Abstract: The organization and management of the intense computation involved in the solution of large, sparse and non-symmetric systems of equations, arising from the discretization of Elliptic Boundary Value Problems (BVPs) by the Hermite Collocation method, on a Distributed-Shared Memory (DSM) multiprocessor environment is the problem considered herein. As the size of the problem directly suggests the usage of iterative methods, we consider the parallel implementation of the Successive Over-Relaxation (SOR) and the preconditioned Bi-Conjugate Gradient Stabilized (Bi-CGSTAB) methods on DSM systems. Using the parallel algorithms we devised in [9, 10], which take advantage of the red-black structure of the Collocation matrix, we address the problem of efficiently managing the whole computation, through the MPI programming model, on a DSM system. The application was developed on a SGI Origin 350 DSM multiprocessor computer and its performance is revealed by the speedup measurements included.

Key-Words: Collocation, SOR, preconditioned Bi-CGSTAB, Red-Black Orderings, MPI, DSM Systems.

1 Introduction

Hermite Collocation is a high order finite element scheme used as a discretizer especially when continuous first derivatives are required. Among other properties, Collocation produces large and sparse systems of equations which poses no pleasant properties (e.g. symmetry). Memory requirements and performance are two of the main factors suggesting the usage of iterative methods on multiprocessor environments. This motivated relevant research in the areas of iterative method analysis and parallel algorithm development. Main issues addressed were concerning both algorithmic (multi-color orderings, domain decomposition/partitioning techniques, parallel preconditioners, etc) and architectural (memory management/distribution, processor architecture, etc) aspects. Particular results, in this direction, concerning the Collocation method may be found in (e.g. [1,3,5,6,14] and in our work in [13,7,9,10]. It is worthwhile to mention that in [8,11], we conducted a performance analysis on a large family of Krylov subspace methods, including GMRES[15] and Bi-CGSTAB[17] as well as several preconditioning schemes, and concluded that from the tested Krylov methods the Backward Gauss-Seidel (B-GS) preconditioned Bi-CGSTAB-P yields fast rates of convergence when it applies to the solution of the Hermite Collocation Poisson system, while at the same time, outperforms all stationary iterative schemes, including SOR, for medium and large size problems. In [9] and [10]:

- We presented a generalized technique for devising efficient parallel algorithms, by constructing a virtual parallel machine ideally fitted for the problem at hand, and, in the sequel, by making use of partitioning techniques, the virtual computation was mapped onto a fixed size parallel architecture

- We effectively used this technique to devise efficient parallel algorithms for the iterative solution of Hermite Collocation systems, by the preconditioned Bi-CGSTAB and SOR methods.

In the work herein, to exploit and enhance the algorithm’s parallelism, we implement the above methods on environments supporting Incremental Parallelization [2,4]. Namely, we implement the already appropriately designed parallel algorithms on DSM machines using the MPI standard [12], which utilize data dependency and memory allocation issues, as the Automatic Parallelization Option (APO) of the MipsPro compiler is unable to do so. Therefore we improve the algorithm’s performance, on DSM machines, by managing the whole computation in order to maximize locality and minimize communication among the processing elements.

This paper is organized as follows: In Section 2 we briefly describe the structure of the collocation system and the iterative methods used, as well as we highlight the computationally intense parts where incremen-
Bi-CGSTAB and SOR iterative methods may be efficacious described by:

Choose $x^{(0)}$.

Section 4, we present speedup measurements from the implementation on a SGI Origin 350 DSM system.

2 SOR / Bi-CGSTAB for Collocation

Let us consider the linear system

$$A x = b$$

where the matrix $A$ is in the well known red-black partitioning form

$$A = \begin{bmatrix} D_R & H_B \\ H_R & D_B \end{bmatrix}.$$  

(2)

Collocation, among other celebrated discretizers such as the Finite Differences, through a particular numbering of equations and unknowns, results to a red-black ordered system when applied on Elliptic BVPs (e.g. [11]). And as it is shown in [10,11], both B-GS preconditioned Bi-CGSTAB and SOR iterative methods may be effectively used for its solution.

