A box-shaped cyclically reduced operator

Chen Greif[*,†] and Robert L. Hocking

Department of Computer Science, The University of British Columbia, Vancouver, BC, Canada V6T 1Z4

SUMMARY

We propose a new procedure of partial cyclic reduction, where we apply a 2^d-color ordering (with $d=2$, 3 the dimension of the problem), and use different operators for different gridpoints according to their color. These operators are chosen so that the gridpoints can be readily decoupled, and we then eliminate all colors but one. This yields a smaller cartesian mesh and box-shaped 9-point (in 2D) or 27-point (in 3D) operators that are easy to analyze and implement. Multi-line and multi-plane orderings are considered, and we perform convergence analysis and numerical experiments that demonstrate the merits of our approach. Copyright © 2010 John Wiley & Sons, Ltd.

1. INTRODUCTION

The technique of cyclic reduction has been studied and analyzed in several settings in the last few decades. Early work demonstrates the advantages of applying this procedure to the discrete Poisson equation associated with finite differences on a uniform mesh; see [1–3] for early papers, as well as [4] and the references therein for a review of history and applications. The nonzero pattern of the matrix in this case allows for efficiently eliminating half of the unknowns, while preserving the block structure. The procedure can be applied repeatedly until a small system that can be easily solved is obtained. Recovering the solution for the eliminated unknowns is straightforward, and the overall computational cost is attractively low.

The cyclic reduction method has also been used in the context of multigrid (MG) methods [5]; see for example the early paper [6]. Reduction-based approaches continue to be of interest to designers of MG methods; see for example the recent paper [7], whose focus is on algebraic MG (AMG).

In the non-symmetric case, for example the discrete convection–diffusion equation, on which we focus in this work, some of the attractive features of the discretized Laplacian are lost and a recursive cyclic reduction may be numerically unstable. But carrying out a small number of steps is still a viable strategy.

Consider, then, the two- or three-dimensional convection–diffusion model problem:

$$-\Delta u + \vec{w} \cdot \nabla u = f. \tag{1}$$

---

[*]Correspondence to: Chen Greif, Department of Computer Science, The University of British Columbia, Vancouver, BC, Canada V6T 1Z4.
[†]E-mail: greif@cs.ubc.ca

The domain $\Omega=(0,1)^d$ is the unit square ($d=2$) or unit cube ($d=3$), subject to the Dirichlet-type boundary conditions. For our analysis we will consider a constant row vector $\vec{w}$, and denote its components by (σ,τ) for 2D and (σ,τ,μ) for 3D. In our experiments we also include the variable coefficients case.

Consider a uniform grid of size h, and let us denote the mesh Reynolds numbers by

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

Then, the +-shaped second-order 5-point discretization of (1) in 2D, after scaling by h^2, is given by

$$F^+ u_{i,j}=au_{i,j}+bu_{i,j-1}+cu_{i-1,j}+du_{i+1,j}+eu_{i,j+1}, \tag{2}$$

where

$$a=4, \quad b=-1-\delta, \quad c=-1-\gamma, \quad d=-1+\gamma, \quad e=-1+\delta.$$

In the early 1990s Elman and Golub offered a thorough analysis of the spectral properties and the convergence behavior of linear systems arising from a procedure of one step of cyclic reduction [8–10]. Using red-black ordering and eliminating all the unknowns corresponding to one of the two colors yields a linear system associated with a diamond-shaped 9-point computational molecule. After scaling by ah^2, this reduced operator is given by

$$R_2 u_{i,j}=(a^2-2be-2cd)u_{i,j}-e^2 u_{i,j+2}-2de\,u_{i+1,j+1}-c^2 u_{i-2,j}$$

$$-d^2 u_{i+2,j}-2bc\,u_{i-1,j-1}-b^2 u_{i,j-2}-2ce\,u_{i-1,j+1}-2bd\,u_{i+1,j-1}. \tag{3}$$

The Schur complement (reduced) matrix obtained after the elimination is only half the dimension of the original (unreduced) matrix. Thus, even though the 5-point operator is replaced by a 9-point operator, the cost of performing matrix–vector products on the reduced system is similar to, or in fact marginally lower than the cost of matrix–vector products on the unreduced system. As a result, when comparing the iterative solution procedure for the original system to that for the cyclically reduced system, the overall performance of solvers depends almost exclusively on the spectral structure of the reduced versus the unreduced operators, and not on the cost of a single iteration. The analysis and numerical experiments in [8–10] show that iterative solvers for the reduced system converge faster, and hence it pays off to perform one step of cyclic reduction in the 2D case.

In the late 1990s, Greif and Varah [11–13] showed that gains can be made for the three-dimensional case as well. However, in 3D the original 7-point operator is replaced by a 19-point operator applied to half of the unknowns, and hence matrix–vector products for the reduced system are more expensive, in contrast to the 2D case. Despite that the improvement in the spectral structure and convergence rates lead to the conclusion that in 3D it still pays off to perform a step of cyclic reduction. After scaling by ah^2 and setting $v=-1-\eta$ and $w=-1+\eta$, the cyclically reduced operator in 3D is given by

$$R_3 u_{i,j,k}=(a^2-2be-2cd-2vw)u_{i,j,k}-v^2 u_{i,j,k-2}-2ev\,u_{i,j+1,k-1}$$

$$-2cv\,u_{i-1,j,k-1}-2dv\,u_{i+1,j,k-1}-2bv\,u_{i,j-1,k-1}-e^2 u_{i,j+2,k}$$

$$-2de\,u_{i+1,j+1,k}-c^2 u_{i-2,j,k}-d^2 u_{i+2,j,k}-2bc\,u_{i-1,j-1,k}$$

$$-b^2 u_{i,j-2,k}-2ew\,u_{i,j+1,k+1}-2cw\,u_{i-1,j,k+1}-2ce\,u_{i-1,j+1,k}$$

$$-2bd\,u_{i+1,j-1,k}-2dw\,u_{i+1,j,k+1}-2bw\,u_{i,j-1,k+1}-w^2 u_{i,j,k+2}. \tag{4}$$

Despite their attractive numerical properties and the computational savings, cyclically reduced operators of the form just described have not been widely used. One possible reason for this is the nature of the computational molecule and the mesh, which presents a computational challenge, for example in handling boundary conditions. Applying one step of cyclic reduction to a cartesian mesh based on using a 5-point operator (in 2D) or 7-point operator (in 3D), yields a non-cartesian mesh with 'holes' corresponding to the locations of the gridpoints of the eliminated color.

Another possible reason for cyclic reduction being less popular than expected is that it is effective in settings where alternative techniques are also extremely effective. For example, for the symmetric positive-definite case represented by the discrete Poisson equation, multigrid methods are optimal in terms of computational work, and preconditioning approaches such as incomplete Cholesky (IC) or modified incomplete Cholesky (MIC) are well understood and easy to implement. Many of the computational advantages of multigrid and incomplete factorizations carry over to the non-symmetric case, and these techniques can be effectively combined with Krylov subspace solvers. An important factor in the success of a method is the ease with which it can be implemented, and this is what motivates us to derive a new variant of cyclic reduction.

The basic idea of our approach is quite simple; let us illustrate it briefly for the 2D problem. We apply a four-color ordering and use different stencils for different colors. Gridpoints that belong to two of the colors are discretized using the +-shaped 5-point operator F^+ defined in (2), and gridpoints that belong to the other two colors are discretized using a different operator (still with 5 points) whose stencil is ×-shaped. We then proceed by eliminating all but one of the colors. The gridpoints that are associated with the remaining color form a subgrid which is also cartesian. Note that this is different than the step of cyclic reduction that leads to (3), which generates a *non-cartesian* subgrid.

In 3D the procedure is more complicated but the idea is similar; further details are provided in the following section. In any case, the result is a grid that contains only $1/2^d$ of the original gridpoints (with $d=2, 3$ the dimension), but this grid is again cartesian and the stencil of the operator associated with it has the shape of a box: 9 points in 2D, 27 points in 3D. Boundary conditions can now be seamlessly incorporated, and implementation is straightforward. One obvious benefit from the fact that the grid contains 1/4 (2D) and 1/8 (3D) of the unknowns of the original grid is that matrix–vector products are cheaper than for the original grid. Recovering the unknowns that were eliminated along the way, if necessary, amounts to trivially solving a sequence of diagonal systems.

We use a block ordering strategy that exploits the structure of the stencil. Specifically, we consider k-line and k-plane orderings in 2D and 3D, respectively, where k is a small integer, e.g. 2 or 3. These orderings are based on lexicographic ordering of blocks of multiple lines, which are themselves lexicographically ordered; see also [14] and the references therein. One advantage of using such orderings is that the main diagonal blocks of the permuted matrix are fairly dense. This, in turn, may accelerate the convergence of block iterative methods; in the block stationary case for M-matrices this is supported by the theory of regular splittings [15]. A detailed description of block orderings is provided in Section 3.

We derive exact analytic expressions for the spectral radius of the block Jacobi iteration matrix with lexicographic ordering (which is equivalent to 1-line or 1-plane ordering), and tight bounds for a 2-line (2-plane) ordering in 2D (3D). We show that the latter orderings are superior to lexicographic ordering. Our experiments suggest that a 3-line ordering is even more effective, although we are not able to obtain tight bounds in this case. Since the matrices in question are consistently ordered with respect to the block partitioning that we consider, the analytic bounds carry over to block Gauss–Seidel and block SOR. In our numerical results we also consider ILU preconditioned GMRES iterations.

The remainder of this paper is organized as follows. In Section 2 we introduce the model convection–diffusion problem and derive our new operator. In Section 3 we discuss the notion of block grids. In Section 4 we apply multi-line and multi-plane orderings for the 2D and 3D cases, respectively, and offer a convergence analysis for block Jacobi. In Section 5 we provide numerical comparisons between our reduced operators, the reduced operators (3) and (4) and the unreduced 5-point and 7-point operators, using stationary as well as non-stationary solvers. Finally, in Section 6 we draw some conclusions.

2. THE PROPOSED PROCEDURE OF CYCLIC REDUCTION

We discretize the problem using a uniform rectangular grid G in Ω. Our method requires the number of gridpoints to be odd; we thus set $G=\{h, 2h, \ldots, (2n+1)h\}^d$ where $h=1/(2n+2)$.
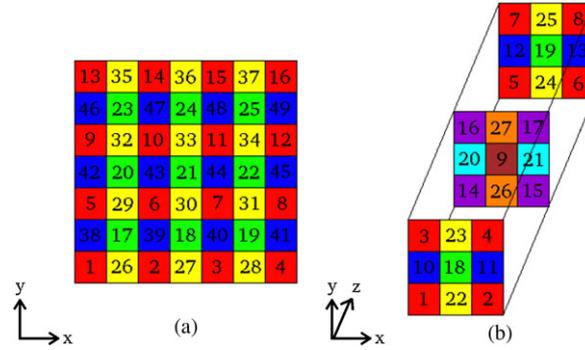
Figure 1. (a) Four-color ordering applied to a 7×7 grid and (b) Eight-color ordering applied to a $3 \times 3 \times 3$ grid. Refer online version for interpretation of color.

