Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2015.my1340

Vol. **8**, No. 4, pp. 515-529
November 2015

Superoptimal Preconditioners for Functions of Matrices

Zheng-Jian Bai[1,*], Xiao-Qing Jin[2], and Teng-Teng Yao[1]

[1] *School of Mathematical Sciences, Xiamen University, Xiamen 361005, P. R. China.*
[2] *Department of Mathematics, University of Macau, Macao, P. R. China.*

Received 18 December 2013; Accepted (in revised version) 26 October 2014

Abstract. For any given matrix $A \in \mathbb{C}^{n\times n}$, a preconditioner $t_U(A)$ called the superoptimal preconditioner was proposed in 1992 by Tyrtyshnikov. It has been shown that $t_U(A)$ is an efficient preconditioner for solving various structured systems, for instance, Toeplitz-like systems. In this paper, we construct the superoptimal preconditioners for different functions of matrices. Let f be a function of matrices from $\mathbb{C}^{n\times n}$ to $\mathbb{C}^{n\times n}$. For any $A \in \mathbb{C}^{n\times n}$, one may construct two superoptimal preconditioners for $f(A)$: $t_U(f(A))$ and $f(t_U(A))$. We establish basic properties of $t_U(f(A))$ and $f(t_U(A))$ for different functions of matrices. Some numerical tests demonstrate that the proposed preconditioners are very efficient for solving the system $f(A)\mathbf{x} = \mathbf{b}$.

AMS subject classifications: 65F10; 65F15; 65L05; 65N22

Key words: Superoptimal preconditioners, functions of matrices, Toeplitz matrix, PCG method.

1. Introduction

Let $U \in \mathbb{C}^{n\times n}$ be a unitary matrix. Define the set

$$\mathcal{M}_U \equiv \Big\{U^*\Lambda U \mid \Lambda \text{ is any } n\times n \text{ diagonal matrix}\Big\},$$

where U^* means the conjugate transpose of U. For any $A \in \mathbb{C}^{n\times n}$, the superoptimal preconditioner $t_U(A)$ is defined to be the minimizer of

$$\min \|I_n - W^{-1}A\|_F$$

over all nonsingular matrices $W \in \mathcal{M}_U$, where $\|\cdot\|_F$ denotes the Frobenius matrix norm.

*Corresponding author. *Email address:* zjbai@xmu.edu.cn (Z. J. Bai), xqjin@umac.mo (X. Q. Jin), yaotengteng718@163.com (T. T. Yao)

 ©2015 Global-Science Press

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

The superoptimal preconditioner was proposed by Tyrtyshnikov [20] in 1992. It was also studied by Chan *et al.* [5, 6]. The superoptimal preconditioner was used to solve some ill-conditioned problems appeared in image deblurring [2] and other structured systems [4, 7, 11, 13]. For more useful properties of $t_U(A)$, we refer to [9, 10, 14].

Recently, Jin *et al.* [16] study the optimal preconditioners for matrix functions. For any given $A \in \mathbb{C}^{n\times n}$, the optimal preconditioner $c_U(A)$ is defined in [8] to be the minimizer of

$$\min_{W\in\mathcal{M}_U} \|A - W\|_F.$$

In this paper, we propose the superoptimal preconditioners for some special functions of matrices. We first recall basic properties of the optimal and superoptimal preconditioners. Here, $\delta(A)$ stands for the diagonal matrix whose diagonal is equal to the diagonal of a matrix $A \in \mathbb{C}^{n\times n}$.

Lemma 1.1. ([10, 13–15]) *Let $A \in \mathbb{C}^{n\times n}$ and $\mathcal{W}(A) = \{\mathbf{x}^*A\mathbf{x} : \mathbf{x} \in \mathbb{C}^n \text{ and } \|\mathbf{x}\|_2 = 1\}$, where $\|\cdot\|_2$ denotes the 2-norm. Then*

(i) *$c_U(A) = U^*\delta(UAU^*)U$ and is uniquely determined by A.*

(ii) *If both A and $c_U(A)$ are nonsingular, i.e., $0 \notin \mathcal{W}(A)$, then the superoptimal preconditioner $t_U(A)$ exists and is given by*

$$t_U(A) = c_U(AA^*)[c_U(A^*)]^{-1} = U^*\delta(UAA^*U^*)\delta^{-1}(UA^*U^*)U,$$

where A^* means the conjugate transpose of A.

(iii) *If both $c_U(A)$ and $t_U(A)$ are nonsingular, then one has*

$$\sigma_k\big([t_U(A)]^{-1}A\big) \le \sigma_k\big([c_U(A)]^{-1}A\big),$$

for $k = 1, \ldots, n$, where the singular values are in increasing order, $\sigma_1 \le \cdots \le \sigma_n$.

(iv) *If A is Hermitian positive definite, then one has*

$$\lambda_k\big([t_U(A)]^{-1}A\big) \le \lambda_k\big([c_U(A)]^{-1}A\big),$$

for $k = 1, \ldots, n$, where the eigenvalues are in increasing order, $\lambda_1 \le \cdots \le \lambda_n$.

(v) *Let $[M]_{kk}$ denote the k-th diagonal entry of a matrix $M \in \mathbb{C}^{n\times n}$. Then*

$$[UAA^*U^*]_{kk} \ge [UAU^*]_{kk} \cdot [UA^*U^*]_{kk} \ge 0,$$

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

In the following sections, we propose two superoptimal preconditioners for some special matrix functions f from $\mathbb{C}^{n\times n}$ to $\mathbb{C}^{n\times n}$: $t_U(f(A))$ and $f(t_U(A))$, where $A \in \mathbb{C}^{n\times n}$ is a given matrix. We discuss properties of both preconditioners $t_U(f(A))$ and $f(t_U(A))$ for different functions of matrices. We also report some numerical experiments for solving the system $f(A)\mathbf{x} = \mathbf{b}$, where $\mathbf{b} \in \mathbb{C}^n$.

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

2. Superoptimal preconditioners for functions of matrices

We construct the superoptimal preconditioners for different functions of matrices: matrix exponential, matrix cosine, matrix sine, and matrix logarithm in this section.

