

Preconditioning for solving Hermite Collocation by the Bi-CGSTAB

E. N. MATHIOUDAKIS*, E. P. PAPADOPOULOU† and Y. G. SARIDAKIS ‡

Applied Mathematics and Computers Laboratory

Department of Sciences

Technical University of Crete

University Campus, 73100 Chania

GREECE

Abstract : - Explicit pre/post conditioning of the large, sparse and non-symmetric system of equations, arising from the discretization of the Dirichlet Poisson's Boundary Value Problem (BVP) by the Hermite Collocation method is the problem considered herein. Using the 2-cyclic (red-black) structure of the Collocation coefficient matrix, we investigate the eigenvalue distribution of its preconditioned analogs emerging from its red-black USSOR (UnSymmetric SOR) splittings. This analysis, coupled with computational efficiency issues, enables us to justify the choice of Gauss-Seidel (GS) preconditioned schemes as efficient and practical ones, when they used to accelerate the rate of convergence of the Bi-CGSTAB iterative Krylov subspace method. Our results are verified by numerical experiments.

Key-Words : - Collocation, 2-cyclic matrices, Gauss-Seidel, SOR, USSOR, preconditioning, Krylov methods, Bi-CGSTAB.

1 Introduction

Many applications, in several fields of science and engineering, are modeled by general Elliptic Boundary Value Problems (BVP) and Finite Elements (FE) is a well established methodology for their numerical solution. Collocation, using Hermite bi-cubic elements, is a competitive high order FE scheme with the following central well known properties :

- avoids numerical integration
- produces direct approximations for the values of the function and its first order partial derivatives at the grid nodes
- the collocation spline representing the solution of the BVP has continuous first derivatives (helpful e.g. in subsurface flow problems)
- the choice of the collocation points directly affects the convergence rate of the method
- the resulting, from the discretization, collocation matrix is large, sparse and enjoys no pleasant properties (such as symmetry, definiteness or diagonal dominance).

Last property directly suggests the usage of iterative methods on multiprocessor environments for the efficient solution of the Collocation system. This direction has been attracted the interest of many researchers and some very interesting results have been produced in the past pertaining to the performance and analysis of the classical SOR-like splitting methods [6-8,11] and the preconditioned Krylov subspace methods [1,2,9]. It is worthwhile to mention that in [9] we conducted a numeri-

cal performance evaluation on several methods from the Krylov subspace family, including GMRES[13] and Bi-CGSTAB[16] as well as several SSOR type preconditioning schemes, and concluded that :

- SSOR-type preconditioned Bi-CGSTAB schemes yielded better performance, in all practical cases, compared to the rest of the preconditioned Krylov subspace methods. The same conclusion was also reached in [1] where a performance evaluation between the GS preconditioned Bi-CGSTAB and GMRES methods was thoroughly performed.
- The Symmetric Gauss-Seidel(SGS) preconditioned Bi-CGSTAB proved to be the most practical, in the sense that there is no need for any relaxation parameter evaluation, without any compromises in the rate of convergence from other SSOR preconditioned Bi-CGSTAB schemes.

In the work herein we attempt a generalized approach to the problem of establishing efficient and practical preconditioned Krylov subspace methods, for the Hermite Collocation, through an eigenvalue analysis on the family of red-black USSOR preconditioners. We remark that :

- The distribution of the eigenvalues of the preconditioned coefficient matrix, although not the only one, is one of the most important factors affecting the convergence rate of preconditioned Krylov methods (e.g. [5]).
- USSOR preconditioning includes both SSOR and SOR type of preconditioning. Effective SOR-type precondi-