2.1. The reduction step in 2D

We will be using a combination of two different discrete operators, as follows. Denote by $F^{\times}$ an $\times$-shaped second-order 5-point discretization of (1)

$$F^{\times}u_{i,j} = au_{i,j} + bu_{i+1,j+1} + cu_{i-1,j+1} + du_{i-1,j-1} + eu_{i+1,j-1}$$

after scaling by $2h^2$. We have in this case

$$a = 4, \quad b = -1 + \gamma + \delta, \quad c = -1 - \gamma + \delta, \quad d = -1 - \gamma - \delta, \quad e = -1 + \gamma - \delta. \tag{5}$$

We note that the values of a through e are different here than those in the Introduction.

We now assign one of four colors to each point in G. Points (i, j) with both i and j odd are called red; green is for i, j even; blue is for i odd and j even; and yellow signifies i even and j odd. We apply a four color ordering to G by ordering the points within each color lexicographically—starting with red, then green, yellow, and blue. This is illustrated in Figure 1(a).

We then discretize (1) by applying F^+ at the yellow and blue points, and $F^{\times}$ at the red and green points.

The resulting $(2n+1)^2 \times (2n+1)^2$ linear system can be written in a block form as

$$\begin{pmatrix} D_1 & B_1 & 0 & 0 \\ B_2 & D_2 & 0 & 0 \\ B_3 & B_4 & D_3 & 0 \\ B_5 & B_6 & 0 & D_4 \end{pmatrix} \begin{pmatrix} u^{\text{red}} \\ u^{\text{green}} \\ u^{\text{blue}} \\ u^{\text{yellow}} \end{pmatrix} = \begin{pmatrix} 2h^2 f^{\text{red}} \\ 2h^2 f^{\text{green}} \\ h^2 f^{\text{blue}} \\ h^2 f^{\text{yellow}} \end{pmatrix},$$

where the matrices $D_1, D_2, D_3,$ and D_4 are diagonal. We apply a block elimination procedure to obtain a reduced (Schur complement) system of size $n^2 \times n^2$, involving only the green points:

$$(D_2 - B_2 D_1^{-1} B_1)u^{\text{green}} = 2h^2(f^{\text{green}} - B_2 D_1^{-1} f^{\text{red}}). \tag{6}$$

The remaining three quarters of the unknowns can be recovered by solving the following three linear systems, whose solution is trivial because D_i, $i = 1, 3, 4$, are diagonal:

$$D_1 u^{\text{red}} = 2h^2 f^{\text{red}} - B_1 u^{\text{green}},$$

$$D_3 u^{\text{blue}} = h^2 f^{\text{blue}} - B_3 u^{\text{red}} - B_4 u^{\text{green}},$$

$$D_4 u^{\text{yellow}} = h^2 f^{\text{yellow}} - B_5 u^{\text{red}} - B_6 u^{\text{green}}.$$

The $n^2 \times n^2$ reduced matrix in (6) is block tridiagonal with tridiagonal blocks:

$$D_2 - B_2 D_1^{-1} B_1 = \text{tri}_{n \times n}[\text{tri}_{n \times n}[-d^2, -2de, -e^2],$$

$$\text{tri}_{n \times n}[-2cd, a^2 - 2bd - 2ce, -2be],$$

$$\text{tri}_{n \times n}[-c^2, -2bc, -b^2]]. \tag{7}$$

The operator for the green points is a 9-point box-shaped operator, as opposed to the diamond-shaped 9-point operator R_2 defined in (3). We observe one of its properties as a discrete differential operator, as follows.

Proposition 2.1
The equation for the discrete differential operator that corresponds to (6) can be posed as a finite difference discretization of the differential equation

$$-\left[\left(1 + \frac{\sigma^2 h^2}{4}\right) u_{xx} + \left(1 + \frac{\tau^2 h^2}{4}\right) u_{yy}\right] + \sigma u_x + \tau u_y = f + O(h^2).$$

Proof
After expanding each $u_{i+\Delta i, j+\Delta j}$ term of (6) using a Taylor series about $u_{i,j}$, the left-hand side of that equation reduces, after dividing by $16h^2$, to

$$\sigma u_x + \tau u_y - \left(1 + \frac{\sigma^2 h^2}{4}\right) u_{xx} - \left(1 + \frac{\tau^2 h^2}{4}\right) u_{yy} - \frac{\sigma \tau h^2}{2} u_{xy} + \frac{2\sigma h^2}{3} u_{xxx} + \frac{2\tau h^2}{3} u_{yyy} + \tau h^2 u_{xxy}$$

$$+ \sigma h^2 u_{xyy} - \frac{h^2}{3} u_{xxxx} - \frac{h^2}{3} u_{yyyy} - h^2 u_{xxyy} + o(h^2).$$

Similarly, the right-hand side of (6) reduces to

$$f - \frac{\sigma h^2}{4} f_x - \frac{\tau h^2}{4} f_y + \frac{h^2}{4} \Delta f + o(h^2).$$

This computation was done using MATLAB's symbolic toolbox. □

Proposition 2.1 shows that the new cyclically reduced operator is in fact a second-order operator for the 2D convection–diffusion equation. A few of the additional $O(h^2)$ terms on the left-hand side represent artificial viscosity, and this suggests that the operator may be better behaved for higher mesh Reynolds numbers, compared with the original operator. In this regard, this operator satisfies properties similar to the ones for the cyclically reduced operators R_2 and R_3 in (3) and (4), respectively.

We now make the point that in fact the new operator is nothing but a rotated version of operator R_2 for a related equation; it is interesting to note that this does not carry over to the three-dimensional case.

Proposition 2.2
Suppose the new cyclic reduction process is applied to the 2D convection–diffusion problem, yielding a box-shaped 9-point operator. Then the operator is identical to the operator obtained by applying one step of cyclic reduction on a grid rotated 45° clockwise with mesh spacing $\sqrt{2}h$.

Proof
The proof of the proposition follows from moving to the rotated grid, where the convection–diffusion equation becomes

$$-\Delta u + \frac{1}{\sqrt{2}}(\sigma - \tau, \sigma + \tau) \cdot \nabla u = f$$

discretizing with mesh spacing $\sqrt{2}h$ using the 5-point operator, and then applying one step of cyclic reduction; see Figure 2 for an illustration, where the original and transformed coordinates are denoted, respectively, by (x, y) and (x', y'). □
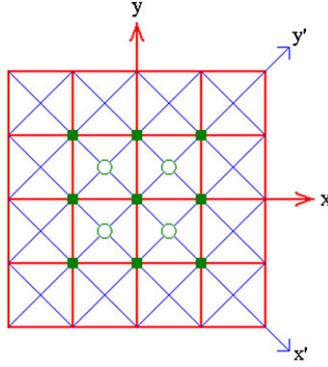
Figure 2. The 2D cyclically reduced operator defined on a mesh with spacing $\sqrt{2}h$ becomes a compact rectangular operator when viewed on a grid rotated $45°$ with mesh spacing h.

2.2. The reduction step in 3D

In three dimensions, our elimination procedure requires the use of the second-order 7-point operator $\hat{F}$, as well as three other second-order operators $\tilde{F}$, $\grave{F}$, and F. After scaling by h^2, $2h^2$, $2h^2$ and $4h^2$, respectively, these operators are

$$\hat{F}u_{i,j,k} = 6u_{i,j,k} + (-1+\gamma)u_{i+1,j,k} + (-1-\gamma)u_{i-1,j,k} + (-1+\delta)u_{i,j+1,k}$$
$$+ (-1-\delta)u_{i,j-1,k} + (-1+\eta)u_{i,j,k+1} + (-1-\eta)u_{i,j,k-1},$$

$$\tilde{F}u_{i,j,k} = 8u_{i,j,k} + (-1+\gamma+\delta)u_{i+1,j+1,k} + (-1-\gamma+\delta)u_{i-1,j+1,k}$$
$$+ (-1-\gamma-\delta)u_{i-1,j-1,k} + (-1+\gamma-\delta)u_{i+1,j-1,k}$$
$$+ 2(-1+\eta)u_{i,j,k+1} + 2(-1-\eta)u_{i,j,k-1},$$

$$\grave{F}u_{i,j,k} = 8u_{i,j,k} + (-1-\gamma+\eta)u_{i-1,j,k+1} + (-1+\gamma+\eta)u_{i+1,j,k+1}$$
$$+ (-1-\gamma-\eta)u_{i-1,j,k-1} + (-1+\gamma-\eta)u_{i+1,j,k-1}$$
$$+ 2(-1+\delta)u_{i,j+1,k} + 2(-1-\delta)u_{i,j-1,k}$$

and

$$Fu_{i,j,k} = au_{i,j,k} + bu_{i+1,j+1,k+1} + cu_{i-1,j+1,k+1} + du_{i-1,j-1,k+1} + eu_{i+1,j-1,k+1}$$
$$+ pu_{i+1,j+1,k-1} + qu_{i-1,j+1,k-1} + ru_{i-1,j-1,k-1} + su_{i+1,j-1,k-1}.$$

Here $a=8$, $b=-1+\gamma+\delta+\eta$, $c=-1-\gamma+\delta+\eta$, $d=-1-\gamma-\delta+\eta$, $e=-1+\gamma-\delta+\eta$, $p=-1+\gamma+\delta-\eta$, $q=-1-\gamma+\delta-\eta$, $r=-1-\gamma-\delta-\eta$, $s=-1+\gamma-\delta-\eta$. See Table I for a classification of gridpoint colors.

We order the mesh G with an eight-color ordering as in Figure 1(b). Equation (1) is then discretized by applying F to red and brown points, $\hat{F}$ to cyan and yellow points, $\grave{F}$ to orange and blue points, and $\tilde{F}$ to green and purple points.

For each color $c \in C \equiv \{$red, brown, orange, blue, purple, green, cyan, yellow$\}$, we obtain a system

$$D_c u^c + \sum_{x \in C, x \neq c} B_{c,x} u^x = \alpha h^2 f^c,$$

where the matrix D_c is diagonal and $\alpha = 1$, 2, or 4, depending on which operator was applied to gridpoints of color c. As seen in Table I, unknowns of a given color only depend on unknowns of a few other colors, so most of the $B_{c,x}$ values are zero.

Table I. Classification of grid colors in terms of i, j, and k, dependence of gridpoints of a given color on gridpoints of other colors, and the associated discrete operators.

