SIAM J. SCI. COMPUT.
Vol. 19, No. 6, pp. 1918–1940, November 1998

008

ITERATIVE SOLUTION OF CYCLICALLY REDUCED SYSTEMS ARISING FROM DISCRETIZATION OF THE THREE-DIMENSIONAL CONVECTION-DIFFUSION EQUATION*

CHEN GREIF† AND JAMES VARAH‡

Abstract. We consider the system of equations arising from finite difference discretization of a three-dimensional convection-diffusion model problem. This system is typically nonsymmetric. We show that performing one step of cyclic reduction, followed by reordering of the unknowns, yields a system of equations for which the block Jacobi method generally converges faster than for the original system, using lexicographic ordering. The matrix representing the system of equations can be symmetrized for a large range of the coefficients of the underlying partial differential equation, and the associated iteration matrix has a smaller spectral radius than the one associated with the original system. In this sense, the three-dimensional problem is similar to the one-dimensional and two-dimensional problems, which have been studied by Elman and Golub. The process of reduction, the suggested orderings, and bounds on the spectral radii of the associated iteration matrices are presented, followed by a comparison of the reduced system with the full system and by details of the numerical experiments.

Key words. finite difference discretization, block iterative schemes, red/black ordering, reduced system

AMS subject classifications. 65F10, 65H22, 65F50

PII. S1064827596296994

1. Introduction.

Consider the following three-dimensional (3D) elliptic problem

$$-\Delta u + (\sigma, \tau, \mu)^T \ \nabla u = p \tag{1.1}$$

on a domain $\Omega \in \mathbb{R}^3$, subject to Dirichlet-type boundary conditions. Standard finite difference discretization of (1.1), for example, seven-point approximation to the 3D Laplacian [7] and upwind or centered difference approximations to the first-order derivatives, leads to a linear system

$$Au = p \ , \tag{1.2}$$

where now u and p denote vectors of a finite size, representing the approximated values of the solution to the continuous problem and the exact values of the right-hand side forcing term, respectively, at the grid points. If $(\sigma, \tau, \mu) \neq 0$ the matrix A is nonsymmetric. It is not necessarily diagonally dominant or symmetrizable by real similarity transformations. However, it does have property A (see, for example, [7, section 9.2]), therefore discretizing the equation using red/black ordering of the grid results in a system whose matrix is of the form $\left(\begin{smallmatrix} B & C \\ D & E \end{smallmatrix}\right)$, where both B and E are diagonal. Thus a cheap process of elimination of points corresponding to one of the colors leads to a so-called reduced system, whose matrix is given by $E - DB^{-1}C$. If

*Received by the editors January 3, 1996; accepted for publication (in revised form) January 24, 1997; published electronically July 27, 1998.
http://www.siam.org/journals/sisc/19-6/29699.html
†Institute of Applied Mathematics, University of British Columbia, Vancouver, BC, Canada V6T 1Z2 (greif@math.ubc.ca).
‡Department of Computer Science, University of British Columbia, Vancouver, BC, Canada V6T 1Z4 (varah@cicsr.ubc.ca).

the original system is sparse, then the reduced system is sparse as well. This process of elimination amounts to performing one step of cyclic reduction [3].

Elman and Golub have conducted an extensive investigation for two-dimensional elliptic problems [3], [4], [5] and have shown that the reduced system for the two-dimensional (2D) case has some valuable properties: it can be symmetrized for a large range of values of the coefficients of the PDE, it has block property A for several orderings so that Young's classical SOR analysis [9] can be applied, and also, for several iterative methods that were studied, the convergence rates of the reduced systems are typically faster than for the analogous original (full) systems. The reader is referred to the above-mentioned references for full details on the analysis of the 2D problem.

In this paper we examine the 3D case and present convergence analysis for the block Jacobi iterative method, applied to the reduced system, based on a certain ordering of the grid. We focus on a model problem with constant coefficients on the unit cube. We mainly refer to the case where the reduced system can be symmetrized, with emphasis on equations with small convection terms; however, we stress that our numerical findings suggest that the reduced system leads to an efficient solution process also for other cases. For the ordering suggested we closely examine the algebraic properties of the underlying matrix. We analyze both upwind and centered difference discretizations.

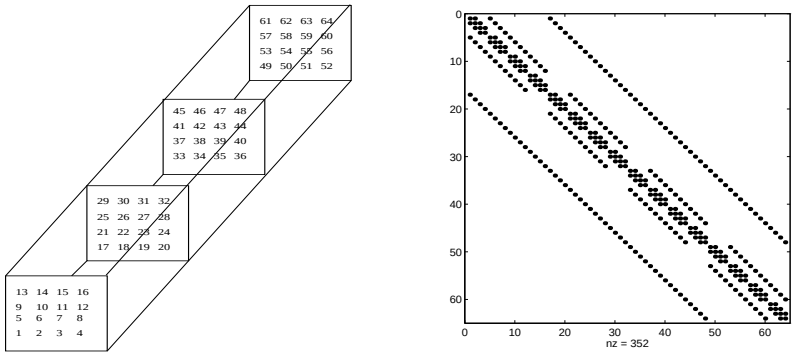
For notational convenience, we use the following rules throughout the paper:

- $E_{s_1,\ldots,s_k}$ will denote a vector whose entries are $s_1,\ldots,s_k$, repeated. For example $E_{01} = (0,1,0,1,0,1,\ldots)$, $E_{1001} = (1,0,0,1,1,0,0,1,\ldots)$, and so on. The size of the vector will be clear from the context where it appears.
- For narrow banded matrices we use the standard "diag," "tri,", "penta," etc. If x, y, and z are vectors of size n, then tri$[x,y,z]$ will denote a tridiagonal matrix whose main diagonal consists of entries of the vector y; the subdiagonal consists of x_1 to x_{n-1}, and the superdiagonal consists of z_2 to z_n. x_n and z_1 do not appear in the matrix in this case.
- The index of a diagonal in a matrix will match the syntax of the Matlab command "diag" : 0 for the main diagonal, positive numbers for superdiagonals, and negative numbers for subdiagonals.
- I_n will stand for the identity matrix of order n.

All the computations were carried out on an HP 735 machine, using Matlab 4.2a and its sparse matrix features.

An outline of the rest of this paper follows. In section 2 we present the system of equations arising from discretization of the problem using natural lexicographic ordering. We analyze the spectrum of the iteration matrix associated with the line Jacobi iterative scheme. In section 3 we describe the process of 3D cyclic reduction. We present the computational molecule and the system resulting from a certain ordering strategy, which we call two-plane ordering. We derive general symmetrization conditions for the reduced system, and we obtain analytical bounds for the spectral radius of the block Jacobi iteration matrix. In all the numerical experiments we have conducted, the spectral radius of the reduced system is smaller than the one for the full system. In section 4 we examine the asymptotic bounds and we address the question of performance and amount of computational work. Then we present numerical results which validate the analysis. In section 5 we summarize our findings and draw some conclusions.

2. The full system. Consider the 3D elliptic problem (1.1) on $\Omega = (0,1) \times (0,1) \times (0,1)$, subject to Dirichlet boundary conditions: $u = r(x,y,z)$ on $\partial\Omega$, and suppose that (1.2) represents the system arising from finite difference discretization, using natural lexicographic ordering of the unknowns. To illustrate this, the ordering for $n = 4$ is depicted in Fig. 2.1(a).



(a) ordering of the unknowns (b) sparsity pattern of matrix

FIG. 2.1. *Natural lexicographic ordering of the unknowns.*

Let F denote the operator for this system, after scaling by h^2, and denote the values of the associated computational molecule by a, b, c, d, e, f, and g, in the following manner: if $u_{i,j,k} = u(ih, jh, kh)$ is a grid point not next to the boundary, then

$$\begin{aligned} F\,u_{i,j,k} = {} & a\,u_{i,j,k} + b\,u_{i,j-1,k} + c\,u_{i-1,j,k} \\ & + d\,u_{i+1,j,k} + e\,u_{i,j+1,k} + f\,u_{i,j,k-1} + g\,u_{i,j,k+1}\ . \end{aligned} \tag{2.1}$$

The matrix A is an nth-order block tridiagonal matrix with respect to $n^2 \times n^2$ blocks:

$$A = \mathrm{tri}[A^{(-1)},\ A^{(0)},\ A^{(1)}]\ , \tag{2.2}$$

where $A^{(0)} = f\ I_{n^2},\ A^{(1)} = g\ I_{n^2}$, and $A^{(0)}$ are themselves block tridiagonal matrices,

$$A^{(0)} = \mathrm{tri}[B^{(-1)}, B^{(0)}, B^{(1)}]\ , \tag{2.3}$$

consisting of $n \times n$ matrices given by

$$B^{(-1)} = b\ I_n\ ;\quad B^{(0)} = \mathrm{tri}[c, a, d]\ ;\quad B^{(1)} = e\ I_n\ . \tag{2.4}$$

The sparsity pattern of A is depicted in Fig. 2.1(b).

Let $h = \frac{1}{n+1}$ denote the mesh size. If we use centered differences for approximating the first-order derivatives, the values of the computational molecule are as follows:

$$\begin{aligned} a = 6\ ;\ b = -1 - \frac{\tau h}{2}\ ;\ c = -1 - \frac{\sigma h}{2}\ ;\ d = -1 + \frac{\sigma h}{2}\ ; \\ e = -1 + \frac{\tau h}{2}\ ;\ f = -1 - \frac{\mu h}{2}\ ;\ g = -1 + \frac{\mu h}{2}\ . \end{aligned} \tag{2.5}$$

For upwind differences, assuming that σ, τ, and μ are positive, we use backward difference approximations, and in this case we have

$$a = 6 + (\sigma + \tau + \mu)h \ ; \ b = -1 - \tau h \ ; \ c = -1 - \sigma h \ ; \quad d = -1 \ ; \ e = -1 \ ; \ f = -1 - \mu h \ ; \ g = -1 \ . \tag{2.6}$$

Let us denote the cell Reynolds numbers by

$$\beta = \frac{\sigma h}{2}, \quad \gamma = \frac{\tau h}{2}, \quad \delta = \frac{\mu h}{2} \ . \tag{2.7}$$

Following is a convergence analysis for the line Jacobi scheme.