2.1. Superoptimal preconditioners for matrix exponential

The matrix exponential function arises in matrix group theory [1], computational mathematics [19], and financial mathematics [17, 18], etc. We study basic properties of the preconditioners $t_U(e^A)$ and $e^{t_U(A)}$. Note that the matrix exponential for a matrix $A \in \mathbb{C}^{n\times n}$ is defined by

$$e^A \equiv \exp(A) \equiv I + A + \frac{A^2}{2!} + \cdots + \frac{A^m}{m!} + \cdots = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

By Lemma 1.1 (ii), we have

$$\begin{aligned}
t_U(e^A) &= U^*\delta(Ue^A(e^A)^*U^*)\delta^{-1}(U(e^A)^*U^*)U \\
&= U^*\delta(Ue^Ae^{A^*}U^*)\delta^{-1}(Ue^{A^*}U^*)U \\
&= U^*\delta\left(U\Big[\sum_{k=0}^{\infty}\frac{A^k}{k!}\Big]U^*U\Big[\sum_{k=0}^{\infty}\frac{(A^*)^k}{k!}\Big]U^*\right)\delta^{-1}\left(U\Big[\sum_{k=0}^{\infty}\frac{(A^*)^k}{k!}\Big]U^*\right)U \\
&= U^*\delta\left(\Big[\sum_{k=0}^{\infty}\frac{(UAU^*)^k}{k!}\Big]\Big[\sum_{k=0}^{\infty}\frac{(UA^*U^*)^k}{k!}\Big]\right)\delta^{-1}\left(\sum_{k=0}^{\infty}\frac{(UA^*U^*)^k}{k!}\right)U \\
&= U^*\delta(e^Be^{B^*})\delta^{-1}(e^{B^*})U = U^*\delta\big(e^B(e^B)^*\big)\delta^{-1}\big((e^B)^*\big)U,
\end{aligned}$$

where $B = UAU^*$. We point out that e^B may be computed much more easily than that of e^A. In particular, if B is a sparse matrix, then e^B could be computed easily. For example, let U be the $n \times n$ Fourier matrix F and B the $n \times n$ Toeplitz matrix given by

$$B = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \ddots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Then $A = F^*BF$ is a dense matrix compared with B. When n is large, the computation of e^B is much easier than the computation of e^A.

If A is normal, then there exists a unitary matrix W such that $A = W^*\Lambda W$ where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. In this case, $B = UAU^* = UW^*\Lambda WU^*$. Hence,

$$\begin{aligned}
t_U(e^A) &= U^*\delta(e^Be^{B^*})\delta^{-1}(e^{B^*})U = U^*\delta(e^{B+B^*})\delta^{-1}(e^{B^*})U, \\
&= U^*\delta\big(UW^*\mathrm{diag}(e^{\lambda_1+\bar{\lambda}_1}, e^{\lambda_2+\bar{\lambda}_2}, \ldots, e^{\lambda_n+\bar{\lambda}_n})WU^*\big)\times \\
&\qquad \delta^{-1}\big(UW^*\mathrm{diag}(e^{\bar{\lambda}_1}, e^{\bar{\lambda}_2}, \ldots, e^{\bar{\lambda}_n})WU^*\big)U,
\end{aligned} \tag{2.1}$$

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

where $\bar{a}$ means the complex conjugate of a complex number $a \in \mathbb{C}$. If A is Hermitian, then (2.1) turns to be

$$\begin{aligned} t_U(e^A) &= U^*\delta(e^B e^{B^*})\delta^{-1}(e^{B^*})U = U^*\delta(e^{2B})\delta^{-1}(e^B)U, \\ &= U^*\delta\big(UW^*\mathrm{diag}(e^{2\lambda_1}, e^{2\lambda_2}, \ldots, e^{2\lambda_n})WU^*\big)\times \\ &\qquad \delta^{-1}\big(UW^*\mathrm{diag}(e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_n})WU^*\big)U. \end{aligned} \tag{2.2}$$

We observe from (2.2) that instead of the series of matrices, we only need to compute the exponential functions. This can largely reduce the cost of constructing $t_U(e^A)$. Of course, when $A = W^*\Lambda W$ is diagonalizable, one can easily obtain $e^{-A} = W^*e^{-\Lambda}W$ and therefore no need to compute the preconditioned $t_U(e^A)$ from the computational point of view. Hence (2.2) is only for a theoretical analysis later (see Theorem 2.1).

Next, we discuss the preconditioner $e^{t_U(A)}$. Let

$$\delta(UAA^*U^*) = \mathrm{diag}(\gamma_1, \gamma_2, \ldots, \gamma_n) \quad \text{and} \quad \delta(UAU^*) = \mathrm{diag}(\eta_1, \eta_2, \ldots, \eta_n). \tag{2.3}$$

Then

$$\begin{aligned} e^{t_U(A)} &= \sum_{k=0}^{\infty} \frac{[t_U(A)]^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!}\Big[U^*\mathrm{diag}\Big(\frac{\gamma_1}{\bar{\eta}_1}, \frac{\gamma_2}{\bar{\eta}_2}, \ldots, \frac{\gamma_n}{\bar{\eta}_n}\Big)U\Big]^k \\ &= U^*\mathrm{diag}\Big(\sum_{k=0}^{\infty} \frac{1}{k!}\frac{\gamma_1^k}{\bar{\eta}_1^k}, \ldots, \frac{1}{k!}\frac{\gamma_n^k}{\bar{\eta}_n^k}\Big)U = U^*\mathrm{diag}\Big(e^{\frac{\gamma_1}{\bar{\eta}_1}}, e^{\frac{\gamma_2}{\bar{\eta}_2}}, \ldots, e^{\frac{\gamma_n}{\bar{\eta}_n}}\Big)U. \end{aligned} \tag{2.4}$$

Also, we can see from (2.4) that the cost of constructing $e^{t_U(A)}$ can be reduced greatly.

We now discuss the bounds for the norm of the superoptimal preconditioners $t_U(e^A)$ and $e^{t_U(A)}$. We first give a bound for $\|t_U(A)\|_2$.

Lemma 2.1. *Let $A \in \mathbb{C}^{n\times n}$. If both A and $c_U(A)$ are nonsingular, then we have*

$$\max_{1\le k\le n} |\eta_k| \le \|t_U(A)\|_2 = \max_{1\le k\le n} \frac{[UAA^*U^*]_{kk}}{\big|[UA^*U^*]_{kk}\big|} \le \frac{\max\limits_{1\le k\le n} \gamma_k}{\min\limits_{1\le k\le n} |\eta_k|} \le \frac{\sigma_{\max}^2(A)}{\min\limits_{1\le k\le n} |\eta_k|},$$

where $\sigma_{\max}(A)$ means the largest singular value of A, and γ_k and η_k $(k = 1, \ldots, n)$ are given by (2.3).

