
An Accelerated Block-Parallel Newton Method via Overlapped Partitioning

Yurong Chen

Lab. of Parallel Computing, Institute of Software, CAS
(<http://www.rdcps.ac.cn/~ychen/english.htm>)

Summary. This paper presents an overlapped block-parallel Newton method for solving large nonlinear systems. The graph partitioning algorithms are first used to partition the Jacobian into weakly coupled overlapping blocks. Then the simplified Newton iteration is directly performed, with the diagonal blocks and the overlapping solutions assembled in a weighted average way at each iteration. In the algorithmic implementation, an accelerated technique has been proposed to reduce the number of iterations. The conditions under which the algorithm is locally and semi-locally convergent are studied. Numerical results from solving power flow equations are presented to support our study.

1 Introduction

This paper considers the problem of solving the large sparse system of nonlinear equations

$$\mathbf{F}(\mathbf{x}) = \mathbf{0}, \quad (1)$$

where $\mathbf{F}(\mathbf{x}) = (f_1, \dots, f_N)^T$ is a nonlinear operator from \mathbb{R}^N to \mathbb{R}^N . Such systems often arise from scientific and computational engineering problems. It is well-known that Newton methods and its variations (see Ortega and Rheinboldt [1970], etc.) coupled with some direct solution technique such as Gaussian elimination are powerful solvers for these systems when one has a sufficiently good initial guess \mathbf{x}_0 and when N is not too large. When the Jacobian is large and sparse, inexact Newton methods (see Dembo et al. [1982], Brown and Saad [1990], Cai and Keyes [2002], etc.) or some kind of nonlinear block-iterative methods (see Zecevic and Siljak [1994], Yang et al. [1997], Chen and Cai [2003], etc.) may be used.

An inexact Newton method (IN) is a generalization of Newton method for solving system (1), in which, each step $\{\mathbf{s}_k\}$ satisfies $\|\mathbf{F}'(\mathbf{x}_k)\mathbf{s}_k + \mathbf{F}(\mathbf{x}_k)\| < \|\mathbf{r}_k\|$, regardless of how $\{\mathbf{s}_k\}$ is determined. In past years, Krylov subspace methods, such as Arnoldi's method (see Saad [1981]), GMRES (see Saad and Schultz [1986]) and so on, have been studied intensively and applied in IN

for solving large scale linear systems approximately. This combined method is called inexact Newton-Krylov methods or nonlinear Krylov subspace projection methods. Many works on parallel Newton-Krylov methods have been done by Gropp et al. [2000], Knoll and Keyes [2003], etc.

Parallel nonlinear block-iterative method is another powerful solver for large sparse nonlinear systems, which chiefly consists of block Newton-type and block quasi-Newton methods. The classical nonlinear block-Jacobi algorithm and nonlinear block-Gauss-Seidel algorithm (see Ortega and Rheinboldt [1970]) are two original versions. A block-parallel Newton method via overlapping epsilon decompositions was presented by Zecevic and Siljak [1994]. Yang et al. [1997] described a parallelizable Jacobi-type block Broyden method, and more recently a partially overlapped Broyden method has been proposed by Chen and Cai [2003].

In this paper, we consider a parallelizable block simplified Newton method via overlapped partitioning, which is essentially an additive Schwarz method (block-Jacobi algorithm) with overlapping. In the implementation, an accelerated technique (see Sect. 2) is proposed for each iteration to reduce the number of iterations. Sect. 3 gives the sufficient conditions under which the new method is locally and semi-locally convergent. The numerical results for solving power flow equations are presented in Sect. 4. Finally, we draw conclusions and discuss the future work on this subject in Sect. 5.

2 The Algorithm

In the following discussion, $\mathbf{x}^* \in \mathbb{R}^N$ is an exact solution of system (1), i.e., $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$. Let \mathbf{x}^0 be an initial guess of \mathbf{x}^* , and suppose the components of \mathbf{x} and \mathbf{F} are conformally partitioned as follows:

$$\{\mathbf{F}\} = \{\mathbf{F}_1, \dots, \mathbf{F}_M\}, \quad \{\mathbf{x}\} = \{\mathbf{x}_1, \dots, \mathbf{x}_M\}, \quad (2)$$

