http://www.diva-portal.org

Postprint

This is the accepted version of a paper presented at *10th Nordic International Conference on Applied Parallel Computing - State of the Art in Scientific and Parallel Computing (PARA), JUN 06-09, 2010, Univ Iceland, Sch Engn & Nat Sci, Reykjavik, ICELAND*.

Citation for the original published paper:

Kjelgaard Mikkelsen, C C., Kågström, B. (2012)
Parallel solution of narrow banded diagonally dominant linear systems.
In: Kristján Jónasson (ed.), *Applied Parallel and Scientific Computing, Pt II* (pp. 280-290).
Springer Berlin/Heidelberg
Lecture notes in computer science
https://doi.org/10.1007/978-3-642-28145-7_28

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-51059

Parallel Solution of Narrow Banded Diagonally Dominant Linear Systems

Carl Christian Kjelgaard Mikkelsen and Bo Kågström

Department of Computing Science and HPC2N
Umeå University, Sweden {spock,bokg}@cs.umu.se

Abstract. ScaLAPACK contains a pair of routines for solving systems which are narrow banded and diagonally dominant by rows. Mathematically, the algorithm is block cyclic reduction. The ScaLAPACK implementation can be improved using incomplete, rather than complete block cyclic reduction. If the matrix is strictly dominant by rows, then the truncation error can be bounded directly in terms of the dominance factor and the size of the partitions. Our analysis includes new results applicable in our ongoing work of developing an efficient parallel solver.

Keywords: Narrow banded, diagonally dominant linear systems, block cyclic reduction, parallel algorithms, ScaLAPACK

1 Introduction

Let $A = [a_{ij}]$ be a real n by n matrix and consider the solution of the linear system

$$Ax = f,$$

where $f \in \mathbb{R}^n$. The matrix A is banded with bandwidth $2k+1$ if

$$|i-j| > k \Rightarrow a_{ij} = 0.$$

The matrix A is diagonally dominant by rows if $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$ for all i. If the inequality is sharp, then A is strictly diagonally dominant by rows. If A is diagonally dominant by rows and nonsingular, then $a_{ii} \neq 0$ and the dominance factor ϵ is defined by

$$\epsilon = \max_i \left\{ \frac{1}{|a_{ii}|} \sum_{j \neq i} |a_{ij}| \right\} \in [0,1].$$

Narrow banded linear systems which are strictly diagonally dominant can be found throughout the physical sciences. In particular, the solution of parabolic PDEs using compact finite difference methods is a rich source of examples.

In general, we cannot assume $\epsilon \ll 1$. However, in this paper we argue that if ϵ is not too close to 1, then incomplete cyclic reduction becomes a viable alternative to the ScaLAPACK algorithm [2], which is a special case of block cyclic reduction.

We begin by stating a few results on matrices which are strictly diagonally dominant by rows in Section 2. We review the cyclic reduction algorithm by R. W. Hockney and G. H. Golub [4] and provide an elementary extension of a result by Heller [3] in Section 3. We state and prove our main results in Section 4.

This paper builds on the analysis of the truncated SPIKE algorithm by Mikkelsen and Manguoglu [5] and it requires a good understanding of the routines PDDBTRF/PDDBTRS [2, 1] from ScaLAPACK, as well as the work of Heller [3]. The truncated SPIKE algorithm (introduced by Polizzi and Sameh [6, 7]) also applies to systems which are banded and strictly diagonally dominant by rows.

2 Basic Properties

The following results (proved in [5]) are central to our analysis.