Proof. We have by Lemma 1.1 (ii),

$$\begin{aligned} \|t_U(A)\|_2 &= \|U^*\delta(UAA^*U^*)\delta^{-1}(UA^*U^*)U\|_2 \\ &= \|\delta(UAA^*U^*)\delta^{-1}(UA^*U^*)\|_2 \\ &= \max_{1\le k\le n} \frac{[UAA^*U^*]_{kk}}{\big|[UA^*U^*]_{kk}\big|}. \end{aligned} \tag{2.5}$$

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

Then from (2.5),

$$\|t_U(A)\|_2 \le \frac{\max\limits_{1\le k\le n}[UAA^*U^*]_{kk}}{\min\limits_{1\le k\le n}\left|[UA^*U^*]_{kk}\right|} \le \frac{\max\limits_{1\le k\le n}\gamma_k}{\min\limits_{1\le k\le n}|\eta_k|} \le \frac{\sigma_{\max}^2(A)}{\min\limits_{1\le k\le n}|\eta_k|}.$$

Moreover, by (2.5) again and Lemma 1.1 (v), we get

$$\begin{aligned}\|t_U(A)\|_2 &= \max_{1\le k\le n}\frac{[UAA^*U^*]_{kk}\cdot\left|[UAU^*]_{kk}\right|}{\left|[UAU^*]_{kk}\right|\cdot\left|[UA^*U^*]_{kk}\right|}\\ &\ge \max_{1\le k\le n}\left|[UAU^*]_{kk}\right| = \max_{1\le k\le n}|\eta_k|.\end{aligned}$$

□

On the bounds for $\|t_U(e^A)\|_2$ and $\|e^{t_U(A)}\|_2$, we have the following results.

Theorem 2.1. *Let $A\in\mathbb{C}^{n\times n}$ be Hermitian. If both A and $c_U(A)$ are nonsingular, then we have*

$$e^{\lambda_{\min}(A)} \le \|t_U(e^A)\|_2 \le e^{2\lambda_{\max}(A)-\lambda_{\min}(A)},$$

where $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ is the largest and smallest eigenvalues of A respectively.

Proof. Since $A\in\mathbb{C}^{n\times n}$ is Hermitian, there exists a unitary matrix $W\in\mathbb{C}^{n\times n}$ such that $A=W^*\Lambda W$, where $\Lambda=\operatorname{diag}(\lambda_1,\lambda_2,\dots,\lambda_n)$. Then we have by (2.2) and Lemma 2.1,

$$\begin{aligned}\|t_U(e^A)\|_2 &= \max_{1\le k\le n}\frac{[UW^*\operatorname{diag}(e^{2\lambda_1},e^{2\lambda_2},\dots,e^{2\lambda_n})WU^*]_{kk}}{\left|[UW^*\operatorname{diag}(e^{\lambda_1},e^{\lambda_2},\dots,e^{\lambda_n})WU^*]_{kk}\right|}\\ &\le \frac{\max\limits_{1\le k\le n}[UW^*\operatorname{diag}(e^{2\lambda_1},e^{2\lambda_2},\dots,e^{2\lambda_n})WU^*]_{kk}}{\min\limits_{1\le k\le n}\left|[UW^*\operatorname{diag}(e^{\lambda_1},e^{\lambda_2},\dots,e^{\lambda_n})WU^*]_{kk}\right|}\\ &\le \frac{e^{2\lambda_{\max}(A)}}{e^{\lambda_{\min}(A)}} = e^{2\lambda_{\max}(A)-\lambda_{\min}(A)}.\end{aligned}$$

Moreover, we get by (2.2) and Lemma 2.1,

$$\begin{aligned}\|t_U(e^A)\|_2 &= \max_{1\le k\le n}\frac{[UW^*\operatorname{diag}(e^{2\lambda_1},e^{2\lambda_2},\dots,e^{2\lambda_n})WU^*]_{kk}}{\left|[UW^*\operatorname{diag}(e^{\lambda_1},e^{\lambda_2},\dots,e^{\lambda_n})WU^*]_{kk}\right|}\\ &\ge \max_{1\le k\le n}\left|[UW^*\operatorname{diag}(e^{\lambda_1},e^{\lambda_2},\dots,e^{\lambda_n})WU^*]_{kk}\right|\\ &\ge e^{\lambda_{\min}(A)}.\end{aligned}$$

□

Theorem 2.2. *Let $A\in\mathbb{C}^{n\times n}$ with γ_k and η_k given by* (2.3). *If both A and $c_U(A)$ are nonsingular, then we have*

$$\|e^{t_U(A)}\|_2 = \max_{1\le k\le n}\left|e^{\frac{\gamma_k}{\eta_k}}\right|.$$

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

Proof. We have by (2.4),

$$\|e^{t_U(A)}\|_2 = \left\|U^*\mathrm{diag}\left(e^{\frac{\gamma_1}{\bar\eta_1}}, e^{\frac{\gamma_2}{\bar\eta_2}}, \ldots, e^{\frac{\gamma_n}{\bar\eta_n}}\right)U\right\|_2 = \max_{1\le k\le n}\left|e^{\frac{\gamma_k}{\bar\eta_k}}\right|. \qquad \square$$

For the determinants of e^A and $e^{t_U(A)}$, we have the following results.

Lemma 2.2. ([1, p.78] [3, p.648]) *For $A \in \mathbb{C}^{n\times n}$,*

$$\det(e^A) = \exp[\mathrm{tr}(A)].$$

Theorem 2.3. *Let $A \in \mathbb{C}^{n\times n}$ with γ_k and η_k given by* (2.3). *If both A and $c_U(A)$ are nonsingular, then we have*

$$\det(e^{t_U(A)}) = \exp\Big[\sum_{k=1}^{n}\frac{\gamma_k}{\bar\eta_k}\Big].$$

If A is Hermitian positive definite, then

$$\det(e^{t_U(A)}) \ge \det(e^A).$$

Proof. We have by Lemma 2.2,

$$\begin{aligned}\det(e^{t_U(A)}) &= \exp\Big[\mathrm{tr}\big(t_U(A)\big)\Big] = \exp\Big[\mathrm{tr}\big(U^*\delta(UAA^*U^*)\delta^{-1}(UA^*U^*)U\big)\Big]\\ &= \exp\Big[\mathrm{tr}\big(\delta(UAA^*U^*)\delta^{-1}(UA^*U^*)\big)\Big] = \exp\Big[\sum_{k=1}^{n}\frac{\gamma_k}{\bar\eta_k}\Big].\end{aligned}$$

If A is Hermitian positive definite, then by Lemma 1.1 (v),

$$\begin{aligned}\det(e^{t_U(A)}) &= \exp\Big[\mathrm{tr}\big(\delta(UAA^*U^*)\delta^{-1}(UA^*U^*)\big)\Big]\\ &= \exp\left[\sum_{k=1}^{n}\frac{[UAA^*U^*]_{kk}}{[UA^*U^*]_{kk}}\right] = \exp\left[\sum_{k=1}^{n}\frac{[UAA^*U^*]_{kk}\cdot[UAU^*]_{kk}}{[UAU^*]_{kk}\cdot[UA^*U^*]_{kk}}\right]\\ &\ge \exp\left[\sum_{k=1}^{n}[UAU^*]_{kk}\right] = \exp\left[\sum_{k=1}^{n}\lambda_k(UAU^*)\right] = \exp\left[\sum_{k=1}^{n}\lambda_k(A)\right]\\ &= \exp[\mathrm{tr}(A)] = \det(e^A),\end{aligned}$$

where $\lambda_k(E)$ denotes the kth eigenvalue of the matrix E for $k = 1, \ldots, n$. $\square$

Before the end of section 2.1, we discuss the Kronecker product and sum of the matrix exponential.

