SIAM J. Sci. Comput.
Vol. 26, No. 6, pp. 1852–1870

FACTORIZED BANDED INVERSE PRECONDITIONERS FOR MATRICES WITH TOEPLITZ STRUCTURE[*]

FU-RONG LIN[†], MICHAEL K. NG[‡], AND WAI-KI CHING[‡]

Abstract. In this paper, we study factorized banded inverse preconditioners for matrices with Toeplitz structure. We show that if a Toeplitz matrix T has certain off-diagonal decay property, then the factorized banded inverse preconditioner approximates T^{-1} accurately, and the spectra of these preconditioned matrices are clustered around 1. In nonlinear image restoration applications, Toeplitz-related systems of the form $I+T^*DT$ are required to solve, where D is a positive nonconstant diagonal matrix. We construct inverse preconditioners for such matrices. Numerical results show that the performance of our proposed preconditioners are superior to that of circulant preconditioners. A two-dimensional nonlinear image restoration example is also presented to demonstrate the effectiveness of the proposed preconditioner.

Key words. inverse preconditioners, Toeplitz-related matrices, nonlinear image restoration

AMS subject classifications. 65F10, 65F15, 65T10

DOI. 10.1137/030601272

1. Introduction.

An $m \times m$ matrix T_m is called a *Toeplitz* matrix if it is constant along its diagonals, i.e.,

$$T_m = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{2-m} & t_{1-m} \\ t_1 & t_0 & t_{-1} & & t_{2-m} \\ \vdots & t_1 & t_0 & \ddots & \vdots \\ t_{m-2} & & \ddots & \ddots & t_{-1} \\ t_{m-1} & t_{m-2} & \cdots & t_1 & t_0 \end{bmatrix}. \tag{1.1}$$

If each t_i in (1.1) is an $n \times n$ matrix, then the corresponding matrix is called a block-Toeplitz matrix. An $mn \times mn$ matrix $T_{m,n}$ is called a block-Toeplitz matrix with Toeplitz-blocks (BTTB) if it is an $m \times m$ block Toeplitz matrix with each block being an $n \times n$ Toeplitz matrix. In this paper, we consider solving linear systems of equations with Toeplitz structure

$$Ax = b \tag{1.2}$$

by the preconditioned conjugate gradient (PCG) method. Here the matrix A is assumed to be symmetric positive definite and takes one of the following forms: (i) a Toeplitz matrix T_m; (ii) a Toeplitz-related matrix $(I_m + T_m^* D_m T_m)$, (iii) a BTTB matrix $T_{m,n}$, or (iv) a BTTB-related matrix $(I_{mn} + T_{m,n}^* D_{mn} T_{m,n})$. Here I_k is the

[*] Received by the editors October 19, 2003; accepted for publication (in revised form) August 17, 2004; published electronically May 20, 2005.
http://www.siam.org/journals/sisc/26-6/60127.html

[†] Department of Mathematics, Shantou University, Shantou, Guangdong 515063, People's Republic of China. This author's research was supported by Natural Science Foundation of China grant 10271070 and Guangdong Provincial Natural Science Foundation of China grant 021244.

[‡] Department of Mathematics, University of Hong Kong, Pokfulam Road, Hong Kong (mng@maths.hku.hk). The research of the second author was supported in part by Hong Kong Research Grants Council (RGC) grant HKU 7130/02P, HKU 7046/03P, and HKU 7035/04P. The research of the third author was supported in part by RGC grant HKU 7126/02P and HKU CRCG grant 10204436.

identity matrix of size k and D_k is a positive diagonal matrix of size k. Toeplitz systems arise in a variety of applications in mathematics and engineering; see, for instance, [8]. Toeplitz-related systems in (ii) and (iv) occur in the nonlinear signal and image restoration [1, 30].

Strang [22] and Olkin [23] independently proposed the use of circulant matrices to precondition Toeplitz matrices in conjugate gradient iterations. An $n \times n$ matrix C_n is called a circulant matrix if it is a Toeplitz matrix ($[C_n]_{ij} = c_{i-j}$) and it satisfies $c_{-j} = c_{n-j}$ for $j = 1, 2, \ldots , n-1$. The following are some important properties of a circulant matrix: (i) all circulant matrices can be diagonalized by the Fourier matrix [12]; (ii) the multiplication of a circulant matrix to a vector and the solution of a circulant system can be done in $O(n \log n)$ operations respectively by using the Fast Fourier Transform (FFT); and (iii) the multiplication of a Toeplitz matrix T_n to any vector y can also be computed by FFT by first embedding T_n into a $2n \times 2n$ circulant matrix. Therefore, the number of operations per iteration of the PCG method is of order $O(n \log n)$. Because of these important properties, circulant preconditioners have been proved to be successful choices under the assumption that the diagonals of T_n are the Fourier coefficients of a positive 2π-periodic continuous function; see, for instance, [8].

Circulant preconditioners have been extended to block systems by a number of researchers. Chan and Olkin [10] and Holmgren and Otto [13] independently proposed using block-circulant matrices in solving noise reduction problems and hyperbolic differential equations, respectively. Block-circulant-circulant-block (BCCB) preconditioners for BTTB matrices have been studied and analyzed by Holmgren and Otto [13], Ku and Kuo [16], Tyrtyshnikov [29], and Chan and Jin [7]. Other preconditioners for Toeplitz systems include fast trigonometric transform–based preconditioners (see [5, 15, 21], for instance), Hartley transform–based preconditioners (see [4], for instance), and banded preconditioners (see [6, 9, 19, 20], for instance).

Circulant preconditioners are not well suited for linear systems with coefficient matrices $(I_m + T_m^* D_m T_m)$ or $(I_{mn} + T_{m,n}^* D_{mn} T_{m,n})$ because the inverse of $(I_m + c(T_m^*) D_m c(T_m))$ or the inverse of $(I_{mn} + c(T_{m,n}^*) D_{mn} c(T_{m,n}))$ cannot be obtained efficiently, where $c(X)$ is the circulant or BCCB approximation of X. To reduce the cost of the preconditioning step, one may use circulant matrices $(I_m + c(T_m^*) c(D_m) c(T_m))$ and BCCB matrices $(I_{mn} + c(T_{m,n}^*) c(D_{mn}) c(T_{m,n}))$ as preconditioners instead. However, the performance of these preconditioners is not good enough in general; see the numerical results in section 4.

Recently, approximate sparse inverse preconditioners have been studied by a number of researchers; see Kolotilina and Yeremin [17], Benzi, Meyer, and Tuma [3], Tang [25], Chow [11], and Nikishin and Yeremin [18]. The most important property in factorization sparse inverse preconditioners (FSIP) is that both the construction of FSIP and the preconditioning steps possess natural parallelism. Such preconditioning technique is well suited for modern massively parallel computers. In this paper, we study factorized banded inverse preconditioners (FBIP) for the linear systems with Toeplitz structure, as mentioned above. We show that if a Toeplitz matrix T has a certain off-diagonal decay property, then the factorized banded inverse preconditioners can approximate T^{-1} accurately. In particular, the spectra of these preconditioned matrices are clustered around 1, and therefore the PCG method when applied to solve these preconditioned systems converges very quickly.

In nonlinear image restoration applications, Toeplitz-related systems of the form $(I + T^* DT)$ are required to solve, where D is a positive nonconstant diagonal matrix. We discuss the construction of preconditioners for such matrices. Numerical results show that the performance of our proposed preconditioners are superior to that of circulant preconditioners. A two-dimensional nonlinear image restoration example is also presented to demonstrate the effectiveness of the proposed preconditioners.