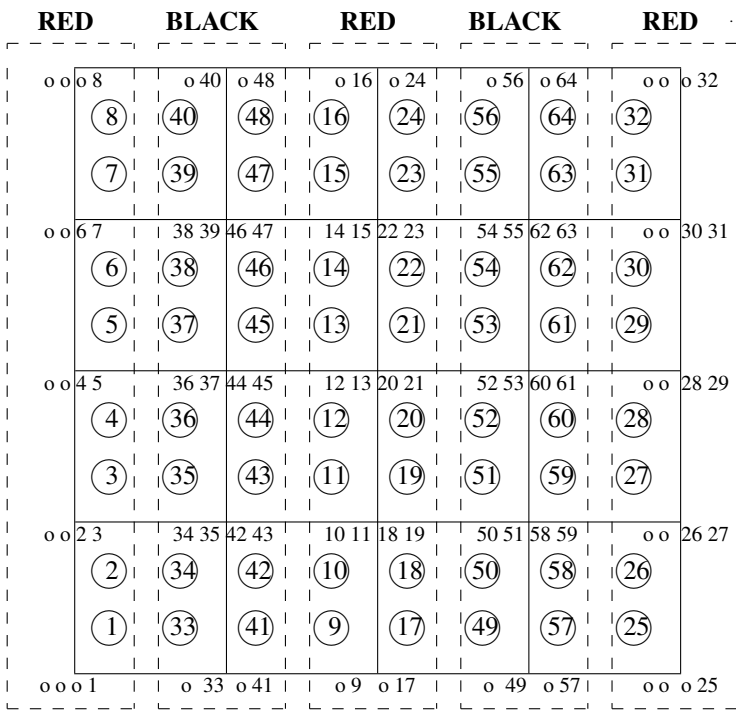


Fig.1 : Red-Black Grouping and Numbering of Unknowns and Equations for $n_s = 4$.

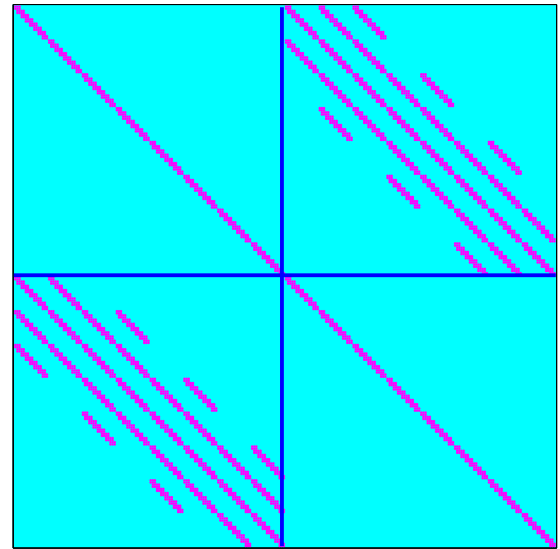


Fig.2 : Block structure of the Red-Black Collocation matrix for $n_s = 8$.

tioning has been observed in [3,4,12].

This paper is organized as follows : After a brief description of the derivation of the 2-cyclic (red-black) Hermite Collocation system in Section 2, we proceed with the eigenvalue analysis of the USSOR preconditioned system in Section 3. In Section 4 we have included numerical experiments to demonstrate our results.

2 Hermite Collocation Poisson system

For the numerical solution of the model Poisson's problem

$$\begin{cases} \nabla^2 u(x, y) = f(x, y) , & (x, y) \in \Omega \\ u(x, y) = g(x, y) , & (x, y) \in \partial\Omega \end{cases} \quad (1)$$

on the rectangular domain $\Omega \equiv (0, 1) \times (0, 1)$, we assume a uniform partition of the intervals $I^x = I^y = [0, 1]$ into $n_s = 2p$ subintervals $I_m^x = I_m^y$, $m = 1, \dots, n_s$. This generates a uniform grid with spacing $h = \frac{1}{n_s}$ and nodal coordinates (x_i, y_j) , where $x_i = (i-1)h$ and $y_j = (j-1)h$, $i, j = 1, \dots, (n_s + 1)$. The Hermite Bi-Cubic finite element approximation seeks an approximate solution $\tilde{u}(x, y)$ in the form

$$u(x, y) \sim \tilde{u}(x, y) = \sum_{i=1}^{\tilde{n}} \sum_{j=1}^{\tilde{n}} \alpha_{i,j} \phi_i(x) \phi_j(y) , \quad (2)$$