Definition 2.1. [12, p.331] *The Kronecker product of $A = [a_{ij}] \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$ is defined by*

$$A \otimes B = [a_{ij}B] \in \mathbb{C}^{mp\times nq}.$$

The Kronecker sum of $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$ is defined by

$$A \oplus B = A \otimes I_m + I_n \otimes B.$$

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

Let U be an $n \times n$ unitary matrix and V be an $m \times m$ unitary matrix. We first recall the following lemma on the operator $c_{U\otimes V}$.

Lemma 2.3. [16, Theorem 2.3] *For $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$, we have*

$$c_{U\otimes V}(A \otimes B) = c_U(A) \otimes c_V(B).$$

On the operator $t_{U\otimes V}$, we have the following theorem.

Theorem 2.4. *Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$. If A, B, $c_U(A)$, and $c_V(B)$ are nonsingular, then we have*

$$t_{U\otimes V}(A \otimes B) = t_U(A) \otimes t_V(B).$$

Proof. By using Lemma 2.3, we have

$$\begin{aligned}
t_{U\otimes V}(A \otimes B) &= c_{U\otimes V}\big((A \otimes B)(A \otimes B)^*\big)\big[c_{U\otimes V}\big((A \otimes B)^*\big)\big]^{-1} \\
&= c_{U\otimes V}(AA^* \otimes BB^*)\big[c_{U\otimes V}(A^* \otimes B^*)\big]^{-1} \\
&= \big(c_U(AA^*) \otimes c_V(BB^*)\big)\big[c_U(A^*) \otimes c_V(B^*)\big]^{-1} \\
&= \Big(c_U(AA^*) \otimes c_V(BB^*)\Big)\Big(\big[c_U(A^*)\big]^{-1} \otimes \big[c_V(B^*)\big]^{-1}\Big) \\
&= \Big(c_U(AA^*)\big[c_U(A^*)\big]^{-1}\Big) \otimes \Big(c_V(BB^*)\big[c_V(B^*)\big]^{-1}\Big) \\
&= t_U(A) \otimes t_V(B).
\end{aligned}$$

□

We recall the following results on the Kronecker product and sum of the matrix exponential.

Lemma 2.4. ([12, p.237] [3, p.755]) *Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$. Then*

$$\begin{aligned}
&e^{A\otimes I} = e^A \otimes I, \qquad e^{I\otimes B} = I \otimes e^B, \\
&e^{A\oplus B} = e^A \otimes e^B, \qquad \mathrm{tr}(e^{A\oplus B}) = \mathrm{tr}(e^A) \cdot \mathrm{tr}(e^B).
\end{aligned}$$

Then we have the following theorem for the superoptimal preconditioners of the matrix exponential.

Theorem 2.5. *Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$. If A, B, $c_U(A)$, and $c_V(B)$ are nonsingular, then we have*

(i) $t_{U\otimes V}(e^{A\oplus B}) = t_U(e^A) \otimes t_V(e^B), \qquad e^{t_{U\otimes V}(A\otimes B)} = e^{t_U(A)\otimes t_V(B)}.$

(ii) $\mathrm{tr}\big(t_{U\otimes V}(e^{A\oplus B})\big) = \sum_{k=1}^{n}(\alpha_k/\bar{\beta}_k) \cdot \sum_{k=1}^{m}(\zeta_k/\bar{\xi}_k)$, *where*

$$\begin{aligned}
&\delta(Ue^Ae^{A^*}U^*) = \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n), \qquad \delta(Ue^AU^*) = \mathrm{diag}(\beta_1, \beta_2, \ldots, \beta_n), \\
&\delta(Ve^Be^{B^*}V^*) = \mathrm{diag}(\zeta_1, \zeta_2, \ldots, \zeta_m), \qquad \delta(Ve^BV^*) = \mathrm{diag}(\xi_1, \xi_2, \ldots, \xi_m).
\end{aligned}$$

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

Proof. The proof of (i) follows directly from Theorem 2.4 and Lemma 2.4. For the equality of (ii), we have by Lemma 1.1 (ii),

$$\begin{aligned}
&\mathrm{tr}\big(t_{U\otimes V}(e^{A\oplus B})\big) = \mathrm{tr}\big(t_U(e^A)\otimes t_V(e^B)\big) = \mathrm{tr}\big(t_U(e^A)\big)\cdot \mathrm{tr}\big(t_V(e^B)\big)\\
=\ & \mathrm{tr}\Big(U^*\delta\big(Ue^A(e^A)^*U^*\big)\delta^{-1}\big(U(e^A)^*U^*\big)U\Big)\\
&\qquad \cdot \mathrm{tr}\Big(V^*\delta\big(Ve^B(e^B)^*V^*\big)\delta^{-1}\big(V(e^B)^*V^*\big)V\Big)\\
=\ & \mathrm{tr}\big(\delta(Ue^Ae^{A^*}U^*)\delta^{-1}(Ue^{A^*}U^*)\big)\cdot \mathrm{tr}\big(\delta(Ve^Be^{B^*}V^*)\delta^{-1}(Ve^{B^*}V^*)\big)\\
=\ & \sum_{k=1}^{n}(\alpha_k/\bar{\beta}_k)\cdot\sum_{k=1}^{m}(\zeta_k/\bar{\xi}_k).
\end{aligned}$$

□

2.2. Superoptimal preconditioners for matrix cosine and matrix sine

The matrix cosine and matrix sine are defined for $A\in\mathbb{C}^{n\times n}$ by

$$\begin{aligned}
\cos(A) &\equiv I - \frac{A^2}{2!}+\frac{A^4}{4!}-\frac{A^6}{6!}+\cdots,\\
\sin(A) &\equiv A - \frac{A^3}{3!}+\frac{A^5}{5!}-\frac{A^7}{7!}+\cdots.
\end{aligned}$$

