# Vectorized Algorithms for Solving Special Tridiagonal Systems 

Kuo-Liang Chung*<br>Department of Information Management, National Taiwan Institute of Technology<br>No. 43 , Section 4, Keelung Road, Taipei, Taiwan 10672, R.O.C.<br>klchung@cs.ntit.edu.tw<br>Wen-Ming Yan<br>Department of Computer Science and Information Engineering<br>National Taiwan University<br>Taipei, Taiwan 10764, R.O.C.

(Received November 1995; revised and accepted September 1996)


#### Abstract

Solving special tridiagonal systems often arise in the fields of engineering and science. This special tridiagonal system is diagonally dominant and circulant near-Toeplitz. This paper presents two fast vectorized algorithms for solving special tridiagonal systems. Both algorithms employ the matrix perturbation technique and have many computational advantages on vector supercomputer. The related error analysis are also given. Some experimental results are illustrated on vector uniprocessor of the CRAY X-MP EA/116se.


Keywords--Circulant near-Toeplitz systems, CRAY X-MP, Diagonally dominant, Error analyses, Tridiagonal matrices, Vectorized algorithms.

## 1. INTRODUCTION

In this paper, we are interested in the solution of the special tridiagonal system

$$
\begin{equation*}
A_{n} \mathrm{x}=\mathrm{b} \tag{1}
\end{equation*}
$$

of order $n$ on vector uniprocessor, where

$$
A_{n}=\left(\begin{array}{cccccc}
\beta_{1} & \beta_{2} & & & & \beta_{3} \\
\alpha & \beta & \gamma & & & \\
& \cdot & \cdot & \cdot & & \\
& & & \alpha & \beta & \gamma \\
\beta_{3}^{\prime} & & & & \beta_{2}^{\prime} & \beta_{1}^{\prime}
\end{array}\right)
$$

and $|\beta|>|\alpha+\gamma|$. Solving (1) arises in many computational problems [1-7], in which it is one of the most time-consuming elements. The availability of vector supercomputers has had a significant impact on scientific computations [8-10]. The motivation of this research is to design efficient vectorized algorithms for solving (1).

[^0]In this paper, two fast vectorized algorithms for solving (1) are presented. Both new algorithms consist of three phases and only differ in the second phase. The first phase is a Toeplitz factorization of a slightly perturbed matrix of $A_{n}$. The second phase is to solve the perturbed problem in a highly vectorized way, but only scale $\times$ vector operations are involved, hence, it leads to a great deal of computational saving. In the third phase, the solution to the original problem is recovered from the solution to the perturbed problem; this is called the update procedure. Some error analyses are also given. In addition, some experimental results are illustrated on CRAY X-MP EA/116se.

Section 2 presents our vectorized algorithms for solving (1) and the related error analyses. The implementations on the CRAY X-MP EA/116se are illustrated in Section 3. Section 4 gives the conclusions.

## 2. VECTORIZATIONS

### 2.1. Toeplitz Factorization

Throughout the remainder of this paper, matrices are represented by uppercase letters, vectors by bold lowercase letters, and scalars by plain lowercase letters. The superscript ${ }^{\top}$ corresponds to the transpose operation; $\|\bullet\|$ denotes the sup-norm of one vector.

In order to avoid memory conflict in CRAY X-MP EA/116se [11], we enlarge (1), then perturb it to $A_{m}^{\prime} \mathbf{x}^{\prime}=\mathbf{b}^{\prime}$, where $m=p q$ ( $p$ denotes the length of one register file in the vector computer, i.e., the vector length $), \mathbf{b}^{\prime}=(\underbrace{b_{1}, b_{2}, \ldots, b_{n}}_{n}, \underbrace{0, \ldots, 0}_{n-m})^{\top}$, and

$$
A^{\prime}=\left(\begin{array}{ccccc}
a & \gamma & n-m & &  \tag{2}\\
\alpha & \beta & \gamma & & \\
& \cdot & \cdot & \cdot & \\
& & & \alpha & \beta
\end{array}\right) \gamma=L^{\prime} U^{\prime}
$$

$$
L^{\prime}=\left(\begin{array}{cccccc}
1 & & & & & \\
-d & 1 & & & & \\
& \cdot & \cdot & & & \\
& & & -d & 1 & \\
& & & & -d & 1
\end{array}\right) \text { and } \quad U^{\prime}=\left(\begin{array}{cccccc}
a & \gamma & & & & \\
& a & \gamma & & & \\
& & \cdot & \cdot & & \\
& & & & a & \gamma \\
& & & & & a
\end{array}\right)
$$

which implies that $a-\gamma d=\beta$ and $-a d=\alpha$. This in turn implies that $d=-(\alpha / a)$ and $a=$ $\left(\beta \pm \sqrt{\beta^{2}-4 \gamma \alpha}\right) / 2$. Since we wish the matrices $L^{\prime}$ and $U^{\prime}$ to be diagonally dominant, we will select the sign so that the absolute value of $a$ is greater than $\max (|\alpha|,|\gamma|)$. That is, when $\beta>|\alpha+\gamma|$, we choose $a=\left(\beta+\sqrt{\beta^{2}-4 \gamma \alpha}\right) / 2$; when $\beta<-|\alpha+\gamma|$, we choose $a=\left(\beta-\sqrt{\beta^{2}-4 \gamma \alpha}\right) / 2$. Since one of our choices always makes $|a|>\max (|\alpha|,|\gamma|)$, hence, the bidiagonal Toeplitz matrices $L^{\prime}$ and $U^{\prime}$ are diagonally dominant. The computation of $a$ and $d$ provides the Toeplitz factorization of the matrix $A^{\prime}$, which can be done in $O(1)$ time.

### 2.2. Solving the Perturbed System

### 2.2.1. The first method

In this section, our first vectorized method for solving $A_{m}^{\prime} \mathbf{x}^{\prime}=\mathbf{b}^{\prime}$ in a highly vectorized way consists of two parts:
(1) vectorization of a lower bidiagonal Toeplitz system, $L_{m}^{\prime} \mathbf{y}^{\prime}=\mathbf{b}^{\prime}$;
(2) vectorization of an upper bidiagonal Toeplitz system, $U_{m}^{\prime} \mathbf{x}^{\prime}=\mathbf{y}^{\prime}$.

We would especially point out that due to our matrix perturbation technique, all the vector operations involved are scaled by a constant, which is very important for the efficient implementation on the vector computer.

For convenience, we first describe the vectorized method for solving the general lower bidiagonal Toeplitz system (LBTS), $L_{m} \mathbf{x}=\mathbf{y}$, where

$$
L_{m}=\left(\begin{array}{cccccc}
r & & & & & \\
s & r & & & & \\
& \cdot & \cdot & & & \\
& & & s & r & \\
& & & & s & r
\end{array}\right)_{m \times m}, \quad \text { for }|r|>|s|
$$

and $\mathbf{y}=\left(y_{1}, y_{2}, \ldots, y_{m}\right)^{\top}$. In Section 2.1, we know that $m=p q$. Therefore, we partition the above LBTS into $p$ LBTS's. Each LBTS can be written as

$$
\begin{equation*}
L_{q} \mathbf{x}^{\prime(i)}=\mathbf{y}^{(i)}, \quad \text { for } i=0,1, \ldots, p-1 \tag{3}
\end{equation*}
$$

where $\mathbf{x}^{\prime(i)}=\left(x_{i q+1}^{\prime}, x_{i q+2}^{\prime}, \ldots, x_{i q+q}^{\prime}\right)^{\top}$ and $\mathbf{y}^{(i)}=\left(y_{i q+1}, y_{i q+2}, \ldots, y_{i q+q}\right)^{\top}$. Our vectorized subroutine for solving these $p$ smaller LBTS's is shown in Appendix 1, where Loop-5 and Loop-20 can be vectorized with vector length $p$. Specifically, only scalar $\times$ vector operations are involved in this subroutine.