with $\tilde{n} = 2(n_s + 1)$. The basis functions $\phi_i(x)$ and $\phi_j(y)$ are the known one dimensional piecewise Hermite cubic polynomials. Based, now, on the basic prop-

erties of the Hermite basis functions, the following four unknowns

$$\begin{cases} a_{2i-1,2j-1} = \tilde{u}(x_i, y_j) \\ a_{2i-1,2j} = \frac{\partial}{\partial y} \tilde{u}(x_i, y_j) \\ a_{2i,2j-1} = \frac{\partial}{\partial x} \tilde{u}(x_i, y_j) \\ a_{2i,2j} = \frac{\partial^2}{\partial x \partial y} \tilde{u}(x_i, y_j) \end{cases} \quad (3)$$

are associated with the mesh point (x_i, y_i) . By imposing the boundary conditions, $8n_s + 4$ unknowns (denoted by \circ on Figure 1), associated with nodes on the boundary $\partial\Omega$, can be determined beforehand. Therefore, the collocation equations needed for the determination of the remaining $n = 4n_s^2$ unknowns are then constructed by forcing the approximate solution $\tilde{u}(x, y)$ to satisfy the BVP in n interior collocation points. These are the four Gauss points in each of the n_s^2 elements $I_{i,j}$, a classical choice for orthogonal spline Collocation.

The Collocation method poses no restrictions in how one orders/numbers equations and unknowns for the construction of the associated system

$$Ax = b , \quad (4)$$

where A is the $n \times n$ Collocation coefficient matrix and

$$x = [x_1 \ x_2 \ \dots \ x_n]^T \equiv [\alpha_{1,1} \ \dots \ \alpha_{\tilde{n},\tilde{n}}]^T$$

is the unknown vector. The block form of the collocation matrix A depends directly on the numbering of unknowns and equations. And as there is an one-to-one correspondence between collocation points (denoted by

with $\omega + \hat{\omega} - \omega\hat{\omega} \neq 0$, the USSOR iteration matrix is given by

$$\mathcal{T}_{\omega, \hat{\omega}} = I - M_{\omega, \hat{\omega}}^{-1} A, \quad (14)$$

where the parameters ω and $\hat{\omega}$ are referred as (over) relaxation parameters. Notice that for specific choices of the relaxation parameters one may recover well known iterative schemes. For example :

- when $(\omega, \hat{\omega}) = (1, 0)$ the USSOR reduces to the Gauss-Seidel method
- when $(\omega, \hat{\omega}) = (0, 1)$ the USSOR reduces to the "backward" Gauss-Seidel method
- when $(\omega, \hat{\omega}) = (\omega, 0)$ the USSOR reduces to the SOR method
- when $(\omega, \hat{\omega}) = (\omega, \omega)$ the USSOR reduces to the SSOR method
- when $(\omega, \hat{\omega}) = (1, 1)$ the USSOR reduces to the SGS method.

Jacobi and USSOR Eigenvalues

Taking into consideration the 2-cyclic [17] structure of the matrix A , it is shown [15] that the eigenvalues τ of the USSOR matrix $\mathcal{T}_{\omega, \hat{\omega}}$ and the eigenvalues μ of the Jacobi iteration matrix \mathcal{T} are related through the eigenvalue relationship

$$[\tau - (1 - \omega)(1 - \hat{\omega})]^2 = \tau(\omega + \hat{\omega} - \omega\hat{\omega})^2 \mu^2. \quad (15)$$

Inspecting now the above equation one may easily observe that :

- upon setting $\tilde{\omega} = \omega + \hat{\omega} - \omega\hat{\omega}$, the spectrums of the USSOR iteration matrix $\mathcal{T}_{\omega, \hat{\omega}}$ and the SOR iteration matrix $\mathcal{T}_{\tilde{\omega}}$ coincide, that is

$$\sigma(\mathcal{T}_{\omega, \hat{\omega}}) = \sigma(\mathcal{T}_{\tilde{\omega}}), \quad (16)$$

and, therefore, the convergence rate of the optimal USSOR method coincides with the rate of convergence of the optimal SOR method