Color	i	j	k	Depends on	Operator
Red	Odd	Odd	Odd	Red, brown	F
Brown	Even	Even	Even	Brown, red	F
Green	Even	Even	Odd	Green, red, brown	$\tilde{F}$
Purple	Odd	Odd	Even	Purple, red, brown	$\tilde{F}$
Blue	Odd	Even	Odd	Blue, red, brown	$\grave{F}$
Orange	Even	Odd	Even	Orange, red, brown	$\grave{F}$
Yellow	Even	Odd	Odd	Yellow, red, green, orange	$\hat{F}$
Cyan	Odd	Even	Even	Cyan, blue, purple, brown	$\hat{F}$

In fact, as soon as u is known for all the brown points it can be found for every other color by inverting a diagonal matrix and performing a few matrix vector products—simply apply

$$u^c = D_c^{-1} \left(\alpha h^2 f^c - \sum_{x \in C, x \neq c} B_{c,x} u^x \right)$$

to the red, green, purple, blue, orange, yellow, and cyan points in that order.

A block elimination procedure completely analogous to the 2D case gives an $n^3 \times n^3$ reduced system involving only the brown unknowns. The reduced matrix

$$A = \text{tri}_{n \times n}[A_1, A_2, A_3] \tag{8}$$

is block triagonal, with each block itself block tridiagonal with tridiagonal blocks. We have

$$A_1 = \text{tri}_{n \times n}[\text{tri}_{n \times n}[-r^2, -2rs, -s^2],$$
$$\text{tri}_{n \times n}[-2qr, -2(qs + pr), -2ps],$$
$$\text{tri}_{n \times n}[-q^2, -2qp, -p^2]];$$
$$A_2 = \text{tri}_{n \times n}[\text{tri}_{n \times n}[-2dr, -2(re + ds), -2es],$$
$$\text{tri}_{n \times n}[-2(cr + qd), a^2 - 2br - 2cs - 2dp - 2eq, -2(pe + bs)],$$
$$\text{tri}_{n \times n}[-2cq, -2(cp + bq), -2bp]];$$
$$A_3 = \text{tri}_{n \times n}[\text{tri}_{n \times n}[-d^2, -2de, -e^2],$$
$$\text{tri}_{n \times n}[-2cd, -2(bd + ce), -2be],$$
$$\text{tri}_{n \times n}[-c^2, -2bc, -b^2]].$$

It is tedious but straightforward to show that as long as $\|(h/2)\vec{w}\|_1 < 1$, the reduced matrices (7) and (8) are strictly diagonally dominant irreducible M-matrices.

A multivariate Taylor expansion similar to the one in 2D shows that the 3D operator has similar properties and can be interpreted as a discretization of the convection–diffusion equation with some artificial viscosity. More precisely, the operator that corresponds to the reduced system for the brown points can be posed as a finite difference discretization of the differential equation

$$-\left[\left(1 + \frac{\sigma^2 h^2}{4}\right) u_{xx} + \left(1 + \frac{\tau^2 h^2}{4}\right) u_{yy} + \left(1 + \frac{\mu^2 h^2}{4}\right) u_{zz} \right] + \sigma u_x + \tau u_y + \mu u_z$$
$$= f + O(h^2).$$

The proof is almost identical to the 2D case, Proposition 2.1, and is omitted.

On the other hand, as opposed to 2D, in 3D there is no clear connection between this operator and the cyclically reduced 3D operator defined in (4). For one, these operators have a different number of components in their stencils. Furthermore, there is no easy way to interpret this operator on a rotated grid.

In Section 4 we will show that when $\vec{w}$ is aligned with the x-axis, both systems are symmetrizable by a diagonal symmetrizer, and are positive definite.

3. BLOCK GRID ORDERINGS

In some cases it may be beneficial to use grid orderings that generate dense diagonal blocks; as stated in the Introduction this is justified by regular splittings [15]. Here multi-line and multi-plane block orderings are effective and easy to analyze; see [14] for an analysis of 2-line and 2-plane orderings applied to the centered differences discretizations of (1).

We explore the effect of k-line and k-plane orderings on the convergence of iterative solvers applied to (7) and (8). For convenience, let us define k-line and k-plane orderings within the context of general block grid orderings.

An ordering on a grid G is a function $i_G : G \rightarrow \{1, 2, \ldots, |G|\}$. A block ordering on a grid G is defined by a set of grid blocks $\{B\}$ partitioning G, an ordering of the grid blocks, and an ordering of the grid elements within each block. The block ordering is given by

$$i_G(p) = i_{B_I}(p) + \sum_{j=1}^{I-1} |B_j|,$$

where B_I is the grid block p belongs to, and $i_{B_I}(p)$ is the order of p with respect to B_I.

Unless stated otherwise, all block orderings considered in this paper will use the lexicographic ordering for both the grid blocks, and the elements within each block. Note, however, that the resulting ordering is not lexicographic; see Figures 3 and 4.

Suppose we work on an $n \times n$ grid G_{2D}, with n an integer multiple of k. The k-line ordering is defined by partitioning G_{2D} into n/k vertical blocks containing k lines of n gridpoints each, as in Figure 3(a) for $k=3$, $n=9$.

Suppose now we work on an $n \times n \times n$ grid G_{3D} with n again an integer multiple of k. The k-plane ordering is defined by partitioning G_{3D} into n^2/k^2 blocks containing nk^2 gridpoints each, as in Figure 4(a) for $k=3$, $n=6$.

Let A denote the matrix obtained by discretizing (1) on a grid G with a block ordering. A has a natural block structure that it inherits from the block ordering of G. For each ordered pair of grid blocks (B_i, B_j), there is a matrix block A_{ij} containing all the dependencies of members of B_i on members of B_j.
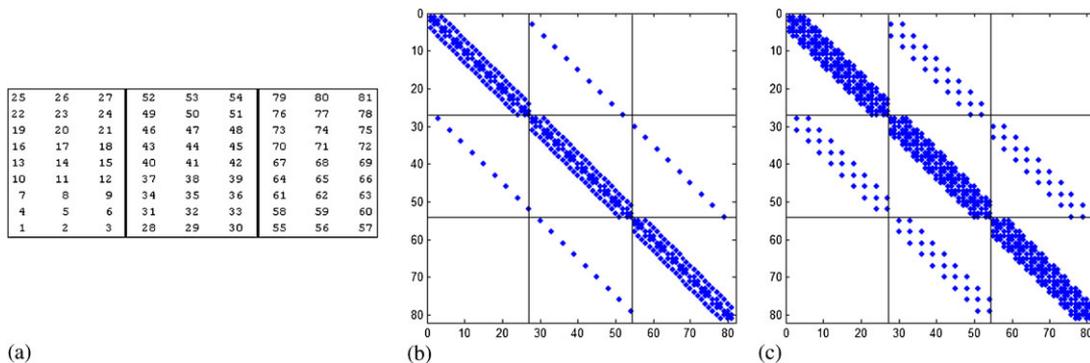
Figure 3. (a) 3-line ordering applied to a 9×9 grid. The grid blocks are separated by lines; (b) The induced block structure of the 5-point operator matrix; and (c) The induced block structure of the 9-point operator matrix.
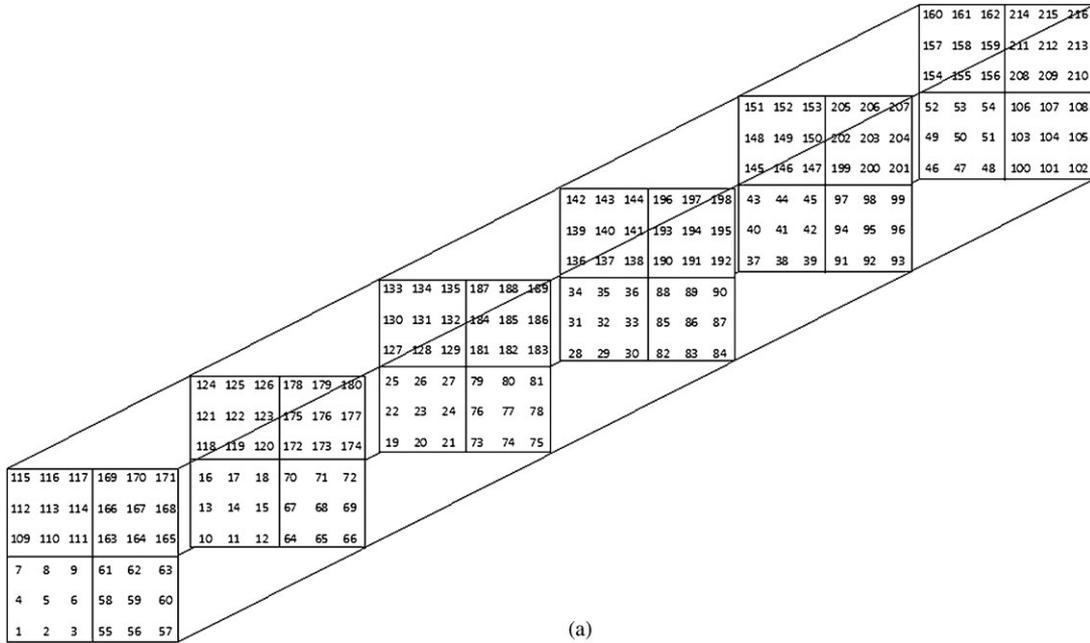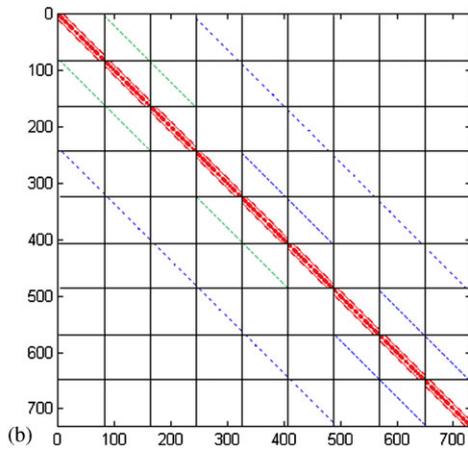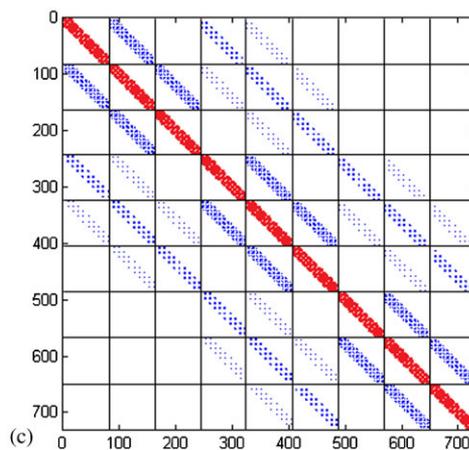
Figure 4. (a) 3-plane ordering applied to a $6 \times 6 \times 6$ grid; (b) The block matrix A_{3L} of the 5-point operator on a $9 \times 9 \times 9$ grid; and (c) The block matrix A_{3L} of the 9-point operator on a $9 \times 9 \times 9$ grid.