The outline of the paper is as follows. In section 2, we briefly introduce the idea of the construction of factorized sparse inverse preconditioners. In section 3, we discuss the sparse pattern of Toeplitz and Toeplitz-related matrix. Then we construct factorized banded inverse preconditioners for these matrices. In section 4, we give numerical examples to demonstrate the effectiveness of our proposed preconditioners. In section 5, the application of the FBIP to nonlinear image restorations is presented in section 5. Finally, concluding remarks are given in section 6.

2. Factorized sparse inverse preconditioners.

In this section, we introduce the general idea of the construction of factorized preconditioners for the linear systems (1.2) where the coefficient matrix A is symmetric positive definite (SPD). The idea is to obtain a sparse lower triangular matrix L such that $A^{-1} \approx L^T L$. Let $\mathcal{S}$ be the given sparse pattern such that

$$\{(i,i) \mid i = 1, \ldots, n\} \subset \mathcal{S}$$

and L be such that

$$L(i,j) = 0 \quad \text{if } (i,j) \notin \mathcal{S}.$$

We briefly introduce the method proposed by Kolotilina and Yeremin [17]. Let $A = GG^T$ be the Cholesky factorization of A. We look for L with sparse pattern $\mathcal{S}$ such that the Frobenius norm

$$\|I - GL\|_F$$

is minimized. It has been shown that L can be obtained by the following algorithm [17].

Algorithm: Construction of FSIP.

Step 1. *Compute $\hat{L}$ with sparse pattern $\mathcal{S}$ such that*

$$[\hat{L}A]_{ij} = \delta_{i,j}, \quad (i,j) \in \mathcal{S}. \tag{2.1}$$

Step 2. *Let $\hat{D} = (\mathrm{diag}(\hat{L}))^{-1}$ and $L = \hat{D}^{\frac{1}{2}}\hat{L}$.*

For factorized sparse inverse preconditioners, we emphasize the following properties:

(i) The computation of L can be done in parallel between rows:

$$\hat{L}(i, \mathcal{S}_i)A(\mathcal{S}_i, \mathcal{S}_i) = [0, \ldots, 0, 1],$$

where $\mathcal{S}_i = \{j \mid (i,j) \in \mathcal{S}\}$ and $A(I,J)$ denotes the submatrix of A containing the rows with index set I and columns with index set J. The preconditioning step has a natural parallelism since the multiplication of $L^T L$ to a residual vector r can be computed in parallel.

(ii) Given a sparse pattern $\mathcal{S}$, the lower triangular matrix L is well defined for SPD matrix A since all principal submatrices of A are also SPD matrices.

(iii) All diagonal entries of LAL^T are one, i.e., $\mathrm{diag}(LAL^T) = I$.

(iv) The matrix L minimizes the functional $\frac{1}{n}\text{trace}(\tilde{L}A\tilde{L}^T)/\det(\tilde{L}A\tilde{L}^T)^{\frac{1}{n}}$ for all $\tilde{L}$ with sparse pattern $\mathcal{S}$. We note that the quantity $\frac{1}{n}\text{trace}(B)/\det(B)^{\frac{1}{n}}$ is a (nonstandard) condition number of B; see, for instance, [2, p. 341]. Thus when the PCG method is applied to solve the preconditioned system $L^TLAx = L^TLb$, the convergence can be much faster than the CG method for the original system without preconditioning.

(v) If $\mathcal{S} = \{(i,j) \mid 1 \le j \le i \le n\}$, then $A^{-1} = L^TL$, i.e., $L = G^{-1}$.

3. Factorized banded inverse preconditioners.

We are interested in solving Toeplitz-related systems $Ax = b$. In general, the coefficient matrix is dense and we have to choose the pattern that is sparser than the original matrix. In many applications, the matrices A have off-diagonal decay property, i.e., $|a_{i,j}|$ decays as $|i-j|$ increases. We find that the inverse of A has also off-diagonal decay property. Therefore it is natural to approximate G^{-1} by some banded lower triangular matrices. As an example, we depict the magnitude of the entries of a 64×64 Toeplitz matrix with diagonals given by $1/(1+j)^{1.1} (j = 0, 1, \dots, 63)$, its inverse, and the lower triangular factor of its inverse in Figure 1.

3.1. One-dimensional case.

We set the sparse pattern to be banded, i.e.,

$$\mathcal{S} = \{(i,j) \mid \max(1, i-k+1) \le j \le i \le n\},$$

where k is the bandwidth of the factor L. Using (2.1), the ith row of $\hat{L}$ can be obtained by solving the linear system

$$\hat{L}(i, i' : i)A(i' : i, i' : i) = [0, \dots, 0, 1], \quad i = 1, \dots, n, \tag{3.1}$$

where $i' = \max(1, i-k+1)$. We note that for each i, the above linear system can be solved in $O(k^3)$ operations, and the cost to obtain the factor L is $O(nk^3)$ operations.

In the following, we focus our discussion on the approximate factorization of Toeplitz matrices. From (3.1) we see that for $i \ge k$ we have

$$\hat{L}(i, i-k+1 : i)A(i-k+1 : i, i-k+1 : i) = [0, \dots, 0, 1], \quad i = k, \dots, n.$$

When $A = T_n$ is an $n \times n$ Toeplitz matrix, we have

$$\hat{L}(i, i-k+1 : i)T_n(1 : k, 1 : k) = [0, \dots, 0, 1], \quad i = k, \dots, n.$$

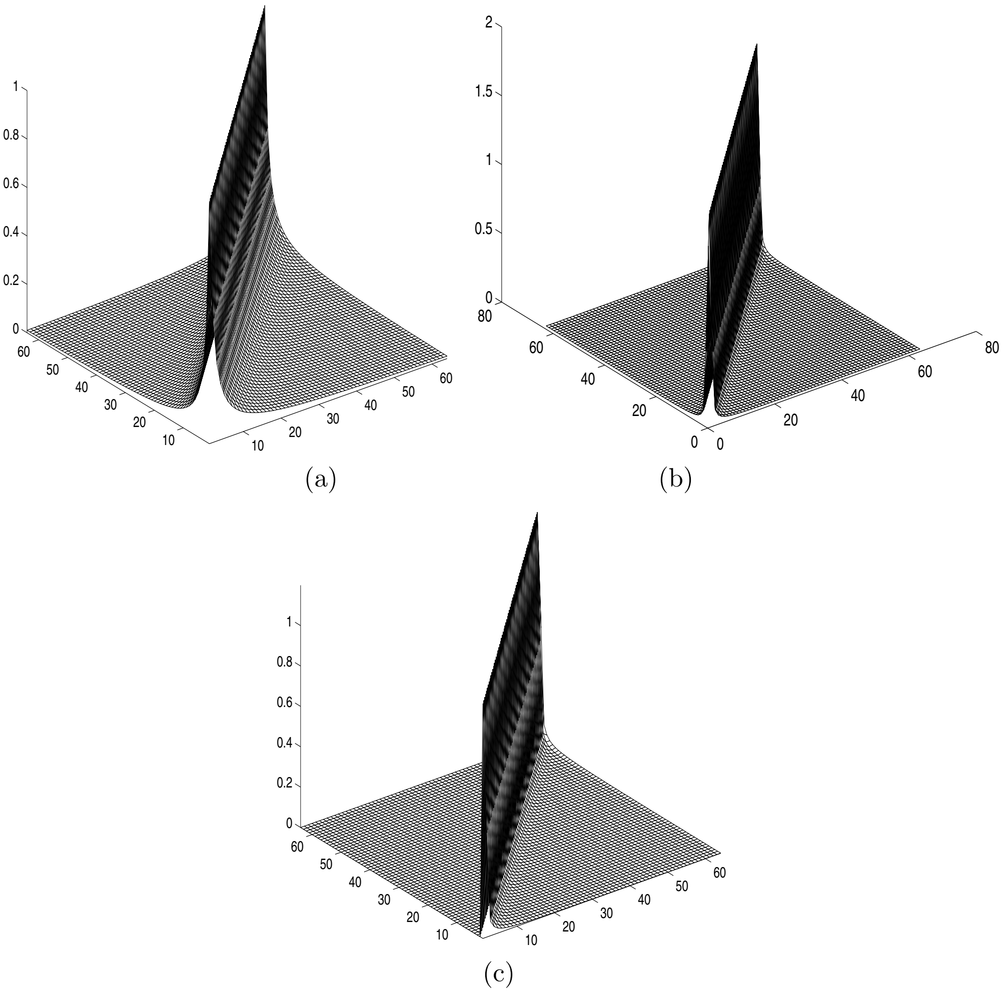
Using displacement structure of T_n, we have

$$\hat{L}(k+j, j+1 : k+j) = \hat{L}(k, 1 : k), \quad j = 1, \dots, n-k,$$

and therefore

$$\hat{L} = \begin{bmatrix} \hat{L}(k,k) & & & & 0 \\ \hat{L}(k,k-1) & \hat{L}(k,k) & & & \\ \vdots & \ddots & & & \\ \hat{L}(k,1) & \dots & \hat{L}(k,k) & & \\ & \ddots & & \ddots & \\ 0 & & \hat{L}(k,1) & \dots & \hat{L}(k,k) \end{bmatrix}$$

$$+ \begin{bmatrix} B_{(k-1)\times(k-1)} & O_{(k-1)\times(n-k+1)} \\ O_{(n-k+1)\times(k-1)} & O_{(n-k+1)\times(n-k+1)} \end{bmatrix},$$



(a)

(b)

(c)

FIG. 1. *The magnitude of the entries of a Toeplitz matrix with diagonal given by* $1/(1+j)^{1.1}$ (a)*, its inverse* (b)*, and the lower triangular factor of its inverse* (c)*.*

where $O_{m\times n}$ is the $m \times n$ zero matrix and $B_{(k-1)\times(k-1)}$ is a lower triangular matrix with (i,j) entry given by

$$B_{(k-1)\times(k-1)}(i,j) = \hat{L}(i,j) - \hat{L}(k,k-i+j), \quad 1 \le j \le i \le k-1.$$

Using the recursive inversion algorithm proposed by Trench [26], the inverse of a Toeplitz matrix of size $k \times k$ can be obtained in $O(k^2)$ operations. We note that the algorithm also computes the inverses of all principal submatrices. Thus, the construction cost of the FBIP for Toeplitz matrices can be reduced from $O(nk^3)$ to $O(k^2)$.

In the rest of this subsection, we discuss the spectra of the preconditioned matrices L^TLT_n and $L^TL(I_n+T_n^*D_nT_n)$, where T_n and D_n are Toeplitz and diagonal matrices, respectively. We first consider the following off-diagonal decay property.

DEFINITION 3.1 (see Strohmer [24]). *Let $A = [a_{i,j}]_{i,j\in\mathcal{I}}$ be a matrix, where the index set is $\mathcal{I} = \mathbb{Z}$, $\mathbb{N}$, or $\{1, 2, \dots, N\}$.*

1. *A belongs to the space $\mathcal{E}_\gamma$ if*

$$|a_{i,j}| \leq ce^{-\gamma|i-j|} \quad \textit{for } \gamma > 0$$

and some constant $c > 0$.

2. *A belongs to the space $\mathcal{Q}_s$ if*

$$|a_{i,j}| \leq c(1 + |i - j|)^{-s} \quad \textit{for } s > 1,$$

and some constant $c > 0$.

With these definitions, we have the following results about the off-diagonal decay of the entries of A^{-1}.

THEOREM 3.2 (see Jaffard [14]). *Let $A : l^2(\mathcal{I}) \to l^2(\mathcal{I})$ be an invertible matrix, where $\mathcal{I} = \mathbb{Z}, \mathbb{N}$ or $\{1, 2, \dots, N\}$.*

1. *If $A \in \mathcal{E}_\gamma$, then $A^{-1} \in \mathcal{E}_{\gamma_1}$ for some $\gamma_1 \in (0, \gamma)$.*
2. *If $A \in \mathcal{Q}_s$, then $A^{-1} \in \mathcal{Q}_s$.*

Here we remark that if A is Hermitian and positive definite, then there exists a constant $c > 0$ such that for all $n > 0$, the entries of the inverse of the submatrix $A_n = A(1 : n, 1 : n)$ satisfy

$$A_n^{-1}(i, j) \leq ce^{-\gamma_1|i-j|} \quad \text{for } 1 \leq i,\ j \leq n, \tag{3.2}$$

or

$$A_n^{-1}(i, j) \leq c(1 + |i - j|)^{-s} \quad \text{for } 1 \leq i,\ j \leq n. \tag{3.3}$$

By using the above results, we have the following main theorem of this paper.

THEOREM 3.3. *Let T_n be a Hermitian Toeplitz matrix with its diagonal entries satisfying*

$$|t_j| \leq ce^{-\gamma|j|} \tag{3.4}$$

for some $c > 0$ and $\gamma > 0$ or

$$|t_j| \leq c(|j| + 1)^{-s} \tag{3.5}$$

for some $c > 0$ and $s > 3/2$. Then for any given $\epsilon > 0$, there exists $K > 0$ such that for all $k > K$,

$$\|\hat{L} - \hat{G}\|_2 \leq \epsilon,$$

where $\hat{L}$ and $\hat{G}$ be the solution of (2.1) *with sparse patterns*

$$\mathcal{S} = \{(i, j) \mid \max(1, i - k + 1) \leq j \leq i \leq n\}$$

for the banded case and

$$\{(i, j) \mid 1 \leq j \leq i \leq n\}$$

for the complete factorization, respectively.

Proof. We prove the result by using the inequality

$$\|\hat{L}-\hat{G}\|_2 \le \sqrt{\|\hat{L}-\hat{G}\|_1\|\hat{L}-\hat{G}\|_\infty}.$$

We first consider the case when (3.4) is satisfied. Now we estimate $\|\hat{L}-\hat{G}\|_\infty$. For $j=1,\ldots,k$

$$\|\hat{L}(j,:)-\hat{G}(j,:)\|_\infty = 0 \tag{3.6}$$

and for $j=k+1,\ldots,n$

$$\|\hat{L}(j,:)-\hat{G}(j,:)\|_\infty = \sum_{i=1}^{j-k}|\hat{G}(j,i)| + \sum_{i=j-k+1}^{j}|\hat{L}(j,i)-\hat{G}(j,i)|. \tag{3.7}$$

Note that for $j=k+1,\ldots,n$, we have

$$\sum_{i=1}^{j-k}|\hat{G}(j,i)| = \sum_{i=1}^{j-k}\left|T_j^{-1}(j,i)\right| \le c\sum_{i=1}^{j-k}e^{-\gamma_1(j-i)} \le c_1e^{-\gamma_1 k}, \tag{3.8}$$

where $c_1 = ce^{\gamma_1}/(e^{\gamma_1}-1)$. Let T_j be partitioned as

$$T_j = \begin{bmatrix} T_{j-k} & U_{(j-k)\times k} \\ U^T_{(j-k)\times k} & T_k \end{bmatrix};$$

we have

$$\hat{L}(j,j-k+1:j)T_k = [0,\ldots,0,1]$$

and

$$\hat{G}(j,1:j-k)U_{(j-k)\times k} + \hat{G}(j,j-k+1:j)T_k = [0,\ldots,0,1].$$

Therefore

$$(\hat{L}(j,j-k+1:j)-\hat{G}(j,j-k+1:j))T_k = -\hat{G}(j,1:j-k)U_{(j-k)\times k}.$$

It follows that

$$\sum_{i=j-k+1}^{j}|\hat{L}(j,i)-\hat{G}(j,i)| \le \|\hat{G}(j,1:j-k)\|_\infty\|U_{(j-k)\times k}T_k^{-1}\|_\infty \le c_2e^{-\gamma_1 k} \tag{3.9}$$

for some constant $c_2>0$. Substituting (3.8) and (3.9) into (3.7) we get

$$\|\hat{L}(j,:)-\hat{G}(j,:)\|_\infty \le (c_1+c_2)e^{-\gamma_1 k} = c_3e^{-\gamma_1 k}, \quad j=k+1,\ldots,n. \tag{3.10}$$

Combining (3.6) with (3.10) we get

$$\|\hat{L}-\hat{G}\|_\infty \le c_3e^{-\gamma_1 k}.$$

Similar to (3.8), we have

$$\sum_{i=j+k}^{n}|\hat{G}(i,j)| \le c_1e^{-\gamma_1 k}, \quad j=1,\ldots,n-k. \tag{3.11}$$

Using (3.11) and (3.9), we have that for $j = 1, \ldots, n$,

$$\|\hat{L}(:,j) - \hat{G}(:,j)\|_1 = \sum_{i=j}^{\min(j+k-1,n)} |\hat{L}(i,j) - \hat{G}(i,j)| + \sum_{i=\min(j+k-1,n)+1}^{n} |\hat{G}(i,j)|$$
$$\leq kc_2 e^{-\gamma_1 k} + c_1 e^{-\gamma_1 k} = (c_2 k + c_1) e^{-\gamma_1 k}.$$

Thus, $\|\hat{L} - \hat{G}\|_1 \leq (c_2 k + c_1) e^{-\gamma_1 k}$. It follows that

$$\|\hat{L} - \hat{G}\|_2 \leq \sqrt{\|\hat{L} - \hat{G}\|_1 \|\hat{L} - \hat{G}\|_\infty} \leq \sqrt{c_3(c_2 k + c_1)} e^{-\gamma_1 k}.$$

We now prove the result for the case when (3.5) is satisfied.

Similarly, there exists a constant $c_1 > 0$ such that

$$\sum_{i=1}^{j-k} |\hat{G}(j,i)| \leq c_1 (1+k)^{-s+1} \quad \text{and} \quad \sum_{i=j+k}^{n} |\hat{G}(i,j)| \leq c_1 (1+k)^{-s+1}.$$

Furthermore, there exists a constant $c_2 > 0$ such that

$$\|\hat{L} - \hat{G}\|_\infty \leq (c_1 + c_2)(1+k)^{-s+1}$$

and

$$\|\hat{L} - \hat{G}\|_1 \leq (c_1 + kc_2)(1+k)^{-s+1} \leq (c_1 + c_2)(1+k)^{-s+2}.$$

It follows that

$$\|\hat{L} - \hat{G}\|_2 \leq (c_1 + c_2)(1+k)^{-s+3/2}.$$

Hence the theorem is proved. □

COROLLARY 3.4. *Let T_n be a Toeplitz matrix with its diagonal entries satisfying* (3.4) *or* (3.5). *Let $D = \text{diag}(d_1, \ldots, d_n)$ with $0 < d_i \leq d$ for $i = 1, 2, \ldots, n$, where d is a constant. Let*

$$A_n = I_n + T_n^* D_n T_n.$$

Then for any given $\epsilon > 0$, there exists $K > 0$ such that for all $k > K$,

$$\|\hat{L} - \hat{G}\|_2 \leq \epsilon,$$

where $\hat{L}$ and $\hat{G}$ be the solution of (2.1) *with $A = A_n$ and sparse patterns*

$$\mathcal{S} = \{(i,j) \mid \max(1, i-k+1) \leq j \leq i \leq n\}$$

for the banded case and

$$\{(i,j) \mid 1 \leq j \leq i \leq n\}$$

for the complete factorization respectively.

Proof. We first prove that if (3.5) holds, then the entries $A_n = (I_n + T_n^* D_n T_n)$ have off-diagonal decay property. For the sake of simplicity, all constants are denoted by c in the proof. Suppose that $l = j - i \geq 3$; then the (i,j)-entry of $T_n^* D_n T_n$ which is given by $\sum_{k=1}^{n} t_{k-i} d_k t_{k-j}$ satisfies

$$
\begin{aligned}
&\left|\sum_{k=1}^{n} t_{k-i} d_k t_{k-j}\right| \\
&\quad \leq \sum_{k=1}^{n} \frac{c}{(|k-i|+1)^s(|k-j|+1)^s} \\
&\quad = \sum_{k=1}^{i+1} \frac{c}{(|k-i|+1)^s(|k-j|+1)^s} + \sum_{k=j-1}^{n} \frac{c}{(|k-i|+1)^s(|k-j|+1)^s} \\
&\qquad + \sum_{k=i+2}^{j-2} \frac{c}{(|k-i|+1)^s(|k-j|+1)^s} \\
&\quad \leq \sum_{k=1}^{i+1} \frac{c}{(l^s|k-i|+1)^s} + \sum_{k=j-1}^{n} \frac{c}{l^s(|k-j|+1)^s} \\
&\qquad + \sum_{k=i+2}^{j-2} \frac{c}{(|k-i|+1)^s(|k-j|+1)^s} \\
&\quad \leq \frac{c}{l^s} + \sum_{k'=2}^{l-2} \frac{c}{(k'+1)^s(l-k'+1)^s} \\
&\quad \leq \frac{c}{l^s} + \sum_{k'=2}^{l-2} \int_0^1 \frac{c}{(k'+x)^s(l-k'-x+1)^s} dx \\
&\quad = \frac{c}{l^s} + \sum_{k'=2}^{l-2} \int_{k'}^{k'+1} \frac{c}{x^s(l-x+1)^s} dx \\
&\quad = \frac{c}{l^s} + \int_2^{l-1} \frac{c}{x^s(l-x+1)^s} dx \\
&\quad = \frac{c}{l^s} + \frac{c}{(l+1)^s} \int_2^{l-1} \left(\frac{1}{x} + \frac{1}{l-x+1}\right)^s dx.
\end{aligned}
$$

Obviously, for $x \in [2, l-1]$, $\frac{1}{x} + \frac{1}{l-x+1} \leq 1$. Note that $s > 3/2$, we have

$$
\int_2^{l-1} \left(\frac{1}{x} + \frac{1}{l-x+1}\right)^s dx \leq \int_2^{l-1} \left(\frac{1}{x} + \frac{1}{l-x+1}\right)^{3/2} dx = 2\sqrt{2}\left(\frac{l+1}{l-1}\right)^{1/2} \frac{l-3}{l+1} < 4.
$$

It follows that the entries of $T_n^* D_n T_n$ also satisfy (3.5):

$$
|[T_n^* D_n T_n]_{i,j}| = O(|i-j|^{-s}) = O((|i-j|+1)^{-s}).
$$

Similarly, if the entries of T_n satisfy (3.4), then

$$
|[T_n^* D_n T_n]_{i,j}| \leq c e^{-\lambda|i-j|} + c|i-j| e^{-\lambda|i-j|} \leq c e^{-\lambda_1|i-j|},
$$

where $0 < \lambda_1 < \lambda$. The rest of the proof is similar to the proof of Theorem 3.3 and therefore we omit it. □

From Theorem 3.3 and Corollary 3.4 we see that if the coefficients of a Hermitian positive definite Toeplitz matrix satisfy (3.4) or (3.5), then for any given $\epsilon > 0$, there exists a K such that for all $k > K$, all eigenvalues of the preconditioned matrix $L^T LT_n$ ($L^T L(I_n + T_n^* D_n T_n)$ lie in the interval $[1-\epsilon, 1+\epsilon]$. It follows that the PCG method, when applied to solving the preconditioned systems $L^T LT_n x = L^T Lb$ and $L^T L(I_n + T_n^* D_n T_n)x = L^T Lb$, converges very quickly; see the numerical results in section 4.

3.2. Two-dimensional case. In this subsection, we consider the sparse pattern for block matrices $A_{m,n}$. Let

$$A_{m,n} = \left[\begin{array}{cccc} A^{(1,1)} & A^{(1,2)} & \dots & A^{(1,m)} \\ A^{(2,1)} & A^{(2,2)} & \dots & A^{(2,m)} \\ \vdots & \vdots & \dots & \vdots \\ A^{(m,1)} & A^{(m,2)} & \dots & A^{(m,m)} \end{array} \right],$$

where the $n \times n$ blocks $A^{(i,j)}$ are defined by

$$[A^{(i,j)}]_{k,l} = a_{k,l}^{(i,j)}, \quad k,l = 1,2,\dots,n.$$

We assume that the matrix $A_{m,n}$ has off-diagonal decay property, i.e., $\|A^{(i,j)}\|_F$ decays as $|i-j|$ increases, where $\|\cdot\|_F$ denotes the Frobenius norm and each block $A^{(i,j)}$ has also the off-diagonal decay property. In this case, we set the factor of the factorized inverse preconditioner to be triangular banded block matrix with each block being a banded matrix. For example,

$$(3.12) \quad L = \left[\begin{array}{cccc|cccc|cccc|cccc}
+ & & & & & & & & & & & & & & & \\
+ & + & & & & & & & & & & & & & & \\
& + & + & & & & & & & & & & & & & \\
& & + & + & & & & & & & & & & & & \\
\hline
+ & + & & & + & & & & & & & & & & & \\
+ & + & + & & + & + & & & & & & & & & & \\
& + & + & + & & + & + & & & & & & & & & \\
& & + & + & & & + & + & & & & & & & & \\
\hline
& & & & + & + & & & + & & & & & & & \\
& & & & + & + & + & & + & + & & & & & & \\
& & & & & + & + & + & & + & + & & & & & \\
& & & & & & + & + & & & + & + & & & & \\
\hline
& & & & & & & & + & + & & & + & & & \\
& & & & & & & & + & + & + & & + & + & & \\
& & & & & & & & & + & + & + & & + & + & \\
& & & & & & & & & & + & + & & & + & +
\end{array} \right].$$

Let p and q be the half bandwidth of each block and the bandwidth of the banded block matrix respectively (for the matrix given by (3.12), $m = n = 4$ and $p = q = 2$). The sparse pattern is set to

$$\mathcal{S} = \cup_{i=1}^{mn} \mathcal{S}_i,$$

where $\mathcal{S}_i$ is the index set corresponding to the ith row. Let

$$i = i_0 n + i_1 \quad \text{with } 0 \le i_0 \le m-1 \text{ and } 1 \le i_1 \le n.$$

Let

$$i_0' = \min(i_0, q-1), \quad i_1' = \max(i_1 - p + 1, 1), \quad \text{and} \quad i_1'' = \min(i_1 + p - 1, n),$$

we define

$$\mathcal{S}_i = \{(i,j) \mid j = i_0 n + i_1' : i\} \cup \left(\cup_{i'=1}^{i_0'} \{(i,j) \mid j = (i_0 - i')n + i_1' : (i_0 - i')n + i_1''\} \right).$$

It is easy to see that the number of nonzero entries in each row of L is not greater than

$$p + (q-1)(2p-1) = 2pq - p - q + 1.$$

Thus, the total cost of computing the factor L is bounded by $O(mn(2pq-p-q+1)^3) = O(mnp^3q^3)$ and the storage requirement is $O(mnpq)$.

Similar to the one-dimensional case, if A is a BTTB matrix $T_{m,n}$, then the factor L is a near BTTB lower triangular matrix and the cost of constructing L only depends on p, q (independent of m and n). More precisely, we have

$$L = \begin{bmatrix} L^{(1,1)} & & & & & 0 \\ L^{(2,1)} & L^{(2,2)} & & & & \\ \cdots & \cdots & \ddots & & & \\ L^{(q,1)} & L^{(q,2)} & \cdots & L^{(q,q)} & & \\ & \ddots & \ddots & \ddots & \ddots & \\ 0 & & L^{(q,1)} & L^{(q,2)} & \cdots & L^{(q,q)} \end{bmatrix},$$

where the blocks $L^{(i,j)}$ are near-Toeplitz matrices ($n' = n-p+2$ and $n'' = n'-p+1$):

$$L^{(i,i)} = \begin{bmatrix} l_{1,1}^{(i,i)} & & & & & & & & 0 \\ l_{2,1}^{(i,i)} & l_{2,2}^{(i,i)} & & & & & & & \\ \cdots & \cdots & \ddots & & & & & & \\ l_{p,1}^{(i,i)} & l_{p,2}^{(i,i)} & \cdots & l_{p,p}^{(i,i)} & & & & & \\ & \ddots & \ddots & \ddots & \ddots & & & & \\ & & l_{p,1}^{(i,i)} & l_{p,2}^{(i,i)} & \cdots & l_{p,p}^{(i,i)} & & & \\ & & & l_{n',n''}^{(i,i)} & l_{n',n''+1}^{(i,i)} & \cdots & l_{n',n'}^{(i,i)} & & \\ & & & & \ddots & \ddots & \ddots & \ddots & \\ 0 & & & & & l_{n,n'-1}^{(i,i)} & l_{n,n'}^{(i,i)} & \cdots & l_{n,n}^{(i,i)} \end{bmatrix}$$

TABLE 1
Construction cost and storage requirement of factorized banded inverse preconditioners.

	1D Case (bandwidth k)		2D Case (bandwidth p, q)	
	Toeplitz	General	BTTB	General
Cost	$O(k^2)$	$O(nk^3)$	$O(p^4q^4)$	$O(mnp^3q^3)$
Storage	$O(k^2)$	$O(nk)$	$O(p^2q^2)$	$O(mnpq)$

and for $i > j$

$$L^{(i,j)} = \begin{bmatrix} l_{1,1}^{(i,j)} & l_{1,2}^{(i,j)} & \cdots & l_{1,p}^{(i,j)} & & & & & 0 \\ l_{2,1}^{(i,j)} & l_{2,2}^{(i,j)} & \cdots & l_{2,p}^{(i,j)} & l_{2,p+1}^{(i,j)} & & & & \\ \vdots & \vdots & & \cdots & \cdots & \ddots & & & \\ l_{p,1}^{(i,j)} & l_{p,2}^{(i,j)} & \cdots & l_{p,p}^{(i,j)} & l_{p,p+1}^{(i,j)} & \cdots & l_{p,2p-1}^{(i,j)} & & \\ & \ddots & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & l_{p,1}^{(i,j)} & l_{p,2}^{(i,j)} & \cdots & l_{p,p}^{(i,j)} & l_{p,p+1}^{(i,j)} & \cdots & l_{p,2p-1}^{(i,j)} \\ & & & l_{n',n''}^{(i,j)} & l_{n',n''+1}^{(i,j)} & \cdots & l_{n',n'}^{(i,j)} & \cdots & l_{n',n}^{(i,j)} \\ & & & & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & & & & & l_{n,n'-1}^{(i,j)} & l_{n,n'}^{(i,j)} & \cdots & l_{n,n}^{(i,j)} \end{bmatrix}.$$

Thus, we need to solve $(2pq - p - q + 1)$ linear systems of size less than or equal to $(2pq-p-q+1)$. Therefore the total cost is $O((2pq-p-q+1)^4) = O(p^4q^4)$. Moreover, the storage requirement is $O(p^2q^2)$.

We end this section by giving a summary of the construction cost and the storage requirement in Table 1.

4. Numerical results.

In this section, numerical examples are given to illustrate the efficiency of PCG methods with FBIP. We compare the performance of the FBIP with that of Chan's circulant preconditioners. We test linear systems $Ax = b$ for A to be (i) a Toeplitz matrix T_m, (ii) a Toeplitz-related matrix $(I_m + T_m^* D_m T_m)$, (iii) a BTTB matrix $T_{m,n}$, and (iv) a BTTB-related matrix $(I_{mn} + T_{m,n}^* D_{mn} T_{m,n})$. In our examples, the right-hand sides are random vectors. In all PCG methods, we use the zero vector as an initial guess, and the stopping criterion is

$$\|r^{(j)}\|_2 / \|r^{(0)}\|_2 \le 10^{-7},$$

where $r^{(j)}$ is the residual vector of the jth iteration. In Tables 2–5, the symbol I denotes the PCG method without a preconditioner, the symbol C denotes the method with circulant preconditioner or BCCB preconditioner proposed by Chan, and the symbols FI_k and $FI_{p,q}$ denote the FBIP introduced in section 3. The symbols ** denote that the number of iterations exceeds 1000. We note that Chan's circulant preconditioner can be defined for all square matrices. In the $(I + T^*DT)$ case, the circulant preconditioner is given by $(I + c(T^*)c(D)c(T))$. For Toeplitz systems and BTTB systems, the numbers of iterations for the PCG methods with different preconditioners are shown in Tables 2–5.

For Toeplitz-related systems $(I_n + T_n^* D_n T_n)x = b$, we generate the diagonal entries of the diagonal matrix D randomly using the formula $100 * (1 + 3 * \text{rand}(n, 1))^2$. (Here we use the Matlab representation.) We remark that it is too expensive to

TABLE 2

Number of iterations of the PCG methods for solving Toeplitz systems. The entries of the Toeplitz matrices are defined by $t_j = 1/(|j|+1)^{1.1}$ *(left),* $t_j = 1/(|j|+1)^{1.6}$ *(middle), and* $t_j = \exp(-0.5j^2)$ *(right), respectively.*

n	I	C	FI_{25}	I	C	FI_{25}	I	C	FI_{25}
64	20	6	5	17	6	4	55	8	2
128	24	7	5	18	6	4	65	7	2
256	28	7	6	19	6	5	66	7	2
512	30	7	6	19	6	5	66	6	2
1024	33	7	7	19	6	5	67	6	2
2048	35	7	7	19	6	5	67	6	2
4096	36	7	8	19	6	5	67	6	2

TABLE 3

Number of iterations of the PCG methods for BTTB systems. The entries of the BTTB matrices are given by $t_j^{(u-v)} = 1/((|u-v|+1)^{1.1}+(|j|+1)^{1.1})$ *(left) and* $t_j^{(u-v)} = \exp(-0.5((u-v)^2+j^2))$ *(right), respectively.*

n	I	C	$FI_{4,4}$	$FI_{5,5}$	$FI_{6,6}$	I	C	$FI_{4,4}$	$FI_{5,5}$	$FI_{6,6}$
16	52	16	8	8	7	236	31	16	11	8
32	89	19	9	9	8	471	28	20	13	9
64	131	21	12	11	10	532	25	21	14	10
128	215	25	17	14	13	556	23	21	14	10

compute the matrix $T_n^* D_n T_n$ explicitly. In our implementation, we set the bandwidth of the FBIP to be k, and we first approximate the matrix T_n by a banded Toeplitz matrix $\hat{T}_n$ with bandwidth $2(2k-1)-1 = 4k-3$ and then construct the FBIP for $(I_n + \hat{T}_n^* D_n \hat{T}_n)$. It is easy to check that by exploiting the banded structure of $\hat{T}_n$, the computation of $\hat{T}_n^* D_n \hat{T}_n$ can be done in $O(nk^2)$ operations (recall that the cost for constructing the FBIP is $O(nk^3)$). For the two-dimensional case, we first obtain an approximation of $T_{m,n}^* D_{mn} T_{m,n}$. Let p and q be the bandwidth parameters of the FBIP. Then we approximate $T_{m,n}$ by a banded-block-Toeplitz-banded-Toeplitz-block matrix $\hat{T}_{m,n}$ with half bandwidths $2p-1$ and $2q-1$, respectively, and then construct the FBIP for $(I_{mn} + \hat{T}_{m,n}^* D_{mn} \hat{T}_{m,n})$ by using the algorithm introduced in subsection 3.2. By exploiting the banded structure of $\hat{T}_{m,n}$, we can obtain $\hat{T}_{m,n}^* D_{mn} \hat{T}_{m,n}$ in $O(mnp^2q^2)$ operations. (Note that the cost for constructing the FBIP is $O(mnp^3q^3)$.) The numbers of iterations for different preconditioners are shown in Tables 4 and 5.

From Tables 2 through 5 we see that the FBIP are more efficient than circulant preconditioners, especially for Toeplitz-related and BTTB-related systems. Moreover, the multiplication of the FBIP to any vector requires $O(nk)$ and $O(nmpq)$ operations for one-dimensional and two-dimensional problems, respectively. (Recall that for circulant preconditioners, the costs are $O(n \log n)$ and $O(mn \log(mn))$, respectively.)

5. Application to nonlinear image restorations. In the literature of image restoration, a blurred image is often modeled as the linear convolution of an original image with the *point spread function* of the blur. However, in practice, image formation systems or image sensors usually incorporate a built-in nonlinearity. For instance, the nonlinearity is introduced in the transformation of light intensity to the output units of the imaging system such as current intensity in photoelectric systems and photographic films. The modeling of sensor nonlinearities was first studied by Andrews and Hunt [1]. In matrix-vector notation, the general space-invariant imaging

Table 4
Number of iterations of the PCG methods for solving Toeplitz-related systems. The entries of the Toeplitz matrices are defined by $t_j = 1/(|j|+1)^{1.1}$ *(left) and* $t_j = \exp(-0.5j^2)$ *(right), respectively.*

n	I	C	FI_{25}	I	C	FI_{25}
64	62	30	7	148	33	2
128	93	32	8	283	34	2
256	138	35	9	423	38	2
512	189	34	10	522	38	2
1024	238	35	11	568	39	2
2048	288	34	13	586	39	2
4096	336	35	15	608	42	2

Table 5
Number of iterations of the PCG methods for solving BTTB-related systems. The entries of the BTTB matrices are given by $t_j^{(u-v)} = 1/((|u-v|+1)^{1.1} + (|j|+1)^{1.1})$ *(left) and* $t_j^{(u-v)} = \exp(-0.5((u-v)^2+j^2))$ *(right), respectively.*

n	I	C	$FI_{4,4}$	$FI_{5,5}$	$FI_{6,6}$	I	C	$FI_{4,4}$	$FI_{5,5}$	$FI_{6,6}$
16	340	72	21	17	16	745	124	46	26	17
32	983	100	28	25	22	**	167	55	30	20
64	**	150	36	34	28	**	215	58	32	21
128	**	206	61	50	41	**	272	59	32	21

system with additive noise can be represented by the nonlinear equation

$$g = s(Af) + \eta, \tag{5.1}$$

where g, f, and η represent the observed image, the original image, and the noise vectors, respectively. Here $s(\cdot)$ denotes a point nonlinearity and the matrix A is a BTTB matrix. In the literature, different nonlinear image restoration algorithms have been proposed and analyzed. For instance, Andrews and Hunt [1] proposed using Taylor series expansion about the mean value of the observed image to approximate (5.1) to a linear equation. An approximate filter for linear image restoration can then be derived. Trussell and Hunt [28] applied the maximum a posteriori probability estimation scheme in nonlinear image restoration algorithms. This approach results in an iterative solution algorithm whose computational complexity is very large. Tekalp and Pavlović [27] also proposed to transform the noisy and blurred image into "the exposure domain" using the inverse of the nonlinear sensor characteristics. A linear minimum mean square error deconvolution filter was derived by using the linear convolutional model in the presence of multiplicative noise in the exposure domain. In this paper, we consider solving nonlinear least squares problems with regularization,

$$\min_f \|g - s(Af)\|_2^2 + \mu\|f\|_2^2, \tag{5.2}$$

to restore the original image. Here $\|\cdot\|_2$ denotes the usual Euclidean norm and μ is a small positive number controlling the degree of regularity of the solution.

In [30], Zervakis and Venetsanopoulos considered using the Gauss–Newton method for the nonlinear least squares problem (5.2). Given an initial guess $f^{(0)}$, for $j = 0, 1, \ldots$, we solve the linear least squares problem

$$f^{(j+1)} = \arg\min_f \{\|g - s(A\tilde{f}) - D_s^{(j)} A(f - f^{(j)})\|_2^2 + \mu\|f\|_2^2\} \tag{5.3}$$

until $\|f^{(j+1)} - f(\mu)\|_2$ is small enough, where $f(\mu)$ is the solution of (5.2) with regularization parameter μ. Here $D_s^{(j)}$ is a diagonal matrix with diagonal entries

$$[D_s^{(j)}]_{kk} = \frac{\partial s}{\partial x}\Big|_{x=\sum_i A_{ki} f_i^{(j)}}.$$

We remark that under the practical assumption on the nonlinear function $s(\cdot)$, the diagonal entries of $D_s^{(j)}$ are always positive values; see Andrews and Hunt [1]. The least squares problem (5.3) is equivalent to

$$[\mu I + A^*(D_s^{(j)})^2 A](f - f^{(j)}) = A^* D_s^{(j)}[g - s(Af^{(j)})] - \mu f,$$

i.e.,

$$[\mu I + A^*(D_s^{(j)})^2 A]f = A^* D_s^{(j)}[g - s(Af^{(j)}) + D_s^{(j)} A f^{(j)}]. \tag{5.4}$$

In the Gauss–Newton method, it is important to choose a good initial guess. We propose the following algorithm to compute $f^{(0)}$:

(1) Solve $g = s(\hat{g})$ (in many practical applications, s^{-1} can be easily obtained).

(2) Choose suitable parameter μ_0 and solve

$$(\mu_0 I + A^* A)f^{(0)} = A^* \hat{g}. \tag{5.5}$$

Our numerical tests show that our initial guess $f^{(0)}$ is quite close to the real solution of the nonlinear least squares problems (5.2) and therefore the Gauss–Newton method converges very fast. (In fact, only one iteration is required in solving the examples tested in the next subsection.) We see that in computing the initial guess and in each Gauss–Newton iteration, we require to solve a BTTB-related systems with diagonal matrix $D = I$ (cf. (5.5)) and $D = (D_s^{(j)})^2$ (cf. (5.4)), respectively. In the PCG methods for solving the system (5.5), we use the zero vector as the initial guess, and the stopping criteria is $\|r^{(i)}\|_2/\|A^*\hat{g}\|_2 < 10^{-7}$, where $r^{(i)}$ is the residual after i iterations. In the PCG methods for solving the system (5.4), we use $f^{(j)}$ as initial guess and the stopping criteria is

$$\frac{\|r^{(i)}\|_2}{\|A^* D_s^{(j)}[g - s(Af^{(j)}) + D_s^{(j)} A f^{(j)}]\|_2} < 10^{-7},$$

where $r^{(i)}$ is the residual after the ith iterations.

5.1. An example. We test two 128 × 128 images: Bridge (Figure 2(a)) and Cameraman (Figure 2(b)). The pointwise nonlinearity employed is of the logarithmic form

$$s(x) = 30\log(x)$$

(tested in [30]), and the blurring function is given by

$$a(x, y) = \exp[-0.5 * (x^2 + y^2)].$$

We construct the observed image by forming the vector $g = s(Af) + \eta$, where f is a vector formed by row ordering the original image. In our tests, the noise η is set to Gaussian white noise with noise-to-signal ratio of 50 dB, 40 dB, 30 dB, and 20 dB, respectively. Observed images for noise-to-signal ratio of 40 dB are shown in

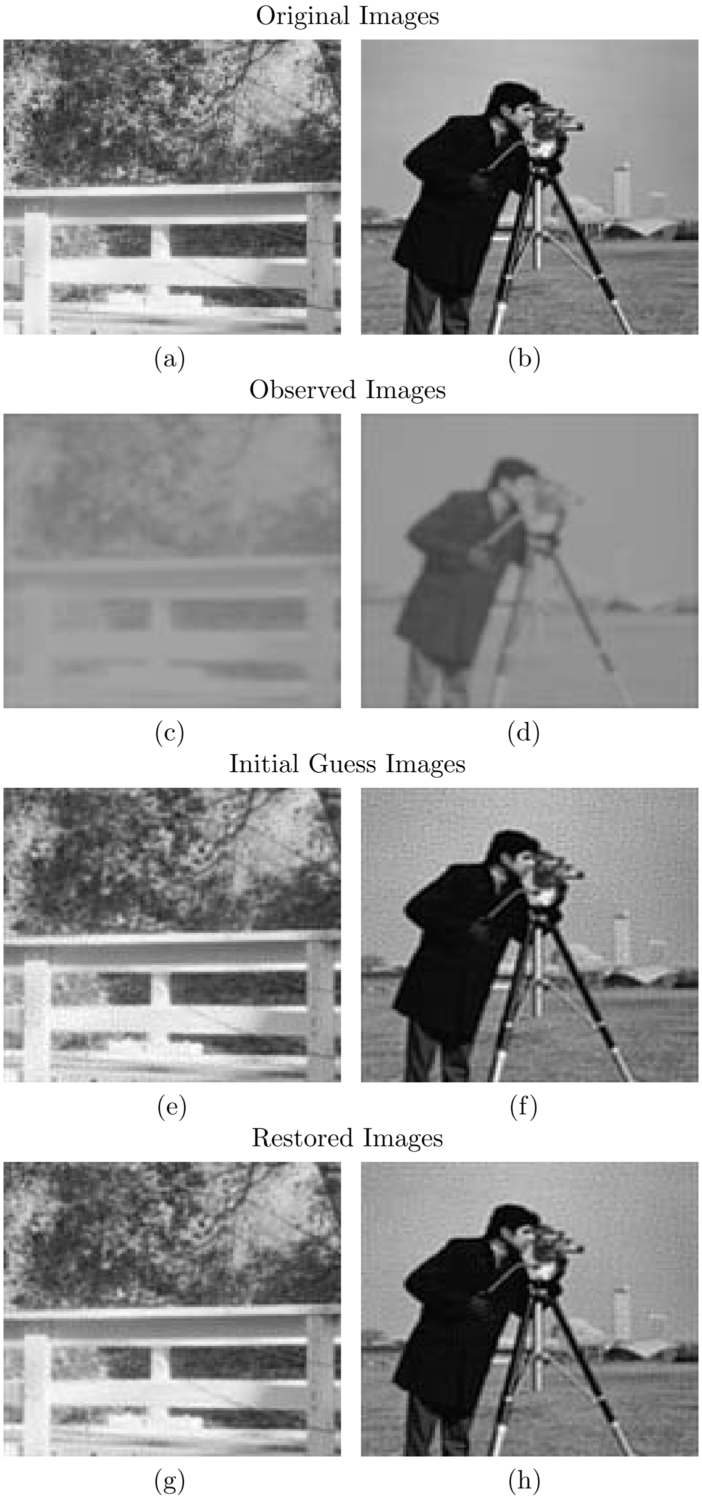


Fig. 2. *Original, observed, initial guess, and restored images of Bridge (left) and Cameraman (right).*

TABLE 6
Number of iterations for solving (5.5).

	Bridge				Cameraman			
	I	C	$FI_{4,4}$	$FI_{5,5}$	I	C	$FI_{4,4}$	$FI_{5,5}$
50 dB	242	79	25	14	120	49	13	8
40 dB	106	46	12	7	75	33	8	5
30 dB	56	27	6	5	43	21	6	5
20 dB	36	19	6	5	28	15	6	4

TABLE 7
Number of iterations for solving (5.4).

	Bridge				Cameraman			
	I	C	$FI_{4,4}$	$FI_{5,5}$	I	C	$FI_{4,4}$	$FI_{5,5}$
50 dB	389	93	17	11	1081	133	30	17
40 dB	199	72	13	8	666	128	27	15
30 dB	111	53	8	5	401	126	23	13
20 dB	82	46	6	4	293	130	20	12

Figure 2(c) and Figure 2(d), respectively. In Figure 2(e) and Figure 2(f), we present our initial guesses for the restored images, i.e., the solutions of (5.5). The optimal regularization parameter μ_0 is chosen such that it minimizes the relative error of $f^{(0)}(\mu_0)$, i.e., it minimizes

$$R_0 = \frac{\|f - f^{(0)}(\mu_0)\|_2}{\|f\|_2}.$$

The restored images are shown in Figure 2(g) and Figure 2(h). Again, the optimal regularization parameter μ is chosen such that

$$R_1 = \frac{\|f - f(\mu)\|_2}{\|f\|_2}$$

is minimized, where $f(\mu)$ is the solution of (5.4). According to the figures, it is clear that the quality of restored images is visually better than that of initial guess images. The R_0 for Bridge and Cameraman images are 0.0532 and 0.0548, respectively, while the R_1 are 0.0512 and 0.0504, respectively. Here the optimal regularization parameters for the restoration of the bridge image are 2×10^{-3} and 7.5×10^{-5} for R_0 and R_1, respectively. Also the optimal regularization parameters for the restoration of the cameraman image are 5×10^{-3} and 3×10^{-4} for R_0 and R_1, respectively.

The number of iterations for the PCG methods with different preconditioners to solving systems (5.5) and (5.4) are listed in Tables 6 and 7, respectively. We see that the PCG method with FBIP converges much faster than the method with circulant preconditioner, especially when the methods are applied to solving (5.4).

6. Concluding remarks. We have considered the solution of Toeplitz-related systems, including one-dimensional and two-dimensional cases. The FBIP are constructed based on the off-diagonal decay property of the concerned matrices. Numerical results show that our new preconditioners are superior to circulant preconditioners.

In the example of nonlinear image restorations, the Gauss–Newton method converges in two iterations to restore the image. We use the block-circulant preconditioner in the first iteration and FBIP in the second iteration. Therefore, the construction of the proposed preconditioner is relatively inexpensive. One issue that arises in some nonlinear optimization problems is that the coefficient matrix changes at each outer iteration (e.g., the matrix D, which is related to certain constraints, changes at each outer iteration). Thus a new preconditioner should be computed. A possible future research work is to develop a scheme such that the factorized banded inverse preconditioner can be easily updated if the coefficient matrix changes at each outer iteration.

REFERENCES

[1] H. Andrews and B. Hunt, *Digital Image Restoration*, Prentice–Hall, Englewood Cliffs, NJ, 1977.

[2] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, UK, 1994.

[3] M. Benzi, C. D. Meyer, and M. Tůma, *A sparse approximate inverse preconditioner for the conjugate gradient method*, SIAM J. Sci. Comput., 17 (1996), pp. 1135–1149.

[4] D. Bini and P. Favati, *On a matrix algebra related to the discrete Hartley transform*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 500–507.

[5] E. Boman and I. Koltracht, *Fast transform based preconditioners for Toeplitz equations*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 628–645.

[6] R. Chan, *Toeplitz preconditioners for Toeplitz systems with nonnegative generating functions*, IMA J. Numer. Anal., 11 (1991), pp. 333–345.

[7] R. H. Chan and X.-Q. Jin, *A family of block preconditioners for block systems*, SIAM J. Sci. Comput., 13 (1992), pp. 1218–1235.

[8] R. H. Chan and M. K. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Rev., 38 (1996), pp. 427–482.

[9] R. Chan and P. Tang, *Constrained minimax approximation and optimal preconditioners for Toeplitz systems*, Numer. Algorithms, 5 (1993), pp. 353–364.

[10] T. Chan and J. Olkin, *Preconditioners for Toeplitz-block matrices*, Numer. Algorithms, 6 (1993), pp. 89–101.

[11] E. Chow, *A priori sparsity patterns for parallel sparse approximate inverse preconditioners*, SIAM J. Sci. Comput., 21 (2000), pp. 1804–1822.

[12] P. Davis, *Circulant Matrices*, John Wiley & Sons, New York, 1979.

[13] S. Holmgren and K. Otto, *Iterative solution methods and preconditioners for block-tridiagonal systems of equations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 863–886.

[14] S. Jaffard, *Propriétés des matrices "bien localisées" près de leur diagonale et quelques applications*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 7 (1990), pp. 461–476.

[15] T. Kailath and V. Olshevsky, *Displacement structure approach to discrete-trigonometric-transform based preconditioners of G. Strang type and of T. Chan type*, Calcolo, 33 (1996), pp. 191–208.

[16] T. Ku and C. C. Jay Kuo, *On the spectrum of a family of preconditioned block Toeplitz matrices*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 948–966.

[17] L. Yu. Kolotilina and A. Yu. Yeremin, *Factorized sparse approximate inverse preconditionings* I. *Theory*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 45–58.

[18] A. Nikishin and A. Yu. Yeremin, *Prefiltration technique via aggregation for constructing low-density high-quality factorized sparse approximate inverse preconditionings*, Numer. Linear Algebra Appl., 10 (2003), pp. 235–246.

[19] D. Noutsos and P. Vassalos, *New band Toeplitz preconditioners for ill-conditioned symmetric positive definite Toeplitz systems*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 728–743.

[20] S. Serra, *Optimal, quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems*, Math. Comp., 66 (1997), pp. 651–665.

[21] S. Serra, *A Korovkin-type theory for finite Toeplitz operators via matrix algebras*, Numer. Math., 82 (1999), pp. 117–142.

[22] G. Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.

[23] J. Olkin, *Linear and Nonlinear Deconvolution Problems*, Ph.D. thesis, Rice University, Houston, TX, 1986.

[24] T. Strohmer, *Four short stories about Toeplitz matrix calculations*, Linear Algebra Appl., 343/344 (2002), pp. 321–344.

[25] W.-P. Tang, *Toward an effective sparse approximate inverse preconditioner*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 970–986.

[26] W. F. Trench, *An algorithm for the inversion of finite Toeplitz matrices*, J. Soc. Indust. Appl. Math., 12 (1964), pp. 515–522.

[27] A. Tekalp and G. Pavlović, *Restoration of Scanned Photographic Images*, Digital Image Restoration, A. Katsaggelos, ed., Springer-Verlag, New York, 1991, pp. 209–240.

[28] H. Trussell and B. Hunt, *Improved methods of maximum a posteriori restoration*, IEEE Trans. Comput., 27 (1979).

[29] E. Tyrtyshnikov, *A unifying approach to some old and new theorems on distribution and clustering*, Linear Algebra Appl., 232 (1996), pp. 1–43.

[30] M. Zervakis and A. Venetsanopoulos, *Iterative Algorithms with Fast Convergence Rates in Nonlinear Image Restoration*, in SPIE 1452, Image Processing Algorithms and Techniques, 1991, pp. 90–103.