Since

$$e^{\mathrm{i}A} = \cos(A)+\mathrm{i}\sin(A),\quad \mathrm{i}\equiv\sqrt{-1},$$

we have

$$\cos(A) = \frac{e^{\mathrm{i}A}+e^{-\mathrm{i}A}}{2},\quad \sin(A)=\frac{e^{\mathrm{i}A}-e^{-\mathrm{i}A}}{2\mathrm{i}}. \tag{2.6}$$

Note that

$$\begin{aligned}
U\cos(A)U^* &= U\Big(I-\frac{1}{2!}A^2+\frac{1}{4!}A^4-\frac{1}{6!}A^6+\cdots\Big)U^*\\
&= I-\frac{B^2}{2!}+\frac{B^4}{4!}-\frac{B^6}{6!}+\cdots=\cos(B),
\end{aligned}$$

where $B=UAU^*$. If both A and $c_U(A)$ are nonsingular, then $t_U(\cos(A))$ is given by Lemma 1.1 (ii),

$$\begin{aligned}
t_U(\cos(A)) &= U^*\delta\big(U\cos(A)\cos(A)^*U^*\big)\delta^{-1}\big(U\cos(A)^*U^*\big)U\\
&= U^*\delta\big(U\cos(A)U^*U\cos(A)^*U^*\big)\delta^{-1}\big(U\cos(A)^*U^*\big)U\\
&= U^*\delta\big(\cos(B)\cos(B)^*\big)\delta^{-1}\big(\cos(B)^*\big)U.
\end{aligned}$$

Similarly, we have

$$t_U(\sin(A)) = U^*\delta\big(\sin(B)\sin(B)^*\big)\delta^{-1}\big(\sin(B)^*\big)U.$$

We see from (2.6) that, if B is very sparse, one may compute $\cos(B)$ and $\sin(B)$ easily since e^B can be computed readily.

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

Moreover, if A is Hermitian, then there exists a unitary matrix W such that $A = W^*\Lambda W$ where $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. As we did in the case of matrix exponential, one can obtain

$$t_U(\cos(A)) = U^*\delta\Big(UW^*\operatorname{diag}\big(\cos^2(\lambda_1), \cos^2(\lambda_2), \ldots, \cos^2(\lambda_n)\big)WU^*\Big)\times \\ \delta^{-1}\Big(UW^*\operatorname{diag}\big(\cos(\lambda_1), \cos(\lambda_2), \ldots, \cos(\lambda_n)\big)WU^*\Big)U,$$

and similarly

$$t_U(\sin(A)) = U^*\delta\Big(UW^*\operatorname{diag}(\sin^2(\lambda_1), \sin^2(\lambda_2), \ldots, \sin^2(\lambda_n))WU^*\Big)\times \\ \delta^{-1}\Big(UW^*\operatorname{diag}\big(\sin(\lambda_1), \sin(\lambda_2), \ldots, \sin(\lambda_n)\big)WU^*\Big)U.$$

Again, let γ_k and η_k be given by (2.3). For preconditioners $\cos(t_U(A))$ and $\sin(t_U(A))$, we have

$$\cos(t_U(A)) = U^*\operatorname{diag}\Big(\cos\Big(\frac{\gamma_1}{\bar{\eta}_1}\Big), \cos\Big(\frac{\gamma_2}{\bar{\eta}_2}\Big), \ldots, \cos\Big(\frac{\gamma_n}{\bar{\eta}_n}\Big)\Big)U,$$
$$\sin(t_U(A)) = U^*\operatorname{diag}\Big(\sin\Big(\frac{\gamma_1}{\bar{\eta}_1}\Big), \sin\Big(\frac{\gamma_2}{\bar{\eta}_2}\Big), \ldots, \sin\Big(\frac{\gamma_n}{\bar{\eta}_n}\Big)\Big)U.$$

Now, we study norm bounds of $t_U(\cos(A))$ and $t_U(\sin(A))$.

Theorem 2.6. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian with eigenvalues $\{\lambda_k\}_{k=1}^n$. Suppose that both A and $c_U(A)$ are nonsingular. Then we have*

(i) *If $-\pi/2 \le \lambda_k \le \pi/2$ for $k = 1, \ldots, n$, then*

$$\max_{1\le k\le n} \cos(\lambda_k) \le \|t_U(\cos(A))\|_2 \le \frac{\max\limits_{1\le k\le n} \cos^2(\lambda_k)}{\min\limits_{1\le k\le n} \cos(\lambda_k)}.$$

(ii) *If $0 \le \lambda_k \le \pi$ for $k = 1, \ldots, n$, then*

$$\max_{1\le k\le n} \sin(\lambda_k) \le \|t_U(\sin(A))\|_2 \le \frac{\max\limits_{1\le k\le n} \sin^2(\lambda_k)}{\min\limits_{1\le k\le n} \sin(\lambda_k)}.$$

Proof. The proof follows from Theorem 2.1. □

For norm bounds of $\cos(t_U(A))$ and $\sin(t_U(A))$, we have the following theorem. The proof of the theorem is straightforward.

Theorem 2.7. *Let $A \in \mathbb{C}^{n\times n}$. If both A and $c_U(A)$ are nonsingular, then we have*

$$\|\cos(t_U(A))\|_2 = \max_k \Big|\cos\Big(\frac{\gamma_k}{\bar{\eta}_k}\Big)\Big|, \qquad \|\sin(t_U(A))\|_2 = \max_k \Big|\sin\Big(\frac{\gamma_k}{\bar{\eta}_k}\Big)\Big|,$$

where γ_k and η_k are given by (2.3).

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

We now recall the following results on the Kroneker product.

Lemma 2.5. [12, p.288] *Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$. Then*

$$\begin{aligned}&\cos(A\otimes I) = \cos(A)\otimes I, \quad \cos(I\otimes B) = I\otimes\cos(B),\\ &\sin(A\otimes I) = \sin(A)\otimes I, \quad \sin(I\otimes B) = I\otimes\sin(B).\end{aligned}$$

Hence, based on Theorem 2.4, we can easily establish the following theorem for the superoptimal preconditioners of matrix cosine and matrix sine.