After solving $\mathbf{x}^{\prime}$ in a vectorized way, we have

$$
\begin{equation*}
L_{m} \mathbf{x}^{\prime}=\mathbf{y}+s \sum_{i=1}^{p-1} x_{i q}^{\prime} \mathbf{e}_{i q+1} \tag{4}
\end{equation*}
$$

where $e_{k}=(\underbrace{0, \ldots, 0,1}_{k}, \underbrace{0, \ldots, 0}_{m-k})^{\top}$ for $1 \leq k \leq m$. Since
$L_{m} \mathbf{e}_{k}=r \mathbf{e}_{k}+s \mathbf{e}_{k+1}$,

$$
1 \leq k<m
$$

we have

$$
\begin{align*}
L_{m}\left(\sum_{j=0}^{t-1} \frac{1}{r}\left(-\frac{s}{r}\right)^{j} \mathbf{e}_{k+j}\right) & =\sum_{j=0}^{t-1}\left(-\frac{s}{r}\right)^{j} \mathbf{e}_{k+j}-\sum_{j=0}^{t-1}\left(-\frac{s}{r}\right)^{j+1} \mathbf{e}_{k+j+1}  \tag{5}\\
& =\mathbf{e}_{k}-\left(-\frac{s}{r}\right)^{t} \mathbf{e}_{k+t}
\end{align*}
$$

Let

$$
\begin{equation*}
\mathbf{x}=\mathbf{x}^{\prime}-\frac{s}{r} \sum_{i=1}^{p-1} x_{i q}^{\prime} \sum_{j=0}^{t-1}\left(-\frac{s}{r}\right)^{j} \mathbf{e}_{i q+j+1} \tag{6}
\end{equation*}
$$

where $t$ denotes the length of the update vector which will be discussed in Section 2.3 , the solution vector $\mathbf{x}$ of (6) can be computed by using the vectorized subroutine as shown in Appendix 2, where Loop- 40 can be vectorized with vector length $p-1$. Only scalar $\times$ vector operations are involved in this subroutine.

By (5) and (6), we have

$$
\begin{aligned}
L_{m} \mathbf{x} & =\mathbf{y}+s \sum_{i=1}^{p-1} x_{i q}^{\prime} \mathbf{e}_{i q+1}-s \sum_{i=1}^{p-1} x_{i q}^{\prime}\left(\mathbf{e}_{i q+1}-\left(-\frac{s}{r}\right)^{t} \mathbf{e}_{i q+t+1}\right) \\
& =\mathbf{y}+s \sum_{i=1}^{p-1} x_{i q}^{\prime}\left(-\frac{s}{r}\right)^{t} \mathbf{e}_{i q+t+1}
\end{aligned}
$$

hence,

$$
\begin{equation*}
\left\|L_{m} \mathbf{x}-\mathbf{y}\right\| \leq|s|\left|\frac{s}{r}\right|^{t}\left\|\mathbf{x}^{\prime}\right\| \tag{7}
\end{equation*}
$$

To estimate $\left\|\mathbf{x}^{\prime}\right\|$, we need the following lemma.

Lemma 1. If

$$
L=\left(\begin{array}{cccccc}
r & & & & & \\
s & r & & & & \\
& \cdot & \cdot & & & \\
& & & s & r & \\
& & & & s & r
\end{array}\right), \quad|r|>|s|
$$

then $\left\|L^{-1} \mathbf{y}\right\| \leq 1 /(|r|-|s|)\|y\|$.
Proof. See Appendix 3.
Using the above similar partition approach, it is easy to design our two vectorized subroutines, VUPPER1 ( $\mathrm{r}, \mathrm{s}$ ) (vs. VLOW1 $(\mathrm{r}, \mathrm{s})$ ) and UPDATEUPPER1 ( $\mathrm{r}, \mathrm{s}, \mathrm{t}$ ) (vs. UPDATELOW1 ( $\mathrm{r}, \mathrm{s}, \mathrm{t}$ )), for solving the general upper bidiagonal Toeplitz system (UBTS), $U_{m} \mathbf{x}=\mathbf{y}$, where

$$
U_{m}=\left(\begin{array}{cccccc}
r & s & & & & \\
& r & s & & & \\
& & \cdot & \cdot & & \\
& & & & r & s \\
& & & & & r
\end{array}\right), \quad \text { for }|r|>|s|
$$

For saving space, we omit the pseudo codes for these two subroutines.
Return to solve $A_{m}^{\prime} \mathbf{x}^{\prime}=\mathbf{b}^{\prime}$. Since $A_{m}^{\prime}=L_{m}^{\prime} U_{m}^{\prime}$, we first solve $L_{m}^{\prime} \mathbf{y}^{\prime}=\mathbf{b}^{\prime}$, where $\mathbf{b}^{\prime}=$ $(b_{1}, b_{2}, \ldots, b_{n}, \underbrace{0, \ldots, 0}_{m-n})^{\top}$, then solve $U_{m}^{\prime} \mathbf{x}^{\prime}=\mathbf{y}^{\prime}$. Using our previous vectorized methods for solving general lower and upper bidiagonal systems, the following algorithm is used to solve $A_{m}^{\prime} \mathbf{x}^{\prime}=\mathbf{b}^{\prime}$.

```
C******the entry array y represents vector b'
C*******the exit array represents vector x'
    CALL VLOW1(1,-d)
    CALL UPDATELOW(1,-d,t1)
    CALL VUPPER1(a,gamma)
    CALL UPDATEUPPER(a,gamma,t2)
```

Lemma 2. Let $c_{1}=1 /(1-|d|)$ and $c_{2}=(1+|d|)|\gamma|(1 /(|a|-|\gamma|))((1+|d|) /(1-|d|))$, then

$$
\left\|A_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{b}^{\prime}\right\| \leq\left(c_{1}|d|^{t_{1}+1}+c_{2}\left|\frac{\gamma}{a}\right|^{t_{2}}\right)\|\mathbf{b}\|
$$

The term $c_{1}|d|^{t_{1}+1}$ will be less than $\xi$, if $t_{1}$ is greater than $\left(\log \xi-\log c_{1}\right) / \log |d|-1$; the term $c_{2}|\gamma / a|^{t_{2}+1}$ will be less than $\xi$, if $t_{2}$ is greater than $\left(\left(\log \xi-\log c_{2}\right) /(\log |\gamma / a|)\right)-1$. Let $c_{3}=$ $((1+|\gamma / a|) /(|a|-|\gamma|))((1+|d|) /(1-|d|))$, then it yields

$$
\left\|\mathbf{x}^{\prime}\right\| \leq c_{3}\|\mathbf{b}\|
$$

## Proof. See Appendix 4.

Lemma 2 will be used in the analysis of the update phase in Section 2.3.
In the following section, based on the product expansion method [12,13], we present the second vectorized method for solving $A_{m}^{\prime} \mathbf{x}^{\prime}=\mathbf{b}^{\prime}$, where $m=n$, since we do not need the partition approach as described in Section 2.2.1.

### 2.2.2. The second method

Now we describe our second vectorized method for solving $L_{m} \mathbf{x}=\mathbf{y}$. All the vector operations involved are scaled only by a constant.