We denote by A_{kL} and A_{kP} the matrices that correspond to k-line and k-plane orderings. For k-line ordering, there are n/k grid blocks of nk gridpoints each. A_{kL} can therefore be referred to as a block matrix with $(n/k) \times (n/k)$ blocks, each of size $nk \times nk$. For the k-plane ordering, there are n^2/k^2 grid blocks containing nk^2 points each. Therefore, A_{kP} is an $(n^2/k^2) \times (n^2/k^2)$ block matrix with blocks of size $nk^2 \times nk^2$.

The sparsity pattern of the matrix depends, of course, also on the discretization used. In Figures 3(b) and (c) we show the matrices corresponding to 3-line ordering, namely A_{3L}, for 5-point and 9-point discretizations. In Figures 4(b) and (c) we show the matrices for 3-plane ordering, A_{3P}, using 7-point and 27-point stencils, respectively.

4. ORDERING STRATEGIES AND BOUNDS ON CONVERGENCE RATES

Our analysis makes extensive use of Kronecker products and their properties, and of known results on spectra of tridiagonal Toeplitz matrices. The following results are used so often in our analysis that they are worth stating explicitly.

Proposition 4.1

Suppose a and c are two scalars of the same sign (or zero). Then, the tridiagonal Toeplitz matrix $T = \text{tri}_{n \times n}[a, b, c]$ is symmetrizable by a real diagonal similarity transformation; that is, there exists a real diagonal matrix S_T such that $S_T^{-1} T S_T$ is symmetric. Furthermore, the eigenvalues of T are given by

$$\lambda_i = b + 2\sqrt{ac}\cos\left(\frac{i\pi}{n+1}\right), \quad i = 1, 2, \ldots, n.$$

Proof

See [8, Lemmas 1 and 2]. □

Proposition 4.2

Suppose A and B are square matrices of sizes $n \times n$ and $m \times m$, respectively, with respective eigenpairs $\{\lambda_i, \vec{x}_i\}$, $i = 1, 2, \ldots, n$ and $\{\mu_j, \vec{y}_j\}$, $j = 1, 2, \ldots, m$. Then $A \otimes B$ is $nm \times nm$ with eigenpairs $\{\lambda_i \mu_j, \ \vec{x} \otimes \vec{y}\}$, $i = 1, 2, \ldots, n, \ j = 1, 2, \ldots, m$.

Proof

See [16, Theorem 4.2.12]. □

Proposition 4.3

If A and B are square matrices symmetrizable by S_A and S_B, that is $S_A A S_A^{-1}$ and $S_B B S_B^{-1}$ are symmetric, then $A \otimes B$ is symmetrizable by $S_A \otimes S_B$.

Proof

If B and D are matrices such that the products AB and CD are defined, then $(A \otimes C)(B \otimes D) = AB \otimes CD$; see, e.g. [16, Lemma 4.2.10]. Thus, by this property and other basic properties of Kronecker products [16, Section 4.2], we have

$$(S_A \otimes S_B)^{-1}(A \otimes B)(S_A \otimes S_B) = (S_A^{-1} \otimes S_B^{-1})(A \otimes B)(S_A \otimes S_B) = (S_A^{-1} A S_A) \otimes (S_B^{-1} B S_B).$$

The identity $(A \otimes B)^{\text{T}} = A^{\text{T}} \otimes B^{\text{T}}$ completes the proof. □

We will also be using the following inequality, whose proof is trivial. Let M and N be symmetric matrices, with M positive definite. Then

$$\rho(M^{-1}N) \leqslant \|M^{-1}\|_2 \|N\|_2 = \frac{\rho(N)}{\lambda_{\min}(M)}. \tag{9}$$

In our analysis we will consider the simple case where convection is aligned with the x-axis, namely $\tau = \mu = 0$. We also assume that $|\gamma| \leqslant 1$, so that A_{kL} and A_{kP} are M-matrices. We note that one-dimensional convection in the y- or z-direction is just as easy to analyze. On the other hand, other situations may be rather difficult to analyze, even when simplifying assumptions (such as $\sigma = \tau = \mu$) are made.

4.1. 2D Case: *k-line ordering*

Consider the block Jacobi splitting $A_{kL} = M_{kL} - N_{kL}$, with M_{kL} consisting of all the blocks $(A_{kL})_{ii}$ discussed in Section 3.

As will be evident from the analysis, it is useful to consider the general setting of a uniform $n_x \times n_y$ grid, to establish the eigenvalues of A_{kL}; we will then resort back to $n_x = n_y = n$.

The matrix A_{1L} is block tridiagonal of size $n_x \times n_x$ with respect to its blocks. Each block is itself tridiagonal of size $n_y \times n_y$. As $\tau = 0$, we have $b = e$ and $c = d$, and thus the matrix of (7) reduces to

$$A_{1L} = \text{tri}_{n_x \times n_x}[\text{tri}_{n_y \times n_y}[-c^2, -2bc, -b^2],$$

$$\text{tri}_{n_y \times n_y}[-2c^2, a^2 - 4bc, -2b^2],$$

$$\text{tri}_{n_y \times n_y}[-c^2, -2bc, -b^2]]$$

$$= a^2 I_{n_x n_y} + \text{tri}_{n_x \times n_x}[1, 2, 1] \otimes \text{tri}_{n_y \times n_y}[-c^2, -2bc, -b^2].$$

Proposition 4.4
The eigenvalues of the matrix A_{kL} are given by

$$\lambda_{i,j}(A_{kL}) = a^2 - 4bc \left[1 + \cos\left(\frac{i\pi}{n_y+1}\right)\right]\left[1 + \cos\left(\frac{j\pi}{n_x+1}\right)\right],$$

$$i = 1, 2, \ldots, n_y, \quad j = 1, 2, \ldots, n_x.$$

Proof
Using Propositions 4.1 and 4.2, the eigenvalues for A_{1L} are easily calculated, and since all the matrices A_{kL} are permutations of each other, the result follows. □

The eigenvalues of M_{kL} can now be found as a simple corollary; we will set $n_x = n_y = n$ from this point on. Since the matrices M_{kL} are *not* permutations of each other, there is an explicit dependence on k.

Corollary 4.5
Let $h = 1/(2(n+1))$. Then, the eigenvalues of M_{kL} are given by

$$\lambda_{i,j}(M_{kL}) = a^2 - 4bc \left[1 + \cos\left(\frac{i\pi}{k+1}\right)\right](1 + \cos(2\pi j h)),$$

$$i = 1, 2, \ldots, k, \quad j = 1, 2, \ldots, n$$

each with multiplicity n/k. The minimal eigenvalue of M_{kL} is given by

$$\lambda_{\min}(M_{kL}) = a^2 - 4bc \left[1 + \cos\left(\frac{\pi}{k+1}\right)\right](1 + \cos(2\pi h)).$$

In addition, M_{kL} is symmetrizable by a real similarity transformation, and the symmetrized matrix is positive definite.

Proof
M_{kL} is a block diagonal with n/k identical blocks of size kn. The structure of each block $(M_{kL})_{ii}$ is that of the matrix A_{1L} discretized on a $k \times n$ grid.

Substituting $k = n_x = n_y$ into Proposition 4.4, there are n/k blocks and the minimal eigenvalue is attained when $i = j = 1$. From (5) it follows that $\lambda_{\min}$ is positive. Since M_{kL} is a Kronecker product of tridiagonal Toeplitz matrices, by Propositions 4.1 and 4.3 it is symmetrizable. □

To find the spectrum of the matrix N_{kL} for $k > 1$, we reorder grid unknowns in a way that exploits the structure of the graph of N_{kL}. In particular, we observe that all graph edges are between boundary nodes of neighboring grid blocks. Let $\partial_{x+} B_i$ and $\partial_{x-} B_i$ denote the boundary columns of a grid block B_i on the right and left sides, respectively. Then for every graph edge e there is an $i \in 1, 2, \ldots, n/k - 1$ such that e contains one node from $\partial_{x+} B_i$ and one node from $\partial_{x-} B_{i+1}$. This also implies that the interior nodes of B_i are not present in any of the edges of the graph of N_{kL}. Therefore, by grouping all the interior nodes together, we can get a permuted matrix with dense and well-structured blocks.

Motivated by these observations, we repartition G_{2D} into n/k new grid blocks $\tilde{B}_i$. The first $(n/k) - 1$ blocks $\tilde{B}_i$ contain the boundary nodes of original blocks,

$$\tilde{B}_i = \partial_{x+} B_i \cup \partial_{x-} B_{i+1}, \quad i = 1, 2, \ldots, \frac{n}{k} - 1.$$

The last block contains all the interior nodes,

$$\tilde{B}_{n/k} = G_{2D} - \bigcup_{i=1}^{(n/k)-1} \tilde{B}_i.$$

The permuted matrix corresponding to this new ordering, which we denote by $\tilde{N}_{kL}$, is block diagonal with n/k diagonal blocks. The first $(n/k)-1$ blocks are of size $2n \times 2n$, whereas the final block is of size $(n^2 - 2n((n/k)-1)) \times (n^2 - 2n((n/k)-1))$.

An example is shown in Figure 5, on a 9×9 grid that was originally ordered using a 3-line ordering.

Proposition 4.6
The matrix N_{kL} is symmetrizable and has the following eigenvalues: for $k=1$,

$$\lambda_{i,j}(N_{1L}) = 4bc(1 + \cos(2\pi i h))\cos(2\pi j h), \quad i,j = 1, 2, \ldots, n.$$

When $1 < k \leqslant n$, N_{kL} has $2n((n/k)-1)$ eigenvalues of the form

$$\lambda_i(N_{kL}) = \pm 2bc(1 + \cos(2\pi i h)), \quad i = 1, 2, \ldots, n$$

each of multiplicity $(n/k)-1$. The remaining $n^2 - 2n((n/k)-1)$ eigenvalues are zero.

Proof
For $k=1$ we have

$$N_{1L} = \mathrm{tri}_{n \times n}[1, 0, 1] \otimes \mathrm{tri}_{n \times n}[c^2, 2bc, b^2]. \tag{10}$$

Symmetrizability follows from Propositions 4.1, 4.3 and Equation (10), and the eigenvalues follow from Propositions 4.1 and 4.2.

For $1 < k \leqslant n$, we note that the first $(n/k)-1$ blocks of the block diagonal permuted matrix $\tilde{N}_{kL}$ are identical and are equal to

$$\mathrm{tri}_{2 \times 2}[c^2, 0, b^2] \otimes \mathrm{tri}_{n \times n}[1, 2, 1].$$

The expression for the nonzero eigenvalues follows from Propositions 4.1 and 4.2, whereas symmetrizeability follows from Propositions 4.1 and 4.3. The final diagonal block of $\tilde{N}_{kL}$ consists entirely of zeros and is of size $n^2 - 2n((n/k)-1)$, showing the existence and multiplicity of zero eigenvalues, as claimed. $\square$

Corollary 4.7
The spectral radii of N_{kL} are given by

$$\rho(N_{1L}) = 4bc(1 + \cos(2\pi h))\cos(2\pi h), \quad \rho(N_{kL}) = 2bc(1 + \cos(2\pi h)), \quad 1 < k < n.$$

We note that the case $k = n$ is trivial, since $N_{nL} = 0$. We are now in a position to make a statement regarding the convergence of block Jacobi.

