Newton-Krylov-Schwarz Methods in CFD

X.-C. Cai
Department of Computer Science
University of Colorado, Boulder, CO 80309 USA

W. D. Gropp
Mathematics and Computer Science Division
Argonne National Laboratory, Argonne, IL 60439 USA

D. E. Keyes
Institute for Computer Applications in Science and Engineering
NASA-LaRC, Hampton, VA 23681 USA,
Department of Computer Science
Old Dominion University, Norfolk, VA 23529 USA, and
Department of Mechanical Engineering
Yale University, New Haven, CT 06520 USA

M. D. Tidriri
Institute for Computer Applications in Science and Engineering
NASA-LaRC, Hampton, VA 23681 USA

Summary

Newton-Krylov methods are potentially well suited for the implicit solution of nonlinear problems whenever it is unreasonable to compute or store a true Jacobian. Krylov-Schwarz iterative methods are well suited for the parallel implicit solution of multidimensional systems of boundary value problems that arise in CFD. They provide good data locality so that even a high-latency workstation network can be employed as a parallel machine. We call the combination of these two methods Newton-Krylov-Schwarz and report numerical experiments on some algorithmic and implementation aspects: the use of mixed discretization schemes in the (implicitly defined) Jacobian and its preconditioner, the selection of the differencing parameter in the formation of the action of the Jacobian, the use of a coarse grid in additive Schwarz preconditioning, and workstation network implementation. Three model problems are considered: a convection-diffusion problem, the full potential equation, and the Euler equations.

1. Introduction

Newton-like methods, together with fully implicit linear solvers, in principle allow a more rapid asymptotic approach to steady states, $f(u) = 0$, than do time-explicit methods or semi-implicit methods based on defect correction. Strict Newton methods have the disadvantage of requiring solutions of linear systems of equations based on the Jacobian, $f_u(u)$, of the true steady nonlinear residual and are often impractical in several respects:

1. Their quadratic convergence properties are realized only asymptotically. In early stages of the nonlinear iteration, continuation or regularization is typically required in order to prevent divergence.
2. Some popular discretizations (e.g., using limiters) of $f(u)$ are nondifferentiable, leaving the Jacobian undefined in a continuous sense.
3. Even if $f_u(u)$ exists, it is often inconvenient or expensive to form either analytically or numerically, and may be inconvenient to store.
4. Even if the true Jacobian is easily formed and stored, it may have a bad condition number.
5. The most popular family of preconditioners for large sparse Jacobians on structured or unstructured two- or three-dimensional grids, incomplete factorization, is difficult to parallelize efficiently.

In this paper we examine how points (3) through (5) may be addressed through Newton-Krylov-Schwarz methods. Our point of view with respect to (1) is that there will usually be an asymptotic regime in which the power of Newton's method is desirable if the storage overhead is not too great. To connect the opening iterations to the asymptotic regime, polyalgorithmic linear solvers for the Newton corrections were shown to be desirable in, for instance, [8]. Regarding (2), we refer to [19] for recent developments. The last three considerations are the most important with respect to parallel CFD. For a variety of reasons, industrial CFD groups are inclining towards the distributed network computing environment characterized by coarse to medium granularity, large memory per node, and very high latency. The all-to-all data dependencies between the unknown fields in a fully implicit method have led to a resurgence of interest in less rapidly convergent methods in high-latency parallel environments. Resisting, we present related investigations that lie along the route to parallel implicit CFD. Sections §2 and §3 briefly review Newton-Krylov and Krylov-Schwarz domain decomposition methods, respectively. Numerical results on three model problems, each focusing on different parts of the overall development of parallel Newton-Krylov-Schwarz methods, are then presented in §4 through §6. It is our intention to bring these developments together in a Navier-Stokes code, as described in the conclusions.