Since $L_{m}=r(I-E)$ with

$$
E=\left(\begin{array}{cccccc}
0 & & & & & \\
-\frac{s}{r} & 0 & & & & \\
& \cdot & \cdot & & & \\
& & & -\frac{s}{r} & 0 & \\
& & & & -\frac{s}{r} & 0
\end{array}\right)
$$

the system $L_{m} \mathbf{x}=\mathbf{y}$ is equal to $(I-E) \mathbf{x}=(1 / r) \mathbf{y}$. $\mathbf{x}$ can be obtained by computing $\mathbf{x}=$ $\left(I+E^{2^{k}}\right) \cdots\left(I+E^{2}\right)(I+E)((1 / r) \mathbf{y})$ because

$$
\begin{aligned}
L_{m} \mathbf{x} & =r(I-E) \mathbf{x} \\
& =r(I-E)(I+E)\left(I+E^{2}\right) \cdots\left(I+E^{2^{k}}\right)\left(\frac{1}{r} \mathbf{y}\right) \\
& =\left(I-E^{2^{k+1}}\right) \mathbf{y}
\end{aligned}
$$

Since

$$
E^{2^{k+1}}=\left(\begin{array}{llll}
\left(-\frac{s}{r}\right)^{2^{k+1}} & & \\
& \cdots & & \\
& & \left(-\frac{s}{r}\right)^{k^{k+1}} & \\
& & & \left(-\frac{s}{r}\right)^{k^{k+1}}
\end{array}\right)
$$

we then have

$$
\begin{equation*}
\left\|L_{m} \mathbf{x}-\mathbf{y}\right\| \leq\left|\frac{s}{r}\right|^{2^{k+1}}\|\mathbf{y}\| \tag{8}
\end{equation*}
$$

The relative residual $\left(\left\|L_{m} \mathbf{x}-\mathbf{y}\right\|\right) /\|\mathbf{y}\|$ will be less than $\xi$, if $k$ is greater than $\log (\log \xi /(\log |s / r|))$ -1 . Therefore, the computation of $\mathbf{x}\left(=\mathbf{x}^{(k+1)}\right)$ can be accomplished by the following iterative formula:

$$
\begin{equation*}
\mathbf{x}^{(i+1)}=\left(I+E^{2^{i}}\right) \mathbf{x}^{(i)}=\mathbf{x}^{(i)}+E^{2^{2}} \mathbf{x}^{(i)}, \quad 0 \leq i \leq k \tag{9}
\end{equation*}
$$

with the initial assignment $\mathbf{x}^{(0)}=(1 / r) \mathbf{y}$.
From (8), we have $\left\|\mathbf{x}^{(i+1)}\right\| \leq\left(1+|s / r|^{2^{i}}\right)\left\|\mathbf{x}^{(i)}\right\|$. Thus,

$$
\begin{align*}
\left\|\mathbf{x}^{(k+1)}\right\| & \leq\left(1+\left|\frac{s}{r}\right|^{2^{k}}\right)\left(1+\left|\frac{s}{r}\right|^{2^{k-1}}\right) \cdots\left(1+\left|\frac{s}{r}\right|\right)\left|\frac{1}{r}\right|\|\mathbf{y}\| \\
& =\frac{1-|s / r|^{2^{k+1}}}{1-|s / r|}\left|\frac{1}{r}\right|\|\mathbf{y}\|  \tag{10}\\
& \leq \frac{1}{|r|-|s|}\|\mathbf{y}\|
\end{align*}
$$

The formal vectorized subroutine for computing $\mathbf{x}^{(k+1)}$ is shown in Appendix 5, where mainly scalar $\times$ vector operations are involved in this subroutine.

Similarly, it is easy to design our vectorized subroutine for solving the general UBTS system, $U_{m} \mathbf{x}=\mathbf{y}$. The corresponding vectorized subroutine is shown in Appendix 6, where mainly scalar $\times$ vector operations are involved in this subroutine.

Using our previous two vectorized subroutines shown in Appendices 5 and 6 for solving general lower and upper bidiagonal systems, respectively, the following procedure is used to solve $A_{m}^{\prime} \mathbf{x}^{\prime}=\mathbf{b}^{\prime}$.

```
CALL VLOW2(1, -d,k1)
CALL VUPPER2(a,gamma,k2)
```

Lemma 3. Let $c_{1}^{\prime}=1$ and $c_{2}^{\prime}=(1+|d|) /(1-|d|)$, then we have

$$
\left\|A_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{b}^{\prime}\right\| \leq\left(c_{1}^{\prime}|d|^{2^{k_{1}+1}}+c_{2}^{\prime}\left|d^{\prime}\right|^{2^{k_{2}+1}}\right)\|\mathbf{b}\|
$$

The term $c_{1}^{\prime}|d|^{2^{k_{1}+1}}$ will be less than $\xi$, if $k_{1}$ is greater than $\log (\log \xi / \log |d|)-1$; the term $c_{2}^{\prime}\left|d^{\prime}\right|^{k_{2}+1}$ will be less than $\xi$, if $k_{2}$ is greater than $\log \left(\left(\log \xi-\log c_{2}^{\prime}\right) / \log |d|\right)-1$. Let $c_{3}^{\prime}=$ $(1 /(|a|-|\gamma|)(1 /(1-|d|))$, then it follows that

$$
\left\|\mathbf{x}^{\prime}\right\| \leq c_{3}^{\prime}\|\mathbf{b}\|
$$

Proof. See Appendix 7.
Lemma 3 will be used in the following section.

### 2.3. Update

After solving the perturbed system $A_{m}^{\prime} \mathbf{x}^{\prime}=\mathbf{b}^{\prime}$ approximately, the approximate solution of $\mathbf{x}$ will be recovered from the perturbed system in this section.

Let $\mathbf{z}^{\prime}=A_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{b}^{\prime}, \overline{\mathbf{z}}=\left(z_{1}^{\prime}, z_{2}^{\prime}, \ldots, z_{n}^{\prime}\right)^{\top}, \overline{\mathbf{x}}=\left(x_{1}^{\prime}, x_{2}^{\prime}, \ldots, x_{n}^{\prime}\right)^{\top}$, and $\mathbf{w}=A_{n} \overline{\mathbf{x}}-\mathbf{b}$. Since

$$
\begin{align*}
w_{i} & =\alpha x_{i-1}^{\prime}+\beta x_{i}^{\prime}+\gamma x_{i+1}^{\prime}-b_{i}=z_{i}^{\prime}, \quad \text { for } 1<i<n \\
w_{1} & =\beta_{1} x_{1}^{\prime}+\beta_{2} x_{2}^{\prime}+\beta_{3} x_{n}^{\prime}-b_{1}  \tag{11}\\
w_{n} & =\beta_{3}^{\prime} x_{1}^{\prime}+\beta_{2}^{\prime} x_{n-1}^{\prime}+\beta_{1} x_{n}^{\prime}-b_{n}
\end{align*}
$$

we have $\left\|\mathbf{w}-w_{1} \mathbf{e}_{1}-w_{n} \mathbf{e}_{n}\right\| \leq\|\overline{\mathbf{z}}\| \leq\left\|\mathbf{z}^{\prime}\right\|$. By Lemma 2, it follows that

$$
\begin{equation*}
\left\|A_{n} \overline{\mathbf{x}}-\mathbf{b}-w_{1} \mathbf{e}_{1}-w_{n} \mathbf{e}_{n}\right\| \leq\left\|\mathbf{z}^{\prime}\right\| \leq\left(c_{1}|d|^{t_{\mathbf{1}}+1}+c_{2}\left|\frac{\gamma}{a}\right|^{t_{2}}\right)\|\mathbf{b}\| \tag{12}
\end{equation*}
$$

The value of $\left\|\mathbf{z}^{\prime}\right\|$ will be very small when $t_{1}$ and $t_{2}$ are large enough. Therefore, the approximate solution of $\mathbf{x}$ to be determined equals $\overline{\mathbf{x}}-\mathbf{p}$, where

$$
\begin{equation*}
A \mathbf{p} \doteq w_{1} \mathbf{e}_{1}+w_{n} \mathbf{e}_{n} \tag{13}
\end{equation*}
$$