Theorem 2.8. *Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$. If A, B, $c_U(A)$, and $c_V(B)$ are nonsingular, then we have*

$$\begin{aligned}&t_{U\otimes V}(\cos(A\otimes I)) = t_U(\cos(A))\otimes I, && \cos(t_{U\otimes V}(A\otimes I)) = \cos(t_U(A))\otimes I,\\ &t_{U\otimes V}(\sin(A\otimes I)) = t_U(\sin(A))\otimes I, && \sin(t_{U\otimes V}(A\otimes I)) = \sin(t_U(A))\otimes I,\\ &t_{U\otimes V}(\cos(I\otimes B)) = I\otimes t_V(\cos(B)), && \cos(t_{U\otimes V}(I\otimes B)) = I\otimes\cos(t_V(B)),\\ &t_{U\otimes V}(\sin(I\otimes B)) = I\otimes t_V(\sin(B)), && \sin(t_{U\otimes V}(I\otimes B)) = I\otimes\sin(t_V(B)).\end{aligned}$$

2.3. Superoptimal preconditioners for matrix logarithm

In this section we discuss the matrix logarithm:

$$\log(I+A) \equiv A - \frac{A^2}{2} + \frac{A^3}{3} - \frac{A^4}{4} + \cdots, \qquad \rho(A) < 1,$$

where $\rho(A)$ is the spectral radius of matrix A.

We observe that

$$\begin{aligned}U\log(I+A)U^* &= U\Big(A - \frac{A^2}{2} + \frac{A^3}{3} - \frac{A^4}{4} + \cdots\Big)U^*\\ &= B - \frac{B^2}{2} + \frac{B^3}{3} - \frac{B^4}{4} + \cdots = \log(I+B), \quad \rho(B) < 1,\end{aligned}$$

where $B = UAU^*$. If both A and $c_U(A)$ are nonsingular, then we have

$$\begin{aligned}&t_U(\log(I+A))\\ &= U^*\delta\Big(\log(I+B)\big[\log(I+B)\big]^*\Big)\delta^{-1}\Big(\big[\log(I+B)\big]^*\Big)U, \quad \rho(B) < 1.\end{aligned}$$

Similarly, if A is Hermitian, then there exists a unitary matrix W such that $A = W^*\Lambda W$ where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ with $|\lambda_k| < 1$ for any k. As we did in sections 2.1 and 2.2, we could obtain

$$\begin{aligned}&t_U(\log(I+A))\\ &= U^*\delta\Big(UW^*\mathrm{diag}\big(\log^2(1+\lambda_1), \log^2(1+\lambda_2), \ldots, \log^2(1+\lambda_n)\big)WU^*\Big)\times\\ &\qquad \delta^{-1}\Big(UW^*\mathrm{diag}\big(\log(1+\lambda_1), \log(1+\lambda_2), \ldots, \log(1+\lambda_n)\big)WU^*\Big)U.\end{aligned}$$

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

Furthermore, for the preconditioner $\log(I + t_U(A))$, again let γ_k and η_k be given by (2.3). When A is Hermitian with $\rho(A) < 1$ and $\gamma_k/|\eta_k| < 1$ for any k, we have

$$\log(I + t_U(A)) = U^* \mathrm{diag}\Big(\log\Big(1 + \frac{\gamma_1}{\eta_1}\Big), \ldots, \log\Big(1 + \frac{\gamma_n}{\eta_n}\Big)\Big)U. \tag{2.7}$$

Thus we immediately have the following theorem by using (2.7).

Theorem 2.9. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian with $\rho(A) < 1$ and let γ_k and η_k be given by* (2.3). *If $\gamma_k/|\eta_k| < 1$ for any k, then we have*

$$\|\log(I + t_U(A))\|_2 = \max_{1\le k\le n} \Big|\log\Big(1 + \frac{\gamma_k}{\eta_k}\Big)\Big|.$$

3. Numerical experiments

In this section, we solve a system of matrix function $f(A)\mathbf{x} = \mathbf{b}$. We propose the preconditioned conjugate gradient (PCG) method with two superoptimal preconditioners $t_U(f(A))$ and $f(t_U(A))$. We discuss the cost for constructing preconditioners and report some numerical results.

3.1. Cost of constructing preconditioners

We compare the computational cost of constructing two superoptimal preconditioners $t_U(f(A))$ and $f(t_U(A))$ for the matrix function $f(A)$. By Lemma 1.1 (ii), for different matrix functions f, we get

$$t_U(f(A)) = U^*\delta(Uf(A)f(A)^*U^*)\delta^{-1}(Uf(A)^*U^*)U,$$
$$f(t_U(A)) = U^*f(\delta(UAA^*U^*)\delta^{-1}(UA^*U^*))U.$$

Given a general unitary matrix U and the matrix $f(A)$, it needs $\mathcal{O}(n^3)$ operations for constructing $t_U(f(A))$ and $f(t_U(A))$. Moreover, the matrix-vector products $[t_U(f(A))]^{-1}\mathbf{y}$ and $[f(t_U(A))]^{-1}\mathbf{y}$ require $\mathcal{O}(n^3)$ operations, where $\mathbf{y}$ is an n-vector.

When U is the $n \times n$ Fourier matrix F, $t_F(A)$, $t_F(f(A))$, and $f(t_F(A))$ are all circulant matrices. Thus the construction of both $t_F(f(A))$ and $f(t_F(A))$ requires $\mathcal{O}(n^2 \log n)$ operations for general matrices A [5]. Then the matrix-vector products $[t_F(f(A))]^{-1}\mathbf{y}$ and $[f(t_F(A))]^{-1}\mathbf{y}$ need $\mathcal{O}(n \log n)$ operations by using the fast Fourier transforms (FFTs) [4, 5, 13].

Let the matrix A be the $n \times n$ Toeplitz matrix T_n given by:

$$T_n \equiv \begin{pmatrix} t_0 & t_{-1} & \cdots & t_{2-n} & t_{1-n} \\ t_1 & t_0 & t_{-1} & \ddots & t_{2-n} \\ \vdots & t_1 & t_0 & \ddots & \vdots \\ t_{n-2} & \ddots & \ddots & \ddots & t_{-1} \\ t_{n-1} & t_{n-2} & \cdots & t_1 & t_0 \end{pmatrix}.$$

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

Suppose that the diagonals $\{t_k\}_{k=-n+1}^{n-1}$ of T_n are the Fourier coefficients of a function g, i.e.,

$$t_k(g) \equiv \frac{1}{2\pi}\int_{-\pi}^{\pi} g(x)e^{-\mathbf{i}kx}dx.$$

Then g is called the generating function of $T_n(g)$ for $1 \le n < \infty$ [13, p.19]. If g is a real-valued function defined on $[-\pi, \pi]$, then the Toeplitz matrix $T_n(g)$ is Hermitian. By using Theorem 1.21 in [13, p.20], we get

$$g_{\min} \le \lambda_{\min}(T_n(g)) \le \lambda_{\max}(T_n(g)) \le g_{\max},$$

where $g_{\min}$, $g_{\max}$, $\lambda_{\min}(\cdot)$, and $\lambda_{\max}(\cdot)$ denote the minimum and maximum values of g on $[-\pi, \pi]$ and the smallest and largest eigenvalues respectively. Moreover, from Theorem 1.13 in [12, p.10], it follows that the eigenvalues of $f(T_n)$ are given by $\{f(\lambda_k)\}_{k=1}^{n}$, where λ_k, $k = 1, \ldots, n$, are the eigenvalues of $T_n(g)$. Thus,

$$f_{\min} \le \lambda_{\min}(f(T_n)) \le \lambda_{\max}(f(T_n)) \le f_{\max}, \tag{3.1}$$

where $f_{\min}$ and $f_{\max}$ denote the minimum and maximum values of f on $[g_{\min}, g_{\max}]$ respectively. For the matrix function $f(T_n)$, it follows that the construction of both $t_F(f(T_n))$ and $f(t_F(T_n))$ only requires $\mathcal{O}(n \log n)$ operations [5]. The matrix-vector products $[t_F(f(T_n))]^{-1}\mathbf{y}$ and $[f(t_F(T_n))]^{-1}\mathbf{y}$ require $\mathcal{O}(n \log n)$ operations by using the FFTs.

3.2. Numerical results

In this section, we report some numerical tests to illustrate the effectiveness of our superoptimal preconditioners $t_F(f(T_n))$ and $f(t_F(T_n))$ for solving the system

$$f(T_n)\mathbf{x} = \mathbf{b} \tag{3.2}$$

by using the PCG method. We compute the matrix functions $f(T_n)$ by the built-in function `funm`. For the Toeplitz matrix exponential, the matrix-vector product $\exp(T_n)\mathbf{x}$ can be calculated efficiently by a fast algorithm provided in [18]. By (2.6), we know that $\cos(A)\mathbf{x}$ and $\sin(A)\mathbf{x}$ can be computed fast. Let the n-vector $\mathbf{b}$ be generated randomly, i.e., $\mathbf{b} = \texttt{randn}(n)$. For comparison purposes, we solve the system (3.2) by the built-in function `pcg` without preconditioner, with the preconditioners $c_F(f(T_n))$, $f(c_F(T_n))$, $t_F(f(T_n))$, or $f(t_F(T_n))$. Here $c_F(f(T_n))$ and $f(c_F(T_n))$ are the optimal preconditioners discussed in [16]. Then we compare the performance of these different preconditioners. In our numerical experiments, we set the largest number of the PCG iterations to be 3×10^4 and the PCG method is stopped when

$$\frac{\|f(T_n)\mathbf{x} - \mathbf{b}\|}{\|\mathbf{b}\|} < 10^{-6}.$$

All the numerical tests are carried out by using `MATLAB` 2010a on a personal computer of 2.0 GHz CPU and 1GB RAM. We repeat our experiments for 10 different random vectors $\mathbf{b}$. In what follows, '`cputime`', '`IT.`', '`Res.`', and '`Cond.`' mean the averaged

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

total CPU time in seconds, the averaged number of the PCG iterations, the averaged relative residual at the final iterate of the algorithm with proposed preconditioners, and the condition numbers of the matrices $f(T_n)$, respectively. We also use I, II, III, VI, and V to denote the PCG method without preconditioner, with $c_F(f(T_n))$, $f(c_F(T_n))$, $t_F(f(T_n))$, and $f(t_F(T_n))$, respectively.

Example 3.1. We consider the matrix cosine $\cos(T_n)$ for different n, where the Toeplitz matrices T_n are generated by the function $g(x) = \frac{\pi}{2}\cos(0.25x)$.

Example 3.2. We consider the matrix sine $\sin(T_n)$ for different n, where the Toeplitz matrices T_n are generated by the function $g(x) = \pi\cos(0.1x)$.

Example 3.3. We consider the matrix logarithm $\log(I + T_n)$ for different n, where the Toeplitz matrices T_n are generated by the function $g(x) = \frac{1}{2}\cos(x) + \frac{1}{2}$.

Tables 1–3 list the numerical results for Examples 3.1–3.3. We observe from Tables 1–3 that our proposed superoptimal preconditioners $t_F(f(T_n))$ and $f(t_F(T_n))$ can effectively reduce the number of iterations and then improve the performance of the PCG method. We also note that the proposed superoptimal preconditioners $t_F(f(T_n))$ and $f(t_F(T_n))$ work a little bit worse than the optimal preconditioners $c_F(f(T_n))$ and $f(c_F(T_n))$ for some examples (see Table 1 (VI), Table 2 (VI), and Table 3 (VI,V)). The numerical experiments show that $f(t_F(T_n))$ is better for Examples 3.1–3.2, and is worse than $t_F(f(T_n))$ for Example 3.3. It needs further research to theoretically determine which superoptimal preconditioners ($t_U(f(A))$ or $f(t_U(A))$) is better for solving the system $f(A)\mathbf{x} = \mathbf{b}$.

Table 1: Numerical results for Example 3.1.

	n	cputime	IT.	Res.	Cond.
I	200	0.2122 s	265.1	6.0259×10^{-7}	3.6706×10^{4}
	400	0.2587 s	556.0	4.6721×10^{-7}	1.4673×10^{5}
	800	0.6881 s	1145.4	6.2358×10^{-7}	5.8674×10^{5}
II	200	0.0463 s	22.1	6.6235×10^{-7}	1.6660×10^{3}
	400	0.0177 s	28.4	4.1695×10^{-7}	5.4677×10^{3}
	800	0.0441 s	37.1	5.5844×10^{-7}	1.8161×10^{4}
III	200	0.0105 s	22.1	6.8617×10^{-7}	1.6671×10^{3}
	400	0.0363 s	28.7	3.6988×10^{-7}	5.4718×10^{3}
	800	0.0330 s	37.1	5.9451×10^{-7}	1.8175×10^{4}
VI	200	0.0468 s	68.6	7.4037×10^{-7}	2.5869×10^{4}
	400	0.0970 s	146.5	5.6862×10^{-7}	1.2762×10^{5}
	800	0.2538 s	292.0	6.6667×10^{-7}	6.3618×10^{5}
V	200	0.0358 s	22	3.3927×10^{-7}	1.6458×10^{3}
	400	0.0195 s	27.8	6.0484×10^{-7}	5.3941×10^{3}
	800	0.0371 s	36.5	5.2896×10^{-7}	1.7903×10^{4}

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

Table 2: Numerical results for Example 3.2.