- upon setting $\omega = 1$ or $\hat{\omega} = 1$ equation (16) reduces to

$$\tau = \mu^2 \quad (17)$$

and, therefore,

$$\sigma(\mathcal{T}_{1, \hat{\omega}}) = \sigma(\mathcal{T}_{\omega, 1}) \quad (18)$$

hence, the spectrums of the forward and backward GS iteration matrices $\mathcal{T}_{1,0}$ and $\mathcal{T}_{0,1}$ as well as the spectrum of the SGS iteration matrix $\mathcal{T}_{1,1}$ coincide, that is

$$\sigma(\mathcal{T}_{1,0}) = \sigma(\mathcal{T}_{0,1}) = \sigma(\mathcal{T}_{1,1}), \quad (19)$$

with, of course, $\sigma(\mathcal{T}_{1,0}) = \{\mu_k^2, \mu_k \in \sigma(\mathcal{T})\}$.

And, moreover, as the eigenvalues μ_k (depicted in Figure 4) of the Jacobi iteration matrix have been analytically evaluated in [8], equation (15) easily determines the spectrum of the USSOR family of methods.

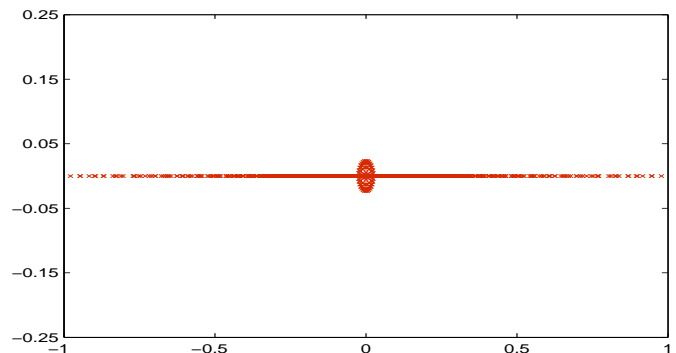


Fig. 4 : Jacobi Eigenvalues for $n_s = 16$.

Effective Preconditioning for the Bi-CGSTAB

A USSOR preconditioned Krylov subspace method is, of course, a Krylov method used to solve the USSOR preconditioned system

$$\tilde{A}_{\omega, \hat{\omega}} x = \tilde{b} \quad (20)$$

where

$$\tilde{A}_{\omega, \hat{\omega}} = M_{\omega, \hat{\omega}}^{-1} A \quad \text{and} \quad \tilde{b} = M_{\omega, \hat{\omega}}^{-1} b. \quad (21)$$

For effective preconditioning we require $\tilde{A}_{\omega, \hat{\omega}} \approx I$ or, equivalently,

$$M_{\omega, \hat{\omega}}^{-1} A \approx I \Leftrightarrow A \approx M_{\omega, \hat{\omega}} \quad (22)$$

and at the same time we allow fast iterations. For this observe the following :

I. Distribution of Eigenvalues. Following [5], for effective convergence properties of the Bi-CGSTAB, the eigenvalues of the preconditioned matrix \tilde{A} (or, their approximations by the eigenvalues of the reduced matrix, namely the Ritz values) should be located in the half complex plane with small imaginary parts and the origin being outside, or towards the boundary, of the convex hull containing them. Moreover it is well known, and apparent by (22), that they also have to be clustered around unity.