2. Newton-Krylov Methods

High-accuracy evaluation of the discrete residuals of d-dimensional flow formulations may require a large number of arithmetic operations. (For instance, a $(d+2)$-dimensional eigendecomposition may be required at each grid point in an Euler code.) Their Jacobians, though block-sparse, have dense blocks and are usually an order of magnitude even more complex to evaluate, whether by analytical or numerical means. Hence, matrix-free Newton-Krylov methods, in which the action of the Jacobian is required only on a set of given vectors, instead of all possible vectors, are natural in this context. To solve the nonlinear system $f(u) = 0$, given u^0, let $u^{l+1} = u^l + \lambda^l \delta u^l$, for $l = 0, 1, \ldots$, until the residual is sufficiently small, where δu^l approximately solves the Newton correction equation $J(u^l)\delta u^l = -f(u^l)$, and parameter λ^l is selected by some line search or trust region algorithm [6]. Krylov methods, such as the method of conjugate gradients for symmetric positive definite systems or GMRES for general nonsingular systems, find the best approximation of the solution in a relatively small-dimensional subspace that is built up

from successive powers of the Jacobian on the initial residual. The Krylov solver used throughout this paper is GMRES [15], because of previous comparisons [10] with other modern Krylov solvers on the same problem class that showed CPU cost differences to be small and unsystematic when well-enough preconditioned that any of the methods were practical.

The action of Jacobian J on an arbitrary Krylov vector w can be approximated by

$$J(u^l)w \approx \frac{1}{\epsilon}\left[f(u^l + \epsilon w) - f(u^l)\right].$$

Finite-differencing with ϵ makes such matrix-free methods potentially much more susceptible to finite word-length effects than ordinary Krylov methods [13]. Steady aerodynamics applications require the solution of linear systems that lack strong diagonal dominance, so it is important to verify that properly-scaled matrix-free methods can be employed in this context.

An approximation to the Jacobian can be used to precondition the Krylov process. Examples are:

1. the Jacobian of a lower-order discretization,
2. the Jacobian of a related discretization that allows economical analytical evaluation of elements,
3. a finite-differenced Jacobian computed with lagged values for expensive terms, and
4. domain decomposition-parallel preconditioners composed of Jacobian blocks on subdomains of the full problem domain.

We consider case (1) in §4, case (2) in §6, and case (4) in §5 and §6. Case (4) can be combined with any of the split-discretization techniques (cases (1)–(3)), in principle.

Left preconditioning of the Jacobian with an operator B^{-1} can be accommodated via

$$B^{-1}J(u^l)w \approx \frac{1}{\epsilon}\left[B^{-1}f((u^l + \epsilon w)) - \tilde{f}(u^l)\right],$$

where $\tilde{f}(u^l) = B^{-1}f(u^l)$ is stored once, and right preconditioning via

$$J(u^l)B^{-1}w \approx \frac{1}{\epsilon}\left[f((u^l + \epsilon B^{-1}w)) - f(u^l)\right].$$

Right preconditioning is preferable when the focus is on comparing different preconditioners, since the residual norm measured as a by-product in GMRES and used in the termination test is independent of any right preconditioning. On the other hand, any left preconditioning changes the by-product residual norm in GMRES. Left preconditioning may be preferable when GMRES is applied in practice as the solver for an inexact Newton method. When the preconditioning B^{-1} is of high quality, the left-preconditioned residual serves as an estimate of the error in the Newton update vector. This leads to a useful termination condition when Newton step acceptance tests are based on $||\delta u||$.

3. Krylov-Schwarz Algorithms

A variety of parallel preconditioners, whose inverse action we denote by B^{-1}, can be induced by decomposing the domain of the underlying PDE, finding an approximate representation of J on each subdomain, inverting locally, and combining the results. Generically, we seek to approximate the inverse of J by a sum of local inverses:

$$B^{-1} = R_0^T J_{0,u^l}^{-1} R_0 + \sum_{k=1}^{K} R_k^T J_{k,u^l}^{-1} R_k \ , \quad \text{where } J_{k,u^l} = \left\{ \frac{\partial f_i(u^l)}{\partial u_j} \right\}$$