2.1. Analysis for the block Jacobi iteration. Consider the splitting $A = D - C$, where D is the one dimensional preconditioner

$$D = \mathrm{diag}[B^{(0)}, \ldots, B^{(0)}] \ . \tag{2.8}$$

The associated block Jacobi iterative scheme is

$$u^{(k+1)} = D^{-1}Cu^{(k)} + D^{-1}p \ . \tag{2.9}$$

We are interested in the spectral radius of the iteration matrix $D^{-1}C$. In general, we will adopt for both the full system and the reduced system the strategy used by Elman and Golub in [3], [4], [5]: we find circumstances in which the matrix can be symmetrized by a real diagonal matrix, say Q; then, we denote $\hat{A} = Q^{-1}AQ$ and look at the splitting of this matrix. We let $\hat{D}$ and $\hat{C}$ denote $Q^{-1}DQ$ and $Q^{-1}CQ$, respectively, and since $\hat{D}^{-1}\hat{C} = Q^{-1}D^{-1}CQ$ it follows that both $D^{-1}C$ and $\hat{D}^{-1}\hat{C}$ have the same spectrum, thus we can continue the analysis with the latter.

Motivated by this strategy, we now present the following symmetrization result for the full system.

THEOREM 2.1. *If $cd > 0$, $be > 0$, and $fg > 0$ then there exists a real nonsingular diagonal matrix Q such that the matrix $\hat{A} = Q^{-1}AQ$ is symmetric.*

Proof. This can be shown by direct substitution, by requiring that the symmetrized matrix be equal to its transpose. Denote the ith entry in the main diagonal of Q by q_i. $q_1 \neq 0$ can be an arbitrary value, and the following algorithm generates the desired diagonal matrix:

for $i = 2$ to n
  $q_i = \sqrt{\frac{c}{d}} \cdot q_{i-1}$

for $\ell = 2$ to n
  for $i = 1$ to n
    $q_{(\ell-1)n+i} = \sqrt{\frac{b}{e}} \cdot q_{(\ell-2)n+i}$

for $j = 2$ to n
  for $\ell = 1$ to n
    for $i = 1$ to n
      $q_{(j-1)n^2+(\ell-1)n+i} = \sqrt{\frac{f}{g}} \cdot q_{(j-2)n^2+(\ell-1)n+i}$

From the above it is clear that the similarity transformation is real and nonsingular only if cd, fg, and be are positive. ∎

The resulting symmetrized matrix is

$$\hat{A} = \mathrm{tri}[\hat{A}^{(-1)}, \hat{A}^{(0)}, \hat{A}^{(1)}] \ , \tag{2.10}$$

where

$$\hat{A}^{(-1)} = \hat{A}^{(1)} = \sqrt{fg}\ I_{n^2}; \quad \hat{A}^{(0)} = \mathrm{tri}[\hat{B}^{(-1)},\ \hat{B}^{(0)},\ \hat{B}^{(1)}] \ , \tag{2.11}$$

with

$$\hat{B}^{(-1)} = \hat{B}^{(1)} = \sqrt{be}\ I_n; \quad \hat{B}^{(0)} = \mathrm{tri}[\sqrt{cd},\ a,\ \sqrt{cd}] \ . \tag{2.12}$$

Theorem 2.1 leads to the following symmetrization result, obtained by expressing the conditions in terms of $\beta,\ \gamma$, and δ.

COROLLARY 2.2. *If $|\beta| < 1,\ |\gamma| < 1$, and $|\delta| < 1$, the coefficient matrix for the centered difference scheme is symmetrizable by a real diagonal similarity transformation. For upwind (backward) schemes, the coefficient matrix is symmetrizable for all $\beta,\ \gamma,\ \delta > 0$.*

Corollary 2.2 can be viewed as a straightforward generalization of the result for the 2D case [3].

We now examine the spectrum of the symmetrized iteration matrix. We start this part of the analysis by quoting a lemma that appears in [3, p. 674], which will be useful in the discussion that follows. We then present the eigenvalues of the iteration matrix.

LEMMA 2.3. *The eigenvalues of the tridiagonal matrix* $\mathrm{tri}[b, a, c]$ *of order n are $\{\lambda_j = a + \mathrm{sign}(c) 2\sqrt{bc} \cdot \cos(j\pi/(n+1)),\ j = 1, \ldots, n\}$. The eigenvectors corresponding to each λ_j are $v^{(j)}$, where $v_k^{(j)} = (b/c)^{k/2} \sin(\pi jk/(n+1))\ ,\ k = 1, \ldots, n$.*

THEOREM 2.4. *The eigenvalues of the iteration matrix $\hat{D}^{-1}\hat{C}$ are given by*

$$\frac{2\sqrt{fg} \cdot \cos(\pi jh) + 2\sqrt{be} \cdot \cos(\pi kh)}{a + 2\sqrt{cd} \cdot \cos(\pi \ell h)}, \qquad 1 \le j, k, \ell \le n \ . \tag{2.13}$$

Proof. Inspired by the techniques used in [3, p. 677], suppose V_n is an $n \times n$ matrix whose columns are the eigenvalues of the tridiagonal matrix $\mathrm{tri}[\sqrt{cd}, a, \sqrt{cd}]$, and let $V = \mathrm{diag}[V_n, \ldots, V_n]$ have n^2 copies of V_n. Then $\tilde{A} = V^{-1}\hat{A}V$ diagonalizes $\hat{D}$. Denote $\tilde{D} = V^{-1}\hat{D}V$ and $\tilde{C} = V^{-1}\hat{C}V$. Since $V^{-1}\tilde{D}^{-1}\tilde{C}V = \hat{D}^{-1}\hat{C}$ we can find the spectrum of $\hat{D}^{-1}\hat{C}$ by examining $\tilde{D}^{-1}\tilde{C}$. The latter is easier to form explicitly, as $\tilde{D}$ is diagonal (as opposed to $\hat{D}$, which is tridiagonal). $\tilde{D}^{-1}\tilde{C}$ has the same nonzero pattern as $\hat{C}$.

Let P_1 denote the permutation matrix that transforms rowwise ordering into columnwise ordering, leaving the ordering of planes (z direction) unchanged. $F_1 = P_1^T \tilde{D}^{-1}\tilde{C}P_1$ is block tridiagonal with respect to $n^2 \times n^2$ blocks, and its superdiagonal and subdiagonal blocks are diagonal matrices, all equal to each other. Each of these $n^2 \times n^2$ matrices looks as follows:

$$\mathrm{diag}\left[\underbrace{\frac{\sqrt{fg}}{a + 2\sqrt{cd}\cos(\pi h)}}_{n \text{ terms}}, \underbrace{\frac{\sqrt{fg}}{a + 2\sqrt{cd}\cos(2\pi h)}}_{n \text{ terms}}, \ldots, \underbrace{\frac{\sqrt{fg}}{a + 2\sqrt{cd}\cos(\pi n h)}}_{n \text{ terms}} \right] \ . \tag{2.14}$$

The diagonal block consists of n identical $n^2 \times n^2$ matrices, each being an nth-order block diagonal matrix whose jth component is the tridiagonal matrix

$$G_j = \text{tri}\left[\frac{\sqrt{be}}{a + 2\sqrt{cd}\cos(\pi jh)}\right] . \tag{2.15}$$

Let V_{n^2} denote a matrix whose columns are the eigenvectors of $\text{diag}[G_1, \ldots, G_n]$, and let V_2 be the block diagonal matrix consisting of n uncoupled copies of V_{n^2}; then $F_2 = V_2^{-1} F_1 V_2$ is still block tridiagonal with respect to $n^2 \times n^2$ blocks, but now the main diagonal block is a *diagonal* matrix: it contains n identical $n^2 \times n^2$ blocks, each being a diagonal matrix whose entries are $\frac{2\sqrt{be}\cos(\pi jh)}{a+2\sqrt{cd}\cos(\pi kh)}$, $1 \le j, k \le n$. Finally, let P_2 be a permutation matrix which transforms rowwise ordering to planewise ordering, leaving the orientation of columns unchanged. Then $F_3 = P_2^{-1} F_2 P_2$ is an n^2-order block diagonal matrix whose components are $n \times n$ symmetric tridiagonal matrices. Their eigenvalues can be found using Lemma 2.3 and are given by (2.13). □

We can now state the following useful result.

COROLLARY 2.5. *For cd, be, $fg > 0$ the spectral radius of the line Jacobi iteration matrix, using the preconditioner defined in* (2.8), *is*

$$\frac{2\sqrt{be}\cdot\cos(\pi h) + 2\sqrt{fg}\cdot\cos(\pi h)}{a + 2\sqrt{cd}\cdot\cos(\pi nh)} . \tag{2.16}$$

3. The reduced system.

3.1. Construction of the reduced system. The construction of the reduced matrix is a process of Gaussian elimination of half of the rows and the columns in the original system. In Fig. 3.1 we number the points that have to do with the block elimination associated with a typical grid point. The central point is indexed in this case by 13.