Lemma 1. *Let A be an m by m matrix which is strictly diagonally dominant by rows with dominance factor ϵ. Let $A = LU$ be the LU factorization of A. Then U is strictly diagonally dominant by rows and ϵ_U, the dominance factor of U, satisfies $\epsilon_U \leq \epsilon$.*

Lemma 2. *Let A be an m by m matrix and let B be an m by n matrix such that $[A, B]$ is strictly diagonally dominant by rows with dominance factor ϵ. Then $C = [I, A^{-1}B]$ is strictly diagonally dominant by rows and ϵ_C, the dominance factor of C, satisfies $\epsilon_C \leq \epsilon$.*

Lemma 3. *Let A be a banded matrix with bandwidth $2k + 1$ which is strictly diagonally dominant by rows with dominance factor ϵ, and let A be partitioned in the block tridiagonal form*

$$A = \begin{bmatrix} D_1 & F_1 & & \\ E_2 & \ddots & \ddots & \\ & \ddots & \ddots & F_{m-1} \\ & & E_m & D_m \end{bmatrix} \tag{1}$$

with block size $\mu = qk$ for an integer $q > 0$. Moreover, let $\begin{bmatrix} U_i \ V_i \end{bmatrix}$ be the solution of the linear system

$$D_i \begin{bmatrix} U_i \ V_i \end{bmatrix} = \begin{bmatrix} E_i \ F_i \end{bmatrix},$$

where E_1, U_1, F_m, and V_m are undefined and should be treated as zero, and let

$$U_i = \left(U_{i,1}^T, U_{i,2}^T, \dots, U_{i,q}^T\right)^T, \quad V_i = \left(V_{i,1}^T, V_{i,2}^T, \dots, V_{i,q}^T\right)^T$$

be a partitioning of U_i and V_i into blocks each consisting of k rows. Then

$$\|U_{i,j}\|_\infty \leq \epsilon^j, \quad \|V_{i,j}\|_\infty \leq \epsilon^{q-(j-1)}, \quad j = 1, 2, \dots, q.$$

3 Block Cyclic Reduction

Mathematically, the algorithm used by PDDBTRF/PDDBTRS is a special case of block cyclic reduction [4] which we briefly review below. In addition, we present an extension of a relevant result by D. Heller on incomplete block cyclic reduction [3].

Let A be an m by m block tridiagonal matrix in the form (1) which is also strictly diagonally dominant by rows, and let D be the matrix given by

$$D = \text{diag}(D_1, D_2, \ldots, D_m)$$

and consider the auxiliary matrix B defined by

$$B = D^{-1}(A - D) = \begin{bmatrix} 0 & D_1^{-1}F_1 & & & \\ D_2^{-1}E_2 & 0 & D_2^{-1}F_2 & & \\ & D_3^{-1}E_3 & \ddots & \ddots & \\ & & \ddots & \ddots & D_{m-1}^{-1}F_{m-1} \\ & & & D_m^{-1}E_m & 0 \end{bmatrix}.$$

The norm of the matrix B measures the significance of the off diagonal blocks of A. Specifically, let $f \in \mathbb{R}^n$ and let x and y be the solutions of the linear systems

$$Ax = f, \quad Dy = f.$$

Then,

$$x - y = x - D^{-1}Ax = (I - D^{-1}A)x = D^{-1}(D - A)x = -Bx,$$

which for all $x \neq 0$ implies that

$$\frac{\|x - y\|_\infty}{\|x\|_\infty} \leq \|B\|_\infty.$$

Therefore, if $\|B\|_\infty$ is sufficiently small, then y is a good approximation of x. The linear systems $D_i y_i = f_i$, $i = 1, 2, \ldots, m$ can be solved concurrently on different processors without any communication. Therefore, the block diagonal linear system $Dy = f$ is even more suitable for parallel computing than the original linear system $Ax = f$.