is the Jacobian of $f(u)$ for i and j in subdomain k ($k > 0$), subscript "0" corresponds to a possible coarse grid, and where R_k is a restriction operator that takes vectors spanning the entire space into the smaller dimensional subspace in which J_k is defined. We use the term "Krylov-Schwarz" to distinguish these methods within the general class of domain decomposition methods. In the parallel computing literature the latter term is now used as a synonym for "data parallelism," whereas in the computational engineering literature it has come to be associated with any algorithm based on traversing a "multiblock" data structure. Meanwhile, in the applied mathematics literature, domain decomposition has become associated with the process of identifying the subdomains in which different dominant balances between terms of the governing equations hold, in the sense of asymptotic analysis.

The simplest of the Schwarz preconditioners is block Jacobi, which can be regarded as a zero-overlap form of additive Schwarz [7]. The convergence rate of block Jacobi can be improved, at higher cost per iteration, with subdomain overlap and (for many problems) by solving an additional judiciously chosen coarse grid system. It is demonstrated numerically in [5] for a variety of nonselfadjoint scalar elliptic problems that additive Schwarz with a nested coarse grid, containing one degree of freedom per subdomain, provides an "optimal" preconditioning, in the sense that the number of iterations required to attain a fixed reduction in residual is bounded by a constant as either the mesh spacing h or the diameter of the subdomains H is indefinitely refined. Multiplicative Schwarz methods improve on additive methods as block Gauss-Seidel improves upon block Jacobi, by roughly a factor of two, with the same serialization penalty. In a situation in which there are more subdomains than processors, hybrid multicolored multiplicative/additive Schwarz is recommended for optimal convergence at a given parallel granularity [2].

Parallelism is not the sole motivation for Schwarz methods. We remark that, given a preconditioner for the global domain, a Krylov-Schwarz method in which the same preconditioner is applied locally on each subdomain may provide a better serial algorithm than Krylov acceleration of the original global preconditioner. Given a problem of size N and a preconditioner with arithmetic complexity $c \cdot N^{\alpha}$, partition the problem into P subproblems of size N/P. The complexity of applying the solver independently to the set of subproblems is $P \cdot c \cdot (N/P)^{\alpha}$. Even in serial, $P^{\alpha-1}$ sets of subdomain iterations iterations can be afforded to coordinate the solutions of the subproblems per single global iteration, while breaking even in total complexity. If $\alpha > 1$, there is "headroom" for the domain-decomposed approach, depending upon the overall spectral properties of the global and multidomain preconditioners. There may still be parallel headroom even if $\alpha = 1$, since the global method may involve too much communication to parallelize efficiently. In addition, a hierarchical data structure is often natural for modeling or implementation

reasons; and memory requirements, cache thrashing, or I/O costs on large problems may demand decomposition anyway.

4. A Convection-Diffusion Problem

The academic nonlinear convection-diffusion Dirichlet problem

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} - \nu\nabla^2 u = 0$$

from [17] is employed for tests of the Newton-Krylov method because, under the assumption of backward Euler time-differencing, $\frac{\partial u}{\partial t} \approx (u^{n+1} - u^n)/\Delta t$, an exact semi-discrete solution can be constructed for u^{n+1}. Specifically, if we set $u^n = (x^2 + y^2 + 1) + \Delta t \cdot (2x(x^2 + y^2 + 1) + 2y - 4\nu)$, then $u^{n+1} = x^2 + y^2 + 1$, and Dirichlet boundary values are set accordingly. This problem is discretized on a Courant-triangulated unit square using a hybrid finite-volume/finite-element first-order approximation [17]. It is extended to second-order upwinding for the convective terms using MUSCL-type approach [22].