The Algorithms

By considering now the classical splitting

$$A = D_A - L_A - U_A$$

(3)

where

$$D_A = \begin{bmatrix} D_R & O \\ O & D_B \end{bmatrix}, \quad L_A = \begin{bmatrix} O & O \\ -H_R & O \end{bmatrix}$$

(4)

and

$$U_A = \begin{bmatrix} O & -H_B \\ O & O \end{bmatrix},$$

(5)

the iterative methods under consideration may be algorithmically described by:

**SOR**

$$\mathcal{M}_\omega = D_A - \omega L_A$$

$$\mathcal{E}_\omega = (1 - \omega)D_A + \omega U_A$$

Choose $x^{(0)}$.

for $i = 1, 2, ...$

$$t = \mathcal{E}_\omega x^{(i)}$$

$$t = t + \omega b$$

Solve $\mathcal{M}_\omega x^{(i+1)} = t$

Check for Convergence

end

**B-GS Preconditioned Bi-CGSTAB [17]**

Choose $x^{(0)}$.

$$r^{(0)} = b - A x^{(0)}$$

Choose $\hat{r}$ (usually $\hat{r} = r^{(0)}$)

for $i = 1, 2, ...$

$$\rho_{i-1} = \hat{r}^T r^{(i-1)}$$

if $\rho_{i-1} = 0$ method fails

if $\rho_{i-1} = 1$

$$p^{(1)} = r^{(0)}$$

else

$$\beta_{i-1} = \frac{\rho_{i-1}}{\rho_{i-2}} \frac{\alpha_{i-1}}{\alpha_{i-1}}$$

$$p^{(i)} = r^{(i-1)} + \beta_{i-1}(p^{(i-1)} - \omega_{i-1} v^{(i-1)})$$

end

$$s = r^{(i-1)} - \alpha_i v^{(i)}$$

if $\| s \|$ is small enough then

$$x^{(i)} = x^{(i-1)} + \alpha_i \hat{p} \quad \text{and stop}$$

Solve $\mathcal{M} z = s$ ;

$$t = A z$$

$$\omega_i = \frac{s^T t}{t^T t}$$

$$x^{(i)} = x^{(i-1)} + \alpha_i \hat{p} + \omega_i z$$

Check for Convergence

if $\omega_i = 0$ stop

$$r^{(i)} = s - \omega_i t$$

end

Notice that:

- The highlighted statements are the computationally intense parts of the algorithms
- The above algorithm for the Bi-CGSTAB implements the Bi-CGSTAB-P version[16], which minimizes the residual, instead of the preconditioned residual, and is equivalent to the postconditioned Bi-CGSTAB. For this case, in [11], we observed that the B-GS postconditioned Bi-CGSTAB yields faster convergence, hence

$$\mathcal{M} = D_A - U_A.$$

(6)

Incorporating the Collocation Structure

Aiming to the overall (serial and parallel) improvement of the computational performance, it is necessary to incorporate the particular structures, of the matrices involved, into the algorithms. For instance, in doing so for the intense

Solve $\mathcal{M} z = s$ ;

$$t = A z$$

one may easily observe its equivalence, and therefore replace it into the algorithm, with the red-black instruction block:

Solve $D_B z_B = s_B$

$$y = H_B z_B$$
Furthermore, upon application of the Hermite Collocation method on Helmholtz-type Dirichlet BVPs, on a uniformly partitioned (into $n_s = 2p$ subintervals in both directions) unit square, the associated with relation (2) matrices $D_R, D_B, H_R, H_B \in \mathbb{R}^{8p^2, 8p^2}$ are defined [10] by:

\[
D_R = \text{diag}[A_2 2A_1 2A_2 \cdots 2A_1 2A_2 - A_2], \quad \text{and for } k = 1, \ldots, p - 1
\]

\[
D_B = 2 \text{diag}[A_1 A_2 \cdots A_1 A_2],
\]

