

---

# The explicit SPIKE algorithm: Iterative solution of the reduced system

Carl Christian Kjelgaard Mikkelsen<sup>1</sup>

Department of Computing Science and HPC2N  
Umeå University  
Sweden  
spock@cs.umu.se

*Dedicated to Ahmed Sameh on the occasion of his 70th birthday*

**Summary.** The explicit SPIKE algorithm applies to narrow banded linear systems which are strictly diagonally dominant by rows. The parallel bottleneck is the solution of the so-called reduced system which is block tridiagonal and strictly diagonally dominant by rows. The reduced system can be solved iteratively using the truncated reduced system matrix as a preconditioner. In this paper we derive a tight estimate for the quality of this preconditioner.

## 1 Introduction

A matrix  $A = [a_{ij}] \in \mathbb{R}^{n \times n}$  is diagonally dominant by rows if

$$\forall i : \sum_{j \neq i} |a_{ij}| \leq |a_{ii}|.$$

If the inequality is sharp, then  $A$  is strictly diagonally dominant by rows. If  $A$  is nonsingular and diagonally dominant by rows or if  $A$  is strictly diagonally dominant by rows, then  $a_{ii} \neq 0$ , and the dominance factor  $\epsilon$  given by

$$\epsilon = \max_i \left\{ \frac{\sum_{j \neq i} |a_{ij}|}{|a_{ii}|} \right\}$$

is well defined.

The matrix  $A$  has lower bandwidth  $b_l$  if  $a_{ij} = 0$  for  $i > j + b_l$  and upper bandwidth  $b_u$  if  $a_{ij} = 0$  for  $j > i + b_u$ . If  $b = \max\{b_l, b_u\} \ll n$ , then we say that  $A$  is narrow banded. Every square banded matrix can be partitioned as a block tridiagonal matrix with square diagonal blocks, i.e.

$$A = \begin{bmatrix} A_1 & C_1 & & & \\ B_1 & \ddots & \ddots & & \\ & \ddots & \ddots & C_{m-1} & \\ & & & B_m & A_m \end{bmatrix}, \quad (1)$$

only the dimension of each diagonal block must be bounded from below by  $b$ . In particular, we do not have to choose the same dimension for each diagonal block, even in the exceptional case where  $b$  divides  $n$ .

The SPIKE algorithms are designed to solve banded systems on a parallel machine. The central idea was introduced by Sameh and Kuck [7] who considered the tridiagonal case and Chen, Kuck and Sameh [1] who studied the triangular case. Lawrie and Sameh [3] applied the algorithm to the symmetric positive definite case while Dongarra and Sameh [2] considered the diagonally dominant case. Polizzi and Sameh [8, 9] introduced the truncated SPIKE algorithm for systems which are strictly diagonally dominant by rows. Recently, Manguoglu, Sameh, and Schenk [4] have combined PARDISO with SPIKE in the PSPIKE package.

The explicit SPIKE algorithm by Dongarra and Sameh [2] can be used to solve narrow banded linear systems which are strictly diagonally dominant by rows. The algorithm extends naturally to systems which are block tridiagonal. Moreover, the analysis is simplified if we focus on the number of diagonal blocks, rather than the bandwidth of the matrix.

In Section 2 we state the explicit SPIKE algorithm for systems which are block tridiagonal and strictly diagonally dominant by rows. The parallel bottleneck is the solution of a reduced system which is block tridiagonal and strictly diagonally dominant by rows. The reduced system can be solved iteratively using the main block diagonal as a preconditioner. We derive a tight estimate for the quality of this preconditioner in Section 3. This is a special case of a more general theorem by Mikkelsen [6].

## 2 The explicit SPIKE algorithm

In this section we state the explicit SPIKE algorithm for systems which are block tridiagonal and strictly diagonally dominant by rows. The validity and the basic analysis of the algorithm hinges on the following lemma.