We illustrate block cyclic reduction in the case of $m = 7$. Let P denote the matrix which represents the usual odd-even permutation σ of the blocks, i.e.

$$\sigma = (1, 3, 5, 7, 2, 4, 6),$$

and define A' by

$$A' := PAP^T = \left[\begin{array}{cccc|ccc} D_1 & & & & F_1 & & \\ & D_3 & & & E_3 & F_3 & \\ & & D_5 & & & E_5 & F_5 \\ & & & D_7 & & & E_7 \\ \hline E_2 & F_2 & & & D_2 & & \\ & E_4 & F_4 & & & D_4 & \\ & & E_6 & F_6 & & & D_6 \end{array}\right] = \left[\begin{array}{c|c} A'_{11} & A'_{12} \\ \hline A'_{21} & A'_{22} \end{array}\right].$$

The Schur complement of A' is the block tridiagonal matrix $A^{(1)}$ given by

$$A^{(1)} := A'_{11} - A'_{21} {A'}_{22}^{-1} A'_{12} = \begin{bmatrix} D_1^{(1)} & F_1^{(1)} & \\ E_2^{(1)} & D_2^{(1)} & F_2^{(1)} \\ & E_3^{(1)} & D_3^{(1)} \end{bmatrix},$$

where

$$D_i^{(1)} = D_{2i} - E_{2i} D_{2i-1}^{-1} F_{2i-1} - F_{2i} D_{2i+1}^{-1} E_{2i+1}$$

and

$$E_i^{(1)} = -E_{2i} D_{2i-1}^{-1} E_{2i-1}, \quad F_i^{(1)} = -F_{2i} D_{2i+1}^{-1} F_{2i+1}.$$

Heller [3] showed that if A is strictly diagonally dominant by rows, then block cyclic reduction is well defined and the new auxiliary matrix

$$B^{(1)} = {D^{(1)}}^{-1} (A^{(1)} - D^{(1)}),$$

with $D^{(1)} = \mathrm{diag}(D_1^{(1)}, D_2^{(1)}, D_3^{(1)})$, satisfies

$$\|B^{(1)}\|_\infty \leq \|B^2\|_\infty \leq \|B\|_\infty^2.$$

In addition, Heller [3] showed that if A is strictly diagonally dominant by rows, then the initial matrix B satisfies $\|B\|_\infty < 1$ and the significance of the off diagonal blocks decays quadratically to zero.

We have found that it is possible to explicitly incorporate the dominance factor into the analysis. For the sake of notational simplicity we define U_i and V_i as the solution of the linear system

$$D_i \begin{bmatrix} U_i & V_i \end{bmatrix} = \begin{bmatrix} E_i & F_i \end{bmatrix},$$

where E_1, U_1, F_m, and V_m are undefined and should be treated as zero. It follows that

$$\|B^{(1)}\|_\infty = \max \|Z_i\|_\infty,$$

where

$$Z_i = (I - U_{2i} V_{2i-1} - V_{2i} U_{2i+1})^{-1} \left[U_{2i} U_{2i-1},\ V_{2i} V_{2i+1} \right].$$

Therefore,

$$\begin{aligned} Z_i &= \begin{bmatrix} U_{2i} & V_{2i} \end{bmatrix} \begin{bmatrix} V_{2i-1} \\ U_{2i+1} \end{bmatrix} Z_i + \begin{bmatrix} U_{2i} & V_{2i} \end{bmatrix} \begin{bmatrix} U_{2i-1} & 0 \\ 0 & V_{2i+1} \end{bmatrix} \\ &= \begin{bmatrix} U_{2i} & V_{2i} \end{bmatrix} \left[\begin{array}{c|cc} V_{2i-1} & U_{2i-1} & 0 \\ U_{2i+1} & 0 & V_{2i+1} \end{array} \right] \left[\begin{array}{c} Z_i \\ \hline I \end{array} \right]. \end{aligned}$$

The right hand side can be estimated using Lemma 2. We have

$$\|Z_i\|_\infty \leq \epsilon^2 \max\{\|Z_i\|_\infty, 1\}. \tag{2}$$

However, if we assume $\|Z_i\|_\infty \geq 1$, then (2) reduces to

$$\|Z_i\|_\infty \leq \epsilon^2 \|Z_i\|_\infty$$

which forces the contradiction $\|Z_i\|_\infty = 0$, simply because $\epsilon < 1$. Therefore, $\|Z_i\|_\infty < 1$, which inserted into (2) yields

$$\|Z_i\|_\infty \leq \epsilon^2.$$

It follows, that

$$\|B^{(1)}\|_\infty \leq \epsilon^2. \tag{3}$$

This estimate is tight and equality is achieved for matrices of the form

$$A = \begin{bmatrix} I_k & \epsilon I_k & & \\ & \ddots & \ddots & \\ & & \ddots & \epsilon I_k \\ & & & I_k \end{bmatrix}.$$

4 Preliminary Analysis of the ScaLAPACK Routine PDDBTRF

The ScaLAPACK routine PDDBTRF can be used to obtain a factorization of a narrow banded matrix A which is diagonally dominant by rows, [1].

Mathematically, the algorithm is block cyclic reduction applied to a special partitioning of the matrix, which is designed to exploit the banded structure. Specifically, the odd numbered blocks are very large, say, of dimension $\mu = qk$, where $q \gg 1$ is a large positive integer, while the even numbered blocks have dimension k.

The large odd numbered diagonal blocks can be factored in parallel without any communication. It is the construction and factorization of the Schur complement $A^{(1)}$ which represents the parallel bottleneck. Obviously, the factorization can be accelerated, whenever the off diagonal blocks can be ignored. Now, while we do inherit the estimate

$$\|B^{(1)}\|_\infty \leq \epsilon^2$$

from the previous analysis, this estimate does not take the banded structure into account. We have the following theorem.

Theorem 1. *Let A be a tridiagonal matrix which is strictly diagonally dominant by rows. Then the significance of the off diagonal blocks of the initial Schur complement is bounded by*

$$\|B^{(1)}\|_\infty \leq \epsilon^{1+q}.$$

Proof. We must show that Z_i given by

$$Z_i = (I - U_{2i}V_{2i-1} - V_{2i}U_{2i+1})^{-1} \left[U_{2i}U_{2i-1},\ V_{2i}V_{2i+1}\right]$$

satisfies

$$\|Z_i\|_\infty \leq \epsilon^{1+q}.$$

We solve this optimization problem by partitioning it into two subproblems, which can be solved by induction.

We begin by making the following very general estimate

$$\|Z_i\|_\infty \leq \frac{\|U_{2i}\|_\infty \|U_{2i-1}^{(b)}\|_\infty + \|V_{2i}\|_\infty \|V_{2i+1}^{(t)}\|_\infty}{1 - \|U_{2i}\|_\infty \|V_{2i-1}^{(b)}\|_\infty - \|V_{2i}\|_\infty \|U_{2i+1}^{(t)}\|_\infty},$$

where the notation $U^{(t)}$ ($U^{(b)}$) is used to identify the matrix consisting of the top (bottom) k rows of the matrix U. This estimate is easy to verify, but it relies on the zero structure of the matrices U_{2i} and V_{2i}. Now, let

$$\alpha = \|U_{2i-1}^{(b)}\|_\infty, \quad \beta = \|V_{2i-1}^{(b)}\|_\infty, \quad \gamma = \|U_{2i+1}^{(t)}\|_\infty, \quad \delta = \|V_{2i+1}^{(t)}\|_\infty,$$

and define an auxiliary function

$$g(x, y) = \frac{\alpha x + \delta y}{1 - \beta x - \gamma y},$$

where the appropriate domain will be determined shortly. If

$$x = \|U_{2i}\|_\infty, \qquad y = \|V_{2i}\|_\infty$$

then by design

$$\|Z_i\|_\infty \leq g(x, y).$$

In general, Lemma 1 implies that

$$\|[U_{2i}, V_{2i}]\|_\infty \leq \epsilon$$

but in the current case of $k = 1$, this follows directly from the definition of strict diagonal dominance. Regardless, we see that the natural domain for g is the closure of the set Ω given by

$$\Omega = \{(x, y) \in \mathbb{R}^2 \, : \, 0 < x \, \wedge \, 0 < y \, \wedge x + y < \epsilon\}.$$

It suffices to show that $g(x, y) \leq \epsilon^q$ for all $(x, y) \in \bar{\Omega}$. It is clear, that g is well defined and $g \in C^\infty(\bar{\Omega})$, simply because $\beta \leq \epsilon$, $\gamma \leq \epsilon$ and $\epsilon < 1$, so that we never divide by zero.

Now, does g assume its maximum within Ω? We seek out any stationary points. We have

$$\frac{\partial g}{\partial x}(x, y) = \frac{\alpha - (\alpha\gamma - \beta\delta)y}{(1 - \beta x - \gamma y)^2}, \quad \text{and} \quad \frac{\partial g}{\partial y}(x, y) = \frac{\delta + (\alpha\gamma - \beta\delta)x}{(1 - \beta x - \gamma y)^2}.$$

Therefore, there are now two distinct scenarios, namely

$$\alpha\gamma - \beta\delta = 0 \quad \text{or} \quad \alpha\gamma - \beta\delta \neq 0.$$

If $\alpha\gamma - \beta\delta = 0$, then there are no stationary points, unless $\alpha = \delta = 0$, in which case $g \equiv 0$ and there is nothing to prove. If $\alpha\gamma - \beta\delta \neq 0$, then

$$(x_0, y_0) = \left(\frac{\alpha}{\alpha\gamma - \beta\delta}, -\frac{\delta}{\alpha\gamma - \beta\delta} \right)$$

is the only candidate, but $(x_0, y_0) \notin \Omega$, simply because

$$x_0 y_0 = -\frac{\alpha\delta}{(\alpha\gamma - \beta\delta)^2} \leq 0$$

is not strictly positive. In both cases, we conclude that the global maximum for g is assumed on the boundary of Ω.

The boundary of Ω consists of three line segments. We examine them one at a time. We begin by defining

$$g_1(y) = g(0, y) = \frac{\delta y}{1 - \gamma y}, \quad y \in [0, \epsilon].$$

Then

$$g_1'(y) = \frac{\delta(1 - \gamma y) - \delta y(-\gamma)}{(1 - \gamma y)^2} = \frac{\delta}{(1 - \gamma y)^2} \geq 0$$

and we conclude that

$$g_1(y) \leq g_1(\epsilon) = g(0, \epsilon) = \frac{\delta\epsilon}{1 - \gamma\epsilon}.$$

Similarly, we define

$$g_2(x) = g(x, 0) = \frac{\alpha x}{1 - \beta x}, \quad x \in [0, \epsilon].$$

Then

$$g_2'(x) = \frac{\alpha(1 - \beta x) - \alpha x(-\beta)}{(1 - \beta x)^2} = \frac{\alpha}{(1 - \beta x)^2} \geq 0$$

which allows us to conclude that

$$g_2(x) \leq g_2(\epsilon) = \frac{\epsilon\alpha}{1 - \epsilon\beta}.$$

Finally, we let $s \in [0, \epsilon]$ and define

$$g_3(s) = g(s, \epsilon - s) = \frac{\alpha s + \delta(\epsilon - s)}{1 - \beta s - \gamma(\epsilon - s)} = \frac{(\alpha - \delta)s + \delta\epsilon}{1 - \gamma\epsilon - (\beta - \gamma)s}.$$

Then

$$g_3'(s) = \frac{(\alpha - \delta)(1 - \gamma\epsilon) + \delta\epsilon(\beta - \gamma)}{(1 - \gamma\epsilon - (\beta - \gamma)s)^2}.$$