(7)

where each $H_j^{(R)} \in \mathbb{R}^{8p^2, 4p}$ is defined by [16]:

\[
H_1^{(R)} = (e_1 - e_2) \otimes A_4
\]

\[
H_2^{(R)} = -(e_{2p-1} + e_{2p}) \otimes A_4,
\]

\[
H_k^{(R)} = (e_{2k-1} + e_{2k} + e_{2k+1} - e_{2k+2}) \otimes A_3
\]

\[
H_{2k+1}^{(R)} = -(e_{2k-1} + e_{2k} - e_{2k+1} + e_{2k+2}) \otimes A_4
\]

while each $H_j^{(B)} \in \mathbb{R}^{8p^2, 4p}$ is defined by [16]:

\[
H_1^{(B)} = (e_1 + e_2 - e_3) \otimes A_3
\]

\[
H_2^{(B)} = -(e_1 - e_2 + e_3) \otimes A_4
\]

\[
H_{2p-1}^{(B)} = (e_{2p-2} + e_{2p-1} + e_{2p}) \otimes A_3
\]

\[
H_{2p}^{(B)} = -(e_{2p-2} + e_{2p-1} - e_{2p}) \otimes A_4
\]

and for $k = 2, \ldots, p - 2$

\[
H_{2k-1}^{(B)} = (e_{2k-2} + e_{2k-1} + e_{2k} - e_{2k+1}) \otimes A_3
\]

\[
H_{2k}^{(B)} = -(e_{2k-2} + e_{2k-1} - e_{2k} + e_{2k+1}) \otimes A_4
\]

In all the above, $e_j$ denotes the $j$-th unit vector in $\mathbb{R}^{2p}$ while the matrices $A_i \in \mathbb{R}^{4p, 4p}$ are as defined in [11]. With this formulation of the matrices, it can be shown [16] that for any vector $v \in \mathbb{R}^{8p^2}$, partitioned conformably as

\[
v = [v_1^T, \ldots, v_{2p}^T]^T
\]

there holds

\[
H_R v = \sum_{j=1}^{2p} H_j^{(R)} \odot v_j,
\]

(8)

where

\[
H_j^{(R)} \odot v_j = \begin{cases} (\sum c_j e_i) \otimes (A_4 v_j) & \text{when } j \text{ odd} \\ (\sum \hat{c}_j e_i) \otimes (A_3 v_j) & \text{when } j \text{ even} \end{cases}
\]

with the constants $c_j$ and $\hat{c}_j$ to be obviously defined for each $j$ by (7). Similarly,

\[
H_B v = \sum_{j=1}^{2p} H_j^{(B)} \odot v_j,
\]

(9)

where

\[
H_j^{(B)} \odot v_j = \begin{cases} (\sum d_j e_i) \otimes (A_3 v_j) & \text{when } j \text{ odd} \\ (\sum \hat{d}_j e_i) \otimes (A_4 v_j) & \text{when } j \text{ even} \end{cases}
\]

From relations (8) and (9) it becomes apparent that, in order to efficiently manage a block matrix-vector computation involving $H_R$ (equiv. $H_B$) in both serial and parallel environments, one has to preprocess the vector by multiplying all of its odd (equiv. even) partitioning blocks by $A_4$ and all of its even (equiv. odd) partitioning blocks by $A_3$.

### 3 MPI Management

In this section, considering the observations made in the previous section, we focus our attention on the MPI

management of the computation involved with the BG-S preconditioned Bi-CGSTAB, as the SOR case can be treated in a similar way and has been treated in some detail in [10] on a similar environment. The whole discussion is based on our work in [9, 10] where, taking into consideration the essential factors of (a) uniform load balancing, (b) minimal idle cycles of processors,
and (c) minimal communication cost, we partitioned a virtual architecture, in an optimal way for general operators, and mapped it on a proposed (basically pipelined) architecture consisting of \( P_j \), \( j = 1, \cdots, N \) processors (depicted in Figure 1 for \( N = 6 \)). Here, since we deal with operators of Helmholtz-type, we follow a detour to arrive at a more balanced computation. Referring to Figure 1 we remark that:

- Processor \( P_1 \), in addition to the computational tasks assign to each processor, has been also assigned the tasks of gathering partially processed data, assemble, in the sequel, the final values for the inner products and other parameters of the algorithm, and finally broadcast (green dashed communication lines) the results to all other processors.