where $\mathbf{F}_i = (f_{(i,1)}, \dots, f_{(i,n_i)})^T : \mathbb{R}^N \rightarrow \mathbb{R}^{n_i}$, $\mathbf{x}_i = (x_{(i,1)}, \dots, x_{(i,n_i)})^T \in \mathbb{R}^{n_i}$ for $i = 1, \dots, M$. Let $S_i = \{(i, 1), \dots, (i, n_i)\}$, then the partition satisfies that $\bigcup_{i=1}^M S_i = \{1, 2, \dots, N\}$ and $S_i \cap S_{i+1} \neq \emptyset$ for $i = 1, \dots, M-1$, which means that the adjacent blocks have partial overlaps. This partition may be obtained by graph-theoretic decomposition algorithms. Several overlapped strategies based on the general graph partitioning scheme included in Chaco, a publicly available graph partitioning software package developed by Hendrickson and Leland [1995], have been chiefly discussed by Chen and Cai [2003].

Let \mathbf{J}^0 be the Jacobian matrix of \mathbf{F} at \mathbf{x}^0 , i.e., $\mathbf{J}^0 = \frac{\partial \mathbf{F}(\mathbf{x})}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\mathbf{x}^0}$, and for $i = 1, \dots, M$ let $\mathbf{J}_i^0 = \frac{\partial \mathbf{F}_i(\mathbf{x})}{\partial \mathbf{x}_i} \Big|_{\mathbf{x}=\mathbf{x}^0} \in \mathbb{R}^{n_i \times n_i}$ be a nonsingular matrix. An algorithm for the Overlapped Block Simplified Newton method is as follows:

Overlapped Block Simplified Newton (OBSN) algorithm.

1. a. Partition \mathbf{J}^0 into M blocks $\mathbf{J}_i^0, i = 1, \dots, M$.

- b. Select weighted average parameter $\alpha = \{\alpha_i\}_{i=1}^{M-1}, 0 \leq \alpha_i \leq 1$.
- 2. For $k = 0, 1, \dots$ until convergence:
 - For $i = 1, \dots, M$, do:
 - a. Solve $\mathbf{J}_i^0 \mathbf{s}_i^k = -\mathbf{r}_i^k$.
 - b. Assemble the solutions: $\mathbf{x}_i^{k+1} = \mathbf{x}_i^k + \hat{\mathbf{s}}_i^k$, where for $j = 1, \dots, n_i$,

$$\hat{s}_{(i,j)}^k = \begin{cases} \alpha_i s_{(i,j)}^k + (1 - \alpha_i) s_{(i+1,j)}^k, & (i, j) \in S_i \cap S_{i+1}; \\ \alpha_{i-1} s_{(i-1,j)}^k + (1 - \alpha_{i-1}) s_{(i,j)}^k, & (i, j) \in S_{i-1} \cap S_i; \\ s_{(i,j)}^k, & \text{others.} \end{cases}$$
 - c. Calculate $\mathbf{r}_i^{k+1} = \mathbf{F}_i(\mathbf{x}^{k+1})$. If $\|\mathbf{r}^{k+1}\|$ is small enough, stop.

Step 2 of the above algorithm can be essentially replaced by the Newton-type iteration:

$$\mathbf{x}^{k+1} = \mathbf{G}(\mathbf{x}^k) = \mathbf{x}^k - (\mathbf{A}(\mathbf{J}_D^0, \alpha))^{-1} \mathbf{F}(\mathbf{x}^k), \tag{3}$$

where \mathbf{J}_D^0 denotes the partially overlapping block diagonal Jacobian and the matrix $(\mathbf{A}(\mathbf{J}_D^0, \alpha))^{-1}$ is determined by $\mathbf{J}_1^{-1}, \dots, \mathbf{J}_M^{-1}$ and α . To obtain local convergence for OBSN, one only needs to prove the convergence of the iteration (3). However, OBSN is proposed here for solving large sparse nonlinear systems in parallel. The reason is that Step 2 of the algorithm is easily parallelizable despite the use of a direct or an iterative solver.