The discretization is general enough to accommodate unstructured triangulated grids in two-dimensions; however, for easy visualization of the effects of inconsistent discretization of the true Jacobian and its preconditioner, four pairs of inconsistent preconditioner/Jacobian discretizations were first applied to the one-dimensional constant coefficient steady submodel,

$$a\frac{\partial u}{\partial x} - \nu\frac{\partial^2 u}{\partial x^2} = 0,$$

with $a > 0$ for a range of Peclet numbers, $\mathrm{Pe} \equiv a\Delta x/\nu$. In this model, the second-derivative term is always approximated by the standard central difference formula $-\frac{\partial^2 u}{\partial x^2}|_i \approx \frac{1}{h^2}(-u_{i-1}+2u_i-u_{i+1}) \equiv Du_i$ wherever it appears in the Jacobian, and whenever the diffusive terms are made a part of the preconditioner. The first-derivative term was variously approximated by either of the second-order formulae, $\frac{\partial u}{\partial x}|_i \approx \frac{1}{4h}(u_{i-2}-5u_{i-1}+3u_i+u_{i+1}) \equiv C_{U2}u_i$ or $\frac{\partial u}{\partial x}|_i \approx \frac{1}{2h}(u_{i-2} - 4u_{i-1} + 3u_i) \equiv C_{U2,FD}u_i$, or by the first-order formula, $\frac{\partial u}{\partial x}|_i \approx \frac{1}{h}(-u_{i-1} + u_i) \equiv C_{U1}u_i$. The last of these, the only diagonally dominant formula, and the formula with the most compact stencil, is used whenever a convective term appears in the preconditioner. Of the first two, the four-point formula, with one stencil point on the downwind side, corresponds to the second-order upwind extension of the cell-upwind scheme developed in [22] and considered further in the two-dimensional nonlinear convection-diffusion cases described below. The three-point formula, with all stencil points on the upwind side is a commonly used finite-difference form, which appears further below only for one-dimensional comparison purposes. We do not consider second-order central discretization of the convective term in the Jacobian, as is sometimes allowed in stationary defect correction methods, since, as was shown in [12], the preconditioned operator has an eigenvalue that tends to zero in the limit of large Peclet number for this choice.

Figure 1, generated with MATLAB, shows the spectra for four Jacobian/preconditioner pairs at each of three cell Peclet numbers: 0.1, 1.0, and 10.0. All Jacobian/preconditioner pairs considered have spectra that stay bounded away from the origin over the full range of Peclet number from zero to infinity. However, preconditioning with the convective