**Lemma 1.** *Let  $G = [E, D, F]$  be a matrix such that  $[D, E, F]$  is strictly diagonally dominant by rows with dominance factor  $\epsilon$ . Then  $G$  is row equivalent to a unique matrix  $K = [U, I, V]$ . Moreover, the matrix  $[U, V]$  satisfies*

$$\| [U, V] \|_{\infty} \leq \epsilon.$$

*Proof.* Mikkelsen and Manguoglu [5] contains an elementary proof.

Now consider the solution of a block tridiagonal linear system

$$Ax = f$$

on a parallel machine with  $p$  processors. Given a small tolerance  $\delta > 0$ , we shall now seek an approximation  $y$ , such that the forward error satisfies

$$[A|f] = \left[ \begin{array}{ccc|ccc|ccc} A_1^{(1)} & C_1^{(1)} & & & & & & & f_1^{(1)} \\ B_2^{(1)} & \ddots & \ddots & & & & & & \vdots \\ & \ddots & \ddots & C_{q-1}^{(1)} & & & & & \vdots \\ & & B_q^{(1)} & A_q^{(1)} & C_q^{(1)} & & & & f_q^{(1)} \\ \hline & & & B_1^{(2)} & A_1^{(2)} & C_1^{(2)} & & & f_1^{(2)} \\ & & & B_2^{(2)} & \ddots & \ddots & & & \vdots \\ & & & & \ddots & \ddots & C_{q-1}^{(2)} & & \vdots \\ & & & & & B_q^{(2)} & A_q^{(2)} & C_q^{(2)} & f_q^{(2)} \\ \hline & & & & & & B_1^{(3)} & A_1^{(3)} & C_1^{(3)} & f_1^{(3)} \\ & & & & & & B_2^{(3)} & \ddots & \ddots & \vdots \\ & & & & & & & \ddots & \ddots & \vdots \\ & & & & & & & & B_q^{(3)} & A_q^{(3)} & C_q^{(3)} & f_q^{(3)} \end{array} \right].$$

Fig. 1. The SPIKE partitioning for  $p = 3$  processors.

$$\|x - y\|_\infty \leq \delta \|x\|_\infty.$$

We assume that  $A$  has  $m = pq$  diagonal blocks and we assign  $q$  consecutive block rows to each processor. The case of  $p = 3$  is illustrated in Figure 1. If  $A$  is strictly diagonally dominant by rows, then we can predivide with the main block diagonal in order to obtain an equivalent linear system

$$Sx = g.$$

The case of  $p = 3$  is displayed in Figure 2. It is from the narrow columns or spikes protruding from the main diagonal that the original algorithm has derived its name. The matrix  $S$  is called the SPIKE matrix; the vector  $g$  is called the modified right hand side. By Lemma 1,

$$\|S - I\|_\infty \leq \epsilon < 1,$$

so  $S$  is strictly diagonally dominant by rows.

The equations within a single block row of each the main partitions lines form a reduced system

$$Rx_r = g_r$$

which can be solved independently. The general structure of the reduced system is given in Figure 3. Once the reduced system has been solved, the solution of the original system can be retrieved by backsubstitution. Specifically, we have

$$x_i^{(j)} = g_i^{(j)} - U_i^{(j)} x_q^{(j-1)} - V_i^{(j)} x_1^{(j+1)}, \quad 1 \leq i \leq q, \quad 1 \leq j \leq p, \quad (2)$$

where  $U_i^{(1)}$ ,  $V_i^{(p)}$ ,  $x_q^{(0)}$ , and  $x_1^{(p+1)}$  are undefined and should be taken as zero.

Suppose for the moment that we have somehow solved the reduced system with a small normwise relative forward error, say,

$$[S|g] = \left[ \begin{array}{cc|cc|cc} I & & V_1^{(1)} & & & g_1^{(1)} \\ & \ddots & \vdots & & & \vdots \\ & & \vdots & & & \vdots \\ & & I & V_q^{(1)} & & g_q^{(1)} \\ \hline & U_1^{(2)} & I & & V_1^{(2)} & g_2^{(2)} \\ & \vdots & & \ddots & \vdots & \vdots \\ & \vdots & & & \vdots & \vdots \\ & U_q^{(2)} & & I & V_q^{(2)} & g_q^{(2)} \\ \hline & & & U_1^{(3)} & I & g_1^{(3)} \\ & & & \vdots & & \vdots \\ & & & \vdots & & \vdots \\ & & & U_q^{(3)} & & I g_q^{(3)} \end{array} \right]$$