To solve $\mathbf{p}$, we try to ignore the first and last equalities of the system $A \mathbf{p} \doteq w_{1} \mathbf{e}_{1}+w_{n} \mathbf{e}_{n}$, then we must solve the recurrence relation: $\alpha p_{i-1}+\beta p_{i}+\gamma p_{i+1}=0$ for $2 \leq i \leq n-1$. From $a-\gamma d=\beta$, $-a d=\alpha$, it follows that $\alpha+\beta d+\gamma d^{2}=0$ and $\gamma+\beta d^{\prime}+\alpha d^{\prime 2}=0$, where $d^{\prime}=-\gamma / a$. Naturally, if we try $\mathbf{p}_{1}=(\underbrace{d, d^{2}, \ldots, d^{t_{3}}}_{t_{3}}, \underbrace{0, \ldots, 0}_{n-t_{3}})^{\top}$, then

$$
\begin{align*}
A_{n} \mathbf{p}_{1} & =(\underbrace{\beta_{1} d+\beta_{2} d^{2}, \alpha d+\beta d^{2}+\gamma d^{3}, \ldots, \alpha d^{t_{3}-1}+\beta d^{t_{3}}}_{t_{3}}, \underbrace{\alpha d^{t_{3}}, 0, \ldots, 0, \beta_{3}^{\prime} d}_{n-t_{3}})^{\top} \\
& =(\underbrace{\beta_{1} d+\beta_{2} d^{2}, 0, \ldots, 0, \alpha d^{t_{3}-1}+\beta d^{t_{3}}}_{t_{3}}, \underbrace{\alpha d^{t_{3}}, 0, \ldots, 0, \beta_{3}^{\prime} d}_{n-t_{3}})^{\top}  \tag{14}\\
& =u \mathbf{e}_{1}+v \mathbf{e}_{n}+\alpha d^{t_{3}}\left(-d \frac{\gamma}{\alpha} \mathbf{e}_{t_{3}}+\mathbf{e}_{t_{3}+1}\right),
\end{align*}
$$

where $u=\left(\beta_{1} d+\beta_{2} d^{2}\right)$ and $v=\beta_{3}^{\prime} d$. Similarly, if we try $\mathbf{p}_{2}=(\underbrace{0, \ldots, 0}_{n-t_{4}}, \underbrace{d^{\prime t_{4}}, \ldots, d^{2}, d^{\prime}}_{t_{4}})^{\top}$, then

$$
\begin{equation*}
A_{n} \mathbf{p}_{2}=u^{\prime} \mathbf{e}_{\mathbf{1}}+v^{\prime} \mathbf{e}_{n}+\gamma d^{t_{4}}\left(-d^{\prime} \frac{\alpha}{\gamma} \mathbf{e}_{n-t_{4}+1}+\mathbf{e}_{n-t_{4}}\right) \tag{15}
\end{equation*}
$$

where $v^{\prime}=\left(\beta_{1}^{\prime} d^{\prime}+\beta_{2}^{\prime} d^{2}\right)$ and $u^{\prime}=\beta_{3} d^{\prime}$.
Let

$$
\begin{equation*}
\mathbf{p}=s \mathbf{p}_{1}+s^{\prime} \mathbf{p}_{2} \tag{16}
\end{equation*}
$$

then by (14) and (15), we have

$$
\begin{equation*}
A_{n} \mathbf{p}=\left(s u+s^{\prime} u^{\prime}\right) \mathbf{e}_{1}+\left(s v+s^{\prime} v^{\prime}\right) \mathbf{e}_{n}+\mathbf{r} \tag{17}
\end{equation*}
$$

where $\mathbf{r}=s \alpha d^{t_{3}}\left(-d(\gamma / \alpha) \mathbf{e}_{t_{3}}+\mathbf{e}_{t_{3}+1}\right)+s^{\prime} \gamma d^{t_{4}}\left(-d^{\prime}(\alpha / \gamma) \mathbf{e}_{n-t_{4}+1}+\mathbf{e}_{n-t_{4}}\right)$.
Comparing with (13), we let $s u+s^{\prime} u^{\prime}=w_{1}$ and $s v+s^{\prime} v^{\prime}=w_{n}$, and it follows that

$$
\begin{align*}
s & =\frac{w_{1} v^{\prime}-w_{n} u^{\prime}}{u v^{\prime}-v u^{\prime}} \\
s^{\prime} & =\frac{u w_{n}-v w_{1}}{u v^{\prime}-v u^{\prime}}  \tag{18}\\
A_{n} \mathbf{p} & =w_{1} \mathbf{e}_{1}+w_{n} \mathbf{e}_{n}+\mathbf{r}
\end{align*}
$$

After determining $\mathbf{p}$, the subroutine for computing $\mathbf{x}(=\overline{\mathbf{x}}-\mathbf{p})$ is shown in Appendix 8 , where the above concerning operations are the well-known prefix-product operations.

Furthermore, combining the vectorized subroutines described in Section 2.2.1, and the above subroutine shown in Appendix 8, our first vectorized algorithm for solving (1) is constituted by the following five subroutines.

```
CALL VLOW1(1,-d)
CALL UPDATELOW (1,-d,t1)
CALL VUPPER1(a,gamma)
CALL UPDATEUPPER(a,gamma,t2)
CALL FINAL(t3,t4)
```

Similarly, combining the vectorized subroutines described in Section 2.2.2, and the subroutine FINAL $\left(t_{3}, t_{4}\right)$, our second vectorized algorithm for solving (1) is shown below.

```
CALL VLOW2(1,-d,k1)
CALL VUPPER2(a,gamma,k2)
CALL FINAL(t3,t4)
```


### 2.4. Error Analyses

### 2.4.1. For the first method

The following theorem gives the error analysis of our first vectorized algorithm for solving (1).
THEOREM 4. Let $c_{4}=1+\left(\left|\beta_{1}\right|+\left|\beta_{2}\right|+\left|\beta_{3}\right|\right) c_{3}, c_{5}=1+\left(\left|\beta_{1}^{\prime}\right|+\left|\beta_{2}^{\prime}\right|+\left|\beta_{3}^{\prime}\right|\right) c_{3}, c_{6}=\left(c_{4}\left|v^{\prime}\right|+\right.$ $\left.c_{5}\left|u^{\prime}\right|\right) /\left(\left|u v^{\prime}-v u^{\prime}\right|\right)$, and $c_{7}=\left(c_{5}|u|+c_{4}|v|\right) /\left(\left|u v^{\prime}-v u^{\prime}\right|\right)$, we have

$$
\frac{\left\|A_{n} \mathbf{x}-\mathbf{b}\right\|}{\|\mathbf{b}\|} \leq\left(c_{1}|d|^{t_{1}+\mathbf{1}}+c_{2}\left|d^{\prime}\right|^{t_{2}}+c_{8}|d|^{t_{3}}+c_{9}\left|d^{\prime}\right|^{t_{4}}\right)
$$

where $c_{8}=c_{6}|\alpha|$ and $c_{9}=c_{7}|\gamma|$.
Proof. See Appendix 9.

For simplicity, by Theorem 4 we define the following notations: $t_{1}(\xi)=\min \left\{t \in N: c_{1}|d|^{t+1}<\right.$ $\xi / 4\} ; t_{2}(\xi)=\min \left\{t \in N: c_{2}\left|d^{\prime}\right|^{t}<\xi / 4\right\} ; t_{3}(\xi)=\min \left\{t \in N: c_{8}|d|^{t}<\xi / 4\right\} ; t_{4}(\xi)=\min \{t \in$ $\left.N: c_{9}\left|d^{\prime}\right|^{t}<\xi / 4\right\}$, where $\xi$ is the upper bound of the relative residual. If $\max \left(t_{1}(\xi), t_{2}(\xi)\right)>n / p$ or $\max \left(t_{3}(\xi), t_{4}(\xi)\right)>n$, then our first vectorized algorithm will break down when the required relative tolerance is sufficiently small and/or the diagonal dominance ratio is sufficiently close to 2 .