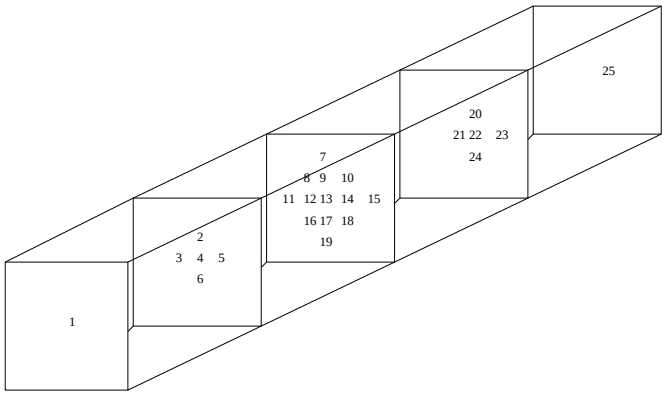


FIG. 3.1. *Points that are affected by block elimination for point #13 (center).*

If we decide that #13 is "black," then the following indices are "red" and are to be eliminated: #4, #9, #12, #14, #17, and #22. The other points are "black," but their corresponding entries in the matrix at row #13 are to be changed after eliminating the "red" points. In terms of the matrix elements, below are the entries

of the matrix A which are affected by the block elimination step:

column :	4	9	12	14	17	22	13
row 4	a						g
row 9		a					b
row 12			a				d
row 14				a			c
row 17					a		e
row 22						a	f
row 13	f	e	c	d	b	g	a

column :	1	2	3	5	6	7	8	10	11	15	16	18	19	20	21	23	24	25
row 4	f	e	c	d	b													
row 9		f				e	c	d						g				
row 12			f				e		c		b				g			
row 14				f				e		d		b				g		
row 17					f						c	d	b				g	
row 22														e	c	d	b	g

Here are the computations performed during the block elimination. The numbers at the first column correspond to the rows that are eliminated:

col.	13	1	2	3	5	6	7	8	10
4 :	$-fa^{-1}g$	$-fa^{-1}f$	$-fa^{-1}e$	$-fa^{-1}c$	$-fa^{-1}d$	$-fa^{-1}b$			
9 :	$-ea^{-1}b$		$-ea^{-1}f$				$-ea^{-1}e$	$-ea^{-1}c$	$-ea^{-1}d$
12 :	$-ca^{-1}d$			$-ca^{-1}f$				$-ca^{-1}e$	
14 :	$-da^{-1}c$				$-da^{-1}f$				$-da^{-1}e$
17 :	$-ba^{-1}e$					$-ba^{-1}f$			
22 :	$-ga^{-1}f$								

col.	11	15	16	18	19	20	21	23	24	25
4 :										
9 :						$-ea^{-1}g$				
12 :	$-ca^{-1}c$		$-ca^{-1}b$				$-ca^{-1}g$			
14 :		$-da^{-1}d$		$-da^{-1}b$				$-da^{-1}g$		
17 :			$-ba^{-1}c$	$-ba^{-1}d$	$-ba^{-1}b$				$-ba^{-1}g$	
22 :						$-ga^{-1}e$	$-ga^{-1}c$	$-ga^{-1}d$	$-ga^{-1}b$	$-ga^{-1}g$

From this, we can see that the typical value on the diagonal of the reduced matrix is

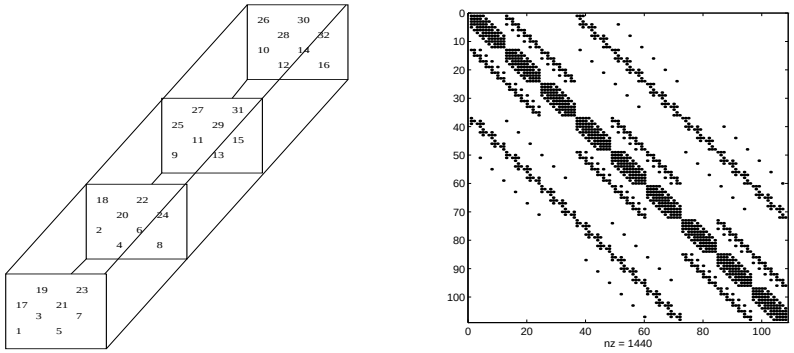
$$a^{-1}(a^2 - 2be - 2cd - 2fg)\ . \tag{3.1}$$

By "typical" we mean an entry that is associated with an *interior* point. That is, a point that is not next to the boundary of the domain. For noninterior points fewer operations of elimination are required, thus the value in their associated diagonal entry will change with respect to (3.1) in the following manner.

- Add $a^{-1}cd$ in case the x coordinate of the associated point is $1/(n+1)$ or $n/(n+1)$.
- Add $a^{-1}be$ in case the y coordinate of the associated point is $1/(n+1)$ or $n/(n+1)$.
- Add $a^{-1}fg$ in case the z coordinate of the associated point is $1/(n+1)$ or $n/(n+1)$.

Let R be the reduced operator, after scaling by ah^2. Then for an interior grid point, $u_{i,j,k}$, we have

$$\begin{aligned} R\,u_{i,j,k} = {} & (a^2 - 2be - 2cd - 2fg)\,u_{i,j,k} - f^2\,u_{i,j,k-2} - 2ef\,u_{i,j+1,k-1} \\ & -2cf\,u_{i-1,j,k-1} - 2df\,u_{i+1,j,k-1} - 2bf\,u_{i,j-1,k-1} - e^2\,u_{i,j+2,k} \\ & -2de\,u_{i+1,j+1,k} - c^2\,u_{i-2,j,k} - d^2\,u_{i+2,j,k} - 2bc\,u_{i-1,j-1,k} \\ & -b^2\,u_{i,j-2,k} - 2eg\,u_{i,j+1,k+1} - 2cg\,u_{i-1,j,k+1} - 2ce\,u_{i-1,j+1,k} \\ & -2bd\,u_{i+1,j-1,k} - 2dg\,u_{i+1,j,k+1} - 2bg\,u_{i,j-1,k+1} - g^2\,u_{i,j,k+2}\ . \end{aligned} \tag{3.2}$$



(a) ordering of the unknowns (b) sparsity pattern of matrix

FIG. 3.2. *Two-plane ordering.*

3.2. Orderings for the reduced system. With the reduced operator for the 3D problem at hand, we can observe that the underlying matrix of the reduced system is sparse, and the number of nonzero entries in each row is typically 19 (fewer for non-interior points). However, the exact sparsity pattern of the reduced matrix depends upon the ordering of the unknowns in the reduced grid. The ordering can significantly affect the convergence of whatever iterative scheme is used. The connection between the ordering of the unknowns and the direction of the flow for the 2D case is discussed in [5]. For the 3D case we have examined several ordering strategies. We mention some of them now.

- The two-line ordering is based on looking at the unknowns at each plane separately and grouping the points by pairs of horizontal lines. This ordering is a straightforward generalization to three dimensions of the two-line ordering used in [4].
- The natural lexicographic ordering is based on ordering the unknowns row-wise, and then planewise (analogous to the ordering discussed in section 2 for the full system).
- Two-plane ordering is based on gathering points from two horizontal lines and two adjacent planes simultaneously. This ordering fits only 3D problems, and is presented in this paper for the first time.
- Block red/black versions of the above.

Of the strategies we have examined, we so far find the two-plane ordering most effective. Its advantage is that in the underlying matrix more nonzero entries (compared to the other strategies we have examined) are clustered next to the main diagonal. For certain classes of matrices (e.g., M-matrices) this will lead to faster convergence (see, for example, [8, p. 91, Cor. 1]). We choose, then, to focus here on this ordering.

For the sake of simplicity we shall assume that n is even. The suggested ordering (for $n = 4$) and the sparsity pattern of the matrix (for $n = 6$) are depicted in Fig. 3.2. The reduced linear system is represented by an $\frac{n}{2}$-order block tridiagonal matrix

$$S = \mathrm{tri}[S_{j,j-1}, S_{j,j}, S_{j,j+1}] \ . \tag{3.3}$$

Each of the above components of S is an $n^2 \times n^2$ matrix, block tridiagonal with respect to $2n \times 2n$ blocks. Let us use superscripts (-1), (0), and (1) to describe subdiagonal blocks, diagonal blocks, and superdiagonal blocks of each of the block

matrices, respectively. Denote by "cntr" the center of the computational molecule, specified by (3.1) and the explanation that follows. The diagonal submatrices $S_{j,j}^{(0)}$ have nonzero entries in diagonals -4 to 4, and their diagonals are given as follows:

(3.4)
$$[-c^2,\ -2cf\cdot E_{10},\ -2ce\cdot E_{1001}-2bc\cdot E_{0110}, -2cg\cdot E_{01}-2ef\cdot E_{1000}-2bf\cdot E_{0010},$$
$$\text{cntr}, -2bg\cdot E_{0100}-2df\cdot E_{10}-2eg\cdot E_{0001},\ -2bdE_{0110}-2de\cdot E_{1001},$$
$$-2dg\cdot E_{01},\ -d^2].$$

The other submatrices contained in $S_{j,j}$ are of an irregular structure. The superdiagonal submatrix $S_{j,j}^{(1)}$ has nonzero entries in diagonals -3 to 1, as follows:

$$[-2cg\cdot E_{10},\ 0,\ -2egE_{1000}-2bg\cdot E_{0010},\ -g^2,\ -2dg\cdot E_{10}]\ , \tag{3.5}$$

and the subdiagonal submatrix $S_{j,j}^{(-1)}$ has nonzero entries in diagonals -1 to 3, as follows:

$$[-2cf\cdot E_{01},\ -f^2,\ -2bf\cdot E_{0100}-2ef\cdot E_{0001},\ 0,\ -2df\cdot E_{01}]\ . \tag{3.6}$$

The components of $S_{j,j+1}$ and $S_{j,j-1}$ are given by

$$S_{j,j+1}^{(-1)} = \text{diag}[-2ef\cdot E_{0100}]\ ; \tag{3.7}$$

$$S_{j,j+1}^{(0)} = \text{penta}[-2ce\cdot E_{0110},\ -2ef\cdot E_{0010},\ -e^2,\ -2eg\cdot E_{0100},\ -2deE_{0110}]\ ; \tag{3.8}$$

$$S_{j,j+1}^{(1)} = \text{diag}[-2eg\cdot E_{0010}]\ ; \tag{3.9}$$

$$S_{j,j-1}^{(-1)} = \text{diag}[-2bf\cdot E_{0001}]\ ; \tag{3.10}$$

$$S_{j,j-1}^{(0)} = \text{penta}[-2bc\cdot E_{1001},\ -2bf\cdot E_{1000},\ -b^2,\ -2bg\cdot E_{0001},\ -2bdE_{1001}]\ ; \tag{3.11}$$

$$S_{j,j-1}^{(1)} = \text{diag}[-2bg\cdot E_{1000}]\ . \tag{3.12}$$

The connection between the index of the points, as given in Fig. 3.2, and the location in terms of x, y, and z coordinates follows: suppose a certain point is indexed by i ($1 \le i \le \frac{n^3}{2}$). If we denote its associated coordinate indices by $ix = \frac{x}{h}$, $iy = \frac{y}{h}$, and $iz = \frac{z}{h}$, and use the term "fix" to mark rounding to the nearest integer towards zero, then

$$ix = \text{fix}\{[(i-1) \bmod (2n)]/2\} + 1, \tag{3.13a}$$

$$iy = \begin{cases} 2\cdot[\text{fix}(\frac{i-1}{n^2})+1], & i \bmod 4 = 0 \text{ or } 1, \\ 2\cdot\text{fix}(\frac{i-1}{n^2})+1, & i \bmod 4 = 2 \text{ or } 3, \end{cases} \tag{3.13b}$$

$$iz = \begin{cases} 2\cdot[\text{fix}(\frac{(i-1) \bmod n^2}{2n})+1], & i \text{ even}, \\ 2\cdot\text{fix}(\frac{(i-1) \bmod n^2}{2n})+1, & i \text{ odd}. \end{cases} \tag{3.13c}$$