In this direction, recall relation (14) and observe that

$$\tilde{A}_{\omega, \hat{\omega}} = M_{\omega, \hat{\omega}}^{-1} A = I - \mathcal{T}_{\omega, \hat{\omega}} \quad (23)$$

hence

$$\sigma(\tilde{A}_{\omega, \hat{\omega}}) = \{1 - \tau_k, \tau_k \in \sigma(\mathcal{T}_{\omega, \hat{\omega}})\}. \quad (24)$$

Therefore, to satisfy the qualifications for effective performance of the Bi-CGSTAB, described earlier, there must hold

$$0 < 1 - Re(\tau_k) \approx 1 \quad \text{and} \quad |Im(\tau_k)| \approx 0. \quad (25)$$

To avoid cumbersome algebraic manipulations, notice that, for $\omega = 1$ or $\hat{\omega} = 1$, the spectrums of the corresponding preconditioned matrices satisfy

$$\sigma(\tilde{A}_{\omega,1}) \equiv \sigma(\tilde{A}_{1,\hat{\omega}}) = \{1 - \mu_k^2, \mu_k \in \sigma(T)\}. \quad (26)$$

Inspecting Figure 5, where we depict these eigenvalues, one may easily verify that the conditions in (25) are readily verified.

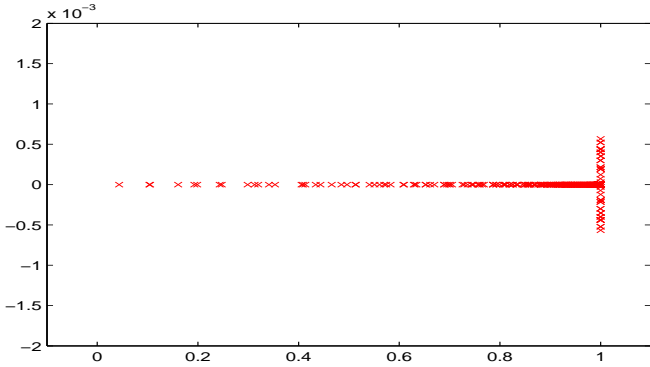


Fig. 5 : The spectrums $\sigma(\tilde{A}_{\omega,1}) \equiv \sigma(\tilde{A}_{1,\hat{\omega}})$ for $n_s = 16$.

II. Error Matrices. Recalling now equation (13), it is evident that the error matrix $\mathcal{E}_{\omega,\hat{\omega}} \doteq A - M_{\omega,\hat{\omega}}$ for the USSOR preconditioning is given by

$$\mathcal{E}_{\omega,\hat{\omega}} = \frac{(\tilde{\omega} - 1)}{\tilde{\omega}} D_A - \frac{(\tilde{\omega} - \omega)}{\tilde{\omega}} L_A - U_{\omega,\hat{\omega}} \quad (27)$$

where, $\tilde{\omega}$ is as in (16) and

$$U_{\omega,\hat{\omega}} \doteq \frac{(\tilde{\omega} - \hat{\omega})}{\tilde{\omega}} U_A + \frac{\omega \hat{\omega}}{\tilde{\omega}} L_A D_A^{-1} U_A \quad (28)$$

is a strictly upper triangular matrix. Recalling relations (10)-(12), $\mathcal{E}_{\omega,\hat{\omega}}$ takes the form

$$\mathcal{E}_{\omega,\hat{\omega}} = \begin{pmatrix} \frac{(\tilde{\omega}-1)}{\tilde{\omega}} D_R & \frac{(\tilde{\omega}-\hat{\omega})}{\tilde{\omega}} H_B - \frac{\omega \hat{\omega}}{\tilde{\omega}} H_R D_R^{-1} H_B \\ \frac{(\tilde{\omega}-\omega)}{\tilde{\omega}} H_R & \frac{(\tilde{\omega}-1)}{\tilde{\omega}} D_B \end{pmatrix}. \quad (29)$$

Therefore, by noticing that for $\omega = 1$ or $\hat{\omega} = 1$ there holds $\tilde{\omega} = 1$, we obtain

$$\mathcal{E}_{1,\hat{\omega}} = \begin{pmatrix} O & (1 - \hat{\omega})H_B - \hat{\omega}H_R D_R^{-1} H_B \\ O & O \end{pmatrix} \quad (30)$$

and

$$\mathcal{E}_{\omega,1} = \begin{pmatrix} O & -\omega H_R D_R^{-1} H_B \\ (1 - \omega)H_R & O \end{pmatrix}. \quad (31)$$

The choice $\omega = 0$ or $\hat{\omega} = 0$, for which the USSOR reduces to backward-GS and GS, respectively, have the following obvious advantages :

- They decrease, in a very simple way, the number of nonzero entries of the corresponding error matrices

$$\mathcal{E}_{1,0} = \begin{pmatrix} O & H_B \\ O & O \end{pmatrix} \quad \mathcal{E}_{0,1} = \begin{pmatrix} O & O \\ H_R & O \end{pmatrix}$$

- They decrease the number of operations involved, in case of explicit pre or post conditioning, since the corresponding preconditioning matrices are readily simplified.

Therefore, different choices of the relaxation parameters will be considered to be *better* only if they result in substantial reduction of iteration count. In the Figures 6 and 7, we graphically illustrate the iteration count of the explicitly pre or post conditioned Bi-CGSTAB method, by $\mathcal{M}_{1,\hat{\omega}}$ and $\mathcal{M}_{\omega,1}$ respectively, as a function of $\hat{\omega}$ and ω . By inspection, now, of said figures, one may easily observe that there are other, than the zero, values of the relaxation parameters $\hat{\omega}$ and ω for which the corresponding preconditioned Bi-CGSTAB attains better convergence. However, the difference in the number of iterations is small and does not justify the usage of other relaxation values. There is always, of course, the issue of efficient implementation of pre/post conditioning which will be addressed in another publication.

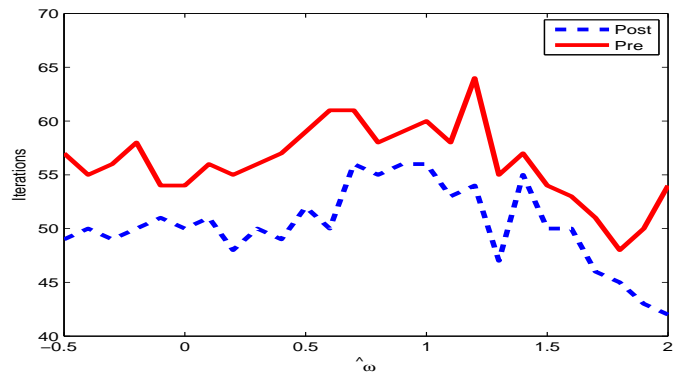


Fig. 6 : Iteration count of the pre/post conditioned Bi-CGSTAB method by $\mathcal{M}_{1,\hat{\omega}}$ for $n_s = 64$.

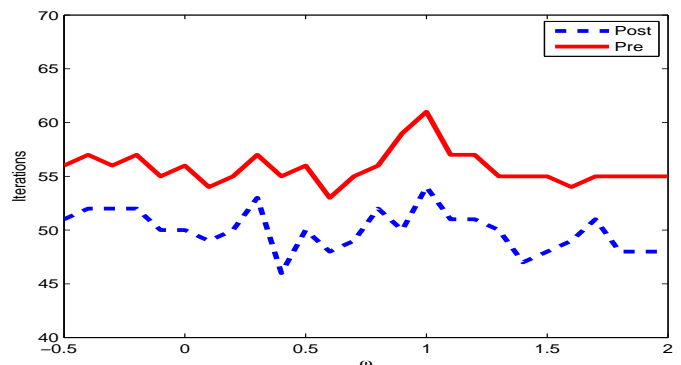


Fig. 7 : Iteration count of the pre/post conditioned Bi-CGSTAB method by $\mathcal{M}_{\omega,1}$ for $n_s = 64$.

4 Numerical Verification

We conclude our work herein by presenting, in Tables T1 and T2, results from numerical experimentation for a test Poisson problem with exact solution defined by

$$u(x, y) = 10 \phi(x) \phi(y), \quad \phi(x) = e^{-100(x-0.1)^2} (x^2 - x).$$

T1		GS preconditioned Bi-CGSTAB			
n_s	Iterations	Time	$\ u - x\ _\infty$	$\ b - Ax\ _\infty$	
16	16	0.029	6.08e-4	4.89e-7	
32	29	0.188	3.35e-5	7.16e-7	
64	59	1.473	1.79e-6	3.43e-7	
128	116	11.86	2.64e-6	1.50e-7	
256	221	102.7	2.49e-6	3.23e-8	

T2		B-GS postconditioned Bi-CGSTAB			
n_s	Iterations	Time	$\ u - x\ _\infty$	$\ b - Ax\ _\infty$	
16	14	0.023	6.07e-4	2.18e-5	
32	22	0.136	3.20e-5	2.29e-5	
64	62	1.510	2.04e-6	2.99e-8	
128	103	10.32	2.15e-6	5.16e-7	
256	207	95.19	2.37e-6	5.55e-8	

In Table T1 we have included the results from the GS preconditioned Bi-CGSTAB, which in our implementation, performs relatively better from the backward-GS preconditioning. The opposite is true for the case of postconditioning. Observe that the backward-GS postconditioned Bi-CGSTAB performs better. Efficient implementation of all schemes is under study.

References :

[1] S. Brill, Hermite Collocation solution of Partial Differential Equations via preconditioned Krylov methods, *Numerical Methods for PDEs*, 17(2), 2001, pp. 120-136.

[2] S. Brill and G. Pinder, Eigenvalue Analysis of a Block Red-Black Preconditioner applied to the Hermite Collocation discretization of Poissons's equation, *Numerical Methods for PDEs*, 17(3), 2001, pp. 204-228.

[3] M. DeLong and J. Ortega, SOR as a preconditioner, *Applied Numerical Mathematics*, 18, 1995, pp. 431-440.

[4] M. DeLong and J. Ortega, SOR as a preconditioner II, *Applied Numerical Mathematics*, 26, 1998, pp. 465-481.

[5] J. Dongarra, I. Duff, D. Sorensen, H. van der Vorst, *Numerical Linear Algebra for High-Performance Computers*, SIAM, 1998.

[6] A. Hadjidimos and Y. Saridakis, MSOR and equivalent 2-step iterative methods for Collocation Matrices, *J. Comp. Appl. Math.* 42(3), 1992, pp. 375-393.

[7] A. Hadjidimos, T. Papatheodorou and Y. Saridakis, Optimal Block Iterative schemes for certain large sparse and non-symmetric linear systems, *Linear Alg. and Appl.* 110, 1988, pp. 285-318.

[8] Y. Lai, A. Hadjidimos, E. Houstis and J. Rice, On the Iterative solution of Hermite Collocation equations, *SIAM J. Matrix Anal. Appl.*, 16, 1995, pp. 254-277

[9] E. Mathioudakis, E. Papadopoulou and Y. Saridakis, Non-Stationary Iterative Schemes for the Solution of Elliptic Collocation Systems, *Procs of the 4nd Hellenic - European conf. on computer mathematics and its applications*, 1998, pp. 1044-1051.

[10] E. Mathioudakis, E. Papadopoulou and Y. Saridakis, Iterative Solution of Elliptic Collocation Systems on a Cognitive Parallel Computer, *Computers and Mathematics with applications*, 48, 2004, pp. 951-970.

[11] T. Papatheodorou, Block AOR Iteration for nonsymmetric matrices, *Mathematics of Computation*, 41, 1983, pp. 511-525.

[12] Y. Saad, Highly parallel preconditioners for general sparse matrices, *IMA Volumes in Mathematics and its Applications*, 60, Springer Verlag, pp. 165-199.

[13] Y. Saad and M. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.*, 7, 1986, pp. 856-869.

[14] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, 2003

[15] Y. Saridakis, On the analysis of the Unsymmetric Successive Overrelaxation Method when applied to p-cyclic matrices, *Numer. Math.*, 49, 1986, pp. 461-473.

[16] H. van der Vorst, Bi-CGSTAB : A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.*, 13, 1992, pp. 631-644.

[17] R. Varga, *Matrix Iterative Analysis*, Prentice Hall, 1962