For most practical problems, increasing the number of blocks will yield a severe increase in the number of iterations for the block-iterative method even the blocks with overlapping. In order to obtain an efficient parallel implementation for OBSN, it is critical to reduce the number of iterations. We therefore propose an accelerated technique based on the zero-nonzero structure of the partitioned Jacobian \mathbf{J}^0 . Suppose the set

$$\bar{S} = \{(i, j) : J_{ij}^0 \neq 0, J_{ij}^0 \notin \mathbf{J}_D^0, i, j = 1, \dots, N\}, \tag{4}$$

to be nonempty, then for all $\langle p, q \rangle \in \bar{S}$, append an updated formula following Step 2.(b) in OBSN as follows:

$$\begin{cases} \mathbf{x}_{(p)}^{k+1} = \mathbf{x}_{(p)}^{k+1} + \gamma \hat{\mathbf{s}}_{(q)}^k, \\ \mathbf{x}_{(q)}^{k+1} = \mathbf{x}_{(q)}^{k+1} + \gamma \hat{\mathbf{s}}_{(p)}^k, \end{cases} \tag{5}$$

where $\gamma \in (0, 1)$ is an accelerated parameter. The algorithm with the updated formula (5) is referred to as AOBSN algorithm.

Note that OBSN is essentially a variation of nonlinear block-multisplitting method presented by Frommer [1989] or additive Schwarz methods (block-Jacobi algorithm) with overlapping. The main difference is that the Jacobian matrix of \mathbf{F} need not be computed at each iteration and there is overlapping only between adjacent blocks in OBSN. Here, the restriction of overlapping just makes the selection of α much easier. The numerical results in Sect. 4 also show that the convergence performance of OBSN is much improved by the formula (5) with γ rather than by α .

3 Local and Semi-local convergence

Let $\|\cdot\|$ be a norm on \mathbb{R}^N and $\bar{\Omega}(\mathbf{x}, r)$ be a close ball of radius r about \mathbf{x} . We can immediately obtain the local convergence theorem for OBSN.

Theorem 1. *Let $\mathbf{F} : D \subset \mathbb{R}^N \rightarrow \mathbb{R}^N$ be Fréchet differentiable at the zero $\mathbf{x}^* \in D$ of \mathbf{F} , and suppose that $\mathbf{J}_i(\mathbf{x}) : \Omega_0 \subset D \rightarrow \mathcal{L}(\mathbb{R}^{n_i})$ is defined on some open neighborhood $\Omega_0 \subset D$ of \mathbf{x}^* and is continuous at \mathbf{x}^* with nonsingular $\mathbf{J}_i(\mathbf{x}^*)$ for $i = 1, \dots, M$. Then there exists a close ball $\Omega = \bar{\Omega}(\mathbf{x}^*, \delta) \subset \Omega_0$, $\delta > 0$, on which for any $\mathbf{x}^0 \in \Omega$ the mapping $\mathbf{G} : \mathbf{x} \in \Omega \rightarrow \mathbf{x} - (\mathbf{A}(\mathbf{J}_D^0, \alpha))^{-1} \mathbf{F}(\mathbf{x}) \in \mathbb{R}^N$ is well-defined and \mathbf{G} has at \mathbf{x}^* the Fréchet derivative*

$$\mathbf{G}'(\mathbf{x}^*) = \mathbf{I} - (\mathbf{A}(\mathbf{J}_D^0, \alpha))^{-1} \mathbf{J}(\mathbf{x}^*). \tag{6}$$

If $\rho(\mathbf{G}'(\mathbf{x}^*)) < 1$, then the sequence $\{\mathbf{x}^k\}$ generated by OBSN is well-defined for any $\mathbf{x}^0 \in \Omega$ and it converges to \mathbf{x}^* .

Proof. Set $\eta = \max_{i=1, \dots, M} \|\mathbf{J}_i(\mathbf{x}^*)^{-1}\|$ and for given $\varepsilon > 0$, $2\eta\varepsilon < 1$, choose $\delta > 0$ such that $\Omega = \bar{\Omega}(\mathbf{x}^*, \delta) \subset \Omega_0$ and $\|\mathbf{J}_i(\mathbf{x}) - \mathbf{J}_i(\mathbf{x}^*)\| \leq \varepsilon$ for any $\mathbf{x} \in \Omega$, $i = 1, \dots, M$. Then $\mathbf{J}_i(\mathbf{x})$ is invertible for all $\mathbf{x} \in \Omega$, and

$$\|(\mathbf{J}_i(\mathbf{x}))^{-1}\| \leq \frac{\eta}{1 - \eta\varepsilon} < 2\eta, \mathbf{x} \in \Omega, i = 1, \dots, M, \tag{7}$$

that is, the mapping $\mathbf{G}(\mathbf{x}) = \mathbf{x} - (\mathbf{A}(\mathbf{J}_D^0, \alpha))^{-1} \mathbf{F}(\mathbf{x})$ is well-defined on Ω for any $\mathbf{x}^0 \in \Omega$. In addition, $\rho(\mathbf{G}'(\mathbf{x}^*)) < 1$ which implies the \mathbf{x}^* is an attractor of the iterative formula (3), so the sequence \mathbf{x}^k generated by OBSN is well-defined for any $\mathbf{x}^0 \in \Omega$ and it converges to \mathbf{x}^* . \square

Furthermore, we can also obtain the semi-local convergence theorem for OBSN from Theorem 12.5.5 in Ortega and Rheinboldt [1970] by the Newton-type iteration (3). The proof is trivial.

Theorem 2. *Let $\mathbf{F} : D \subset \mathbb{R}^N \rightarrow \mathbb{R}^N$ be Fréchet differentiable and Lipschitz continuous with Lipschitz constant τ on a close ball $\bar{\Omega}(\mathbf{x}^0, r) \subset D$. Also suppose that there exist κ, η , and μ with $h = \kappa\tau\eta(1 - \mu)^2 \leq \frac{1}{2}$ such that*

$$\|(\mathbf{A}(\mathbf{J}_D^0, \alpha))^{-1}\| \leq \kappa, \tag{8}$$

$$\|(\mathbf{A}(\mathbf{J}_D^0, \alpha))^{-1} \mathbf{F}(\mathbf{x}^0)\| \leq \eta, \tag{9}$$

$$\|\mathbf{I} - (\mathbf{A}(\mathbf{J}_D^0, \alpha))^{-1} \mathbf{J}(\mathbf{x}^0)\| \leq \mu < 1, \tag{10}$$

Set

$$r_- = \frac{1 - \mu}{\kappa\tau} (1 - \sqrt{1 - 2h}), \tag{11}$$

$$r_+ = \frac{1 - \mu}{\kappa\tau} (1 + \sqrt{1 - 2h}). \tag{12}$$

If $r \geq r_-$, then the sequence $\{\mathbf{x}^k\}$ generated by OBSN keeps in $\bar{\Omega}(\mathbf{x}^0, r)$ and it converges to the unique root \mathbf{x}^* of \mathbf{F} in $\bar{\Omega}(\mathbf{x}^0, r')$ with $r' = \min\{r, r_+\}$.

4 Numerical Results

In this section, OBSN and AOBSN are applied to the load flow problem in power systems. The relative importance of α, γ on the convergence of (A)OBSN for the IEEE 118-bus system (Problem 1) are studied, and some values are suggested for obtaining good performance with (A)OBSN. Then, the parallel numerical results for the IEEE 662-bus system (Problem 2) on a PC cluster are presented.

For an n -bus system without the slack bus, the load flow problem is described by a system of nonlinear algebraic equations:

$$\mathbf{F}(\mathbf{x}_1, \dots, \mathbf{x}_n; \mathbf{P}, \mathbf{Q}) = \mathbf{0} \quad (13)$$

where

$$\mathbf{F}_i = (F_{P_i}, F_{Q_i})^T, \quad \mathbf{x}_i = (f_i, e_i)^T, \quad (14)$$

$$F_{P_i} = P_i - \operatorname{Re}(E_i \sum_{k=1}^n Y_{ik}^* E_k^*), \quad (15)$$

$$F_{Q_i} = Q_i - \operatorname{Im}(E_i \sum_{k=1}^n Y_{ik}^* E_k^*). \quad (16)$$

In the above equations, $E_i = e_i + \mathbf{j}f_i$ represents the unknown complex node voltage, $P_i + \mathbf{j}Q_i$ represents the injected power and $Y_{ik} = G_{ik} + \mathbf{j}B_{ik}$ represents the admittance. For PV buses, where the voltage magnitude V_i is fixed, equation (16) is replaced by $F_{Q_i} = V_i^2 - (e_i^2 + f_i^2)$.

In practice, good initial guess can be easily obtained in the load flow computation, especially in the load flow track case. So an approximate start \mathbf{x}^0 which is obtained by adding random values ranging from -0.01 to 0.01 to the solutions is considered to evaluate the cases where good initial approximations are provided. In all tests, Gaussian elimination was used to solve the linear subproblems exactly and the nonlinear error tolerance was 10^{-3} .