3.3. Symmetrization of the reduced matrix. For the reduced system, we again use the idea of symmetrizing first, and then examining the symmetrized matrix. A bound for spectral radius of the iteration matrix can be obtained by using (see [3], [4])

$$\rho(\hat{D}^{-1}\hat{C}) \le ||\hat{D}^{-1}||_2||\hat{C}||_2 = \frac{\rho(\hat{C})}{\lambda_{min}(\hat{D})} \ . \tag{3.14}$$

For a certain ordering, the underlying matrix is merely a symmetrically permuted version of a matrix that is associated with another ordering. Thus, we can examine the symmetrization conditions for the matrix associated with two-plane ordering *without loss of generality*. The result that follows shows that the 2D and 3D reduced systems are similar in the sense of symmetrization conditions (see [3]).

THEOREM 3.1. *The reduced matrix S can be symmetrized with a real diagonal similarity transformation if and only if the products bcde, befg, and cdfg are positive.*

Proof. Our aim is to find a real diagonal matrix Q, so that $Q^{-1}SQ$ is symmetric. Suppose $Q = \text{diag}[Q_{1,1},\ Q_{1,2},\ \ldots,\ Q_{1,\frac{n}{2}},\ Q_{2,1},\ Q_{2,2},\ \ldots,\ Q_{2,\frac{n}{2}},\ \ldots,\ Q_{\frac{n}{2},\ \frac{n}{2}}]$, where each matrix $Q_{j,l},\ 1 \le j \le \frac{n}{2},\ 1 \le l \le \frac{n}{2}$, is a diagonal $2n \times 2n$ matrix whose entries are denoted by

$$Q_{j,l} = \text{diag}[q_1^{(j,l)}, \ldots, q_{2n}^{(j,l)}] \ .$$

We start with considering the diagonal block matrices $S_{j,j}^{(0)}$. We require that $Q_{j,l}^{-1} S_{j,j}^{(0)} Q_{j,l}$ be symmetric. Notice that $S_{j,j}^{(0)}$ is a notation that corresponds to $\frac{n}{2}$ submatrices, but in order to avoid additional notation, we do not add a script that would indicate the location of these matrices in each block, as it can be understood from the double index used for the submatrices of Q. A straightforward computation for the entries of $S_{j,j}^{(0)}$, for $1 \le j \le \frac{n}{2}$ and for all the $\frac{n}{2}$ submatrices in each block $S_{j,j}$, leads to the following conditions:

$$\left(\frac{q_i^{(j,l)}}{q_{i-4}^{(j,l)}}\right)^2 = \frac{c^2}{d^2} \ , \qquad i = 5, \ldots, 2n \ , \tag{3.15}$$

$$\left(\frac{q_i^{(j,l)}}{q_{i-3}^{(j,l)}}\right)^2 = \frac{cf}{dg} \ , \qquad i = 4, 6, 8, \ldots, 2n \ , \tag{3.16}$$

$$\left(\frac{q_i^{(j,l)}}{q_{i-2}^{(j,l)}}\right)^2 = \frac{ce}{bd} \ , \qquad 1 \le i \le 2n \ , \quad i \bmod 4 = 2 \text{ or } 3 \ , \tag{3.17}$$

$$\left(\frac{q_i^{(j,l)}}{q_{i-2}^{(j,l)}}\right)^2 = \frac{bc}{de} \ , \qquad 1 \le i \le 2n \ , \quad i \bmod 4 = 0 \text{ or } 1 \ , \tag{3.18}$$

$$\left(\frac{q_i^{(j,l)}}{q_{i-1}^{(j,l)}}\right)^2 = \frac{cg}{df} \ , \qquad i = 3, 5, 7, \ldots, 2n-1 \ , \tag{3.19}$$

$$\left(\frac{q_i^{(j,l)}}{q_{i-1}^{(j,l)}}\right)^2 = \frac{ef}{bg}\ , \qquad 1 \le i \le 2n\ ,\quad i \bmod 4 = 2\ , \tag{3.20}$$

$$\left(\frac{q_i^{(j,l)}}{q_{i-1}^{(j,l)}}\right)^2 = \frac{bf}{eg}\ , \qquad 1 \le i \le 2n\ ,\quad i \bmod 4 = 0\ . \tag{3.21}$$

The restriction that the diagonal similarity transformation be real leads to the following conditions:

- Equations (3.16) and (3.19) imply $cdfg > 0$.
- Equations (3.17) and (3.18) imply $bcde > 0$.
- Equations (3.20) and (3.21) imply $befg > 0$.

We choose $q_1^{(j,l)},\ 1 \le j,l \le \frac{n}{2}$ arbitrarily, and use equations (3.19), (3.20), and (3.21) to determine $q_i^{(j,l)},\ 2 \le i \le 2n,\ 1 \le j,l \le \frac{n}{2}$. In so doing, we must make sure that equations (3.15)–(3.18) are consistent with equations (3.19)–(3.21). Indeed there is full consistency. Below we present a proof of this for entries whose index, i, satisfies $i \bmod 4 = 0$. The same procedure can be done for the other values of i. Since $(i-1) \bmod 4 = 3$, applying (3.19) for $i-1$ (rather than i) and multiplying it by (3.21) we obtain

$$\left(\frac{q_i^{(j,l)}}{q_{i-2}^{(j,l)}}\right)^2 = \frac{bf}{eg}\cdot\frac{cg}{df} = \frac{bc}{de}\ ,$$

which is exactly (3.18). Next, since $(i-2) \bmod 4 = 2$, we can use (3.20) to conclude that

$$\left(\frac{q_{i-2}^{(j,l)}}{q_{i-3}^{(j,l)}}\right)^2 = \frac{ef}{bg}\ ,$$

and combining this equation with (3.18) we obtain

$$\left(\frac{q_i^{(j,l)}}{q_{i-3}^{(j,l)}}\right)^2 = \frac{bc}{de}\cdot\frac{ef}{bg} = \frac{cf}{dg}\ ,$$

which is identical to equation (3.16). Finally, $(i-3) \bmod 4 = 1$, therefore (3.19) implies

$$\left(\frac{q_{i-3}^{(j,l)}}{q_{i-4}^{(j,l)}}\right)^2 = \frac{cg}{df}\ .$$

Multiplying this by equation (3.16) yields

$$\left(\frac{q_i^{(j,l)}}{q_{i-4}^{(j,l)}}\right)^2 = \frac{cf}{dg}\cdot\frac{cg}{df} = \frac{c^2}{d^2}\ ,$$

which is identical to (3.15). This completes the proof of consistency for indices which correspond to equation (3.21). For indices which satisfy either (3.19) or (3.20) the process is completely analogous, and we omit the algebraic details.

The same procedure repeats when we move to consider the off-diagonal matrices of the main block diagonals, namely $S_{j,j}^{(\pm 1)}$, and the off-diagonal block matrices $S_{j,j\pm 1}$. For the former we have the following equation, which determines the ratio between $q_i^{(j,l+1)}$ and $q_i^{(j,l)}$, and thus defines the values of $q_i^{(j,l+1)}$, using $q_i^{(j,l)}$:

$$\left(\frac{q_i^{(j,l+1)}}{q_i^{(j,l)}}\right)^2 = \frac{f^2}{g^2}\ , \qquad 1 \le i \le 2n\ . \tag{3.22}$$

Notice that (3.22) establishes conditions for values that were previously considered arbitrary, namely $q_1^{(j,l)},\ l > 1$. In other words, at this stage only $q_1^{(j,1)},\ 1 \le j \le \frac{n}{2}$, are left arbitrary. The new condition is followed by four additional equations which can all be obtained from combinations of the group of equations (3.15)–(3.22), and thus are consistent and do not impose any additional restrictions.

Last, for the off-diagonal block matrices $S_{j,j\pm 1}$ we have

$$\left(\frac{q_i^{(j+1,l)}}{q_i^{(j,l)}}\right)^2 = \frac{b^2}{e^2}\ , \qquad 1 \le i \le 2n\ . \tag{3.23}$$

Equation (3.23) defines the values of $q_i^{(j+1,l)},\ 1 \le j \le \frac{n}{2}-1$, given $q_i^{(j,l)}$. Imposing this condition leaves only $q_1^{(1,1)}$ arbitrary ($Q^{-1}SQ$ remains the same for any matrix S as long as the entries of Q are determined up to a multiplicative constant).

As in the other cases, there are additional conditions (four for $S_{j,j\pm 1}^{(0)}$ and two for $S_{j,j\pm 1}^{(\pm 1)}$) that are consistent with the equations presented so far and do not contribute any additional data. We omit the algebraic details. We stress, though, that for each of the equations one must make sure whether there are conditions that need to be imposed in order to have Q be real. It turns out that all the conditions are contained in the set of the three conditions that were imposed when equations (3.15)–(3.21) were discussed. □

In order to construct the matrix Q we have to use equations (3.19)–(3.23). Notice that from the proof it is clear how the symmetrized matrix $Q^{-1}SQ$ looks and there is no need to construct Q and actually perform the similarity transformation. Moreover, the symmetrizer might contain very large values, thus using it might cause numerical difficulties. Elman and Golub showed in [3] that in the one-dimensional case, for the equation $-u'' + \sigma u' = f$ with Dirichlet boundary conditions, if the first entry of the symmetrizer is 1, and the cell Reynolds number is between 0 and 1, the last entry of the symmetrizer goes to e^{σ} as n goes to infinity, and thus is very large for large values of the underlying PDE coefficient. The same phenomenon will occur in multidimensional problems. Thus we stress that we are using the symmetrized matrix only as an analytical tool; in actual numerical computation, we always use the nonsymmetric system.