Theorem 4.8
The spectral radius of the block Jacobi iteration matrix obeys the bound

$$\rho_{kL} \leqslant \tilde{\rho}_{kL} = \frac{2bc(1 + \cos(2\pi h))}{a^2 - 4bc\left(1 + \cos\left(\dfrac{\pi}{k+1}\right)\right)(1 + \cos(2\pi h))}[1 + \delta_{1k}(2\cos(2\pi h) - 1)].$$

In the case $k = 1$, the bound is attained, that is $\rho_{1L} = \tilde{\rho}_{1L}$.

Proof
By Corollary 4.5, M_{kL} can be symmetrized, and its symmetrized version is also positive definite. By Proposition 4.6 N_{kL} can also be symmetrized, and hence (9) can be used. Thus, by Corollaries 4.5, 4.7 and Equation (9), the stated bound holds.
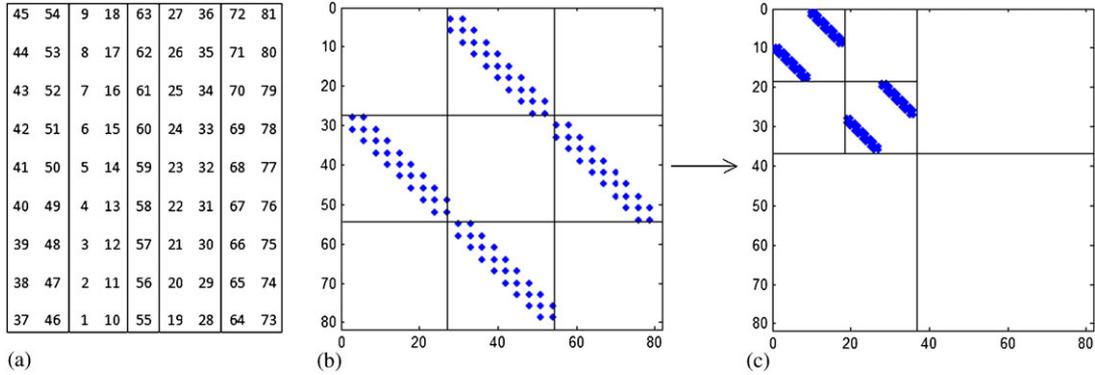
Figure 5. When unknowns are reordered as in (a), the matrix N_{kL}, whose sparsity pattern is shown in (b), takes on the block diagonal form depicted in (c).

For $k=1$ we can obtain an *exact* expression, as follows. We have

$$M_{1L} = a^2 I_{n^2} - 2I_n \otimes \text{tri}_{n \times n}[c^2, 2bc, b^2]$$

and N_{1L} given in the Kronecker product form by (10). If $\vec{x}$ and $\vec{y}$ are the dominant eigenvectors of $\text{tri}_{n \times n}[1, 0, 1]$ and $\text{tri}_{n \times n}[c^2, 2bc, b^2]$, by Proposition 4.2, $\vec{x} \otimes \vec{y}$ is both the dominant eigenvector of N_{1L} and the eigenvector corresponding to the minimal eigenvalue of M_{1L}. Therefore, the bound is attained. □

The following proposition shows that multiplying k by any integer $\geqslant 2$ improves the convergence of block Jacobi. In the proof and the remark that follows, matrix inequalities are used in an elementwise sense. That is, given matrices X and Y of the same dimensions we write $X \geqslant Y$ if $X_{ij} \geqslant Y_{ij}$ for all i and j.

Proposition 4.9
For any given $m > k \geqslant 1$, if k divides m then $\rho_{mL} < \rho_{kL}$.

Proof
Our proof makes use of the theory of nonnegative matrices and regular splittings [15, Chapters 2–3]. Two results are of particular importance. First, the inverse of an irreducible diagonally dominant M-matrix is elementwise positive [15, p. 85, Corollary 1]. Second, the Perron–Frobenius theorem states that when any of the entries of a nonnegative irreducible matrix are increased, the spectral radius also increases [15, Theorem 2.1].

When k divides m, the graph of N_{mL} is a subgraph of N_{kL}, and hence as long as the same ordering is applied to both, elementwise inequalities between the two matrices hold. Hence, $N_{kL} \geqslant N_{mL} \geqslant 0$ when N_{kL} and N_{mL} are permuted by reordering grid unknowns lexicographically. As the matrix A_{kL} is an irreducibly diagonally dominant M-matrix, we have $A_{kL}^{-1} > 0$. The claimed result now follows from [15, p. 90, Theorem 3.15]. □

Remark
We expect Proposition 4.9 to generalize to $\rho_{mL} < \rho_{kL}$ whenever $m > k$. However, if k does not divide m, the graph of N_{mL} is *not* a subgraph of N_{kL}, and there seems to be no permutation matrix P such that $P N_{mL} P^T \leqslant P N_{kL} P^T$. Therefore, the technique described in the proof of the proposition does not apply.

Table II shows experimentally the tightness of our bound. In the case $k=2$ the bound seems to become arbitrarily tight as $h \to 0$. For $k=3$, however, the quality of the bound is quite poor. As the above remark states, we expect the convergence of block Jacobi to improve as k gets larger, since a greater portion of the matrix A_{kL} is included in M_{kL}. However, while we expect that

Table II. Spectral radii ρ_{2L} and ρ_{3L} of the 2-line and 3-line iteration matrices, and the bounds $\tilde{\rho}_{2L}$ and $\tilde{\rho}_{3L}$. For the 2-line ordering the bound grows tight for large n, but not for the 3-line ordering. The values of the mesh Reynolds numbers in this example are $\gamma=0.5$ and $\delta=0$.

n	ρ_{2L}	$\tilde{\rho}_{2L}$	ρ_{3L}	$\tilde{\rho}_{3L}$
6	0.339	0.383	0.302	0.455
12	0.400	0.415	0.345	0.501
18	0.415	0.422	0.356	0.511
24	0.421	0.425	0.360	0.515
30	0.423	0.426	0.362	0.517

$\rho_{(k+1)L}<\rho_{kL}$ and this is confirmed experimentally for $k=2$ in Table II, for the bounds we actually get $\tilde{\rho}_{(k+1)L}>\tilde{\rho}_{kL}$ for $1<k<n$.

We can use the results of Theorem 4.8 to determine exactly the optimal parameter for block SOR with the 1-line ordering. In addition, we can determine the optimal parameter for 2-line ordering to good degree of accuracy. As A_{kL} is block tridiagonal with respect to $nk \times nk$ blocks, the analysis of Young [17, Chapter 14, Sections 5.2 and 14.3] applies.

Comparison of convergence rates (2D)

Let us compare the convergence of block Jacobi applied to our reduced system, the cyclically reduced system associated with operator (3), and the unreduced system. When referring to quantities of interest we use the superscripts $\square$, $\diamond$, and $+$, respectively, motivated by the shapes of the corresponding computational molecules.

A bound on the spectral radius of the 2D cyclically reduced matrix associated with the difference equation (3), with a diagonal one-line ordering strategy, is given by [8]

$$\rho_{1L}^{\diamond}\leqslant\tilde{\rho}_{1L}^{\diamond}=\frac{\left(\sqrt{1-\gamma^2}+\sqrt{1-\delta^2}\right)^2}{8-\left(\sqrt{1-\gamma^2}+\sqrt{1-\delta^2}\right)^2+2\sqrt{\left(1-\gamma^2\right)\left(1-\delta^2\right)}(1-\cos(\pi h))} \tag{11}$$

valid for $|\gamma|,|\delta|<1$. The diagonal ordering is equivalent to a 1-line ordering if the grid is viewed at a $45°$ angle. It was observed experimentally that $\tilde{\rho}_{1L}^{\diamond} \to \rho_{1L}^{\diamond}$ as $h \to 0$ with (γ,δ) fixed.

Setting $\delta=0$, we expand (11) as well as the results of Theorem 4.8 in a Taylor series about $h=0$.

$$\tilde{\rho}_{1L}^{\diamond} = 1-\left(\frac{\pi^2}{4}+\frac{\sigma^2}{4}\right)h^2+o(h^2),$$

$$\rho_{1L}^{\square} = 1-\left(4\pi^2+\frac{\sigma^2}{2}\right)h^2+o(h^2),$$

$$\rho_{2L}^{\square} \leqslant \tilde{\rho}_{2L}^{\square} = 1-(4\pi^2+\sigma^2)h^2+o(h^2).$$

As observed in Table II, it appears that $\tilde{\rho}_{2L}^{\square} \to \rho_{2L}^{\square}$ as $h \to 0$ with γ fixed.

For block Jacobi applied to the $+$-shaped 5-point operator (2), we have

$$\rho_{1L}^{+} = 1-\left(\pi^2+\frac{\sigma^2}{8}\right)h^2+o(h^2)$$

and

$$\rho_{2L}^{+} \leqslant \tilde{\rho}_{2L}^{+} = 1-\left(\pi^2+\frac{\sigma^2}{4}\right)h^2+o(h^2)$$

with the latter bound observed experimentally to grow arbitrarily tight as $h \to 0$ with γ fixed [14].

The asymptotic convergence rate is $R_\infty = -\log(\rho)$. Noting that $\log(1+ch^2)=ch^2+o(h^2)$, we have for $\sigma^2 \gg \pi^2$ and $h \ll 1$ the following hierarchy of relative asymptotic convergence rates:

$$(R_\infty)_{2L}^\square \approx 2(R_\infty)_{1L}^\square \approx 4(R_\infty)_{1L}^\diamond \approx 4(R_\infty)_{2L}^+ \approx 8(R_\infty)_{1L}^+.$$

We confirm this experimentally in Section 5.

4.2. 3D Case: k-plane ordering

Consider the block Jacobi splitting $A_{kP} = M_{kP} - N_{kP}$ with M_{kP} consisting of the blocks $(A_{kP})_{ii}$ of the k-plane matrix discussed in Section 3. As previously stated, we assume that $\vec{w}$ is aligned with the x-axis. Thus, we have $\delta = \eta = 0$, which means $b = e = p = s$ and $c = d = q = r$. As we did for the 2D case, it is useful to consider the general setting of a uniform $n_x \times n_y \times n_z$ grid to establish the eigenvalues of A_{kP}; we will then resort back to $n_x = n_y = n_z$.

A_{1P} is a block tridiagonal matrix of size $n_y \times n_y$ with respect to its blocks. Each block is itself a block tridiagonal matrix of size $n_x \times n_x$ with respect to $n_z \times n_z$ tridiagonal blocks.