4.1 Influence of α, γ on Convergence

For the power flow equations (13), using the similarity of the structure of matrix $\mathbf{B} = (B_{ik})_{n \times n}$ and the Jacobian \mathbf{J}^0 (see Zecevic and Siljak [1994]), we applied the partially overlapped partitioning to \mathbf{B} to achieve the partition of \mathbf{J}^0 which reduces the problem dimension by a factor of 2. The linear-KL partitioning coupled with the boundary-linear strategy (see Chen and Cai [2003]) was chosen as the partition scheme for (A)OBSN in this study. The choice was observed to be better compared to scattered, spectral partitioning and multilevel-KL schemes (see Hendrickson and Leland [1995]) coupled with the boundary-linear strategy. For simplicity, we only considered the case for $\alpha_1 = \dots = \alpha_{M-1} = \alpha$.

Fig. 1 shows the influence of α on the convergence of OBSN and AOBSN ($\alpha = 0.5$) for Problem 1 in ten differently approximate start cases. From the figure, we can see that the influence of α on the convergence of OBSN is more sensitive to \mathbf{x}_0 than that on AOBSN. In addition, the rate of convergence of OBSN is significantly improved by the updated formula (5) with the accelerated parameter $\gamma = 0.5$. Note that the influence of α on the convergence of AOBSN is much less than that of γ , so α can be fixed (for example, $\alpha = 0.5$) if AOBSN is used to solve the power flow equations.

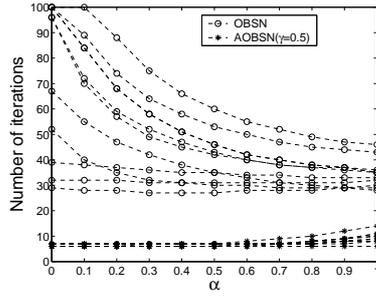


Fig. 1. Influence of α on the convergence of (A)OBSN for Problem 1 (ten approximate starts, $M = 8$)

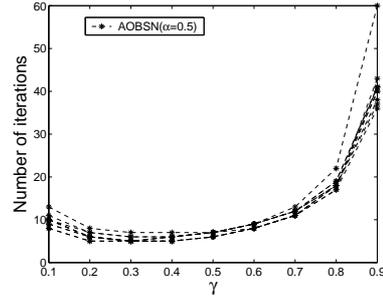


Fig. 2. Influence of γ on the convergence of AOBSN for Problem 1 (ten approximate starts, $M = 8$)

Fig. 2 shows the influence of accelerated parameter γ on the convergence of AOBSN ($\alpha = 0.5$) for Problem 1 in the same ten start cases. The figure also shows that the convergence performance of OBSN is significantly improved by the accelerated technique with $\gamma = 0.1, \dots, 0.8$. Similar conclusions can be drawn for other ten initial approximations obtained even by a larger disturbed parameter as well.

4.2 Parallel Implementation of AOBSN for Problem 2

Using the above scheme, we partition the matrix B of Problem 2 into 2, 4, 8, 16, 32 and 64 blocks (see Fig. 3 for some cases). Fig. 4 shows the number of iterations of three algorithms mentioned above for Problem 2. We can see that AOBSN has much better convergent performance than the Block Simplified Newton method (BSN) and OBSN, and its number of iterations is less sensitive to the number of blocks. It should be pointed out that AOBSN usually requires more iterations than Newton methods and its simplified version. However, by virtue of the reduction of dimensionality, AOBSN can result in significant computational savings.

In the parallel implementation, we assigned the individual blocks or a group of blocks into per processor in an adequate load balancing way. All

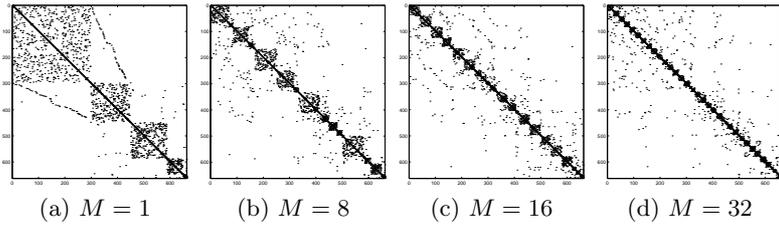


Fig. 3. Zero-nonzero structure of B and the partitioned B for Problem 2

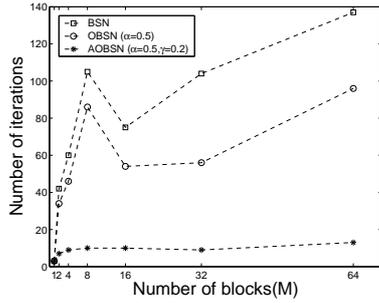


Fig. 4. Comparison of 3 methods for Problem 2 (approximate start)