However, for some cases, our first vectorized algorithm works well. For example, for conventional B-spline curve fitting [14,15], the corresponding system is of (1) with $\beta_{1}=5, \beta_{2}=1$, $\beta=4, \beta_{3}=0, \alpha=1, \gamma=1, \beta_{1}^{\prime}=5, \beta_{2}^{\prime}=1$, and $\beta_{3}^{\prime}=0$. We have $t_{1}(\xi)=11, t_{2}(\xi)=12$, $t_{3}(\xi)=13$, and $t_{4}(\xi)=13$ for $\xi=10^{-6}$; we have $t_{1}(\xi)=13, t_{2}(\xi)=14, t_{3}(\xi)=15$, and $t_{4}(\xi)=15$ for $\xi=10^{-7}$. For closed B-spline curve fitting [14], the corresponding system is of (1) with $\beta_{1}=4$, $\beta_{2}=1, \beta=4, \beta_{3}=1, \alpha=1, \gamma=1, \beta_{1}^{\prime}=4, \beta_{2}^{\prime}=1$, and $\beta_{3}^{\prime}=1$. We have $t_{1}(\xi)=11, t_{2}(\xi)=12$, $t_{3}(\xi)=14$, and $t_{4}(\xi)=14$ for $\xi=10^{-6}$; we have $t_{1}(\xi)=13, t_{2}(\xi)=14, t_{3}(\xi)=15$, and $t_{4}(\xi)=15$ for $\xi=10^{-7}$. Furthermore, let us examine the other two cases. If the system is of (1) with $\beta_{1}=3$, $\beta_{2}=1, \beta=3, \beta_{3}=1, \alpha=1, \gamma=1, \beta_{1}^{\prime}=3, \beta_{2}^{\prime}=1$, and $\beta_{3}^{\prime}=1$, we have $t_{1}(\xi)=16, t_{2}(\xi)=17$, $t_{3}(\xi)=19$, and $t_{4}(\xi)=19$ for $\xi=10^{-6}$. If the system with small diagonal dominance ratio is of (1) with $\beta_{1}=2.1, \beta_{2}=1, \beta=2.1, \beta_{3}=1, \alpha=1, \gamma=1, \beta_{1}^{\prime}=2.1, \beta_{2}^{\prime}=1$, and $\beta_{3}^{\prime}=1$, we have $t_{1}(\xi)=52, t_{2}(\xi)=60, t_{3}(\xi)=68$, and $t_{4}(\xi)=68$ for $\xi=10^{-6}$.

### 2.4.2. For the second method

In what follows, we will discuss the error analysis of our second vectorized algorithm for solving (1).

By Lemma 2 and (11), we have

$$
\begin{equation*}
\left|w_{1}\right| \leq c_{4}^{\prime}\|b\| \quad \text { and } \quad\left|w_{n}\right| \leq c_{5}^{\prime}\|b\| \tag{19}
\end{equation*}
$$

where $c_{4}^{\prime}=1+\left(\left|\beta_{1}\right|+\left|\beta_{2}\right|+\left|\beta_{3}\right|\right) c_{3}^{\prime}$ and $c_{5}^{\prime}=1+\left(\left|\beta_{1}^{\prime}\right|+\left|\beta_{2}^{\prime}\right|+\left|\beta_{3}^{\prime}\right|\right) c_{3}^{\prime}$. By (18) and (19), we have

$$
\begin{equation*}
|s| \leq c_{6}^{\prime}\|\mathbf{b}\| \quad \text { and } \quad\left|s^{\prime}\right| \leq c_{7}^{\prime}\|\mathbf{b}\| \tag{20}
\end{equation*}
$$

where $c_{6}^{\prime}=\left(c_{4}^{\prime}\left|v^{\prime}\right|+c_{5}^{\prime}\left|u^{\prime}\right|\right) /\left(\left|u v^{\prime}-v u^{\prime}\right|\right)$ and $c_{7}^{\prime}=\left(c_{5}^{\prime}|u|+c_{4}^{\prime}|v|\right) /\left(\left|u v^{\prime}-v u^{\prime}\right|\right)$. By the variation of (9a) and (20), we have the following result.

Theorem 5.

$$
\frac{\left\|A_{n} \mathbf{x}-\mathbf{b}\right\|}{\|\mathrm{b}\|} \leq\left(|d|^{2^{k_{1}+1}}+c_{2}^{\prime}\left|d^{\prime}\right|^{2^{k_{2}+1}}+c_{8}^{\prime}|d|^{t_{3}}+c_{9}^{\prime}\left|d^{\prime}\right|^{t_{4}}\right)
$$

where $c_{8}^{\prime}=c_{6}^{\prime}|\alpha|$ and $c_{9}^{\prime}=c_{7}^{\prime}|\gamma|$.
For simplicity, by (20) we define the following notations: $t_{1}(\xi)=\min \left\{t \in N:|d|^{2^{t+1}}<\xi / 4\right\}$; $t_{2}(\xi)=\min \left\{t \in N: c_{2}^{\prime}\left|d^{\prime}\right|^{2^{t+1}}<\xi / 4\right\} ; t_{3}(\xi)=\min \left\{t \in N: c_{8}^{\prime}|d|^{t}<\xi / 4\right\} ; t_{4}(\xi)=\min \{t \in N:$ $\left.c_{9}^{\prime}\left|d^{\prime}\right|^{t}<\xi / 4\right\}$, where $\xi$ is the upper bound of the relative residual.

If $\max \left(t_{1}(\xi), t_{2}(\xi)\right)>n$ or $\max \left(t_{3}(\xi), t_{4}(\xi)\right)>n$, then our second vectorized algorithm will break down when the required relative tolerance is sufficiently small and/or the diagonal dominance ratio is sufficiently close to 2 .

However, for some cases, our second vectorized algorithm works well. For the conventional B-spline curve fitting, we have $t_{1}(\xi)=3, t_{2}(\xi)=3, t_{3}(\xi)=13$, and $t_{4}(\xi)=13$ for $\xi=10^{-6}$; we have $t_{1}(\xi)=3, t_{2}(\xi)=3, t_{3}(\xi)=15$, and $t_{4}(\xi)=15$ for $\xi=10^{-7}$. For closed B-spline curve fitting, we have $t_{1}(\xi)=3, t_{2}(\xi)=3, t_{3}(\xi)=13$, and $t_{4}(\xi)=13$ for $\xi=10^{-6}$; we have $t_{1}(\xi)=3$, $t_{2}(\xi)=3, t_{3}(\xi)=15$, and $t_{4}(\xi)=15$ for $\xi=10^{-7}$. Furthermore, let us examine the previous two cases with diagonal dominance ratios $3 / 2$ and $2.1 / 2$, respectively. For the first case, we have $t_{1}(\xi)=3, t_{2}(\xi)=4, t_{3}(\xi)=19$, and $t_{4}(\xi)=19$ for $\xi=10^{-6}$. For the second case, we have $t_{1}(\xi)=5, t_{2}(\xi)=5, t_{3}(\xi)=65$, and $t_{4}(\xi)=65$ for $\xi=10^{-6}$.

## 3. EXPERIMENTAL RESULTS

The machine used in the numerical experiments is the CRAY X-MP EA/116se. The machine has a register-register architecture without the cache memory and has one vector processor which contains eight 64 -bit vector registers of length 64 . Memory is divided into 16 banks and each bank contains 1 M 64 -bit words. Each bank requires 14 cycle time (one cycle time needs 8.5 nanoseconds) before it is ready for another request. The peak performance is 235 MFLOPS (millions of floating-point operations per second). In order to avoid memory conflicts for some $m$, the value of $m(\geq 512)$ can be selected as $64(2 z+1)$, where $z$ is the smallest positive integer such that it satisfies that $64(2 z+1) \geq n$. Here $m=64(2\lceil n / 128\rceil+1)$.

All of our testing data, $\mathbf{b}$ 's are generated by a random number generator, a function call $\operatorname{ranf}()$, and each entry of $\mathbf{b}$ is ranged from 0 to 1 . Our two vectorized algorithms are coded by CRAY Fortran 77 language. The operating system used here is UNICOS 6.1.6 and the compiler is called CF77.