Assuming that the ordering the gridpoints goes in a $z - x - y$ fashion without loss of generality, we have

$$A_{1P} = \text{tri}_{n_y \times n_y}[B, C, B],$$

where

$$B = \text{tri}_{n_x \times n_x}[\text{tri}_{n_z \times n_z}[-c^2, -2c^2, -c^2],$$
$$\text{tri}_{n_z \times n_z}[-2bc, -4bc, -2bc],$$
$$\text{tri}_{n_z \times n_z}[-b^2, -2b^2, -b^2]];$$
$$C = \text{tri}_{n_x \times n_x}[\text{tri}_{n_z \times n_z}[-2c^2, -4c^2, -2c^2],$$
$$\text{tri}_{n_z \times n_z}[-4bc, a^2 - 8bc, -4bc],$$
$$\text{tri}_{n_z \times n_z}[-2b^2, -4b^2, -2b^2]].$$

Using Kronecker products, we have

$$A_{1P} = a^2 I_{n_x n_y n_z} - \text{tri}_{n_y \times n_y}[1, 2, 1] \otimes \text{tri}_{n_x \times n_x}[c^2, 2bc, b^2] \otimes \text{tri}_{n_z \times n_z}[1, 2, 1]. \tag{12}$$

Proposition 4.10
The eigenvalues of the matrix A_{kP} are

$$\lambda_{i,j,k} = a^2 - 8bc \left[1 + \cos\left(\frac{i\pi}{n_x+1}\right)\right]\left[1 + \cos\left(\frac{j\pi}{n_y+1}\right)\right]\left[1 + \cos\left(\frac{\ell\pi}{n_z+1}\right)\right],$$

$$i = 1, 2, \ldots, n_x, \quad j = 1, 2, \ldots, n_y, \quad \ell = 1, 2, \ldots, n_z.$$

Proof
This follows from (12) and Propositions 4.1 and 4.2. □

Once again, the eigenvalues of M_{kP} follow as a simple corollary. From this point we assume $n_x = n_y = n_z = n$.

Corollary 4.11
The eigenvalues of the matrix M_{kP} are given by

$$\lambda_{i,j,k}(M_{kP}) = a^2 - 8bc \left[1 + \cos\left(\frac{i\pi}{k+1}\right)\right]\left[1 + \cos\left(\frac{j\pi}{k+1}\right)\right](1 + \cos(2\pi\ell h)),$$

$$i = 1, 2, \ldots, k, \quad j = 1, 2, \ldots, k, \quad \ell = 1, 2, \ldots, n$$

each with multiplicity n^2/k^2.
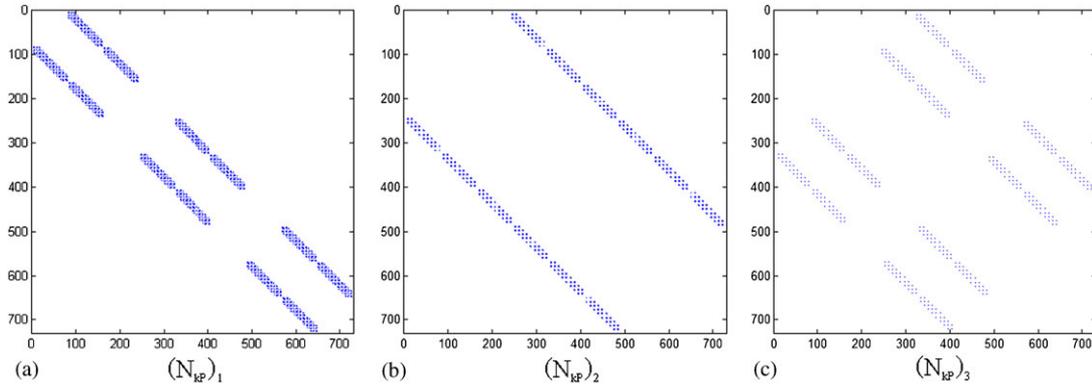
Figure 6. Partition of N_{kP} into (a) $(N_{kP})_1$; (b) $(N_{kP})_2$; and (c) $(N_{kP})_3$. The grid is $9 \times 9 \times 9$ with $k=3$.

The minimal eigenvalue of M_{kP} is

$$\lambda_{\min}(M_{kP}) = a^2 - 8bc \left[1 + \cos\left(\frac{\pi}{k+1}\right) \right]^2 (1 + \cos(2\pi h)).$$

Finally, M_{kP} is symmetrizable, and the symmetrized matrix is positive definite.

Proof
M_{kP} is block diagonal, with each block equal to the matrix A_{kP} discretized on a $k \times k \times n$ grid. The result follows from applying Proposition 4.10 to a $k \times k \times n$ subgrid. As M_{kP} is a Kronecker product of tridiagonal Toeplitz matrices, by Propositions 4.1 and 4.3 it is symmetrizable. By inspecting the values of a, b and c, it readily follows that $\lambda_{\min}$ is positive. □

To obtain a bound on the spectral radius of N_{kP} when $k>1$ we split this matrix into three pieces

$$N_{kP} = (N_{kP})_1 + (N_{kP})_2 + (N_{kP})_3,$$

where the eigenvalues of the three matrices on the right can be computed exactly. The sparsity patterns of these matrices are depicted in Figure 6.

The choice of splitting is motivated by the observation that the edges in the graph of N_{kP} are of three distinct types. The description of these categories is facilitated by the introduction of some notation regarding the faces and edges of each grid block B_i.

We denote by $\partial_{x+} B_i$ and $\partial_{x-} B_i$ the faces of B_i whose outward facing unit normals are $(1,0,0)$ and $(-1,0,0)$, respectively. The faces in the y-direction are denoted in the same fashion. The edge where the faces with normals $(1,0,0)$ and $(0,1,0)$ meet is denoted by $\partial_{x+y+} B_i$, with obvious modifications for the other edges.

The edges in the graph of N_{kP} can be partitioned into those between $\partial_{x+} B_1$ and $\partial_{x-} B_2$ for grid blocks B_1 and B_2 sharing a common face of constant x (type 1), those between $\partial_{y+} B_1$ and $\partial_{y-} B_2$ for grid blocks B_1 and B_2 sharing a common face of constant y (type 2), and those between $\partial_{x+y+} B_1$, $\partial_{x+y-} B_2$, $\partial_{x-y-} B_3$, and $\partial_{x-y+} B_4$ for grid blocks B_1, B_2, B_3, and B_4 sharing a common edge (type 3). The matrix $(N_{kP})_i$ is then defined to be the matrix whose graph contains all N_{kP}'s graph edges of type i.

There exist three orderings i_1, i_2, and i_3 of G_{3D}, each bringing one of the pieces of N_{kP} to block diagonal form. As justified in Section 3, each of these orderings is specified by a new choice of grid blocks for G_{3D}.

The ordering i_1 is defined by the choice of grid blocks

$$\tilde{B}_i = \partial_{x+} B_i \cup \partial_{x-} B_{i+1}, \quad i = 1, 2, \ldots, \frac{n}{k}\left(\frac{n}{k} - 1\right)$$

(where B_i is the ith block of the k-plane ordering) and

$$\tilde{B}_{\frac{n}{k}(\frac{n}{k}-1)+1} = G_{3D} - \bigcup_{i=1}^{\frac{n}{k}(\frac{n}{k}-1)} \tilde{B}_i.$$

The permuted matrix $(\tilde{N}_{kP})_1$ is block diagonal with $\frac{n}{k}(\frac{n}{k}-1)$ blocks of size $2kn \times 2kn$, and one block of zeroes of size $(n^3 - 2n^2(\frac{n}{k}-1)) \times (n^3 - 2n^2(\frac{n}{k}-1))$. For i_2 we do the same thing with grid blocks $\tilde{B}_i = \partial_{y+} B_i \cup \partial_{y-} B_{i+1}$, obtaining a matrix with identical structure.

The ordering i_3 is defined by the choice of grid blocks

$$\grave{B}_i = \partial_{x+y+} B_i \cup \partial_{x-y+} B_{i+1} \cup \partial_{x+y-} B_{i+\frac{n}{k}} \cup \partial_{x-y-} B_{i+\frac{n}{k}+1}, \quad i = 1, 2, \ldots, \left(\frac{n}{k}-1\right)^2,$$

and

$$\grave{B}_{(\frac{n}{k}-1)^2+1} = G_{3D} - \bigcup_{i=1}^{(\frac{n}{k}-1)^2} \grave{B}_i.$$

The block diagonal matrix $(\grave{N}_{kP})_3$ has $((n/k)-1)^2$ diagonal blocks of size $4n \times 4n$, followed by a zero block of size $(n^3 - 4n((n/k)-1)^2) \times (n^3 - 4n((n/k-1)^2))$.

Example 4.12 describes these orderings and shows the permuted $(N_{kP})_i$ for an example grid and a particular k.

Example 4.12
Consider a $9 \times 9 \times 9$ grid with $k=3$. In Figure 7(a) we look at a cross-section with a constant z of the graph of N_{kP}, where graph vertices are placed in the same location as their corresponding grid cells. Graph edges of type 1 (i.e. edges between grid blocks sharing a common face of constant x) are shown with thin lines, graph edges of type 2 (i.e. edges between grid blocks sharing a common face of constant y) are shown in bold, and graph edges of type 3 (i.e. graph edges between grid blocks sharing a common edge) are shown with wavy lines. The ordering that brings $(N_{kP})_2$ into block diagonal form is shown. Note that the full graph of N_{kP} actually consists of nine layered copies of Figure 7(a), with edges between each layer and the layers directly above and below it in the z direction. Figure 7(b) illustrates this point with a constant x slice of the graph.

As stated in Section 3, to define a new block ordering it suffices to specify new grid blocks. The ordering shown uses as grid blocks the six common y-faces (each containing 54 nodes), as well
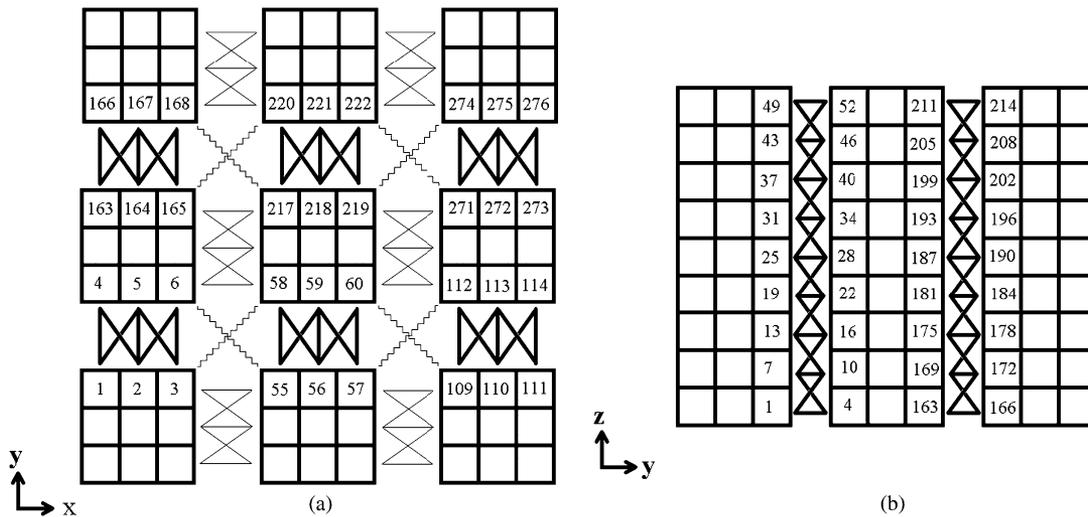
Figure 7. (a) A z=constant slice of the graph of N_{kP} for a $9 \times 9 \times 9$ grid with $k=3$ and (b) A x=constant slice.
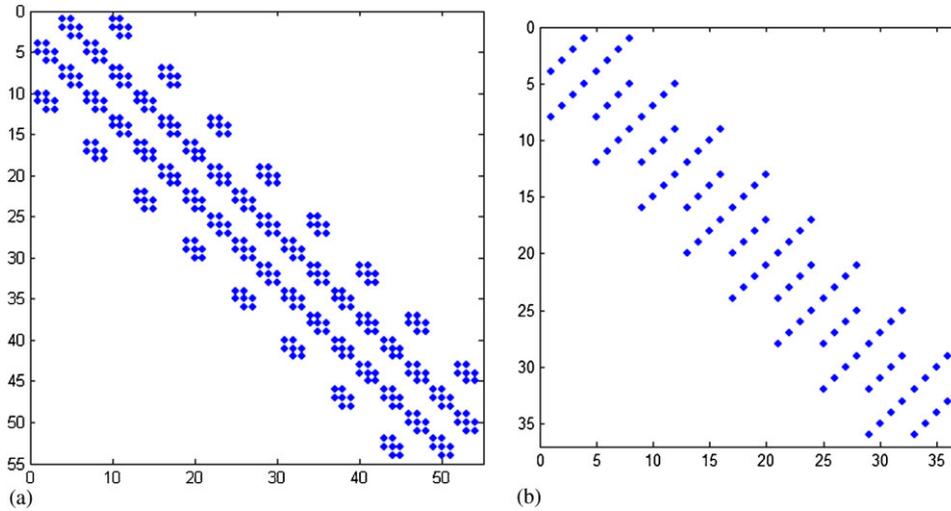
Figure 8. (a) Nonzero blocks of permuted $(N_{kP})_1$ and $(N_{kP})_2$ and (b) Nonzero blocks of permuted $(N_{kP})_3$. Grid is $9 \times 9 \times 9$, $k=3$.

as one final zero block consisting of all nodes that have no type 2 edges. An analogous ordering bringing $(N_{kP})_1$ into block diagonal form does the same thing using faces of constant x.

Finally, the four sets of common edges, each containing 36 nodes (together with a zero block of all nodes with no type 3 edges) bring $(N_{kP})_3$ to block diagonal form. If m is one of the first 36 nodes in this ordering (i.e. $i_3(m) \in \{1, 2, \ldots, 36\}$) then in terms of the ordering i_2 shown in Figure 7, we have

$$i_2(m) = c_{i_3(m)-1(\text{mod}\,4)} + 6\lfloor i_3(m) - 1 \rfloor,$$

where $c_0 = 3$, $c_1 = 55$, $c_2 = 6$, and $c_3 = 58$.

Figure 8(a) shows one of the six identical diagonal blocks of the permuted $(N_{kP})_2$ for this example. From the above description it follows that the structure for the permuted $(N_{kP})_1$ is identical, since it can be obtained by switching the roles of x and y. On the other hand, in Figure 8(b) we show that the permuted $(N_{kP})_3$ has a different structure, because the nonzero values of this matrix correspond to the edges between corners, rather than edges between faces.

Lemma 4.13
N_{kP} is symmetrizable. The eigenvalues of N_{1P} are given by

$$\lambda_{ij\ell}(N_{1P}) = 8bc[(1+\cos(2\pi i h))(1+\cos(2\pi j h))-1][1+\cos(2\pi \ell h)],$$

$$i, j, \ell = 1, 2, \ldots, n.$$

The spectral radius of N_{1P} is thus given by

$$\rho(N_{1P}) = 8bc\cos(2\pi h)[1+\cos(2\pi h)][2+\cos(2\pi h)].$$

When $k > 1$, a bound on $\rho(N_{kP})$ is given by

$$\rho(N_{kP}) \leqslant 2bc\left[1 + 4\left(1 + \cos\left(\frac{\pi}{k+1}\right)\right)\right][1+\cos(2\pi h)].$$

Proof
We have

$$N_{1P} = (\text{tri}_{n \times n}[1, 2, 1] \otimes \text{tri}_{n \times n}[c^2, 2bc, b^2] - 4bc I_{n^2}) \otimes \text{tri}_{n \times n}[1, 2, 1].$$

The claim on its eigenvalues now follows from Propositions 4.1 and 4.2. The spectral radius is attained at $i = j = \ell = 1$. Symmetrizability follows from Propositions 4.1 and 4.3.

To prove symmetrizability when $k>1$ we define the matrix N_{kP}^{lex} to be the permutation of N_{kP} obtained by reverting back to the lexicographic ordering. Next we note that the graph of N_{kP}^{lex} is a subgraph of the graph of N_{1P} and the former graph is obtained from the latter by removing edges in symmetric pairs. That is, whenever edge (v, w) is removed edge (w, v) is also removed. Therefore, the matrix N_{kP}^{lex} is obtained from N_{1P} by setting various matrix entries $N_{1P:i,j}$ to zero but setting the corresponding $N_{1P:j,i}$ to zero as well. Hence, the same matrix that symmetrizes N_{1P} symmetrizes N_{kP}^{lex}.

To prove the claim on the spectral radius for $k>1$, we follow the procedure of block ordering and permutations described in detail in the text following Corollary 4.11 and in Example 4.12. Each nonzero block of the block diagonal permuted $(N_{kP})_1$ is the matrix

$$\text{tri}_{n\times n}[1,2,1]\otimes\text{tri}_{2\times 2}[-c^2,0,-b^2]\otimes\text{tri}_{k\times k}[1,2,1].$$

By Propositions 4.1 and 4.2, the eigenvalues of $(N_{kP})_1$ are

$$\pm 4bc\left(1+\cos\left(\frac{i\pi}{k+1}\right)\right)(1+\cos(2\pi i h)),\quad i=1,2,\ldots,k,\quad j=1,2,\ldots,n$$

each of multiplicity $(n/k)((n/k)-1)$ and 0 of multiplicity $(n^3-2n^2((n/k)-1))$. Clearly,

$$\rho((N_{kP})_1)=4bc\left(1+\cos\left(\frac{\pi}{k+1}\right)\right)(1+\cos(2\pi h)). \tag{13}$$

Similarly, we have

$$\rho((N_{kP})_2)=\rho((N_{kP})_1). \tag{14}$$

We also note that each nonzero block of $(N_{kP})_3$ is the matrix

$$\text{tri}_{n\times n}[1,2,1]\otimes\text{tri}_{2\times 2}[1,0,1]\otimes\text{tri}_{2\times 2}[-c^2,0,-b^2].$$

By Propositions 4.1 and 4.2, the eigenvalues of $(N_{kP})_3$ are

$$\pm 2bc(1+\cos(2\pi i h)),\quad i=1,2,\ldots,n$$

each of multiplicity $((n/k)-1)^2$ and 0 of multiplicity $n^3-4n((n/k)-1)^2$. Clearly,

$$\rho((N_{kP})_3)=2bc(1+\cos(2\pi h)). \tag{15}$$

The bound on $\rho(N_{kP})$ now follows from (13)–(15), and the observation that $\rho(N_{kP})\leqslant\rho((N_{kP})_1)+\rho((N_{kP})_2)+\rho((N_{kP})_3)$. $\square$

Theorem 4.14
The spectral radius of the block Jacobi iteration matrix obeys the bound

$$\rho_{kP}\leqslant\tilde{\rho}_{kP}=\frac{2bc\left[1+4\left(1+\cos\left(\frac{\pi}{k+1}\right)\right)\right](1+\cos(2\pi h))}{a^2-8bc\left[1+\cos\left(\frac{\pi}{k+1}\right)\right]^2(1+\cos(2\pi h))}$$

$$\times\left[1+\delta_{1k}\left(\frac{4}{5}\cos(2\pi h)(2+\cos(2\pi h))-1\right)\right].$$

Furthermore, the bound is exact for $k=1$, that is $\tilde{\rho}_{1P}=\rho_{1P}$.

Proof
By Corollary 4.11, M_{kP} is symmetrizable and positive definite. By Lemma 4.13 N_{kP} is also symmetrizable. Therefore, (9) may be used, and the bound follows from Lemma 4.13 and Corollary 4.11.

Table III. Spectral radii ρ_{2P} and ρ_{3P} of the 2-plane and 3-plane iteration matrices, and the bounds $\tilde{\rho}_{2P}$ and $\tilde{\rho}_{3P}$. For the 2-plane ordering the bound grows tight for large n, but not for the 3-line ordering. We have taken $\gamma = 0.5$ and $\delta = 0 = \eta = 0$.

n	ρ_{2P}	$\tilde{\rho}_{2P}$	ρ_{3P}	$\tilde{\rho}_{3P}$
6	0.430	0.521	0.372	0.726
12	0.524	0.554	0.454	0.784
18	0.547	0.561	0.475	0.797
24	0.556	0.564	0.483	0.802
30	0.556	0.565	0.487	0.805
36	0.562	0.566	0.489	0.806
42	0.564	0.566	0.490	0.807

For $k = 1$ we have

$$M_{1P} = a^2 I_{n^3} - 4bc I_{n^2} \otimes \text{tri}_{n \times n}[1, 2, 1]$$

and

$$N_{1P} = (\text{tri}_{n \times n}[1, 2, 1] \otimes \text{tri}_{n \times n}[c^2, 2bc, b^2] - 4bc I_{n^2}) \otimes \text{tri}_{n \times n}[1, 2, 1].$$

If $\vec{x}$ and $\vec{y}$ denote the dominant eigenvectors of $\text{tri}_{n \times n}[1, 2, 1]$ and $\text{tri}_{n \times n}[b^2, 2ab, a^2]$, then from Proposition 4.2 the vector $\vec{x} \otimes \vec{y} \otimes \vec{x}$ is both the dominant eigenvector of N_{1P} and the eigenvector corresponding to the minimal eigenvalue of M_{1P}. The bound is therefore attained in this case. $\square$

The results of Proposition 4.9 and the remark that followed apply to this case in exactly the same way as in 2D.