Therefore, g_3 is either a constant or strictly monotone. In either case

$$g_3(s) \leq \max\{g_3(0), g_3(\epsilon)\}.$$

We can now conclude that for all $(x,y) \in \bar{\Omega} : g(x,y) \leq \max\{g(\epsilon,0), g(0,\epsilon)\}$. We will only show that

$$g(\epsilon,0) = \frac{\alpha\epsilon}{1-\beta\epsilon} = \frac{\epsilon\|U_{2i-1}^{(b)}\|_\infty}{1-\epsilon\|V_{2i-1}^{(b)}\|_\infty} \leq \epsilon^{1+q}$$

simply because the other case is similar. The proof is by induction on q, i.e. the size of the odd numbered partitions. Let

$$\left[\begin{array}{c|cccc|c} c_1 & a_1 & b_1 & & & \\ & c_2 & a_2 & b_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & c_q & a_q & b_q \end{array}\right]$$

be a representation of the $(2i-1)$th block row of the original matrix A. Using Gaussian elimination without pivoting we obtain the matrix

$$\left[\begin{array}{c|ccccc|c} c_1' & a_1' & b_1' & & & & \\ c_2' & & a_2' & b_2 & & & \\ c_3' & & & a_3' & b_3 & & \\ \vdots & & & & \ddots & \ddots & \\ c_q' & & & & & a_q' & b_q \end{array}\right].$$

Now, let V be the set given by

$$V = \left\{ j \in \{1,2,\ldots,q\} \,:\, \frac{\epsilon|c_j'|/|a_j'|}{1-\epsilon|b_j|/|a_j'|} \leq \epsilon^{1+j}, \quad j = 1,2,\ldots,q \right\}.$$

We claim that $V = \{1,2,\ldots,q\}$. We begin by showing that $1 \in V$. Let

$$x = |c_1'|/|a_1'|, \quad y = |b_1|/|a_1'|.$$

Then $(x,y) \in \bar{\Omega}$ and it is straightforward to show that

$$\frac{\epsilon x}{1-\epsilon y} \leq \epsilon^2.$$

Now, suppose that $j \in V$ for some $j < q$. Does $j+1 \in V$? We have

$$a_{j+1}' = a_{j+1} - c_{j+1}\frac{b_j}{a_j'}, \quad c_{j+1}' = -c_{j+1}\frac{c_j'}{a_j'},$$

which implies

$$\begin{aligned} \frac{\epsilon|c_{j+1}'|/|a_{j+1}'|}{1-\epsilon|b_{j+1}|/|a_{j+1}'|} &= \frac{\epsilon|c_{j+1}'|}{|a_{j+1}'|-\epsilon|b_{j+1}|} = \frac{\epsilon|c_{j+1}|(|c_j'|/|a_j'|)}{|a_{j+1}-c_{j+1}(b_j/a_j')|-\epsilon|b_{j+1}|} \\ &\leq \frac{\epsilon|c_{j+1}|(|c_j'|/|a_j'|)}{|a_{j+1}|-|c_{j+1}||b_j|/|a_j'|-\epsilon|b_{j+1}|} \\ &= \frac{\epsilon(|c_{j+1}|/|a_{j+1}|)(|c_j'|/|a_j'|)}{1-(|c_{j+1}|/|a_{j+1}|)(|b_j|/|a_j'|)-\epsilon(|b_{j+1}|/|a_{j+1}|)}. \end{aligned}$$

We simplify the notation by introducing

$$\nu = |c_j'|/|a_j'|, \quad \mu = |b_j|/|a_j'|, \quad x = |c_{j+1}|/|a_{j+1}|, \quad y = |b_{j+1}|/|a_{j+1}|,$$

and defining

$$h(x,y) = \frac{\epsilon x \nu}{1 - x\mu - \epsilon y}.$$

By the strict diagonal dominance of A, we have $(x,y) \in \bar{\Omega}$ and it is easy to see that $h(x,y) \leq h(\epsilon, 0)$. Therefore

$$\frac{\epsilon x \nu}{1 - x\mu - \epsilon y} \leq \frac{\epsilon^2 \nu}{1 - \epsilon\mu} = \epsilon \left(\frac{\epsilon \nu}{1 - \epsilon \mu} \right) \leq \epsilon \cdot \epsilon^{1+j} = \epsilon^{1+(j+1)}$$

and $j+1 \in V$. By the well ordering principle, $V = \{1, 2, \ldots, q\}$.