- Assuming that \( k = 2p/N \) is an even integer (other cases can be treated similarly), each processor \( P_j \) computes on \( k \) red and \( k \) black subvectors of size \( 4p \). More specifically, to each processor \( P_j \) we assign all the necessary tasks to compute the solution subvectors

\[
x^{(R)}_l, \quad l = (j - 1)k + 1, \cdots, jk
\]

and

\[
x^{(B)}_l, \quad l = 2p + (j - 1)k + 1, \cdots, 2p + jk .
\]

- The communication between processors \( P_j \) and \( P_{j+1} \) in order to compute the matrix-vector product \( H_Bz_B \) is depicted in Figure 2, while for the \( H_Bz_B \) is depicted in Figure 3.

\[
x^{(R)}_{jk+1}, \quad z^{(R)}_{jk}
\]

\[
x^{(B)}_{2p+jk+1}, \quad z^{(B)}_{2p+jk}, \quad z^{(B)}_{2p+jk+2}
\]

**Figure 2 : Communication needed to compute \( H_Bz_B \)**

- The total communication between processors, needed in each iteration step of the B-GS preconditioned Bi-CGSTAB, as well as the communication scheduling (for the black and the red cycles respectively) is being depicted in Figure 4.

\[
\begin{align*}
\text{enddo} \\
\text{Send to } P_{j+1} & \left[ \begin{array}{c}
y_{jk-1} \\
y_j 
\end{array} \right] \\
\text{Send to } P_{j-1} & \left[ \begin{array}{c}
y_{(j-1)k+1} \\
y_{(j-1)k+2} 
\end{array} \right] \\
\text{Send to } P_{j+1} & \left[ \begin{array}{c}
y_{jk+1} \\
y_{jk+2} 
\end{array} \right] \\
t_{c_1} & \leftarrow \text{Receive } \left[ \begin{array}{c}
y_{(j-1)k-1} \\
y_{(j-1)k} 
\end{array} \right] \text{ from } P_{j-1} \\
t_{c_2} & \leftarrow \text{Receive } \left[ \begin{array}{c}
y_{(j+1)k-1} \\
y_{(j+1)k} 
\end{array} \right] \text{ from } P_{j+1} \\
t_{c_3} & \leftarrow \text{Receive } \left[ \begin{array}{c}
y_{j+1k+1} \\
y_{j+1k+2} 
\end{array} \right] \text{ from } P_{j+1} \\
t_{c_4} & \leftarrow \text{Receive } \left[ \begin{array}{c}
y_{(j+1)k+1} \\
y_{(j+1)k+2} 
\end{array} \right] \text{ from } P_{j+1} \\
tm_1 & = tc_1 + tc_2 \\
tm_2 & = y_{(j-1)k+1} - y_{(j-1)k+2} \\
& \quad \text{Solve } D_B z_B = s_B \\
y & = H_B z_B \\
& \quad \text{Solve } D_R z_R = s_R - y \\
y & = H_R z_R 
\end{align*}
\]

**Figure 4 : Communication Scheduling**

- In the local memory of each processor, before iteration starts, we store copies of the matrices \( A_3 \) and \( A_4 \), as well as copies of the \( LU \)-factored matrices \( A_1 \) and \( A_2 \). We also store the appropriate parts of the initial \( x^{(0)} \) and the RHS vectors.

Taking into consideration all the above, the program code each processor \( P_j \), \( j = 1, \cdots, N \) executes for the intense matrix-vector operations (see section 2)

\[
\begin{align*}
\text{Solve } D_B z_B & = s_B \\
& \quad \text{Solve } D_R z_R = s_R - y \\
y & = H_R z_R 
\end{align*}
\]

takes the specific form:

**Black Cycle**

\[
\begin{align*}
do & \quad l = 2p + (j - 1)k + 1 \quad \text{to} \quad 2p + jk - 1 , 2 \\
& \quad \text{Solve } 2A_1 z^{(B)}_l = s^{(B)}_l \\
& \quad y_{l-2p} = A_3 z^{(B)}_l \\
& \quad \text{Solve } 2A_2 z^{(B)}_{l+1} = s^{(B)}_{l+1} \\
& \quad y_{l+1-2p} = A_4 z^{(B)}_{l+1} \\
& \quad \text{enddo}
\end{align*}
\]
do \( l = (j-1)k + 1 \) to \( jk-3, 2 \)
\[
\begin{align*}
\text{tm}_3 & \leftarrow y_l \\
y_l & \leftarrow \text{tm}_2 - \text{tm}_1 \\
\text{tm}_1 & \leftarrow y_{j+1} + \text{tm}_3 \\
\text{tm}_2 & \leftarrow y_{j+2} - y_{j+3} \\
y_{j+1} & \leftarrow \text{tm}_1 + \text{tm}_2
\end{align*}
\]
enddo
\[
\begin{align*}
\text{tm}_3 & \leftarrow y_{jk-1} + y_{jk} \\
y_{jk-1} & \leftarrow \text{tm}_2 - \text{tm}_1 \\
y_{jk} & \leftarrow \text{tm}_3 + tc_3 - tc_4
\end{align*}
\]

Red Cycle

do \( l = (j-1)k + 1 \) to \( jk-1, 2 \)
\[
\begin{align*}
\text{Solve} & \quad 2A_{2}z_{i}^{(R)} = s_{i}^{(R)} - y_{l} \\
y_{l} & \leftarrow A_{4}z_{i}^{(R)} \\
\text{Solve} & \quad 2A_{1}z_{i+1}^{(R)} = s_{i+1}^{(R)} - y_{l+1} \\
y_{l+1} & \leftarrow A_{3}z_{i+1}^{(R)}
\end{align*}
\]
enddo

\[
\begin{align*}
[tc_1] & \leftarrow \text{Receive} \left[ y_{(j-1)k} \right] \text{from } \mathcal{P}_{j-1} \\
\text{Send to} & \quad \mathcal{P}_{j+1} \left[ y_{jk} \right] \\
\text{Send to} & \quad \mathcal{P}_{j-1} \left[ y_{(j-1)k+1} \right] \\
[tc_2] & \leftarrow \text{Receive} \left[ y_{jk+1} \right] \text{from } \mathcal{P}_{j+1}
\end{align*}
\]
\[
\begin{align*}
\text{tm}_1 & \leftarrow tc_1 \\
\text{tm}_2 & \leftarrow y_{jk-1} + \text{tm}_1 \\
\text{tm}_1 & \leftarrow y_{jk} - tc_2 \\
y_{jk-1} & \leftarrow \text{tm}_1 + \text{tm}_2 \\
y_{jk} & \leftarrow \text{tm}_1 - \text{tm}_2
\end{align*}
\]

4 Realization on a DSM computer

SGI Origin 350 is a Shared-Distributed cache coherent - nonuniform memory access (ccNUMA) architecture machine, consisting of eight R16000@600MHz type processors with 4 MB Level 2 cache memory each. The total memory is 4 GB and the operating system is IRIS version 6.5. The applications are developed in double precision Fortran code using the MPI standard for MipsPro compilers version 7.4, which also incorporate the scientific library LAPACK.

The Figures 5 and 6, below, present the speedup measurements of the parallel algorithms, using up to eight processors, for B-GS preconditioned Bi-CGSTAB and SOR methods respectively with discretization of \( n_s \) = 64, \( n_s \) = 128 and \( n_s \) = 256 subintervals.

![Figure 5: Speedup measurements for Bi-CGSTAB.](image)

![Figure 6: Speedup measurements for SOR.](image)

As seen in the above figures the speedup is almost linear for up to 4 processors for both methods. For the 8 processors, available in our case, finer discretization yielded almost scalable performance. As a note, we add that the discretization for \( n_s \) = 256 corresponds to solving a linear system with 262,144 equations and unknowns.
Acknowledgments:
The authors wish to thank Professor Y.G. Saridakis for the tensor product formulation of the problem and many helpful discussions.

References:


