$$|(\exp(A))_{ij}| \le \exp(\rho) I_{|i-j|}(2\rho). \tag{Iserles' bound}$$

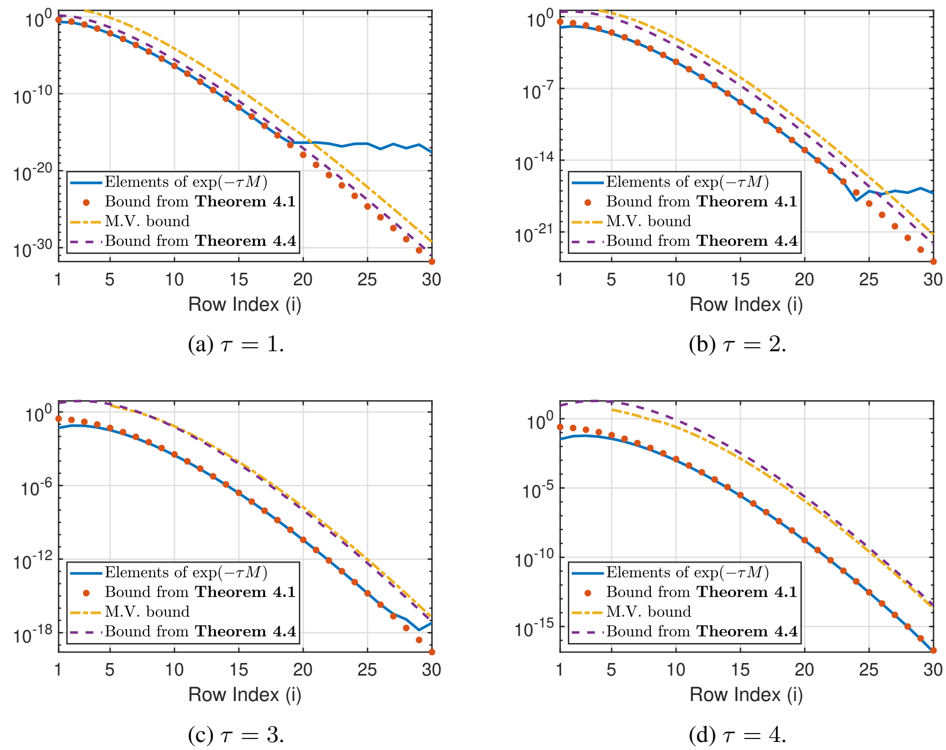
In the first example we compare the bound in Theorem 4.4 with Iserles' bound. We consider the two matrices

$$\mathbf{T}_{200}^{(1)} = \mathrm{tridiag}(1, -2, 1) \quad \text{and} \quad \mathbf{T}_{200}^{(2)} = \mathrm{tridiag}(1, -2, 2).$$

In Figure 6.1 we display the elements $(\exp(\mathbf{T}_{200}^{(r)}))_{30:70,50}$ and their bounds from Theorem 4.4, along with Iserles' bound. The result tells us that the bound from Theorem 4.4 works better in estimating the elements. One of the reasons why the bound from the theorem is more effective may be that, in this bound, the sign of the elements of the main diagonal and the elements off the main diagonal, even if they have smaller absolute values, affect the bound, while in Iserles' bound only the maximum absolute value of the elements of the matrix is considered.

ETNA
Kent State University and
Johann Radon Institute (RICAM)






(a) $\tau = 1$.
(b) $\tau = 2$.
(c) $\tau = 3$.
(d) $\tau = 4$.

FIG. 6.2. *The elements* $|(\exp(-\tau M))_{1:30,1}|$ *and their bounds from Theorem 4.4 and Theorem 4.1, along with the M.V. bound for* $\tau \in \{1, 2, 3, 4\}$*, for* $M = M_1 - \lambda_{\min}(M_1)I$*, where* $M_1 = \mathrm{tridiag}(-1, 4, -1) \in \mathbb{R}^{200\times 200}$.