In view of Theorem 1 we make the following conjecture.

Conjecture 1. The auxiliary matrix corresponding to the initial Schur complement generated by the ScaLAPACK routine PDDBTRF satisfies

$$\|B^{(1)}\|_\infty \leq \epsilon^{1+q},$$

where $\mu = qk$ is the size of the odd number partitions and $\epsilon < 1$ is the dominance factor.

This is one possible generalization of the case $q = 1$ to the case $q > 1$ and it does reduce to Theorem 1 in the case of $k = 1$. The proof of Conjecture 1 for the general case is ongoing work. So far, we have derived the following results.

Theorem 2. *If A is strictly diagonally dominant by rows and banded with bandwidth $(2k+1)$ then*

$$\|B^{(1)}\|_\infty \leq \frac{\epsilon^{1+q}}{1 - \epsilon^2}.$$

Proof. By definition

$$Z_i = [U_{2i},\ V_{2i}] \left\{ \begin{bmatrix} V_{2i-1} \\ U_{2i+1} \end{bmatrix} Z_i + \begin{bmatrix} \tilde{U}_{2i-1} & 0 \\ 0 & \tilde{V}_{2i+1} \end{bmatrix} \right\}.$$

Therefore

$$\|Z_i\|_\infty \leq \epsilon \{ \epsilon \|Z_i\|_\infty + \epsilon^q \} = \epsilon^2 \|Z_i\|_\infty + \epsilon^{1+q}$$

and the proof follows immediately from the fact that $\epsilon < 1$.

It is the elimination of the singularity at $\epsilon = 1$ which is proving difficult in the case of $k > 1$. Specifically, it is the decomposition into two separate subproblems which is difficult to achieve for $k > 1$.

In the case of matrices which are both banded and triangular Conjecture 1 is trivially true.

Theorem 3. *If A is a strictly upper (lower) triangular banded matrix with dominance factor ϵ and upper (lower) bandwidth k, then*

$$\|B^{(1)}\|_\infty \leq \epsilon^{1+q},$$

where $\mu = qk$ is the size of the odd numbered partitions.

In general, we have the following theorem.

Theorem 4. *Dropping the off diagonal blocks in the initial Schur complement is equivalent to replacing the original matrix A with a perturbed matrix $A + \Delta A$ for which*

$$\|\Delta A\|_\infty \leq \epsilon^{1+q}\|A\|_\infty.$$

Proof. By definition

$$\left[E_i^{(1)},\, F_i^{(1)}\right] = -\left[E_{2i}D_{2i-1}^{-1}E_{2i-1},\, F_{2i}D_{2i+1}^{-1}F_{2i+1}\right] = -\left[E_{2i}U_{2i-1},\, F_{2i}V_{2i+1}\right].$$

However, the zero structure of E_{2i} and F_{2i} implies that

$$\left[E_{2i}U_{2i-1},\, F_{2i}V_{2i+1}\right] = \left[E_{2i}\tilde{U}_{2i-1},\, F_{2i}\tilde{V}_{2i+1}\right],$$

where

$$\tilde{U}_{2i-1} = \begin{bmatrix} 0 \\ U_{2i-1}^{(b)} \end{bmatrix}, \quad \tilde{V}_{2i+1} = \begin{bmatrix} V_{2i+1}^{(t)} \\ 0 \end{bmatrix}$$

isolate the bottom k rows of U_{2i-1} and the top k rows of V_{2i+1}. By Lemma 3

$$\|U_{2i-1}^{(b)}\|_\infty \leq \epsilon^q, \quad \|V_{2i+1}^{(t)}\|_\infty \leq \epsilon^q.$$

It follows that

$$\left[E_i^{(1)},\, F_i^{(1)}\right] = -D_{2i}\left[U_{2i},\, V_{2i}\right]\begin{bmatrix} \tilde{U}_{2i-1} & 0 \\ 0 & \tilde{V}_{2i+1} \end{bmatrix}$$

which implies

$$\left\|\left[E_i^{(1)},\, F_i^{(1)}\right]\right\|_\infty \leq \epsilon^{1+q}\|D_{2i}\|_\infty \leq \epsilon^{1+q}\|A\|_\infty$$

and the proof is complete.