The entries of the symmetrizer can be determined up to sign. For $be,\ cd,\ fg > 0$, a symmetrization operator that preserves the sign of the matrix entries is
$-b^2 \to -be$; $-c^2 \to -cd$; $-d^2 \to -cd$; $-e^2 \to -be$; $-f^2 \to -fg$; $-g^2 \to -fg$; $-2bf \to -2\sqrt{befg}$; $-2cf \to -2\sqrt{cdfg}$; $-2df \to -2\sqrt{cdfg}$; $-2ef \to -2\sqrt{befg}$; $-2bg \to -2\sqrt{befg}$; $-2cg \to -2\sqrt{cdfg}$; $-2dg \to -2\sqrt{cdfg}$; $-2eg \to -2\sqrt{befg}$; $-2bc \to -2\sqrt{bcde}$; $-2bd \to -2\sqrt{bcde}$; $-2ce \to -2\sqrt{bcde}$; $-2de \to -2\sqrt{bcde}$.

The value of the computational molecule corresponding to the center point is unchanged under the symmetrization operation. In terms of the cell Reynolds numbers, Theorem 3.1 leads to the following symmetrization result:

COROLLARY 3.2. *The reduced matrix S can be symmetrized with a real diagonal similarity transformation for any $\beta,\ \gamma,\ \delta > 0$ if one uses the upwind (backward) schemes, and for either $|\beta|,\ |\gamma|,\ |\delta| < 1$ or $|\beta|,\ |\gamma|,\ |\delta| > 1$ if one uses centered difference schemes.*

Proof. For upwind schemes $cdfg = (1 + 2\beta)(1 + 2\delta)$, which is always positive for positive values of $\beta,\ \gamma$, and δ. The same is true for $befg = (1 + 2\gamma)(1 + 2\delta)$ and $bcde = (1+2\beta)(1+2\gamma)$. For centered difference schemes $cdfg = (1-\beta^2)(1-\delta^2)$, $befg = (1-\gamma^2)(1-\delta^2)$, and $bcde = (1-\beta^2)(1-\gamma^2)$. For $cdfg$ to be positive, we require that either $|\beta| < 1$ and $|\delta| < 1$ or $|\beta| > 1$ and $\delta| > 1$. If $|\beta| < 1$ and $|\delta| < 1$, then $befg > 0$ implies $|\gamma| < 1$ and then $bcde > 0$ holds as well. If $|\beta|, |\delta| > 1$, then an analogous argument yields $|\gamma| > 1$. ∎

We can now compare the results of Corollaries 2.2 and 3.2 and see that the reduced system is symmetrizable whenever the full system is, but the opposite is not true. The significant observation here is that if all the convection coefficients of the underlying PDE are large, the reduced system can be symmetrized, whereas the full system cannot.

We end this subsection with an observation regarding the conditions for the reduced matrix to be an M-matrix. The result given in the following lemma can be viewed as a generalization of the result for the 2D case, given in [3]. This is another point of similarity between the 2D and 3D problems.

LEMMA 3.3. *If $be,\ cd,\ fg > 0$ then both the reduced matrix and the symmetrized reduced matrix are diagonally dominant M-matrices.*

Proof. The reduced matrix is a symmetric permutation of the Schur complement of the full system, which is a diagonally dominant M-matrix in the above circumstances, and thus is a diagonally dominant M-matrix (see, for example, [1, Thm. 6.10]). For the symmetrized reduced matrix we do the following: since be, cd, and fg are positive, $a^2 - 2be - 2cd - 2fg$ is the minimum of the diagonal entries of the matrix. This value is associated with all the interior grid points. By means of simple algebra it can be shown that all the diagonal entries of the matrix are positive. It suffices to look at the worst case, which occurs when the associated grid point is interior, and all 19 components of the computational molecule are active. The symmetrized reduced matrix is strictly diagonally dominant if

$$a^2 - 2be - 2fg - 2cd \geq 2be + 2cd + 2fg + 8\sqrt{befg} + 8\sqrt{cdfg} + 8\sqrt{bcde}\ , \tag{3.24}$$

which is equivalent to

$$a^2 \geq 4(\sqrt{be} + \sqrt{cd} + \sqrt{fg})^2\ . \tag{3.25}$$

Strict diagonal dominance applies because the diagonal entries which are associated with noninterior points are bigger than the left-hand side of (3.24). For centered schemes, $cd = 1 - \gamma^2$, and so $0 < cd < 1$. Analogously, $0 < be,\ fg < 1$. Since $a = 6$, it follows that (3.25) holds. For the upwind case, the condition (3.25) reads

$$\begin{aligned} 36 + 24(\beta + \gamma + \delta) + 4(\beta + \gamma + \delta)^2 \geq 4\ [3 + 2(\beta + \gamma + \delta) \\ +2\sqrt{(1 + 2\beta)(1 + 2\gamma)} + 2\sqrt{(1 + 2\beta)(1 + 2\delta)} + 2\sqrt{(1 + 2\gamma)(1 + 2\delta)}\ ]\ , \end{aligned} \tag{3.26}$$

which after simplification becomes

$$(\sqrt{1+2\beta}-\sqrt{1+2\gamma})^2+(\sqrt{1+2\beta}-\sqrt{1+2\delta})^2 \\ +(\sqrt{1+2\gamma}-\sqrt{1+2\delta})^2+(\beta+\gamma+\delta)^2\geq 0\ , \tag{3.27}$$

and holds for all $\beta,\gamma,\delta>0$.

Under the conditions stated in the lemma all the off-diagonal entries are nonpositive. Thus by [8, p. 85] the matrix is a diagonally dominant M-matrix. □

3.4. Bounds for solving the reduced system. We now estimate the spectral radius of the iteration matrix of the block Jacobi iteration, associated with partitioning of the matrix into $2n\times 2n$ blocks. Suppose $S=D-C$, where D is the block diagonal matrix whose blocks were denoted earlier by $S^{(0)}_{j,j}$. We shall restrict our interest to the case where cd, be, and fg are positive. In this case S is symmetrizable. As in the case of the full system, let $\hat{D}=Q^{-1}DQ$ and $\hat{C}=Q^{-1}CQ$. We start by examining the spectrum of $\hat{D}$.

DEFINITION 3.4. *An* interior block *is a $2n\times 2n$ block $\hat{S}^{(0)}_{j,j}$ whose associated set of grid points consists of points whose y and z coordinates are not $1/(n+1)$ or $n/(n+1)$. No restriction is imposed on x coordinates.*

Definition 3.4 is useful, as we are eventually interested in the minimal eigenvalue of $\hat{D}$; since noninterior blocks differ only on their diagonals, and these are larger algebraically than the diagonals of interior blocks if $be,\ cd,\ fg>0$, the latter have larger minimal eigenvalues, compared to interior blocks. Let us define a few auxiliary matrices and constants:

$$\begin{aligned} &r=-2\sqrt{befg};\ s=-2\sqrt{cdfg};\ R_n=\mathrm{tri}[E_{01},0,E_{10}];\\ &S_n=\mathrm{tri}[E_{10},0,E_{01}];\ T_n=\mathrm{tri}[1,0,1];\ U_n=\mathrm{penta}[1,0,0,0,1];\\ &V_n=\mathrm{septa}[E_{10},0,0,0,0,0,E_{01}];\ Z_n=\mathrm{tri}[s,r,s]. \end{aligned} \tag{3.28}$$

The subscript n stands for the order of the matrices. Notice that all the above matrices are symmetric (see the Introduction for explanation of the notation). The matrix U_n^2 has 1s on its fourth superdiagonal and subdiagonal, 2s on its main diagonal except for the first two entries and last two entries where the values are 1, and zeros elsewhere. If we define

$$W_n=(a^2-2be-2fg)\cdot I_n-2\sqrt{bcde}\cdot U_n-cd\cdot U_n^2 \tag{3.29}$$

and

$$X_n=-2\sqrt{cdfg}\cdot(R_n+V_n)-2\sqrt{befg}\cdot S_n\ , \tag{3.30}$$

then an interior block of $\hat{D}$ is given by

$$\hat{S}^{(0)}_{j,j}=W_{2n}+X_{2n}\ . \tag{3.31}$$

We now examine the eigenvalues of W_{2n} and X_{2n}.

LEMMA 3.5. *The eigenvalues of W_{2n} are given by*

$$a^2-2be-2fg-4\sqrt{bcde}\cdot\cos(\pi jh)-4cd\ \cos^2(\pi jh)\ ,\qquad j=1,\dots,n\ , \tag{3.32}$$

each with algebraic multiplicity of 2.

Proof. Since

$$U_{2n} = T_n \bigotimes I_2 \ , \tag{3.33}$$

this matrix's eigenvalues are $\{2\cos(\pi jh)\}_{j=1}^{n}$, each of algebraic multiplicity 2. W_n is a polynomial in U_n, therefore it has the eigenvalues stated in (3.32). ∎

LEMMA 3.6. *The matrix X_{2n} has the following eigenvalues:*

$$\lambda_j^{\pm} = \pm[2\sqrt{befg} + 4\sqrt{cdfg}\cdot\cos(\pi jh)], \qquad j = 1, \ldots, n \ . \tag{3.34}$$

Proof. Suppose $Y_2 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Then

$$X_{2n} = Z_n \bigotimes Y_2 \ . \tag{3.35}$$

By the theory of Kronecker products (see, for example, [2, p. 146 ff.]), it follows that the eigenvalues of X_{2n} are all the combinations of products of the eigenvalues of Z_n and Y_2, and thus are given by $[r + 2s\ \cos(\pi jh)]\cdot(\pm 1),\ 1 \le j \le n$. ∎

We can now use Lemmas 3.5 and 3.6 to establish the following theorem.