**6.2. Bound for the entries of the exponential of a banded Hermitian matrix.** In [8, Section 4], the authors provided a bound for the entries of $\exp(-\tau M)$, where $M$ is a Hermitian positive semi-definite matrix. To recall the result, let $M$ have eigenvalues in the interval $[0, 4\rho]$, and for indices $i \neq j$, let $\xi = \lceil |i - j|/\beta \rceil$, where $\beta$ is the half bandwidth of $M$. The bound established in [8], called the “M.V. bound”, is given by:

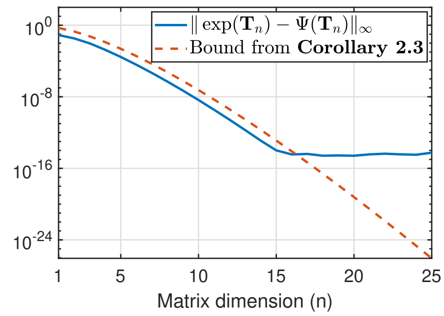
i) For $\rho\tau \geq 1$ and $\sqrt{4\rho\tau} \leq \xi \leq 2\rho\tau$,

$$|(\exp(-\tau M))_{ij}| \leq 10 \exp\left(-\frac{\xi^2}{5\rho\tau}\right).$$

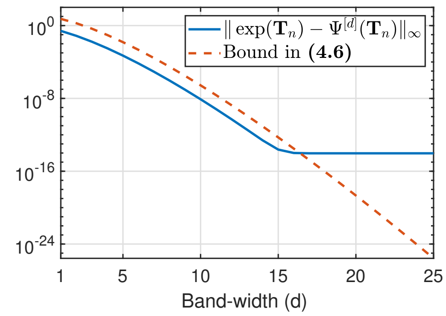
ii) For $\xi \geq 2\rho\tau$,

$$|(\exp(-\tau M))_{ij}| \leq 10 \frac{\exp(-\rho\tau)}{\rho\tau} \left(\frac{e\rho\tau}{\xi}\right)^{\xi}.$$

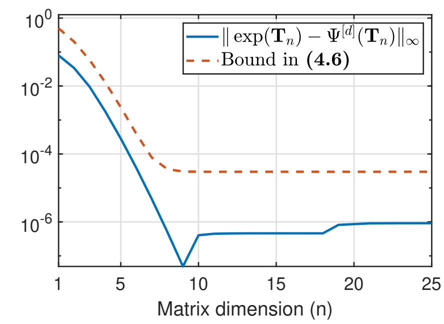
In Figure 6.2 we illustrate the entries $|(\exp(-\tau M))_{1:30,1}|$, where $M \in \mathbb{R}^{200\times 200}$, and their bounds from Theorem 4.4, from Theorem 4.1, and the M.V. bound for $\tau \in \{1, 2, 3, 4\}$. In this example, the matrix $M$ is as follows: first, we take $M_1 = \mathrm{tridiag}(-1, 4, -1)$ and then $M = M_1 - \lambda_{\min}(M_1)I$, where $\lambda_{\min}(M_1)$ is the smallest eigenvalue of $M_1$. Finally, we set $\rho = (\lambda_{\max}(M_1) - \lambda_{\min}(M_1))/4$ (as done in [8, Example 4.3]), and since $M$ is a tridiagonal matrix, we set $\beta = 1$. The result tells us that the bound from Theorem 4.1 estimates the entries with higher accuracy than the other bounds, and the bound from Theorem 4.4 performs



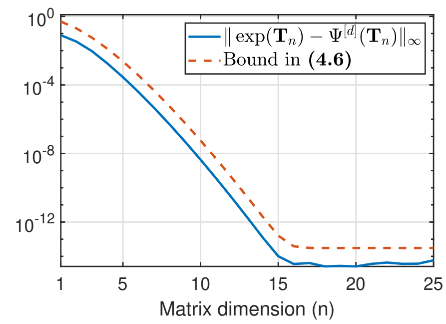

(a) $n = 1 : 25$.




(b) $n = 200$ and $d = 1 : 25$.




(c) $n = 1 : 25$ and $d = 8$.




(d) $n = 1 : 25$ and $d = 16$.

FIG. 6.3. *(a) Comparison of* $\|\exp(\mathbf{T}_n) - \Psi(\mathbf{T}_n)\|_\infty$ *with the bound from Corollary 2.3. The other figures provide a comparison of* $\|\exp(\mathbf{T}_n) - \Psi^{[d]}(\mathbf{T}_n)\|_\infty$ *with the bound in (4.6), where* $\mathbf{T}_n = \mathrm{tridiag}(1, -2, 1)$.

better than the M.V. bound in most cases. As we can see, for different original matrices, the high accuracy of the estimation from Theorem 4.1 is preserved, which can be considered an advantage of this bound.