**Fig. 2.** The SPIKE matrix corresponding to  $p = 3$  processors.

$$[R|g_r] = \left[ \begin{array}{cc|cc|cc|cc} I & V_q^{(1)} & & & & & g_q^{(1)} \\ U_1^{(2)} & I & & V_1^{(2)} & & & g_1^{(2)} \\ \hline U_q^{(2)} & & I & V_q^{(2)} & \ddots & & g_q^{(2)} \\ & & U_1^{(3)} & I & & \ddots & g_1^{(3)} \\ \hline & & \ddots & & \ddots & & \vdots \\ & & & \ddots & & & \vdots \\ & & & & & V_1^{(p-1)} & \vdots \\ \hline & & & U_q^{(p-1)} & & I & V_q^{(p-1)} & g_q^{(p-1)} \\ & & & & & U_q^{(p)} & I & g_1^{(p)} \end{array} \right]$$

**Fig. 3.** The general structure of the reduced system.

$$\|x_r - y_r\| \leq \delta \|x_r\|_\infty$$

In view of equation (2) it is natural to partition  $y_r$  conformally with  $x_r$ , i.e.

$$y_r = (x_q^{(1)T}, x_1^{(2)T}, \dots, x_q^{(p-1)T}, x_1^{(p)T})^T.$$

and define a vector  $y \in \mathbb{R}^n$  using

$$y_i^{(j)} = g_i^{(j)} - U_i^{(j)} y_q^{(j-1)} - V_i^{(j)} y_1^{(j+1)}, \quad 1 \leq i \leq q, \quad 1 \leq j \leq p$$

Then

$$x_i^{(j)} - y_i^{(j)} = - \begin{bmatrix} U_i^{(j)} & V_i^{(j)} \end{bmatrix} \begin{bmatrix} x_q^{(j-1)} - y_q^{(j-1)} \\ x_1^{(j+1)} - y_1^{(j+1)} \end{bmatrix},$$

and it follows immediately that

$$[T|g_r] = \left[ \begin{array}{cc|cc|c} I & V_q^{(1)} & & & g_q^{(1)} \\ U_1^{(2)} & I & & & g_1^{(2)} \\ \hline & I & V_q^{(2)} & & g_q^{(2)} \\ & U_1^{(3)} & I & & g_1^{(3)} \\ \hline & & & \ddots & \vdots \\ & & & & \vdots \\ \hline & & & & I & V_q^{(p-1)} & g_q^{(p-1)} \\ & & & & U_q^{(p)} & I & g_1^{(p)} \end{array} \right]$$

**Fig. 4.** The structure of the truncated reduced system.

$$\|x - y\|_\infty \leq \epsilon \|x_r - y_r\|_\infty \leq \epsilon \delta \|x_r\|_\infty \leq \delta \|x\|_\infty.$$

It is clear that we must solve the reduced system accurately in order to achieve a small forward normwise relative error. We now consider the solution of the reduced system.

The reduced system matrix  $R$  is block tridiagonal and strictly diagonally dominant by rows. The inequality

$$\|R - I\|_\infty \leq \epsilon < 1$$

is inherited from the SPIKE matrix  $S$ . Frequently, but not universally, the off diagonal blocks are insignificant and can be dropped. This phenomenon is exploited heavily in the truncated SPIKE algorithm by Polizzi and Sameh [8, 9]. Let  $T$  denote the main block diagonal of  $R$ , see Figure 4. Mikkelsen and Manguoglu [5] showed that

$$\|T - R\|_\infty \leq \epsilon^q$$

when  $A$  is banded and strictly diagonally dominant by rows.