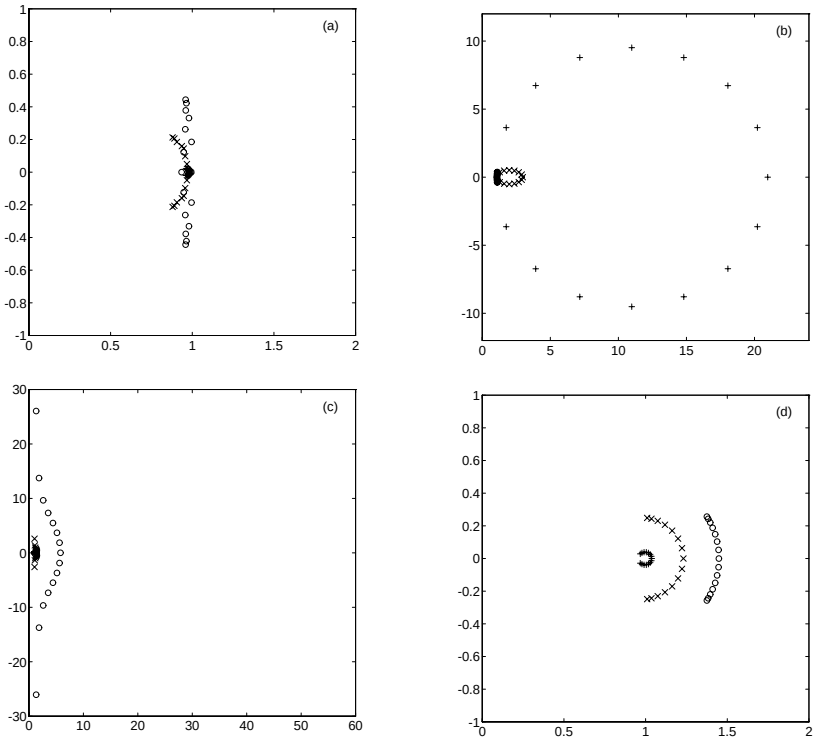

FIG. 1. *Spectra, in the right-half of the complex plane, of inconsistently preconditioned convection-diffusion operators in one-dimension, for Peclet numbers of 0.1 ('+'), 1.0 ('×'), and 10.0 ('∘'). (a)* $C_{U2} + D$ *preconditioned by* $C_{U1} + D$*. (b)* $C_{U2} + D$ *preconditioned by* C_{U1}*. (c)* $C_{U2} + D$ *preconditioned by* D*. (d)* $C_{U2,FD} + D$ *preconditioned by* $C_{U1} + D$*.* Note differences in scales.

operator alone obviously fails as $\mathrm{Pe} \to 0$ and preconditioning with the diffusive operator alone fails as $\mathrm{Pe} \to \infty$, in the sense that their spectra contain some elements that become arbitrarily large in the respective limits while others are clustered near unity, resulting in intolerably large condition numbers. In spite of the first of these unsuitable limits, upwind Euler discretizations for the left-hand side are sometimes applied to full Navier-Stokes residuals on the right-hand side in stationary defect correction methods. Both of the inconsistent Jacobian/preconditioner pairs that contain first-order upwind convection and standard diffusion together in the preconditioner lead to spectra that stay bounded in a small region of the complex plane near unity as either $\mathrm{Pe} \to 0$ or $\mathrm{Pe} \to \infty$. In fact, in case (d), when the convective term in the Jacobian contains no nonzero coefficients on the downwind side, the preconditioned spectrum clusters at a single point $(\frac{3}{2}, 0)$ in the infinite Peclet limit. (Both (a) and (d) have preconditioned spectra cluster that at a single point $(1, 0)$ in the zero Peclet limit.) Therefore, it appears possible to accelerate an inconsistent discretization pair and obtain good conditioning. It should be noted that since the preconditioned operators are generally non-normal, no conclusions should be drawn about the performance of Krylov-accelerated version of these methods on the basis of the exact spectra alone. It has been shown [18] that, in the presence of finite-precision arithmetic, the spectrum itself may be misleading for non-normal operators, and the *pseudo*-spectrum is more revealing. For present purposes, we merely show the spectra, and rely on the actual iteration counts as evidence of successful application of the inconsistent discretization of Fig. 1(a) in its two-dimensional, variable-coefficient generalization.

Before presenting the results of inconsistently preconditioned convection-diffusion Jacobians, we explore another axis of discretization parameter space, namely that of matrix-free approximation of the Jacobian-vector product, as described in §2. Our aim is to validate the existence of a range of the differencing parameter ϵ in the approximation of the Jacobian-vector product in which ϵ is simultaneously small enough for the Jacobian-vector product to be accurately estimated by just the first two terms of the Taylor series

$$f(u^l + \epsilon v) \approx f(u^l) + \epsilon J(u^l)v,$$

and large enough to avoid catastrophic cancellation for moderately ill-conditioned J.

Two techniques for choosing the scalar ϵ were investigated:

$$\begin{aligned} \epsilon &= \sqrt{\varepsilon_{mach}} \cdot (\|u^l\|_2 + 1), \text{ and} \\ \epsilon &= \sqrt{\varepsilon_{mach}} \cdot \frac{|(u^l, v)|}{\|v\|_2^2}. \end{aligned}$$