**6.3. Error bound of the approximation of the exponential of a tridiagonal matrix.** In Section 2.1, Corollary 2.3, we provided a bound for $\|\exp(\mathbf{T}_n) - \Psi(\mathbf{T}_n)\|_\infty$, where $\Psi(\mathbf{T}_n)$ is the approximation method presented in (2.4). We also introduced another approximation method in (4.4), where we approximate $\exp(\mathbf{T}_n)$ by the $d$-banded matrix $\Psi^{[d]}(\mathbf{T}_n)$, with the corresponding error bound given in (4.6). In this experiment, we test these approximation methods and the established error bounds. In this example, we consider $\mathbf{T}_n = \mathrm{tridiag}(1, -2, 1) \in \mathbb{R}^{n\times n}$, which arises in the numerical solution of the heat equation.

In Figure 6.3a, we compare $\|\exp(\mathbf{T}_n) - \Psi(\mathbf{T}_n)\|_\infty$ with the bound from Corollary 2.3 for $n = 1 : 25$, while in the other subfigures, we compare $\|\exp(\mathbf{T}_n) - \Psi^{[d]}(\mathbf{T}_n)\|_\infty$ with the bound in (4.6) for different values of the matrix dimension $n$ and bandwidth $d$. As observed, both bounds are not far from the error arising from the approximation methods.

**6.4. Banded approximation method.** In this example, we examine the numerical test of the banded approximation method presented in Section 4.2 and the original method introduced in (2.4) for the symmetric case and in Theorem 3.2 for the nonsymmetric case. Here, we consider tridiagonal matrices with different dimensions, where their entries are chosen randomly in MATLAB using `randn(1)`.

**ETNA**
**Kent State University and**
**Johann Radon Institute (RICAM)**



TABLE 6.1

*Computation time (seconds) and error in MATLAB for the "Original methods" presented in Theorem 3.2 for nonsymmetric matrices and the method in (2.4) for symmetric matrices, along with "Banded Approx.", which corresponds to the method presented in (4.5) for nonsymmetric matrices and the method in (4.4) for the symmetric case, where for both d* $= 25$ *and the absolute error of the methods is measured in the infinity-norm with* `expm`*, for tridiagonal matrices with randomly chosen entries.*

| Type of matrix | Matrix size | Original Method Time | Original Method Error | Banded Approx. Time | Banded Approx. Error | `expm` Time |
|---|---|---|---|---|---|---|
| Symmetric | 500 | 0.01 | $2.16 \cdot 10^{-14}$ | 0.01 | $2.16 \cdot 10^{-14}$ | 0.02 |
| | 1,000 | 0.03 | $1.68 \cdot 10^{-13}$ | 0.03 | $1.68 \cdot 10^{-13}$ | 0.04 |
| | 3,000 | 0.21 | $2.74 \cdot 10^{-13}$ | 0.21 | $2.74 \cdot 10^{-13}$ | 0.7 |
| | 5,000 | 0.55 | $1.07 \cdot 10^{-11}$ | 0.55 | $1.07 \cdot 10^{-11}$ | 2.3 |
| | 7,000 | 1.21 | $1.75 \cdot 10^{-12}$ | 1.31 | $1.75 \cdot 10^{-12}$ | 5.74 |
| | 9,000 | 2.02 | $1.96 \cdot 10^{-12}$ | 2.23 | $1.96 \cdot 10^{-12}$ | 11.51 |
| | 11,000 | 3.44 | $1.94 \cdot 10^{-12}$ | 3.44 | $1.94 \cdot 10^{-12}$ | 21.69 |
| Nonsymmetric | 500 | 0.01 | $2.33 \cdot 10^{-13}$ | 0.01 | $2.34 \cdot 10^{-13}$ | 0.07 |
| | 1,000 | 0.03 | `NaN` | 0.03 | $8.22 \cdot 10^{-16}$ | 0.13 |
| | 3,000 | 0.21 | `NaN` | 0.21 | $6.35 \cdot 10^{-16}$ | 2.16 |
| | 5,000 | 0.88 | $4.78 \cdot 10^{-14}$ | 0.92 | $4.78 \cdot 10^{-14}$ | 10.36 |
| | 7,000 | 1.91 | $3.31 \cdot 10^{-16}$ | 1.87 | $3.32 \cdot 10^{-16}$ | 18.09 |
| | 9,000 | 3.02 | `NaN` | 3.05 | $5.22 \cdot 10^{-16}$ | 45.63 |
| | 11,000 | 4.51 | `NaN` | 4.59 | $6.07 \cdot 10^{-16}$ | 92.96 |