In this paper we consider the significance of the off diagonal blocks *relative* to the main block diagonal. To this end we define an auxiliary matrix  $\mathcal{B}$  by

$$\mathcal{B} = T^{-1}(T - R)$$

Now, let  $x_{tr}$  be the solution of the truncated reduced system

$$Tx_{tr} = g_r.$$

Then

$$T(x_r - x_{tr}) = (R - (T - R))x_r - g_r = (Rx_r - g_r) - (T - R)x_r = -(T - R)x_r$$

from which it immediately follows, that if  $x_r \neq 0$ , then

$$\frac{\|x_{tr} - x_r\|_\infty}{\|x_r\|_\infty} \leq \|\mathcal{B}\|_\infty.$$

We have already understood the need to solve the reduced system with a forward normwise relative error of at most  $\delta$ . If  $\|\mathcal{B}\|_\infty \leq \delta$ , then we simply drop the off diagonal blocks and approximate  $x_r$  with  $x_{tr}$ . If  $\|\mathcal{B}\|_\infty > \delta$ , then we can solve the

reduced system iteratively using the main block diagonal as a preconditioner. If we use the stationary iteration

$$Tx_r^{(i)} = (T - R)x_r^{(i-1)} + g_r, \quad i = 1, 2, \dots,$$

where  $x_r^{(0)} = 0$ , then

$$\|x_r - x_r^{(i)}\|_\infty \leq \|\mathcal{B}\|_\infty^i \|x_r\|_\infty$$

and we can stop the iteration whenever

$$\|\mathcal{B}\|_\infty^i \leq \delta.$$

In the next section we establish a tight upper bound on the central parameter  $\|\mathcal{B}\|_\infty$ . The SPIKE and the PSPIKE packages both apply BiCG, rather than the stationary iteration. Nevertheless, the size of  $\|\mathcal{B}\|_\infty$  remains an interesting question.

### 3 The main result

Our purpose is to establish Theorem 1.

**Theorem 1.** *The auxiliary matrix  $\mathcal{B}$  satisfies*

$$\|\mathcal{B}\|_\infty \leq \epsilon^q,$$

where  $q$  is the number of diagonal blocks assigned to each processor and equality is possible.

We shall reduce the problem of proving Theorem 1 to a single application of the following theorem.

**Theorem 2 (Mikkelsen [6]).** *Let  $G_k$  be a representation of  $2k - 1$  consecutive block rows of a block tridiagonal matrix  $A$  which is strictly diagonally dominant by rows with dominance factor  $\epsilon$ , i.e.*

$$G_k = \left[ \begin{array}{ccc|ccc|ccc} B_{1-k} & A_{1-k} & C_{1-k} & & & & & & \\ & \ddots & \ddots & & & & & & \\ & & \ddots & & & & & & \\ & & & B_{-1} & A_{-1} & C_{-1} & & & \\ \hline & & & B_0 & A_0 & C_0 & & & \\ \hline & & & & B_1 & A_1 & C_1 & & \\ & & & & & \ddots & \ddots & & \\ & & & & & & \ddots & & \\ & & & & & & & B_{k-1} & A_{k-1} & C_{k-1} \end{array} \right]$$

Then  $G_k$  is row equivalent to a unique matrix  $K_k$  of the form

$$K_k = \left[ \begin{array}{ccc|c|c|c} \mathcal{U}_{1-k}^{(k)} & I & & & & \mathcal{V}_{1-k}^{(k)} \\ \vdots & \ddots & & & & \vdots \\ \vdots & & \ddots & & & \vdots \\ \mathcal{U}_{-1}^{(k)} & & & I & & \mathcal{V}_{-1}^{(k)} \\ \hline \mathcal{U}_0^{(k)} & & & & I & \mathcal{V}_0^{(k)} \\ \hline \mathcal{U}_1^{(k)} & & & & & \mathcal{V}_1^{(k)} \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ \mathcal{U}_{k-1}^{(k)} & & & & & \mathcal{V}_{k-1}^{(k)} \end{array} \right] \quad (3)$$

where the spikes decay exponentially as we move towards the main block row. Specifically, if we define

$$Z_i^{(k)} = \begin{bmatrix} \mathcal{U}_{-i}^{(k)} & \mathcal{V}_{-i}^{(k)} \\ \mathcal{U}_i^{(k)} & \mathcal{V}_i^{(k)} \end{bmatrix}, \quad 0 < i < k,$$

and

$$Z_0^{(k)} = [\mathcal{U}_0^{(k)}, \mathcal{V}_0^{(k)}]$$

then

$$\|Z_i^{(k)}\|_\infty \leq \epsilon^{k-i}, \quad 0 \leq i < k.$$

*Proof.* The existence and uniqueness of  $K_k$  follows immediately from Lemma 1. The central inequality can be established using the well ordering principle. The details can be found in a report by Mikkelsen and Kågström [6].

We now move to prove the estimate given by Theorem 1. It is straightforward to verify that equality is achieved for matrices  $A$  given by equation (1) where

$$B_i = O_k, \quad A_i = I_k, \quad C_i = \epsilon I_k,$$

and  $O_k$  is the  $k$  by  $k$  zero matrix,  $I_k$  is the  $k$  by  $k$  identity matrix and  $\epsilon < 1$ .

In order to prove the general inequality it suffices to consider the interaction between two neighboring partitions. This follows immediately from the properties of the infinity norm. Let  $G_k$  be a compact representation of  $2k$  block rows drawn from the original matrix  $A$ , i.e.

$$G_k = \left[ \begin{array}{ccc|ccc|c} B_{-k} & A_{-k} & C_{-k} & & & & \\ & \ddots & \ddots & & & & \\ & & \ddots & & & & \\ & & & B_{-1} & A_{-1} & C_{-1} & \\ \hline & & & & B_1 & A_1 & C_1 \\ & & & & & \ddots & \\ & & & & & & B_k & A_k & C_k \end{array} \right]$$

and let  $H_k$  be a compact representation of the corresponding rows of the associated SPIKE matrix. Then  $G_k \sim H_k$  and  $H_k$  has the form

$$H_k = \left[ \begin{array}{c|c|c|c} U_{-k} & I & & V_{-k} \\ \vdots & \ddots & & \vdots \\ \vdots & & & \vdots \\ U_{-1} & & I & V_{-1} \\ \hline & & U_1 & I \\ & & \vdots & \ddots \\ & & \vdots & \ddots \\ & & U_k & I \\ & & & V_k \end{array} \right]. \quad (4)$$

Our task is to show that the auxiliary matrix  $Z_k$  given by

$$Z_k = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} = \begin{bmatrix} I & V_{-1} \\ U_1 & I \end{bmatrix}^{-1} \begin{bmatrix} U_{-1} & 0 \\ 0 & V_1 \end{bmatrix} \quad (5)$$

satisfies

$$\|Z_k\|_\infty \leq \epsilon^k, \quad k = 1, 2, \dots, q.$$

We continue to reduce  $H_k$  using row operations. We repartition  $H_k$  in order to focus our attention on the two central block rows, i.e.

$$G_k \sim \left[ \begin{array}{c|c|c|c} U_{-k} & I & & V_{-k} \\ \vdots & \ddots & & \vdots \\ \vdots & & & \vdots \\ U_{-1} & & I & V_{-1} \\ \hline & & U_1 & I \\ & & \vdots & \ddots \\ & & \vdots & \ddots \\ & & U_k & I \\ & & & V_k \end{array} \right].$$

Then we predivide with the central 2 by 2 block matrix and obtain

$$G_k \sim \left[ \begin{array}{c|c|c|c} U_{-k} & I & & V_{-k} \\ \vdots & \ddots & & \vdots \\ \vdots & & & \vdots \\ \hline Z_{11} & & I & Z_{12} \\ Z_{21} & & I & Z_{22} \\ \hline & & \vdots & \ddots \\ & & \vdots & \ddots \\ & & U_k & I \\ & & & V_k \end{array} \right]$$

and it is clear that there exists a matrix  $K_k$  such that  $G_k \sim K_k$  and



$$K_k = \begin{array}{c|c|c|c} \begin{array}{c} \mathcal{U}_{-q} \\ \vdots \\ \mathcal{U}_{-2} \end{array} & \begin{array}{c} I \\ \ddots \\ I \end{array} & & \begin{array}{c} \mathcal{V}_{-q} \\ \vdots \\ \mathcal{V}_{-2} \end{array} \\ \hline \begin{array}{c} \mathcal{U}_{-1} \\ \mathcal{U}_1 \end{array} & & \begin{array}{c} I \\ I \end{array} & \begin{array}{c} \mathcal{V}_{-1} \\ \mathcal{V}_1 \end{array} \\ \hline \begin{array}{c} \mathcal{U}_2 \\ \vdots \\ \mathcal{U}_q \end{array} & & \begin{array}{c} I \\ \ddots \\ I \end{array} & \begin{array}{c} \mathcal{V}_2 \\ \vdots \\ \mathcal{V}_q \end{array} \end{array} \quad (6)$$

and the matrix  $Z_k$  satisfies

$$Z_k = \begin{bmatrix} \mathcal{U}_{-1} & \mathcal{V}_{-1} \\ \mathcal{U}_1 & \mathcal{V}_1 \end{bmatrix}. \quad (7)$$

At this point we have reduced the problem of proving Theorem 1 to a straightforward application of Theorem 2.

### 4 Conclusion

The explicit SPIKE algorithm by Dongarra and Sameh [2] extends naturally to systems which are block tridiagonal and strictly diagonally dominant by rows. Moreover, the analysis of the method is simplified by focusing on the number of diagonal blocks rather than the bandwidth. The parallel bottleneck remains the solution of the reduced system  $Rx_r = g_r$  which is strictly diagonally dominant and block tridiagonal. The significance of the off diagonal blocks can be measured using the auxiliary matrix  $\mathcal{B}$  given by

$$\mathcal{B} = T^{-1}(T - R) = I - T^{-1}R,$$

where  $T$  denotes the main block diagonal of  $R$ . If  $\|\mathcal{B}\|_\infty$  is sufficiently small, then we can ignore the off diagonal blocks and approximate  $x_r$  with the solution of the truncated reduced system  $Tx_{tr} = g_r$ . In general, we can solve the reduced system iteratively using the main block diagonal  $T$  as a preconditioner and the convergence rate is controlled by the size of  $\|\mathcal{B}\|_\infty$ . Our main contribution is Theorem 1 which establishes a tight upper bound on  $\|\mathcal{B}\|_\infty$ .

### References

1. S.C. Chen, D.J. Kuck and A. Sameh, Practical parallel band triangular system solvers, ACM Trans. Math. Software, 4, (1978), pp. 270-277.
2. J.J. Dongarra and A. Sameh, On some parallel banded system solvers, Parallel Comput., vol. 1 (1984), pp. 223-235.
3. D.H. Lawrie and A. Sameh, The computation and communication complexity of a parallel banded system solver, ACM Trans. Math, Software, vol. 10 (1984), pp. 185-195.

4. M. Manguoglu, A. Sameh and O. Schenk, PSPIKE: A Parallel Hybrid Sparse Linear System Solver, In Proceedings of EURO-PAR-2009, LNCS 5704 (2009), pp. 797-808.
5. C. C. K. Mikkelsen and M. Manguoglu, Analysis of the truncated SPIKE algorithm, SIMAX, vol. 30, no. 4 (2008), pp. 1500-1519.
6. C. C. K. Mikkelsen and Bo Kågström, Analysis of incomplete cyclic reduction for narrow banded and strictly diagonally dominant linear systems. Tech. Rep. UMINF 11.07, Department of Computing Science, Umeå University (2011). Submitted to PPAM-2011.
7. A. Sameh and D.J. Kuck, On stable parallel linear systems solvers, J. ACM, vol. 25 (1978), pp. 81-91.
8. E. Polizzi and A. Sameh, A parallel hybrid banded system solver: The SPIKE algorithm, Parallel Comput., vol. 32 (2006), pp. 177-194.
9. E. Polizzi and A. Sameh, SPIKE: A parallel environment for solving banded linear systems, Comput. Fluids, vol. 36 (2007), pp. 113-120.