THEOREM 3.7. *The eigenvalues of interior blocks $S_{j,j}^{(0)}$ are given by*

$$\begin{aligned} \mu_j^{\pm} &= a^2 - 2be - 2fg - 4\sqrt{bcde}\cdot\cos(\pi jh) - 4cd\ \cos^2(\pi jh) \\ &\quad \pm[2\sqrt{befg} + 4\sqrt{cdfg}\cdot\cos(\pi jh)], \qquad j = 1, \ldots, n. \end{aligned} \tag{3.36}$$

Proof. Any four given matrices $A,\ B,\ C,\ D$ with the appropriate sizes satisfy [2]

$$(A \bigotimes B)(C \bigotimes D) = (AC) \bigotimes (BD) \ . \tag{3.37}$$

Applying this, using (3.33) and (3.35), we have

$$U_{2n}X_{2n} = (T_nZ_n) \bigotimes (I_2Y_2), \tag{3.38a}$$

$$X_{2n}U_{2n} = (Z_nT_n) \bigotimes (Y_2I_2). \tag{3.38b}$$

Since $I_2Y_2 = Y_2I_2 = Y_2$ and $T_nZ_n = Z_nT_n$ we conclude that

$$X_{2n}U_{2n} = U_{2n}X_{2n} \ , \tag{3.39}$$

hence X_{2n} and U_{2n} commute, which means that X_{2n} and W_{2n} have common eigenvectors (they can be easily computed using Lemma 2.3), can be simultaneously diagonalized, and the eigenvalues of $S_{j,j}^{(0)}$ are the sum of the eigenvalues of X_{2n} and W_{2n}, given in (3.36). ∎

We remark that another way of analyzing the spectrum of X_{2n} is by using the relation $(X_{2n})^2 = (Z_{2n})^2$.

THEOREM 3.8. *For $be,\ cd,\ fg > 0$ the eigenvalues of $\hat{D}$ are positive. The eigenvalues given in* (3.36) *are also eigenvalues of $\hat{D}$, each of multiplicity $(\frac{n}{2} - 2)^2$. The rest of the eigenvalues of $\hat{D}$ are all clustered in the interval* $[\min_j(\mu_j),\ \max_j(\mu_j) + be + cd + fg]$. *The minimal eigenvalue of $\hat{D}$, namely* $\min_j(\mu_j)$, *is given by*

$$\eta = a^2 - 2be - 2fg - 2\sqrt{befg} - 4(\sqrt{bcde} + \sqrt{cdfg})\cdot\cos(\pi h) - 4cd\ \cos^2(\pi h) \ . \tag{3.40}$$

Proof. The eigenvalues of $\hat{D}$ are positive by combining Lemma 3.3 [8, p. 85, Cor. 2] and [8, Thm. 3.12]. For the eigenvalues (3.36), the multiplicity is determined by counting the number of interior blocks. Since noninterior blocks differ from interior blocks only in their diagonal, and only by 0 to $be+cd+fg$ in each diagonal, we conclude that the minimal eigenvalue is obtained in interior blocks. Using equality of the ℓ_2-norm and the spectral radius of symmetric matrices, and the triangular inequality, results in the upper bound for the eigenvalues, namely $\max_j(\mu_j)+be+cd+fg$. ∎

Next, we consider the spectral radius of $\hat{C}$. Since it is equal to the spectral radius of $-\hat{C}$, we shall refer below to submatrices of the latter. We start with the part contained in $\mathrm{diag}[\hat{S}_{1,1}, \hat{S}_{2,2}, \ldots, \hat{S}_{\frac{n}{2},\frac{n}{2}}]$. For a given value of j $(1 \le j \le \frac{n}{2})$, denote the part of $-\hat{C}$ that is contained in $\hat{S}_{j,j}$ by $\tilde{S}_{j,j} = \mathrm{tri}[\hat{S}_{j,j}^{(-1)}, 0, \hat{S}_{j,j}^{(-1)\ T}]$. $\tilde{S}_{j,j}$ is an $n^2 \times n^2$ block tridiagonal matrix. All the block matrices $\hat{S}_{j,j},\ 1 \le j \le \frac{n}{2}$, are identical to each other, thus we can carry out the analysis for a single block. Each block $\hat{S}_{j,j}^{(-1)}$ is of size $2n \times 2n$ and is obtained by applying the symmetization operator to (3.6).

Define $Y_{2n}(r,s)$ to be the following $2n \times 2n$ matrix:

$$
Y_{2n}(r,s) = \begin{pmatrix}
0 & r & 0 & s & & & & & & \\
0 & 0 & 0 & 0 & 0 & & & & & \\
0 & s & 0 & r & 0 & s & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & & & \\
0 & r & 0 & s & 0 & r & 0 & s & & \\
 & & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\
 & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & & & & & \ddots & \ddots & \ddots & \ddots & \ddots \\
 & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & s & 0 & r \\
 & & & & & & & & & 0
\end{pmatrix} \tag{3.41}
$$

and let $W(r,s) = \mathrm{tri}[Y_{2n}(r,s), 0, Y_{2n}^T(r,s)]$ be an $n^2 \times n^2$ matrix, consisting of $n/2$ blocks, each of size $2n \times 2n$. Then

$$
\tilde{S}_{j,j} = \mathrm{tri}[\hat{S}_{j,j}^{(-1)}, 0, \hat{S}_{j,j}^{(-1)\ T}] = \mathrm{tri}[-fg\ I_{2n}, 0, -fg\ I_{2n}] + W(-2\sqrt{befg}, -2\sqrt{cdfg}). \tag{3.42}
$$

Since $\tilde{S}_{j,j}$ is symmetric, we can estimate its spectral radius using

$$
\begin{aligned}
\rho(\tilde{S}_{j,j}) &= ||\mathrm{tri}[-fg\ I_{2n}, 0, -fg\ I_{2n}] + W(-2\sqrt{befg}, -2\sqrt{cdfg})||_2 \\
&\le ||\mathrm{tri}[-fg\ I_{2n}, 0, -fg\ I_{2n}]||_2 + ||W(-2\sqrt{befg}, -2\sqrt{cdfg})||_2\ .
\end{aligned} \tag{3.43}
$$

The eigenvalues of the $n^2 \times n^2$ matrix $\mathrm{tri}[-fg\ I_{2n}, 0, -fg\ I_{2n}]$ are $-2fg\ \cos(\frac{\pi j}{\frac{n}{2}+1})$, $1 \le j \le \frac{n}{2}$, each with algebraic multiplicity $2n$. Since it is a symmetric matrix, we obtain

$$
||\mathrm{tri}[-fg\ I_{2n}, 0, -fg\ I_{2n}]||_2 = \rho(\mathrm{tri}[-fg\ I_{2n}, 0, -fg\ I_{2n}]) = 2fg\ \cos\left(\frac{\pi}{\frac{n}{2}+1}\right)\ . \tag{3.44}
$$

To estimate $||W||_2$ in (3.43), it is easier to work with W^2, and use $||W||_2 = ||W^2||_2^{1/2}$. In fact we have the following lemma.

LEMMA 3.9. *The matrix W^2 has the following eigenvalues:* 0*, of multiplicity* $2n$*, and*

$$4befg + 16cdfg \cdot \cos^2(\pi jh) + 16\sqrt{bcde} \cdot fg \cdot \cos(\pi jh) \quad , \qquad 1 \le j \le n \ , \tag{3.45}$$

each of multiplicity $n-2$.

Proof. Forming W^2 in terms of Y_{2n} and Y_{2n}^T, we get

$$W^2 = \text{penta}[(Y_{2n})^2, 0, Y_{2n}^T Y_{2n} + Y_{2n} Y_{2n}^T, 0, (Y_{2n}^T)^2] \ , \tag{3.46}$$

except the first and last diagonal block entries are given by $Y_{2n}^T Y_{2n}$ and $Y_{2n} Y_{2n}^T$, respectively. However, a straightforward computation shows that $(Y_{2n})^2 = (Y_{2n}^T)^2 = 0$. From that it follows that W^2 is actually block diagonal. In terms of U_n (see (3.28)) we have

$$Y_{2n} Y_{2n}^T + Y_{2n}^T Y_{2n} = r^2 \cdot I_{2n} + 2rs \cdot U_{2n} + s^2 \cdot U_{2n}^2 \ , \tag{3.47}$$

thus, the eigenvalues of this matrix, with r and s as in (3.28), are the ones given in (3.45). Each eigenvalue is of algebraic multiplicity 2.

Next, we consider the matrices $Y_{2n} Y_{2n}^T$ and $Y_{2n}^T Y_{2n}$, which appear in the first and last diagonal block entries of W^2. The matrix $Y_{2n}^T Y_{2n}$ can be permuted so that the n rows containing nonzero entries are first, followed by the zero rows. Doing so, we obtain a matrix of the type $\left(\begin{smallmatrix} \tilde{X} & 0 \\ 0 & 0 \end{smallmatrix}\right)$, where $\tilde{X}$ is a symmetric pentadiagonal $n \times n$ matrix given by $\text{penta}[s^2, 2rs, r^2 + 2s^2, 2rs, s^2]$, except the first and the last entries in the main diagonal are $r^2 + s^2$.

Again, we use an auxiliary matrix that has been introduced in (3.28), and we have

$$\tilde{X} = r^2 \cdot I_n + 2rs \cdot T_n + s^2 \cdot T_n^2 \ . \tag{3.48}$$

From this it follows that the eigenvalues of $Y_{2n}^T Y_{2n}$ and $Y_{2n} Y_{2n}^T$ are thus exactly the ones given by (3.45), each of multiplicity 1, plus the eigenvalue 0, of multiplicity n.

We can now find the eigenvalues of W^2 by assembling all the eigenvalues of all blocks. Since there are $\frac{n}{2} - 2$ blocks in W^2 that are equal to $Y_{2n}^T Y_{2n} + Y_{2n} Y_{2n}^T$, the result in the statement of the lemma follows. ∎

Having the eigenvalues of W^2 at hand, we can now use equations (3.44) and (3.45) to obtain a bound for the part of $-\hat{C}$ that is contained in $\text{diag}[\hat{S}_{1,1}, \hat{S}_{2,2}, \ldots, \hat{S}_{\frac{n}{2},\frac{n}{2}}]$.