In Table 6.1 we present the results obtained by MATLAB, including the computation time and the error of the original methods, as well as the computation time and the absolute error of the banded approximation with bandwidth $d = 25$, for symmetric and nonsymmetric tridiagonal matrices. The `expm` function in MATLAB has been used to calculate the absolute error of each method with the infinity-norm. As can be seen, for symmetric matrices of all matrix sizes, both methods provide a good approximation with very small error. However, in the case of the nonsymmetric matrix, due to the structure of the matrix $\mathbf{F}_n$ in Theorem 3.2, some of its entries become `Inf` in MATLAB for certain values of $a$ and $c$ (i.e., values larger than `1e309`). Similarly, in $\Psi(\mathbf{T}_n)$, some entries become extremely small (i.e., smaller than `1e` $-$ `324`), which, when used in the corresponding positions of the Hadamard product, result in `NaN`. Consequently, the error in the table is displayed as `NaN`. This problem is solved in the banded approximation method.

**6.5. Crank–Nicolson method.** Here, we compare the Crank–Nicolson method with the method presented in (5.5) for two different initial conditions to test the efficiency and stability of the method in this paper.

**Periodic initial condition.** Consider the one-dimensional heat equation on $(0, 1)$ with homogeneous boundary conditions, where $a = 1$, and the initial condition

$$u(x, 0) = \sin(\pi x). \tag{6.1}$$

The exact solution of this heat equation with these conditions is $u(x, t) = e^{-\pi^2 t} \sin(\pi x)$.

In Figure 6.4 the infinity-norms of the errors of the method presented in Section 5.1, the method in (5.5), and the Crank–Nicolson method are displayed. We illustrate the methods for $\Delta x = 0.05$ and $d = 8$, with $\Delta t = 0.04$ (left) and $\Delta t = 0.05$ (right) as time steps. As shown in the figure, the errors in both methods are approximately similar because both methods are of order 2 in space.

ETNA
Kent State University and
Johann Radon Institute (RICAM)



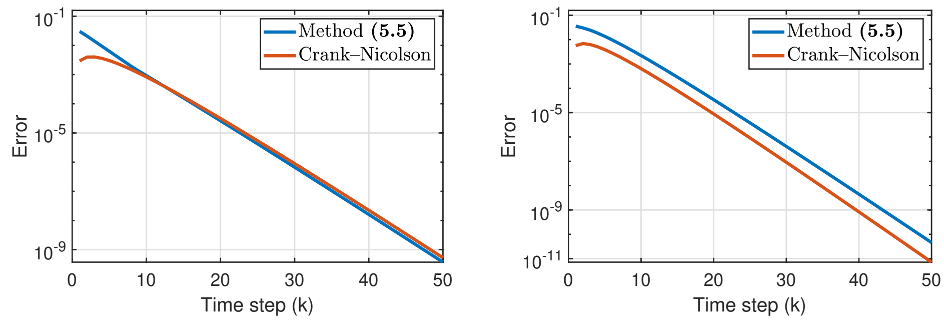



FIG. 6.4. *The absolute infinity-error of the method in* (5.5) *and the Crank–Nicolson method at* $t_k = k\Delta t$ *for* $\Delta t = 0.04$ *(left) and* $\Delta t = 0.05$ *(right), with* $\Delta x = 0.05$ *and* $d = 8$*, where the initial condition is given by* (6.1)*.*

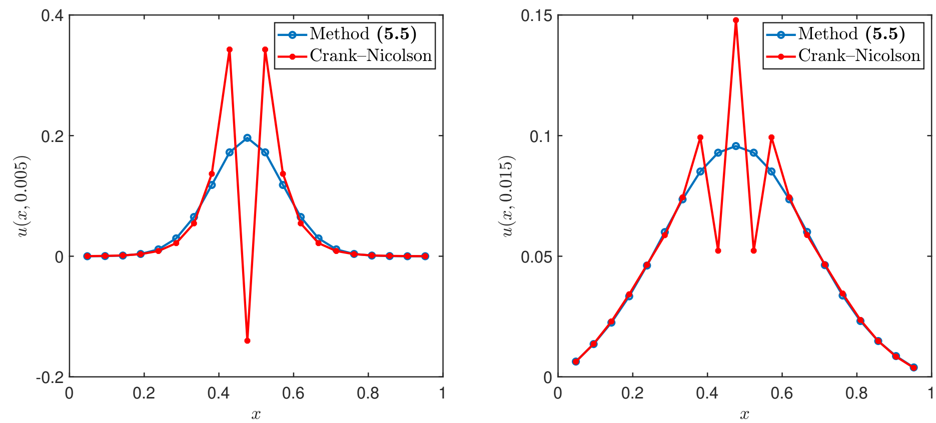



FIG. 6.5. *The numerical solutions of the one-dimensional heat equation with the initial condition* (6.2)*, using the Crank–Nicolson method and the method presented in* (5.5)*, at* $t = 0.005$ *(left) and* $t = 0.015$ *(right), with* $\mu = 2.205$*,* $\Delta x = 1/21$*, and* $d = 8$*.*

**Discontinuous initial condition.** We consider the initial condition

$$u(x,0) = \begin{cases} 1, & \text{if } x = 0.5, \\ 0, & \text{otherwise.} \end{cases} \tag{6.2}$$

In Figure 6.5 the results of the Crank–Nicolson method are compared with those of the method proposed in (5.5). From [25] we know that the Crank–Nicolson method suffers from oscillations when $\mu > 1$. Here, we set $\Delta x = 1/21$, $\Delta t = 0.005$, and $\mu = 2.205$ with $d = 8$. As seen in Figure 6.5, the Crank–Nicolson method exhibits oscillatory behavior. As stated in [25], when $\mu > 1$, the Crank–Nicolson method does not preserve the maximum principle, while, according to the stability property in Theorem 5.1, the proposed method provides non-oscillatory solutions.

**6.6. Gauss–Runge–Kutta method of order 4.** In this experiment, we compare the Gauss–Runge–Kutta method of order 4 (GRK4) with the method presented in (5.5). The

**ETNA**
Kent State University and
Johann Radon Institute (RICAM)



Butcher tableau for GRK4 is given by:

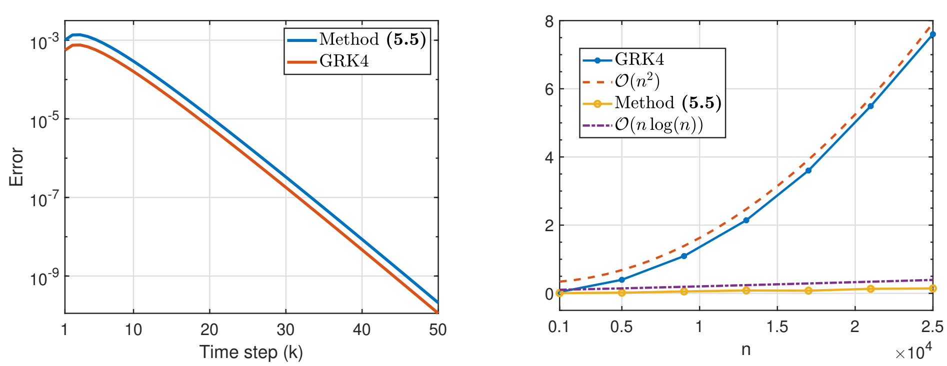
$$\begin{array}{c|cc} \frac{3-\sqrt{3}}{6} & \frac{1}{4} & \frac{3-2\sqrt{3}}{12} \\ \frac{3+\sqrt{3}}{6} & \frac{3+2\sqrt{3}}{12} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