Table III shows experimentally the tightness of our bound. In the case $k = 2$, the bound seems to become arbitrarily tight as $h \to 0$. For $k = 3$ (and in fact for $k > 2$ in general) the bound is of low quality, but the computed spectral radius for $k = 3$ is smaller compared with the one for $k = 2$. This situation is the same as in 2D.

Comparison of convergence rates (*3D*)

A bound on the spectral radius of the block Jacobi iteration matrix obtained by applying a 2-plane ordering to the 19-point 3D cyclically reduced operator (4) is given by [13]

$$\rho_{2P}^{19pt} \leqslant \tilde{\rho}_{2P}^{19pt} = 1 - (\tfrac{10}{9}\pi^2 + \tfrac{1}{6}\mu^2 + \tfrac{1}{6}\tau^2 + \tfrac{1}{6}\sigma^2)h^2 + o(h^2).$$

Expanding the results of Theorem 4.14 in a Taylor series gives

$$\rho_{1P}^{27pt} = 1 - \left(4\pi^2 + \frac{\sigma^2}{3}\right)h^2 + o(h^2)$$

and

$$\rho_{2P}^{27pt} \leqslant \tilde{\rho}_{2P}^{27pt} = 1 - \tfrac{4}{7}(4\pi^2 + \sigma^2)h^2 + o(h^2).$$

For the 7-point operator, we have [14]

$$\rho_{1P}^{7pt} = 1 - (\tfrac{3}{4}\pi^2 + \tfrac{1}{16}\sigma^2 + \tfrac{1}{16}\tau^2 + \tfrac{1}{16}\mu^2)h^2 + o(h^2)$$

and

$$\rho_{2P}^{7pt} \leqslant \tilde{\rho}_{2P}^{7pt} = 1 - (\tfrac{1}{2}\pi^2 + \tfrac{1}{8}\sigma^2 + \tfrac{7}{64}\tau^2 + \tfrac{7}{64}\mu^2)h^2 + o(h^2).$$

We have found experimentally that all the above bounds grow arbitrarily tight as $h \to 0$ with mesh Reynolds numbers held constant. Noting again $-\log(1 + ch^2) = ch^2 + o(h^2)$, we construct Table IV showing the expected relative asymptotic convergence rates, valid for $h \ll 1$ and $\sigma^2 \gg \pi^2$.

Table IV. Relative asymptotic convergence rates for different combinations of discretization scheme and grid ordering. The number in row i and column j is the ratio of the asymptotic convergence rate of method i to method j, using Taylor expansions.

	7-pt $1P$	7-pt $2P$	19-pt $2P$	27-pt $1P$	27-pt $2P$
7-pt $1P$	1				
7-pt $2P$	2	1			
19-pt $2P$	2.67	1.34	1		
27-pt $1P$	5.35	2.67	2	1	
27-pt $2P$	9.14	4.57	3.42	1.71	1

Table V. Number of block Jacobi iterations it takes to reduce the 2-norm of the relative residual by a factor of 10^{-4} for constant coefficients. The initial guess is the zero vector. The right-hand side was constructed so that the solution was a vector of all ones. Results are shown for $\sigma = 60$ in 2D and $\sigma = 30$ in 3D, and $\tau = \mu = 0$.

	$n = 129$	$n = 257$	$n = 513$
(a) 2D			
+, 1L	890	3319	12 299
+, 2L	450	1664	6154
◇, 1L	457	1680	6210
◇, 2L	234	845	3109
□, 1L	243	891	3307
□, 2L	127	451	1659
(b) 3D	$n = 17$	$n = 33$	$n = 65$
7-pt, $1P$	75	287	1098
7-pt, $2P$	42	148	554
19-pt, $1P$	46	170	643
19-pt, $2P$	33	112	416
27-pt, $1P$	15	56	208
27-pt, $2P$	11	35	125

Computational cost. For the operators we consider, it is straightforward to estimate the computational cost per iteration. Due to the process of reduction, our cyclically reduced operators entail a lower computational cost of a single iteration, compared with the other operators. For example, in the 2D case the new operator is nine point, as is the operator R_2 in (3), but we work on a grid with only $1/4$ of the unknowns rather than $1/2$. In the 3D case our operator has 27 points in its stencil (as opposed to 19 points that the operator R_3 in (4) has) but we solve for only $1/8$ of the unknowns rather than $1/2$. Solving for the unknowns that were eliminated in the process entails a negligible cost, as it involves a small number of diagonal system solves.

5. NUMERICAL EXPERIMENTS

In this section we present an experimental examination of our approach. We show results for block Jacobi, for which we have carried out a detailed analysis, and then briefly explore the performance of modern Krylov solvers. Specifically, we apply GMRES, preconditioned with incomplete LU. All experiments were done using the MATLAB. In both 2D and 3D we have looked at three linear systems arising from discretization of (1), corresponding to the second-order centered difference discretization, the reduced system arising from cyclic reduction represented by (3) and (4), and the system arising from our new approach.

Table VI. Constant coefficient problem: number of iterations for GMRES preconditioned with ILU(0.01) to bring the norm of the relative residual down to 10^{-4}. $\sigma=60$, $\tau=\mu=0$, with $h=1/(n+1)$.

	$n=129$	$n=257$	$n=513$
(a) 2D			
1-line			
+	20	46	97
◇	11	26	55
□	7	15	34
2-line	$n=129$	$n=257$	$n=513$
+	18	36	76
◇	12	27	63
□	6	14	31
(b) 3D			
1-plane	$n=17$	$n=33$	$n=65$
7-point	5	9	21
19-point	4	8	16
27-point	2	4	8
2-plane	$n=17$	$n=33$	$n=65$
7-point	5	10	21
19-point	4	7	15
27-point	2	4	8

Table VII. Variable coefficient problem: number of iterations for GMRES preconditioned with ILU(0.01) to bring the norm of the relative residual down by a factor of 10^{-4}.

	$n=129$	$n=257$	$n=513$
(a) 2D			
1-line			
+	40	65	115
◇	18	30	57
□	14	23	42
2-line	$n=129$	$n=257$	$n=513$
+	28	45	85
◇	20	37	63
□	15	25	43
(b) 3D			
1-plane	$n=17$	$n=33$	$n=65$
7-point	7	12	23
19-point	5	9	14
27-point	3	6	11
2-plane	$n=17$	$n=33$	$n=65$
7-point	7	13	24
19-point	5	8	14
27-point	3	6	11

5.1. Test problem 1: constant coefficients

Table V shows the number of iterations required for the convergence of block Jacobi. Then, we show the results of applying GMRES, preconditioned with ILUTP to each system in turn. A variety of values of σ, τ and μ were tried, for several different grid sizes. The results of a few of these experiments are shown in Table VI.

In the tables we use +, ◇, and □, respectively to denote the 5-point operator, the operator R_2, and the new operator, as per the shapes of the corresponding computational molecules.

The results validate our analysis. The reduced system is solved within approximately half the number of iterations of the cyclically reduced system represented by operator (4). For block stationary schemes, the 2-line/plane orderings are provably superior in terms of speed of convergence, and this is manifested also in the experiments. For GMRES the benefit of using these orderings in place of lexicographic ordering is less obvious in terms of iteration counts.

5.2. Test problem 2: variable coefficients

We now modify problem (1) so that the vector $\vec{w}$ is a function of the spatial coordinates. Table VII shows the results for a 2D circular flow problem, $\vec{w} = 20((1/2) - y, (1/2) + x)$, and for a 3D sink problem, $\vec{w} = 20((1/2) - x, (1/2) - y, (1/2) - z)$. The results lead to very similar conclusions as those for the constant coefficient case.

6. CONCLUSIONS

We have introduced new cyclically reduced operators for the discrete convection–diffusion equation in 2D and 3D. The operators are derived by applying multi-color orderings (2^d colors in d dimensions) and assigning different discretizations to different colors, so that an elimination of all but one color is possible. This results in simple stencils on a cartesian grid: a 9-point operator in the 2D case and a 27-point operator in the 3D case. The operators maintain the same good spectral properties that other cyclically reduced operators have, and at the same time, the structure of their stencil makes them easy to implement. We have analyzed the effect of block orderings in conjunction with our new operator, and have found improvements in convergence rates.

REFERENCES

1. Buneman O. A compact non-iterative Poisson solver. *Report 294*, Stanford University, Institute for Plasma Research, 1969.
2. Buzbee BL, Golub GH, Nielson CW. On direct methods for solving Poisson's equations. *SIAM Journal on Numerical Analysis* 1970; **7**:627–656.
3. Hockney RW. A fast direct solution of Poisson's equation using Fourier analysis. *Journal of the Association for Computing Machinery* 1965; **12**:95–113.
4. Gander W, Golub GH. *Cyclic Reduction—History and Applications*. Scientific Computing (Hong Kong, 1997). Springer: Singapore, 1997; 73–85.
5. Trottenberg U, Oosterlee CW, Schüller A. *Multigrid*. Academic Press Inc.: San Diego, CA, 2001.
6. Ries M, Trottenberg U, Winter G. A note on MGR methods. *Linear Algebra and its Applications* 1983; **49**:1–26.
7. MacLachlan S, Manteuffel T, McCormick S. Adaptive reduction-based AMG. *Numerical Linear Algebra with Applications* 2006; **13**(8):599–620.
8. Elman HC, Golub GH. Iterative methods for cyclically reduced nonselfadjoint linear systems. *Mathematics of Computation* 1990; **54**(190):671–700.
9. Elman HC, Golub GH. Iterative methods for cyclically reduced nonselfadjoint linear systems. II. *Mathematics of Computation* 1991; **56**(193):215–242.
10. Elman HC, Golub GH. Line iterative methods for cyclically reduced discrete convection-diffusion problems. *SIAM Journal on Scientific and Statistical Computing* 1992; **13**(1):339–363.
11. Greif C, Varah J. Iterative solution of cyclically reduced systems arising from discretization of the three-dimensional convection-diffusion equation. *SIAM Journal on Scientific Computing* 1998; **19**(6):1918–1940.
12. Greif C. Reduced systems for three-dimensional elliptic equations with variable coefficients. *SIAM Journal on Matrix Analysis and Applications* 1999; **21**(1):29–44.
13. Greif C, Varah J. Block stationary methods for nonsymmetric cyclically reduced systems arising from three-dimensional elliptic equations. *SIAM Journal on Matrix Analysis and Applications* 1999; **20**(4):1038–1059.
14. Golub GH, Greif C, Varah JM. Block orderings for tensor-product grids in two and three dimensions. *Numerical Algorithms* 2002; **30**(2):93–111.
15. Varga RS. *Matrix Iterative Analysis*. Prentice-Hall Inc.: Englewood Cliffs, NJ, 1962.
16. Horn RA, Johnson CR. *Topics in Matrix Analysis*. Cambridge University Press: Cambridge, 1991.
17. Young DM. *Iterative Solution of Large Linear Systems*. Academic Press: New York, 1971.