Let $A_{i 1}=A_{i}(2.5), A_{i 2}=A_{i}(2.7)$, and $A_{i 3}=A_{i}(3)$ for $i=1,2$, where

$$
A_{1}(\beta)=\left(\begin{array}{cccccc}
\beta & 1 & & & & \\
1 & \beta & 1 & & & \\
& \cdot & \cdot & \cdot & & \\
& & & 1 & \beta & 1 \\
& & & & 1 & \beta
\end{array}\right)
$$

and

$$
A_{2}(\beta)=\left(\begin{array}{cccccc}
\beta & 1 & & & & 1 \\
1 & \beta & 1 & & & \\
& \cdot & \cdot & \cdot & & \\
& & & 1 & \beta & 1 \\
1 & & & & 1 & \beta
\end{array}\right)
$$

In addition, we let

$$
A_{14}=\left(\begin{array}{cccccc}
5 & 1 & & & & \\
1 & 4 & 1 & & & \\
& \cdot & \cdot & \cdot & & \\
& & & 1 & 4 & 1 \\
& & & & 1 & 5
\end{array}\right)
$$

and

$$
A_{24}=\left(\begin{array}{ccccccc}
5 & 1 & & & & & 1 \\
1 & 4 & 1 & & & & \\
& \cdot & \cdot & \cdot & & \\
& & & 1 & 4 & 1 \\
1 & & & & 1 & 5
\end{array}\right)
$$

which are derived from the conventional B-spline curve fitting problem and have been discussed in Section 2.4. When running our two algorithms on CRAY X-MP EA/116se, the performance is illustrated in Table 1 and Table 2, respectively.

Table 1. First algorithm's performance on CRAY X-MP EA/116se.

| $n$ | $p$ | $q$ | time | $r_{11}$ | $r_{12}$ | $r_{13}$ | $r_{14}$ | $r_{21}$ | $r_{22}$ | $r_{23}$ | $r_{24}$ |
| :---: | :---: | ---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 2048 | 64 | 33 | 0.18 | $10^{-8}$ | $10^{-10}$ | $10^{-11}$ | $10^{-11}$ | $10^{-8}$ | $10^{-10}$ | $10^{-11}$ | $10^{-11}$ |
| 4096 | 64 | 65 | 0.27 | $10^{-8}$ | $10^{-10}$ | $10^{-11}$ | $10^{-11}$ | $10^{-8}$ | $10^{-10}$ | $10^{-11}$ | $10^{-11}$ |
| 8192 | 64 | 129 | 0.45 | $10^{-8}$ | $10^{-10}$ | $10^{-11}$ | $10^{-11}$ | $10^{-8}$ | $10^{-10}$ | $10^{-11}$ | $10^{-11}$ |
| 16384 | 64 | 257 | 0.81 | $10^{-8}$ | $10^{-10}$ | $10^{-11}$ | $10^{-11}$ | $10^{-8}$ | $10^{-10}$ | $10^{-11}$ | $10^{-11}$ |

Table 2. Second algorithm's performance on CRAY X-MP EA/116se.

| $n$ | time | $r_{11}$ | $r_{12}$ | $r_{13}$ | $r_{14}$ | $r_{21}$ | $r_{22}$ | $r_{23}$ | $r_{24}$ |
| :---: | :--- | :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 128 | 0.043 | $10^{-10}$ | $10^{-12}$ | $10^{-14}$ | $10^{-14}$ | $10^{-10}$ | $10^{-12}$ | $10^{-14}$ | $10^{-14}$ |
| 256 | 0.066 | $10^{-10}$ | $10^{-12}$ | $10^{-14}$ | $10^{-14}$ | $10^{-10}$ | $10^{-12}$ | $10^{-14}$ | $10^{-14}$ |
| 512 | 0.11 | $10^{-10}$ | $10^{-12}$ | $10^{-14}$ | $10^{-14}$ | $10^{-10}$ | $10^{-12}$ | $10^{-14}$ | $10^{-14}$ |
| 1024 | 0.20 | $10^{-10}$ | $10^{-12}$ | $10^{-14}$ | $10^{-14}$ | $10^{-10}$ | $10^{-12}$ | $10^{-14}$ | $10^{-14}$ |

In Tables 1 and 2 , the symbols $n, p$, and $q \times q$ denote the size of $\mathbf{b}$, the number of the blocks, the size of one block. The symbol 'time' in terms of millisecond, represents the time spent in the first vectorized algorithm. The symbol $r_{i j}$ for $1 \leq i \leq 2$ and $1 \leq j \leq 4$ represents the sup-norm of the relative residual corresponding to $A_{i j}$. In Table 1, by Theorem 4, we select $t_{1}=t_{2}=t_{3}=t_{4}=30$; in Table 2, by Theorem 5, we select $k_{1}=k_{2}=5$ and $t_{3}=t_{4}=30$.

## 4. CONCLUSIONS

This paper has presented two fast vectorized algorithms for solving special tridiagonal systems and has analyzed the error analyses. Due to our matrix perturbation technique, all the vector operations involved in the two algorithms are scaled by a constant, which is very important for the efficient implementation on the CRAY X-MP. Some experimental results demonstrate the performance of our two vectorized algorithms. Our results can be applied to solve the quadratic B-spline curve fitting problem [16,17], the parabolic PDE [4] problem, and so on, since these problems belong to the type of special tridiagonal Toeplitz systems.

Using the same matrix perturbation method proposed in this paper, the parallel algorithms for solving (1) on hypercubes [18] and the B-spline surface fitting [19], respectively, have been developed. In addition, the results of this paper can also be applied to solve the diagonally dominant block tridiagonal system to achieve better performance. It is interesting to employ the other parallel tridiagonal solvers [20,21].

## APPENDIX 1

```
    SUBROUTINE VLOW1 ( \(\mathrm{r}, \mathrm{s}\) )
    real \(r, s\)
C*****the entry array y represents vector y
C*****the exit array y represents vector \(x^{\prime}\)
    do \(5 \mathrm{i}=0, \mathrm{p}-1\)
        \(\mathrm{y}(\mathrm{i} * \mathrm{q}+1)=(1 / \mathrm{r}) * \mathrm{y}(\mathrm{i} * \mathrm{q}+1)\)
5 continue
    do \(10 \quad \mathrm{j}=2, \mathrm{q}\)
        do \(20 \mathrm{i}=0, \mathrm{p}-1\)
            \(y(i * q+j)=(1 / r) *(y(i * q+j)-s * y(i * q+j-1))\)
        continue
    continue
```


## APPENDIX 2

        SUBROUTINE UPDATELOW1 \((r, s, t)\)
        real \(r, s\)
        integer \(t\)
    C******the entry array y represents vector $x^{\prime}$
C******the exit array y represents vector x
temp=1
do $30 \mathrm{j}=0, \mathrm{t}-1$
temp=temp*(-s/r)

```
        do 40 i=1,p-1
                        y(i*q+j+1)=y(i*q+j+1)+y(iq)*temp
            continue
continue
```


## APPENDIX 3

Let $L^{-1} \mathbf{y}=\mathbf{x}$, then $L \mathbf{x}=\mathbf{y}$, i.e.,

$$
\begin{align*}
r x_{1} & =y_{1},  \tag{3a}\\
s x_{i-1}+r x_{i} & =y_{i}, \quad i>1 . \tag{3~b}
\end{align*}
$$

If $\|\mathbf{x}\|=\left|x_{1}\right|$, then by (3a), we have $\|\mathbf{y}\| \geq\left|y_{1}\right|=\left|r x_{1}\right|=|r|\|\mathbf{x}\|$. If $\|\mathbf{x}\|=\left|x_{i}\right|$ for some $i>1$, then by (3b) and the triangular inequality, we have

$$
\begin{aligned}
\|\mathbf{y}\| \geq\left|y_{i}\right| & \geq\left|r x_{i}\right|-\left|s x_{i-1}\right| \\
& \geq|r|\|\mathbf{x}\|-|s|\|\mathbf{x}\| \\
& =(|r|-|s|)\|\mathbf{x}\|
\end{aligned}
$$

We complete the proof and have $\|\mathbf{x}\| \leq(1 /(|r|-|s|))\|\mathbf{y}\|$.

## APPENDIX 4

Taking the triangular inequality on both sides of (6), since $|s / r|<1$, it yields

$$
\begin{equation*}
\|\mathbf{x}\| \leq\left(1+\left|\frac{s}{r}\right|\right)\left\|\mathbf{x}^{\prime}\right\| \tag{4a}
\end{equation*}
$$

Applying Lemma 1 to (3), we have $\left\|{\mathbf{x}^{\prime}}^{(i)}\right\| \leq 1 /(|r|-|s|)\left\|\mathbf{y}^{(i)}\right\|$. Thus, we obtain

$$
\begin{aligned}
\left\|\mathbf{x}^{\prime}\right\| & =\max _{1 \leq i \leq p}\left\|\mathbf{x}^{\prime(i)}\right\| \\
& \leq \max _{1 \leq i \leq p} \frac{1}{|r|-|s|}\left\|\mathbf{y}^{(i)}\right\| \\
& \leq \frac{1}{|r|-|s|}\|\mathbf{y}\|
\end{aligned}
$$

Furthermore, by (7) we have

$$
\begin{equation*}
\left\|L_{m} \mathbf{x}-\mathbf{y}\right\| \leq\left.|s| \frac{s}{r}\right|^{t} \frac{1}{|r|-|s|}\|\mathbf{y}\| \tag{4b}
\end{equation*}
$$

by (4a), we have

$$
\begin{equation*}
\|\mathbf{x}\| \leq \frac{1+|s / r|}{|r|-|s|}\|\mathbf{y}\| \tag{4c}
\end{equation*}
$$

By (4b), (4c), and their variations, the bound of the sup-norm of the residual vector $A_{m}^{\prime} \mathbf{x}^{\prime}-\mathrm{b}^{\prime}$ is derived by

$$
\begin{aligned}
\left\|A_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{b}^{\prime}\right\| & \leq\left\|L_{m}^{\prime}\left(U_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{y}^{\prime}\right)\right\|+\left\|L_{m}^{\prime} \mathbf{y}^{\prime}-\mathbf{b}^{\prime}\right\| \\
& \leq(1+|d|)\left\|U_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{y}^{\prime}\right\|+\left\|L_{m}^{\prime} \mathbf{y}^{\prime}-\mathbf{b}^{\prime}\right\| \\
& \leq\left.(1+|d|)|\gamma| \frac{\gamma}{a}\right|^{t_{2}} \frac{1}{|a|-|\gamma|}\left\|\mathbf{y}^{\prime}\right\|+|d|^{t_{1}+1} \frac{1}{1-|d|}\left\|\mathbf{b}^{\prime}\right\| \\
& \leq\left((1+|d|)|\gamma|\left|\frac{\gamma}{a}\right|^{t_{2}} \frac{1}{|a|-|\gamma|} \frac{1+|d|}{1-|d|}+|d|^{t_{1}+1} \frac{1}{1-|d|}\right)\|\mathbf{b}\| .
\end{aligned}
$$

Let $c_{1}=1 /(1-|d|)$ and $c_{2}=(1+|d|)|\gamma|(1 /(|a|-|\gamma|))((1+|d|) /(1-|d|))$, the above bound is simplified by

$$
\begin{equation*}
\left\|A_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{b}^{\prime}\right\| \leq\left(c_{1}|d|^{t_{1}+1}+c_{2}\left|\frac{\gamma}{a}\right|^{t_{2}}\right)\|\mathbf{b}\| \tag{4~d}
\end{equation*}
$$

The term $c_{1}|d|^{t_{1}+1}$ will be less than $\xi$, if $t_{1}$ is greater than $\left(\log \xi-\log c_{1}\right) / \log |d|-1$; the term $c_{2}|\gamma / a|^{t_{2}+1}$ will be less than $\xi$, if $t_{2}$ is greater than $\left(\log \xi-\log c_{2}\right) /(\log |\gamma / a|)-1$.

By (4c) and its variation, we have

$$
\begin{aligned}
\left\|\mathbf{x}^{\prime}\right\| & \leq \frac{1+|\gamma / a|}{|a|-|\gamma|}\left\|\mathbf{y}^{\prime}\right\| \\
& \leq \frac{1+|\gamma / a|}{|a|-|\gamma|} \frac{1+|d|}{1-|d|}\left\|\mathbf{b}^{\prime}\right\|
\end{aligned}
$$

Let $c_{3}=((1+|\gamma / a|) /(|a|-|\gamma|))((1+|d|) /(1-|d|))$, then it yields

$$
\left\|\mathbf{x}^{\prime}\right\| \leq c_{3}\|\mathbf{b}\|
$$

## APPENDIX 5

```
SUBROUTINE VLOW2(r,s,k)
integer k
real r,s
p=1
temp=-s/r
C*****Vector operation
    x[1:n]=(1/r)*x[1:n]
    do 10 i=0,k
C*****Compute (15) in a vectorized way
    x[1+p:n]=x[1+p:n]+temp*x[1:n-p]
        temp=temp*temp
        p=p*2
    continue
```


## APPENDIX 6

```
SUBROUTINE VUPPER2(r,s,k)
integer k
real r,s
p=1
temp=-s/r
x[1:n]=(1/r)*x[1:n]
do 10 i=0,k
        x[1:n-p]=x[1:n-p]+temp*x[1+p:n]
        temp=temp*temp
        p=p*2
continue
```

10

## APPENDIX 7

By (8), (10), and their variations, the bound of the sup-norm of the residual vector $A_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{b}^{\prime}$ isderived by

$$
\begin{aligned}
\left\|A_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{b}^{\prime}\right\| & \leq\left\|L_{m}^{\prime}\left(U_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{y}^{\prime}\right)\right\|+\left\|L_{m}^{\prime} \mathbf{y}^{\prime}-\mathbf{b}^{\prime}\right\| \\
& \leq(1+|d|)\left\|U_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{y}^{\prime}\right\|+\left\|L_{m}^{\prime} \mathbf{y}^{\prime}-\mathbf{b}^{\prime}\right\| \\
& \leq(1+|d|)\left|d^{\prime}\right|^{k_{1}+1}\left\|\mathbf{y}^{\prime}\right\|+|d|^{k^{k_{2}+1}}\left\|\mathbf{b}^{\prime}\right\| \\
& \leq(1+|d|)\left|d^{\prime}\right|^{\mathbf{k}^{k_{1}+1}} \frac{1}{1-|d|}\left\|\mathbf{b}^{\prime}\right\|+|d|^{2^{k_{2}+1}}\left\|\mathbf{b}^{\prime}\right\| \\
& \leq\left((1+|d|)\left|d^{\prime}\right|^{k^{k_{1}+1}} \frac{1}{1-|d|}+|d|^{2^{k_{2}+1}}\right)\|\mathbf{b}\| .
\end{aligned}
$$

Let $c_{1}^{\prime}=1$ and $c_{2}^{\prime}=(1+|d|) /(1-|d|)$, then the bound is simplified by

$$
\begin{equation*}
\left\|A_{m}^{\prime} \mathbf{x}^{\prime}-\mathbf{b}^{\prime}\right\| \leq\left(c_{1}^{\prime}|d|^{2^{k_{1}+1}}+c_{2}^{\prime}\left|d^{\prime}\right|^{2^{k_{2}+1}}\right)\|\mathbf{b}\| \tag{7a}
\end{equation*}
$$

The term $c_{1}^{\prime}|d|^{2^{k_{1}+1}}$ will be less than $\xi$, if $k_{1}$ is greater than $\log (\log \xi / \log |d|)-1$; the term $c_{2}^{\prime}\left|d^{\prime}\right|^{2^{k_{2}+1}}$ will be less than $\xi$, if $k_{2}$ is greater than $\log \left(\left(\log \xi-\log c_{2}^{\prime}\right) / \log |d|\right)-1$.