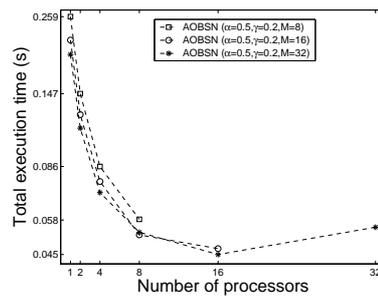


Fig. 5. Parallel computing time of AOBSN for Problem 2

numerical tests were run on an SMP-Cluster (36 nodes, CPU of per node: $4 \times$ Intel Xeon PIII700MHz, Myrinet, MPI). The programs were written in C using double precision floating point numbers. Fig. 5 shows the parallel computation time of AOBSN in 8, 16 and 32 block cases. The total execution time reduces with the number of processors and reaches its minimal value when the number of processors is 8 or 16 in 8 or 16 block cases, respectively. The communication time approximately increases with the number of processors and exceeds the computation time with 32 processors in 32 block case, which indicates that Problem 2 is not sufficiently large to be efficiently mapped onto more than 16 processors on the SMP-Cluster.

5 Conclusions and Discussion

This paper has presented an overlapped block-parallel Newton method (OBSN) for solving large nonlinear systems. In the implementation, an accelerated version (AOBSN) is also proposed to reduce the number of iterations. The numerical results of solving power flow equations confirm that AOBSN is indeed effective, particularly for problems where a good initial guess is available.

As mentioned in the previous sections, OBSN and AOBSN are nonlinear methods depending on several parameters, including the partition scheme,

M, α, γ , etc. In this paper, the relative importance of α, γ on the convergence of (A)OBSN were studied for the power flow problem, and some values for obtaining good results with (A)OBSN were suggested. A theoretical study on how these parameters influence the convergence of the algorithms will be carried out in our future work.

Acknowledgement. The author gratefully thanks Dr. B. Hendrickson and Dr. R. Leland of Sandia National Laboratories for allowing him to use the Chaco package for some of the numerical tests performed in this work.

References

- P. N. Brown and Y. Saad. Hybrid Krylov methods for nonlinear system of equations. *SIAM J. Sci. Stat. Comput.*, 11(3):450–481, 1990.
- X. C. Cai and D. E. Keyes. Nonlinear preconditioned inexact Newton algorithms. *SIAM J. Sci. Comput.*, 24(1):183–200, 2002.
- Y. Chen and D. Cai. Inexact overlapped block Broyden methods for solving nonlinear equations. *Appl. Math. Comput.*, 136(2/3):215–228, 2003.
- R. Dembo, S. Eisenstat, and T. Steihaug. Inexact Newton methods. *Siam J. Numer. Anal.*, 19(2):400–408, 1982.
- A. Frommer. Parallel nonlinear multisplitting methods. *Numer. Math.*, 56:269–282, 1989.
- W. D. Gropp, D. E. Keyes, L. C. McInnes, and M. D. Tidriri. Globalized Newton-Krylov-Schwarz algorithms and software for parallel implicit CFD. *Int. J. High Performance Computing Applications*, 14:102–136, 2000.
- B. Hendrickson and R. Leland. The Chaco user’s guide, version 2.0. Technical report, Sandia National Laboratories, Albuquerque, NM, July 1995. Tech. Rep. SAND 95-2344.
- D. A. Knoll and D. E. Keyes. Jacobian-free Newton-Krylov methods: A survey of approaches and applications. *Int. J. High Performance Computing Applications*, 2003. to appear.
- J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, 1970.
- Y. Saad. Krylov subspace methods for solving unsymmetric linear systems. *Math. Comp.*, 37:105–126, 1981.
- Y. Saad and M. Schultz. GMRES: A generalized minimum residual algorithm for solving non-symmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7:856–869, 1986.
- G. Yang, L. C. Dutton, and M. Fortin. Inexact block Jacobi-Broyden methods for solving nonlinear systems of equations. *SIAM J. Sci. Comput.*, 18(5):1367–1392, 1997.
- A. I. Zecevic and D. D. Siljak. A block-parallel Newton method via overlapping epsilon decompositions. *SIAM J. Matrix Analysis and Applications*, 15(3):824–844, 1994.