Conjecture 1 is interesting in precisely those cases where the relative backward error bound given in Theorem 4 is small, but not small enough, to satisfy the demands of the user. If the conjecture is correct, then the sequence of Schur complements generated by the ScaLAPACK algorithm will satisfy

$$\|B^{(k)}\|_\infty \leq (\epsilon^{1+q})^{2^{k-1}}, \quad k = 1, 2, \ldots$$

Therefore, if ϵ is not too close to 1, then a few steps of cyclic reduction will permit us to drop the off diagonal blocks, thus facilitating a parallel solve. In addition, if we increase the number of processors by a factor of 2, then we must replace q with $q' \approx q/2$, but the accuracy can be maintained by executing a single extra step of cyclic reduction.

5 Future Work

We have shown that incomplete cyclic reduction is applicable to tridiagonal ($k = 1$) linear systems which are diagonally dominant by rows and we identified the worst case behavior. Ongoing work includes extending our analysis to the general case of $k > 1$. In addition, we are developing a parallel implementation of incomplete, rather than complete cyclic reduction for narrow banded systems. The factorization phase will feature an explicit calculation of the auxiliary matrices, in order to determine the minimal number of reduction steps necessary to achieve a given accuracy.

Acknowledgments. The authors would like to thank Lars Karlsson as well as Meiyue Shao from the Umeå Research Group in Parallel and Scientific Computing for their input during several related discussions. The work is supported by eSSENCE, a collaborative e-Science programme funded by the Swedish Research Council within the framework of the strategic research areas designated by the Swedish Government. In addition, support has been provided by the Swedish Foundation for Strategic Research under the frame program A3 02:128 and the EU Mål 2 project UMIT.

References

1. Arbenz, P., Cleary, A., Dongarra, J., Hegland, M.: A comparison of parallel solvers for diagonally dominant and general narrow-banded linear systems. In: EuroPar '99 Parallel Processing, P. Amestoy, Ph. Berger, M. Dayd, I. Duff., V. Frayss, L. Giraud, D. Ruiz, eds., pp. 1078-1087, Springer, Berlin (1999)
2. Blackford, L.S., Choi, J., Cleary, A., D'Azevedo, E., Demmel, J., Dhillon, I., Dongarra, J., Hammarling, S., Henry, G., Petitet, A., Stanley, K., Walker, D., Whaley, R. C.: ScaLAPACK User's Guide, SIAM, USA (1997)
3. Heller, D.: Some aspects of the cyclic reduction algorithm for block tridiagonal linear systems. SIAM J. Numer. Anal., vol. 13, no. 4, pp. 484-496 (1976)
4. Hockney, R.W.: A fast direct solution of Poisson's equation using Fourier analysis. J. Assoc. Comput. Mach., vol. 12, no. 1, pp. 95-113 (1965)
5. Mikkelsen, C.C.K., Manguoglu, M.: Analysis of the truncated SPIKE algorithm. SIAM J. Matrix Anal. Appl., vol. 30, no. 4, pp. 1500-1519 (2008)
6. Polizzi, E., Sameh, A.: A parallel hybrid banded system solver: The SPIKE algorithm. Parallel Comput., vol. 32, no. 2, pp. 177-194 (2006)
7. Polizzi, E., Sameh, A.: SPIKE: A parallel environment for solving banded linear systems. Comput. Fluids, vol. 36, no. 1, pp. 113-120 (2007)