	n	cputime	IT.	Res.	Cond.
I	200	0.0459 s	273.0	6.1295×10^{-7}	3.9531×10^{4}
	400	0.1239 s	574.0	5.7994×10^{-7}	1.5809×10^{5}
	800	1.4712 s	1185.9	6.2790×10^{-7}	6.3231×10^{5}
II	200	0.0079 s	22.4	4.4536×10^{-7}	1.7349×10^{3}
	400	0.0110 s	28.5	4.4570×10^{-7}	5.6965×10^{3}
	800	0.0653 s	37.2	6.7472×10^{-7}	1.8925×10^{4}
III	200	0.0087 s	22.4	4.4719×10^{-7}	1.7350×10^{3}
	400	0.0116 s	28.9	3.6912×10^{-7}	5.6969×10^{3}
	800	0.0667 s	37.9	7.4601×10^{-7}	1.8927×10^{4}
VI	200	0.0200 s	70.1	6.1063×10^{-7}	2.7438×10^{4}
	400	0.0419 s	147.2	6.7253×10^{-7}	1.3541×10^{5}
	800	0.4882 s	304	7.3726×10^{-7}	6.7521×10^{5}
V	200	0.0063 s	22.3	4.8221×10^{-7}	1.7313×10^{3}
	400	0.0128 s	28.8	3.8972×10^{-7}	5.6837×10^{3}
	800	0.0728 s	37.3	7.5441×10^{-7}	1.888×10^{4}

Table 3: Numerical results for Example 3.3.

	n	cputime	IT.	Res.	Cond.
I	200	0.0746 s	175.9	4.4723×10^{-7}	1.1350×10^{4}
	400	0.1880 s	343.7	6.8922×10^{-7}	4.5173×10^{4}
	800	0.8215 s	678.2	7.7740×10^{-7}	1.8024×10^{5}
II	200	0.0098 s	18	6.4025×10^{-7}	7.8105×10^{2}
	400	0.0355 s	23.5	4.4678×10^{-7}	2.5373×10^{3}
	800	0.0510 s	30.5	4.5541×10^{-7}	8.3814×10^{3}
III	200	0.0294 s	18.4	4.2871×10^{-7}	7.9661×10^{2}
	400	0.0157 s	24.2	6.3259×10^{-7}	2.5946×10^{3}
	800	0.0539 s	31.7	3.8683×10^{-7}	8.5827×10^{3}
VI	200	0.0213 s	48.1	3.1752×10^{-7}	4.3108×10^{4}
	400	0.1206 s	85.0	4.9731×10^{-7}	1.6947×10^{5}
	800	0.2602 s	181.4	7.4509×10^{-7}	6.7168×10^{5}
V	200	0.0375 s	52.5	2.5866×10^{-7}	9.4390×10^{3}
	400	0.1006 s	97.5	5.8469×10^{-7}	4.5797×10^{4}
	800	0.2889 s	206.5	5.7851×10^{-7}	2.2504×10^{5}

Acknowledgments This work was partially supported by the National Natural Science Foundation of China grant 11271308, the Fundamental Research Funds for the Central Universities (No. 20720150001) and the research grant MYRG098(Y2-L3)-FST13-JXQ from University of Macau. The authors would like to thank Mr. Zhao Zhi for discussion and to thank the editor and anonymous reviewer for very useful suggestions and comments.

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press

References

[1] A. Baker, *Matrix Groups – An Introduction to Lie Group Theory*, Springer, New York, 2002.
[2] F. Di Benedetto, C. Estatico, and S. Serra-Capizzano, *Superoptimal preconditioner conjugate gradient iteration for image deblurring*, SIAM J. Sci. Comput., vol. 26 (2005), pp. 1012–1035.
[3] D. Bernstein, *Matrix Mathematics – Theory, Facts, and Formulas*, Princeton University Press, Princeton, 2009.
[4] R. Chan and X. Jin, *An Introduction to Iterative Toeplitz Solvers*, SIAM, Philadelpha, 2007.
[5] R. Chan, X. Jin, and M. Yeung, *The circulant operator in the Banach algebra of matrices*, Linear Algebra Appl., vol. 149 (1991), pp. 41–53.
[6] R. Chan, X. Jin, and M. Yeung, *The spectra of super-optimal circulant preconditioned Toeplitz systems*, SIAM J. Numer Anal., vol. 28 (1991), pp. 871–879.
[7] R. Chan and M. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Review, vol. 38 (1996), pp. 427–482.
[8] T. Chan, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Statist. Comput., vol. 9 (1988), pp. 766–771.
[9] C. Cheng, X. Jin, S. Vong, and W. Wang, *A note on spectra of optimal and superoptimal preconditioned matrices*, Linear Algebra Appl., vol. 422 (2007), pp. 482–485.
[10] C. Cheng, X. Jin and Y. Wei, *Stability properties of superoptimal preconditioner from numerical range*, Numer. Linear Algebra Appl., vol. 13 (2006), pp. 513–521.
[11] C. Estatico, *A class of filtering superoptimal preconditioners for highly ill conditioned linear systems*, BIT, vol. 42 (2002), pp. 753–778.
[12] N. Higham, *Functions of Matrices – Theory and Computation*, SIAM, Philadelpha, 2008.
[13] X. Jin, *Preconditioning Techniques for Toeplitz Systems*, Higher Education Press, Beijing, 2010.
[14] X. Jin and Y. Wei, *A short note on singular values of optimal and superoptimal preconditioned matrices*, International J. Comput. Math. vol. 84 (2007), pp. 1261–1263.
[15] X. Jin and Y. Wei, *A survey and some extensions of T. Chan's preconditioner*, Linear Algebra Appl., vol. 428 (2008), pp. 403–412.
[16] X. Jin, Z. Zhao, and S. Tam, *Optimal preconditioners for functions of matrices*, Linear Algebra Appl., vol. 457 (2014), pp. 224–243.
[17] S. Lee, X. Liu, and H. Sun, *Fast exponential time integration scheme for option pricing with jumps*, Numer. Linear Algebra Appl., vol. 19 (2012), pp. 87–101.
[18] S. Lee, H. Pang, and H. Sun, *Shift-invert Aronldi approximation to the Toeplitz matrix exponential*, SIAM J. Sci. Comput., vol. 32 (2010), pp. 774–792.
[19] C. Moler and C. Van Loan, *Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later*, SIAM Review, vol. 45 (2003), pp. 3–49.
[20] E. Tyrtyshnikov, *Optimal and superoptimal circulant preconditioners*, SIAM J. Matrix Anal. Appl., vol. 13 (1992), pp. 459–473.

https://doi.org/10.4208/nmtma.2015.my1340 Published online by Cambridge University Press