From (10) and its variation, we have

$$
\begin{aligned}
\left\|\mathbf{x}^{\prime}\right\| & \leq \frac{1}{|a|-|\gamma|}\left\|\mathbf{y}^{\prime}\right\| \\
& \leq \frac{1}{|a|-|\gamma|} \frac{1}{1-|d|}\|\mathbf{b}\| .
\end{aligned}
$$

Let $c_{3}^{\prime}=(1 /(|a|-|\gamma|))(1 /(1-|d|))$, then the simplified bound is given by

$$
\begin{equation*}
\left\|\mathbf{x}^{\prime}\right\| \leq c_{3}^{\prime}\|\mathbf{b}\| \tag{7b}
\end{equation*}
$$

## APPENDIX 8

```
SUBROUTINE FINAL(t3,t4)
integer t3,t4
temp=d
do 10 i=1,t3
    x[i]=x[i]-temp*s
    temp=temp*d
continue
C*****dp represents d' in the context
    temp=dp
    do 20 i=n,n+1-t4
C*****sp represents s' in the context
            x[i]=x[i]-temp*sp
            temp=temp*dp
                continue
```


## APPENDIX 9

By $\mathbf{x}=\overline{\mathbf{x}}-\mathbf{p}$, (16), (17), and (18), we have

$$
\begin{aligned}
A_{n} \mathbf{x}-\mathbf{b}= & A_{n} \overline{\mathbf{x}}-\mathbf{b}-A_{n} \mathbf{p} \\
= & A_{n} \overline{\mathbf{x}}-\mathbf{b}-w_{1} \mathbf{e}_{1}-w_{n} \mathbf{e}_{n} \\
& -s \alpha d^{t_{3}}\left(-d \frac{\gamma}{\alpha} \mathbf{e}_{t_{3}}+\mathbf{e}_{t_{3}+1}\right)-s^{\prime} \gamma d^{\prime t_{4}}\left(-d^{\prime} \frac{\alpha}{\gamma} \mathbf{e}_{n-t_{3}+1}+\mathbf{e}_{n-t_{4}}\right) .
\end{aligned}
$$

Immediately, it yields

$$
\begin{align*}
\left\|A_{n} \mathbf{x}-\mathbf{b}\right\| & \leq\left\|\mathbf{z}^{\prime}\right\|+\left\|s \alpha d^{t_{3}}\left(-d \frac{\gamma}{\alpha} \mathbf{e}_{t_{3}}+\mathbf{e}_{t_{3}+1}\right)+s^{\prime} \gamma d^{t_{4}}\left(-d^{\prime} \frac{\alpha}{\gamma} \mathbf{e}_{n-t_{4}+1}+\mathbf{e}_{n-t_{4}}\right)\right\|  \tag{9a}\\
& \leq\left(c_{1}|d|^{t_{1}+1}+c_{2}\left|\frac{\gamma}{a}\right|^{t_{2}}\right)\|\mathbf{b}\|+\left|s \alpha d^{t_{3}}\right|+\left|s^{\prime} \gamma d^{t_{4}}\right|
\end{align*}
$$

By (11) and Lemma 2, we have

$$
\begin{equation*}
\left|w_{1}\right| \leq c_{4}\|b\| \quad \text { and } \quad\left|w_{n}\right| \leq c_{5}\|b\| \tag{9b}
\end{equation*}
$$

where $c_{4}=1+\left(\left|\beta_{1}\right|+\left|\beta_{2}\right|+\left|\beta_{3}\right|\right) c_{3}$ and $c_{5}=1+\left(\left|\beta_{1}^{\prime}\right|+\left|\beta_{2}^{\prime}\right|+\left|\beta_{3}^{\prime}\right|\right) c_{3}$.
By (18) and (9b), we have

$$
\begin{equation*}
|s| \leq c_{6}\|\mathbf{b}\| \quad \text { and } \quad\left|s^{\prime}\right| \leq c_{7}\|\mathbf{b}\| \tag{9c}
\end{equation*}
$$

where $c_{6}=\left(c_{4}\left|v^{\prime}\right|+c_{5}\left|u^{\prime}\right|\right) /\left(\left|u v^{\prime}-v u^{\prime}\right|\right)$ and $c_{7}=\left(c_{5}|u|+c_{4}|v|\right) /\left(\left|u v^{\prime}-v u^{\prime}\right|\right)$. By (9a) and (9c), it follows that

$$
\begin{equation*}
\frac{\left\|A_{n} \mathbf{x}-\mathbf{b}\right\|}{\|\mathbf{b}\|} \leq\left(c_{1}|d|^{t_{1}+1}+c_{2}\left|d^{\prime}\right|^{t_{2}}+c_{8}|d|^{t_{3}}+c_{9}\left|d^{\prime}\right|^{t_{4}}\right) \tag{9~d}
\end{equation*}
$$

where $c_{8}=c_{6}|\alpha|$ and $c_{9}=c_{7}|\gamma|$.

## REFERENCES

1. J.J. Dongarra, C.B. Moler, J.R. Bunch and G.W. Stewart, LINPACK User's Guide, SIAM Press, Pennsylvania, (1979).
2. D. Fischer, G. Golub, O. Hald, C. Levia and O. Widlund, On Fourier-Toeplitz methods for separable elliptic problems, Mathematics of Computation 28 (126), 349-368 (1974).
3. O.B. Widlund, On the use of fast methods for separable finite difference equations for the solution of general elliptic problems, In Sparse Matrices and Their Applications, (Edited by D.J. Rose and R.A. Willoughby), pp. 121-131, Plenum Press, New York, (1972).
4. G.D. Smith, Numerical Solution of Partial Differential Equations: Finite Difference Methods, Oxford University Press, (1985).
5. R. Roberts and C.T. Mullis, Digital Signal Processing, Addison-Wesley, Reading, MA, (1987).
6. R.W. Hockney and C.R. Jesshope, Parallel Computers 2, Adam Hilger, Bristol, (1988).
7. R.F. Boisvert, Algorithms for special tridiagonal systems, SIAM J. Sci. Stat. Comput., 12, 423-442 (1991).
8. J. Ortega, Introduction to Parallel and Vector Solution of Linear Systems, Plenum Press, New York, (1988).
9. J.J. Dongarra, I.S. Duff, D.C. Sorensen and H.A. van der Vorst, Solving Linear Systems on Vector and Shared Memory Computers, SIAM Press, Pennsylvania, (1991).
10. G. Golub and J. Ortega, Scientific Computing: An Introduction to Parallel Computing, Academic Press, New York, (1993).
11. Supercomputer Programming (I): Advanced Fortran: Architecture, Vectorization, and Parallel Computing, Working manual for CRAY X-MP EA/116se, (1991).
12. O. Axelsson and V. Eijkhout, A note on the vectorization of scalar recursions, Parallel Comput. 3, 73-83 (1986).
13. H.S. Stone, An efficient parallel algorithm for the solution of a tridiagonal linear system of equations, J. ACM. 20, 27-38 (1973).
14. R.H. Bartels, J.C. Beatty and B.A. Barsky, An Introduction to Splines for Use in Computer Graphics and Geometric Modeling, Morgan Kaufmann, California, (1987).
15. K.L. Chung and L.J. Shen, Vectorized algorithm for B-spline curve fitting on CRAY X-MP EA/16se, Presented at the IEEE Conf. on Supercomputing, November 1992, pp. 166-169.
16. B. Pham, Quadratic B-splines for automatic curve and surface fitting, Comput. \& Graphics 13, 471-475 (1989).
17. K.L. Chung and W.M. Yan, Computing quadratic B-spline curve fitting on CRAY X-MP, In Proc. National Computer Symposium, Chiayi, R.O.C., December 1993, pp. 401-407.
18. K.L. Chung, W.M. Yan and J.G. Wu, A parallel algorithm for solving special tridiagonal systems on ring networks, Computing 56 (4), 385-395 (1996).
19. K.L. Chung and W.M. Yan, Parallel B-spline surface fitting on mesh-connected computers, J. Parallel and Distributed Computing 35, 205-210 (1996).
20. R.W. Hockney, A fast direct solution of Poisson's equation using Fourier analysis, J. ACM. 12, 95-113 (1965).
21. H.H. Wang, A parallel method for tridiagonal equations, ACM Trans. Math. Software 7 (2), 170-183 (1981).

[^0]:    *Author to whom all correspondence should be addressed.
    The authors would like to thank the help of the referees and Prof. E. Y. Rodin which led to an improved version of this paper. This research was supported in part by the National Science Council of R.O.C. under Grants NSC85-2213-E011-009, NSC85-2121-M011-002, and NCHC86-08-015.