LEMMA 3.10. *The spectral radius of* $\text{diag}[\tilde{S}_{1,1}, \tilde{S}_{2,2}, \ldots, \tilde{S}_{\frac{n}{2},\frac{n}{2}}]$ *is bounded by*

$$\xi = 2fg \ \cos\left(\frac{\pi}{\frac{n}{2}+1}\right) + \sqrt{4befg + 16 \cdot cdfg \cdot \cos^2(\pi h) + 16\sqrt{bcde} \cdot fg \cdot \cos(\pi h)}. \tag{3.49}$$

As a last step, we estimate the spectral radius of the part of $-\hat{C}$ that is contained in $\{\hat{S}_{j,j\pm1}\}$. Denote this matrix by $\tilde{C}$. Then

$$\tilde{C} = \text{tri}[\hat{S}_{j,j-1}^{(-1)}, 0, \hat{S}_{j,j-1}^{(-1)\ T}] + \text{tri}[\hat{S}_{j,j-1}^{(1)}, 0, \hat{S}_{j,j-1}^{(1)\ T}] + \text{tri}[\hat{S}_{j,j-1}^{(0)}, 0, \hat{S}_{j,j-1}^{(0)\ T}] \ . \tag{3.50}$$

Using the fact that all the terms in the right-hand side of (3.50) are symmetric matrices, we use the equality between the spectral radius and the ℓ_2-norm of each of them, and then use the triangular inequality.

LEMMA 3.11. *The spectral radius of* $\text{tri}[\hat{S}^{(-1)}_{j,j-1}, 0, \hat{S}^{(-1)\ T}_{j,j-1}] + \text{tri}[\hat{S}^{(1)}_{j,j-1}, 0, \hat{S}^{(1)\ T}_{j,j-1}]$ *is* $2\sqrt{befg}$. *Moreover, its eigenvalues are either* $0,\ +2\sqrt{befg}$, *or* $-2\sqrt{befg}$.

Proof. The square of the matrix is a diagonal matrix whose entries are either zeros or $4befg$. □

LEMMA 3.12. *The spectral radius of the* $\frac{n^3}{2} \times \frac{n^3}{2}$ *matrix* $\text{tri}[-be \cdot I_{n^2}, 0, -be \cdot I_{n^2}]$ *is* $2be \cdot \cos(\frac{\pi}{\frac{n}{2}+1})$.

Proof. Let Z denote the $\frac{n}{2} \times \frac{n}{2}$ matrix $\text{tri}[1, 0, 1]$. Then $\text{tri}[-be \cdot I_{n^2}, 0, -be \cdot I_{n^2}] = -be \cdot (Z \bigotimes I_{n^2})$. □

LEMMA 3.13. *The spectral radius of* $\text{tri}[\hat{S}^{(0)}_{j,j-1}, 0, \hat{S}^{(0)\ T}_{j,j-1}] - \text{tri}[-be \cdot I_{n^2}, 0, -be \cdot I_{n^2}]$ *is given by*

$$2\sqrt{befg} + 4\sqrt{bcde} \cdot \cos(\pi h)\ . \tag{3.51}$$

Proof. Let $\tilde{C}_1$ denote the matrix that permutes the rows of the matrix given in the statement of the lemma so that indices whose modulus 4 are 0 or 1 are indexed first, in increasing order, and indices whose modulus 4 are 2 or 3 are indexed later. Let $\tilde{C}_2$ be a permutation of $\tilde{C}_1$ such that the rows and columns indexed $n^3/4 - n^2/2 + 1$ to $n^3/4 + n^2/2$ (n^2 such rows and columns) become rows/columns $n^3/2 - n^2 + 1$ to $n^3/2$, and the rest of the rows/columns are shifted accordingly. The last n^2 rows and columns of $\tilde{C}_2$ are zeros. If $\tilde{C}_3$ denotes the upper left $(n^3/2 - n^2) \times (n^3/2 - n^2)$ submatrix of $\tilde{C}_2$, then it is a block matrix of the form $\left(\begin{smallmatrix} 0 & \tilde{U} \\ \tilde{U} & 0 \end{smallmatrix}\right)$, and each of the matrices $\tilde{U}$ is a block diagonal matrix consisting of $n \times n$ tridiagonal matrices given by $-2 \cdot tri[\sqrt{bcde}, \sqrt{befg}, \sqrt{bcde}]$. Thus by Lemma 2.3 the spectral radius of this submatrix is the one given in (3.51). It is also the spectral radius of the original matrix, as the rest of its eigenvalues are 0. □

We can now combine the results obtained in Lemmas 3.9–3.13 to obtain the following result.

THEOREM 3.14. *Let*

$$\phi = 4\sqrt{befg} + 4\sqrt{bcde} \cdot \cos\left(\frac{\pi}{n+1}\right) + 2be\cos\left(\frac{\pi}{\frac{n}{2}+1}\right) \tag{3.52}$$

and let ξ *be as in* (3.49). *Then the spectral radius of* $\hat{C}$ *is bounded by* $\xi + \phi$.

Combining Theorems 3.8 and 3.14 we establish the main result of this section.

THEOREM 3.15. *The spectral radius of the reduced iteration matrix satisfies*

$$\rho(D^{-1}C) = \rho(\hat{D}^{-1}\hat{C}) \leq \frac{\phi + \xi}{\eta}\ , \tag{3.53}$$

where η, ξ, *and* ϕ *are defined by* (3.40), (3.49), *and* (3.52) *respectively.*

4. Comparison of computational work and numerical experiments.

4.1. Asymptotic behavior of the bounds. Expanding (3.53) and (2.16) in Taylor expansions about the origin yields the following corollary.

COROLLARY 4.1. *For* h *sufficiently small, the spectral radius of the reduced system is bounded by*

$$1 - \left(\frac{10}{9}\pi^2 + \frac{1}{6}\sigma^2 + \frac{1}{6}\tau^2 + \frac{1}{6}\mu^2\right)\ h^2 + o(h^2)\ . \tag{4.1}$$

The spectral radius of the iteration matrix of the analogous full system is bounded by

$$1 - \left(\frac{3}{4}\pi^2 + \frac{1}{16}\sigma^2 + \frac{1}{16}\tau^2 + \frac{1}{16}\mu^2 \right) \; h^2 + o(h^2) \; . \tag{4.2}$$

Corollary 4.1 shows that the asymptotic bound is always smaller for the reduced system—each term of the $O(h^2)$ terms in (4.1) is smaller than its analogous term in (4.2) for all values of σ, τ, and μ. The $O(h^2)$ terms are significant since they indicate the rate of convergence. In fact, for the smallest magnitude of the cell Reynolds numbers, namely 0, we have a ratio of 40/27, and as the PDE coefficients become large (with h being fixed and sufficiently close to 0), this ratio goes to 8/3. This means that asymptotically, the reduced system is superior to the full system for any value of the PDE coefficients for which our analysis applies, and, roughly speaking, the improvement in convergence rate varies between 1.5 and about 2.66.

4.2. Comparison of computational work. We can estimate and compare the computational work involved as follows: the preconditioner D in the case of the full system is a tridiagonal matrix, whereas the semibandwidth of the preconditioner for the reduced system is equal to 4, and there are up to eight nonzero entries in each of its rows. This means that the amount of work per step is bigger for the reduced system, in comparison to a full system of the same size. Notice, however, that for a given problem, the reduced system has only half the unknowns.

Below L and U denote the matrices associated with the LU decomposition of a given preconditioner (for the full system as well as for the reduced system). For large enough n, we can take into account the leading power of the number of nonzeros. For the reduced system we have approximately $4n^3$ nonzeros in D, $\frac{9}{2}n^3$ nonzeros in $L+U$, and $\frac{11}{2}n^3$ nonzeros in C. For the full system the number of nonzero entries of both D and $L+U$ is approximately $3n^3$ and C has approximately $4n^3$ nonzeros.

After performing the LU decomposition, the linear solve done on each step is approximately as expensive as the sum of the number of nonzeros in $L+U$ and the number of nonzeros in C (see, for example, [4]). Since the preconditioners are matrices whose bandwidths do not depend on n, the LU factorization, which is done once and for all, can be considered as an additional single iteration, as far as computational work is concerned (see [6, p. 151] for an operation count of LU factorization for narrow banded matrices). We conclude that each iteration of the reduced system costs roughly 10/7 that of the full system.

However, using the asymptotic formulas, the number of iterations required for the full system is larger than that required for the reduced system by a factor of at least 40/27, thus solving the reduced system is always more efficient. In practice, in our numerical experiments we have experienced much more dramatic savings than those indicated by the analysis.

Indeed, we report here that in an extensive set of numerical experiments we have conducted for problems with small and moderate sized convection terms, the iteration counts and the run time are significantly better for the reduced system. We find solving the reduced system up to 5 times faster when discretizing using centered differences with cell Reynolds numbers smaller than 2 in magnitude, and as the convection terms become larger, there exists a region where the full system fails to converge, whereas the reduced system still does. For very large convection terms neither system converges. For upwind discretization we observe a more modest rate of improvement. In this case, for all $\beta, \gamma, \delta > 0$, both systems converge, and the gain for

the reduced system is about 1.3 to 1.5 for small cell Reynolds numbers, and becomes smaller as the convection terms grow larger.

4.3. Comparison of spectral radii. In order to demonstrate the superiority of the reduced system, we use the following strategy: first, we present numerical results which show that for a relatively small n the actual spectral radius of the reduced iteration matrix is smaller for both upwind and centered schemes. Then, we show that the bound is very tight and becomes tighter as n grows larger, thus it can be used for our analysis. To complete the argument, we examine the asymptotic bounds, which illustrate the superiority of the reduced system in the limit $h \to 0$.

In Fig. 4.1 a comparison for several cross sections of the cell Reynolds numbers in a coarse mesh is given. For all the graphs we took $n = 6$ ($6^3 = 216$). In graphs (a) and (b) we take $\beta = \gamma = \delta$. Notice that when β, γ, and δ approach 1, the spectral radius goes to 0. In this case the performance of the reduced system is not significantly better than the performance of the full system. In the case of large cell Reynolds numbers, notice that there are cases where the full iteration matrix has a spectral radius larger than 1, whereas the analogous reduced system has a spectral radius smaller than 1, thus there is no convergence for the full system. The cell Reynolds numbers in this case are in the range where the full system cannot be symmetrized by a real nonsingular diagonal similarity transformation. In graphs (c) and (d) we fix two of the three cell Reynolds numbers. It is interesting to notice that for $\beta = \gamma = \delta = 0$, which corresponds to the symmetric case, the spectral radius is larger than for convection terms that are moderate in size, which illustrates that having the nonsymmetry is an advantage. The same phenomenon has been observed by Elman and Golub in the 2D case [3].

Graph (d) is of particular interest, as it corresponds to a nonsymmetrizable case. It suggests that the reduced system is superior to the full system also in cases in which our analysis does not apply. In this case, for most of the values given in the graph the full system does not converge, whereas the reduced system converges fairly fast, with a spectral radius smaller than 0.5.

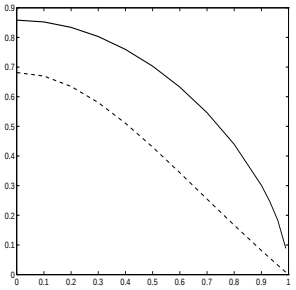
In all the graphs, the spectral radius of the reduced system is smaller, for all the tested values of β, γ, and δ. An extensive set of numerical experiments for systems with other values of n and other cross sections of the cell Reynolds numbers resulted in superiority of the reduced system in all cases.

TABLE 4.1

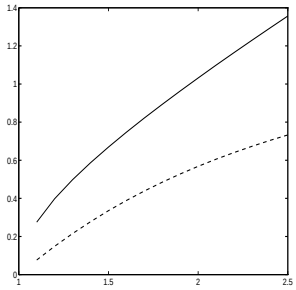
Comparison between the computed spectral radius and the bound, for upwind schemes (left) and centered differences (right), with $\beta = \gamma = \delta = 0.5$.

n	Radius	Bound	Ratio	Radius	Bound	Ratio
4	0.382	0.600	1.57	0.301	0.463	1.54
6	0.552	0.683	1.24	0.426	0.521	1.22
8	0.640	0.725	1.13	0.489	0.549	1.12
10	0.689	0.748	1.09	0.523	0.565	1.08
12	0.719	0.762	1.06	0.544	0.574	1.06
14	0.738	0.771	1.04	0.558	0.580	1.04

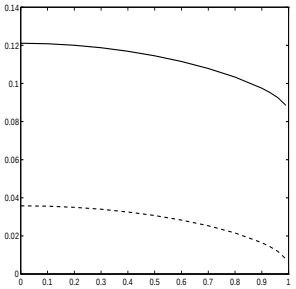
4.4. Comparison of bounds. Next, we examine the quality of the bound for the reduced system. In Table 4.1 we compare the bound with the actual spectral radius, for several values of n. In the experiment whose results follow we took $\beta = \gamma = \delta = 0.5$. We remark that qualitatively similar results are obtained for other values of the cell Reynolds numbers within the region $(0,1) \times (0,1) \times (0,1)$. Notice



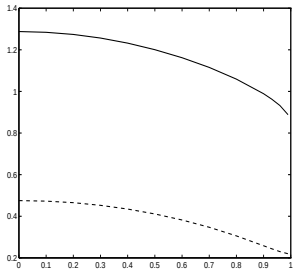
(a) centered, $0 \le \beta = \gamma = \delta \le 0.99$



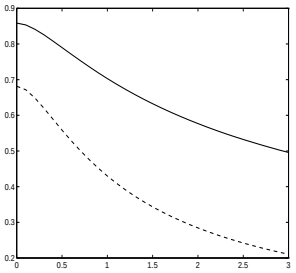
(b) centered, $1.1 \le \beta = \gamma = \delta \le 2.5$



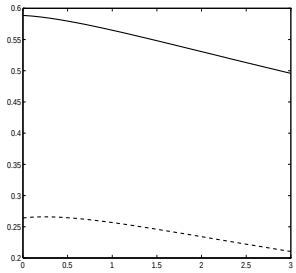
(c) centered, $0 \le \beta \le 0.99,\ \gamma = \delta = 0.99$



(d) centered, $0 \le \beta = \gamma \le 0.99,\ \delta = 3$



(e) upwind, $0 \le \beta = \gamma = \delta \le 3$



(f) upwind, $0 \le \beta \le 3,\ \gamma = \delta = 3$

Full system: solid line; reduced system: broken line

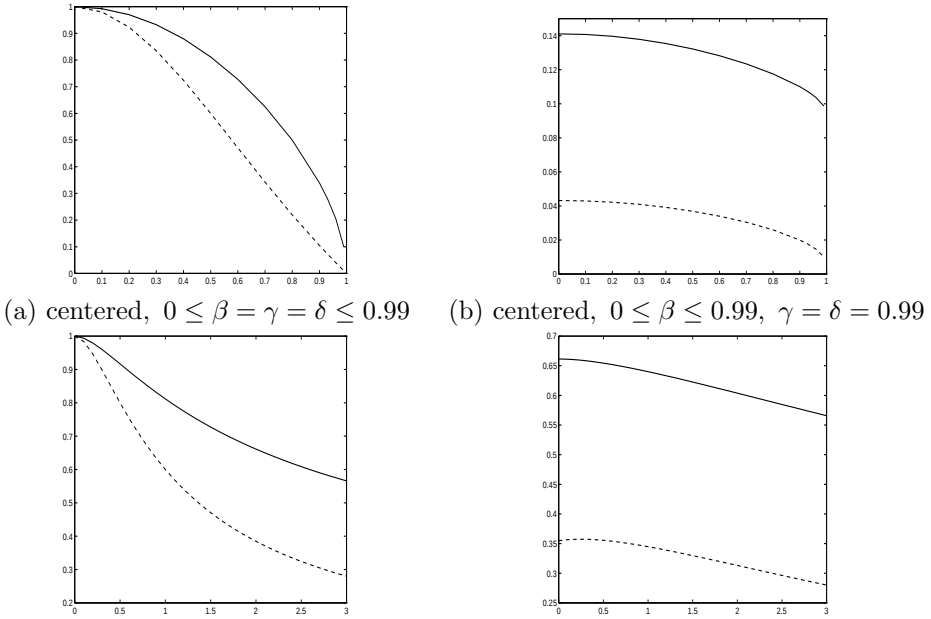
FIG. 4.1. *Spectral radii of the full system vs. the reduced system.*

that as n grows larger, the bound becomes tighter, which suggests that the bound is asymptotic to the spectral radius as h goes to 0.

We now compare the asymptotic bounds which are obtained by letting $h = 0$ in expressions (2.16) and (3.53). In Fig. 4.2 we again look at a few cross sections of the cell Reynolds numbers (cases where our analysis applies). Graphs (a) and (b) correspond to centered differences. These graphs correspond to Fig. 4.1a and 4.1c, respectively. Indeed, the similarity between the graphs is evident, even though for small cell Reynolds numbers the asymptotic bound is much closer to 1 than the actual spectral radius. This could be anticipated by looking at the Taylor expansions (4.1) and (4.2).

In graphs (c) and (d) we examine upwind differences. These graphs are analogous to Fig. 4.1e and 4.1f, respectively. The same qualitative behavior is observed and the

superiority of the reduced system is evident.



(a) centered, $0 \le \beta = \gamma = \delta \le 0.99$ (b) centered, $0 \le \beta \le 0.99$, $\gamma = \delta = 0.99$

(c) upwind, $0 \le \beta = \gamma = \delta \le 3$ (d) upwind, $0 \le \beta \le 3$, $\gamma = \delta = 3$

Full system: solid line; reduced system: broken line

FIG. 4.2. *Asymptotic bounds of the spectral radius of the full system vs. the reduced system.*

5. Concluding remarks. We have examined the effectiveness of one step of cyclic reduction for a 3D problem with constant coefficients. We have presented an ordering strategy which is unique to 3D problems and takes the structure of the reduced grid into account. This ordering strategy is very effective compared to other orderings that we have tried. We have proved that the performance of the block Jacobi scheme for the reduced system is better than the performance for the analogous full system. We have developed bounds for spectral radii and have shown that the bounds are tight for the reduced system for the region in which both systems can be symmetrized. We have discussed properties that both the 2D and the 3D problems share. We have derived symmetrization conditions for the matrix and shown that the region of PDE coefficients for which the full system can be symmetrized does not include the region of large convection terms, for which the reduced system can be symmetrized. Finally, we have conducted numerical experiments which show that the reduced system is superior.

The results that are presented in this work lead us to believe that cyclic reduction can serve as an efficient technique for solving 3D elliptic problems, which are very complicated in nature. The nature of the convergence analysis—the fact that the matrices are split into a sum of several submatrices for which the eigenvalues are known explicitly—has made it possible to obtain tight bounds. Moreover, it makes it possible to find bounds for other splittings, without much additional effort. This, and other aspects of the iterative solution of the 3D problem using one step of cyclic

reduction, including the variable coefficient case and preconditioners for the reduced system, are currently under investigation.

Acknowledgments. We would like to thank Gene Golub for introducing us to this topic, pointing out several references, and holding some helpful discussions with us about different aspects of the problem.

REFERENCES

[1] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, UK, 1994.
[2] S. Barnett, *Matrices - Methods and Applications*, Clarendon Press, Oxford, 1990.
[3] H. C. Elman and G. H. Golub, *Iterative methods for cyclically reduced non-self-adjoint linear systems*, Math. Comp., 54 (1990), pp. 671–700.
[4] H. C. Elman and G. H. Golub, *Iterative methods for cyclically reduced non-self-adjoint linear systems* II, Math. Comp., 56 (1991), pp. 215–242.
[5] H. C. Elman and G. H. Golub, *Line iterative methods for cyclically reduced discrete convection-diffusion problems*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 339–363.
[6] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed., The Johns Hopkins University Press, Baltimore, MD, 1989.
[7] L. A. Hageman and D. M. Young, *Applied Iterative Methods*, Academic Press, New York, 1981.
[8] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.
[9] D. M. Young, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.