In this example, we again consider the heat equation with the same conditions as in the previous example in Section 6.5, i.e., the initial condition $u(x,0) = \sin(\pi x)$ and the exact solution $u(x,t) = e^{-\pi^2 t}\sin(\pi x)$. By considering the GRK4 method for solving (5.3) with initial condition $\mathrm{U}_0 = (\sin(\pi\Delta x), \ldots, \sin(n\pi\Delta x))^T$, we have

$$\mathbf{U}_G^{(k+1)} = \mathbf{U}_G^{(k)} + \frac{\Delta t}{2} V_1^{(k)} + \frac{\Delta t}{2} V_2^{(k)},$$

where $V_1^{(k)}$ and $V_2^{(k)}$ are computed as

$$\begin{aligned} V_1^{(k)} &= \mathbf{K}_n \left( \mathbf{U}_G^{(k)} + \frac{\Delta t}{4} V_1^{(k)} + \Delta t \frac{3-2\sqrt{3}}{12} V_2^{(k)} \right), \\ V_2^{(k)} &= \mathbf{K}_n \left( \mathbf{U}_G^{(k)} + \Delta t \frac{3+2\sqrt{3}}{12} V_1^{(k)} + \frac{\Delta t}{4} V_2^{(k)} \right). \end{aligned} \tag{6.3}$$




(a) $n = 19$ and $k = 1:50$.

(b) $n = 1000:4000:25000$ and $k = 10$.

FIG. 6.6. *Comparison of the Gauss–Runge–Kutta method of order 4 (GRK4) with the method presented in* (5.5)*, where in both methods,* $\Delta x = \frac{1}{n+1}$ *and* $\Delta t = 0.04$.

In (6.3), we need to solve a linear system of size $(2n) \times (2n)$, where $V_1^{(k)}$ and $V_2^{(k)}$ are unknowns that must be determined at each step. Although here GRK4 provides slightly higher accuracy than the method (5.5) (see Figure 6.6a), both methods use a semi-discretization of the second-order spatial derivative and thus have order $\mathcal{O}((\Delta x)^2)$. The overall complexity of GRK4 is approximately $\mathcal{O}(kn^2)$, whereas the complexity of the method (5.5) is $\mathcal{O}(2kn\log(n))$ (see Figure 6.6b). Here, we set $d = n-1$ (i.e., a full matrix) in the method (5.5). Figure 6.6 illustrates the results for both methods with $\Delta x = 1/(n+1)$ and $\Delta t = 0.04$. In Figure 6.6a, we display the absolute errors for $n = 19$ and time steps $k = 1:50$. Figure 6.6b displays the computational time required for both methods at time $t_{10} = 0.4$ ($k = 10$) and dimensions $n = 1000:4000:25000$.

**7. Conclusion.** We analyzed an existing method for approximating the exponential of a symmetric tridiagonal matrix using Bessel functions of the first kind and presented a new and sharp error bound for this approximation. We extended this method to nonsymmetric tridiagonal matrices using the Hadamard product and derived an error bound for it. To reduce the time complexity and provide a dimension-independent approach, we established bounds for the entries of the matrix exponential of a tridiagonal matrix and used them to approximate the matrix exponential by a banded matrix. By using these results, we proposed a method for solving the heat equation, which unconditionally preserves the maximum principle. Finally, we presented numerical results for the methods introduced here and the existing ones, which demonstrated the efficiency of our approach and the superiority of the obtained bounds.

REFERENCES

[1] A. H. AL-MOHY AND N. J. HIGHAM, *A new scaling and squaring algorithm for the matrix exponential*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 970–989.

[2] G. B. ARFKEN AND H. J. WEBER, *Mathematical Methods for Physicists*, 6th ed., Elsevier Academic Press, San Diego, 2005.

[3] Z. BAI, J. DEMMEL, J. DONGARRA, A. RUHE, AND H. VAN DER VORST, EDS., *Templates for the Solution of Algebraic Eigenvalue Problems*, SIAM, Philadelphia, 2000.

[4] M. BENZI AND P. BOITO, *Decay properties for functions of matrices over $C^*$-algebras*, Linear Algebra Appl., 456 (2014), pp. 174–198.

[5] M. BENZI, P. BOITO, AND N. RAZOUK, *Decay properties of spectral projectors with applications to electronic structure*, SIAM Rev., 55 (2013), pp. 3–64.

[6] M. BENZI AND N. RAZOUK, *Decay bounds and $\mathcal{O}(n)$ algorithms for approximating functions of sparse matrices*, Electron. Trans. Numer. Anal., 28 (2007/08), pp. 16–39.
https://etna.ricam.oeaw.ac.at/vol.28.2007-2008/pp16-39.dir/pp16-39.pdf

[7] M. BENZI AND G. H. GOLUB, *Bounds for the entries of matrix functions with applications to preconditioning*, BIT, 39 (1999), pp. 417–438.

[8] M. BENZI AND V. SIMONCINI, *Decay bounds for functions of Hermitian matrices with banded or Kronecker structure*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 1263–1282.

[9] L. BERGAMASCHI, M. CALIARI, AND M. VIANELLO, *Efficient approximation of the exponential operator for discrete 2D advection-diffusion problems*, Numer. Linear Algebra Appl., 10 (2003), pp. 271–289.

[10] L. BERGAMASCHI AND M. VIANELLO, *Efficient computation of the exponential operator for large, sparse, symmetric matrices*, Numer. Linear Algebra Appl., 7 (2000), pp. 27–45.

[11] A. BÖTTCHER, J. GUTIÉRREZ-GUTIÉRREZ, AND P. M. CRESPO, *Mass concentration in quasicommutators of Toeplitz matrices*, J. Comput. Appl. Math., 205 (2007), pp. 129–148.

[12] P. M. CRESPO AND J. GUTIÉRREZ-GUTIÉRREZ, *On the elementwise convergence of continuous functions of Hermitian banded Toeplitz matrices*, IEEE Trans. Inform. Theory, 53 (2007), pp. 1168–1176.

[13] A. CORTINOVIS, D. KRESSNER, AND S. MASSEI, *Divide-and-conquer methods for functions of matrices with banded or hierarchical low-rank structure*, SIAM J. Matrix Anal. Appl., 43 (2022), pp. 151–177.

[14] J. GUTIÉRREZ-GUTIÉRREZ, P. M. CRESPO, AND A. BÖTTCHER, *Functions of the banded Hermitian block Toeplitz matrices in signal processing*, Linear Algebra Appl., 422 (2007), pp. 788–807.

[15] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.

[16] E. HAIRER, C. LUBICH, AND G. WANNER, *Geometric Numerical Integration*, 2nd ed., Springer, Berlin, 2006.

[17] N. J. HIGHAM, *Function of Matrices*, SIAM, Philadelphia, 2008.

[18] R. HOSSEINI AND M. TATARI, *A diagonal splitting method for solving semidiscretized parabolic partial differential equations*, Numer. Methods Partial Differential Equations, 36 (2020), pp. 268–283.

[19] A. ISERLES, *How large is the exponential of a banded matrix?*, New Zealand J. Math., 29 (2000), pp. 177–192.

[20] R. KRESS, *Numerical Analysis*, Springer, New York, 1998.

[21] D. KRESSNER AND R. LUCE, *Fast computation of the matrix exponential for a Toeplitz matrix*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 23–47.

[22] S. T. LEE, H.-K. PANG, AND H.-W. SUN, *Shift-invert Arnoldi approximation to the Toeplitz matrix exponential*, SIAM J. Sci. Comput., 32 (2010), pp. 774–792.

[23] Y. L. LUKE, *Inequalities for generalized hypergeometric functions*, J. Approximation Theory, 5 (1972), pp. 41–65.

[24] C. B. MOLER AND C. F. VAN LOAN, *Nineteen dubious ways to compute the exponential of a matrix*, SIAM Rev., 20 (1978), pp. 801–836.

Reprinted and updated as: *Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later*, SIAM Rev., 45 (2003), pp. 3–49.

[25] K. W. MORTON AND D. F. MAYERS, *Numerical Solution of Partial Differential Equations*, Cambridge University Press, Cambridge, 2005.

[26] I. NÅSELL, *Inequalities for modified Bessel functions*, Math. Comp., 28 (1974), pp. 253–256.

[27] S. NOSCHESE, L. PASQUINI, AND L. REICHEL, *Tridiagonal Toeplitz matrices: properties and novel applications*, Numer. Linear Algebra Appl., 20 (2013), pp. 302–326.

[28] A. QUARTERONI, F. SALERI, AND P. GERVASIO, *Scientific Computing with MATLAB and Octave*, Springer, Heidelberg, 2014.

[29] G. STRANG AND S. MACNAMARA, *Functions of difference matrices are Toeplitz plus Hankel*, SIAM Rev., 56 (2014), pp. 525–546.

[30] L. N. TREFETHEN AND J. A. C. WEIDEMAN, *The exponentially convergent trapezoidal rule*, SIAM Rev., 56 (2014), pp. 385–458.