The first choice is simply the square root of the machine epsilon (or unit roundoff) multiplied by the norm of the current solution vector, so that the size of an assumed order-unity norm perturbation vector is not buried by a potentially large u^l. Conversely, by adding unity to $||u^l||$, the perturbation vector is kept large enough for there to be significance in the difference of the two residual vectors in the limit as $||u^l||$ becomes small. The second technique takes into account the magnitude of v, the Krylov vector, as well as that of u^l. The second was considered preferable to the first. We note here that GMRES may have an advantage over other Krylov methods in the matrix-free context in that the vectors v that arise in GMRES have unit two-norm, but may have widely varying scale in other Krylov methods for nonsymmetric systems. Right preconditioning spoils the perfect unit two-norm, however, and the second technique retains an advantage in this context. For an extended discussion of matrix-free applications of the Jacobian in the Krylov context, see [21].

TABLE 1

Iteration counts for unpreconditioned solution via GMRES of the nonlinear convection-diffusion problem, and the discrete two-norm of the overall error between the algebraic and exact solution for the finest grid case, without preconditioning.

	$\Delta t = 10^{-2}$		$\Delta t = 10^{-1}$	
h^{-1}	A_{U2}^{ex}	A_{U2}^{mf}	A_{U2}^{ex}	A_{U2}^{mf}
4	4	4	5	5
8	4	4	8	11
16	5	5	14	20
32	7	7	23	36
64	10	10	38	64
128	14	14	68	113
Error	1.12(-8)	1.88(-8)	3.20(-5)	3.55(-5)

Table 1 shows the iteration count for a 10^{-5} reduction in residual of the unpreconditioned Newton correction equation, in both explicit (superscript ex) and matrix-free (superscript mf) implementations. Two different time-step sizes are considered, with the

larger ($\Delta t = 0.1$) corresponding to a worse conditioned system. Deterioration of the convergence of the matrix-free method without preconditioning is evident on the finer grids; however, the error in the converged solution does not much suffer.

TABLE 2

Iteration counts for ILU(0)/GMRES preconditioned solution of the nonlinear convection-diffusion problem, and the discrete two-norm of the overall error between the algebraic and exact solution for the finest grid case, with inconsistent preconditioning.

	$\Delta t = 10^{-2}$		$\Delta t = 10^{-1}$	
h^{-1}	$(A^{ex}_{U2})(A^{ex}_{U1})^{-1}$	$(A^{mf}_{U2})(A^{ex}_{U1})^{-1}$	$(A^{ex}_{U2})(A^{ex}_{U1})^{-1}$	$(A^{mf}_{U2})(A^{ex}_{U1})^{-1}$
4	1	2	1	4
8	1	2	1	4
16	1	2	2	4
32	2	2	3	5
64	2	2	3	5
128	2	2	3	5
Error	5.42(-11)	2.03(-9)	2.96(-5)	3.35(-5)

Table 2 shows the iteration count for the same cases in the presence of inconsistent preconditioning. For these tests, a global ILU(0) preconditioning was created from the convection-diffusion operator with first-order upwind convection. (The diagonal dominance of this system protects the ILU(0) factorization from breakdown.) The iteration count for the matrix-free method with preconditioning still deteriorates relative to the explicit Jacobian case. However, the explicit preconditioner is created from the explicitly available matrix, which is second-order accurate. The absolute extent of this deterioration is only a couple of extra Jacobian-vector products. In applications, in which one is faced with a choice between many extra stationary defect correction steps at limited CFL versus a few more costly accelerated Newton correction steps at high CFL, the matrix-free Newton method may prevail.

The numerical experiments in this section considered only a global preconditioner based on approximate factorization. In the next two sections, Schwarz-like domain decomposition preconditioners are considered, instead.

5. A Full Potential Problem

The full potential equation for the velocity potential, Φ, is

$$\nabla \cdot (\rho(||\nabla\Phi||)\nabla\Phi) = 0,$$

where the density is given in terms of the potential by

$$\rho = \rho_\infty \left(1 + \frac{\gamma - 1}{2} M_\infty^2 (1 - \frac{q^2}{q_\infty^2})\right)^{\frac{1}{(\gamma-1)}},$$

where $q = ||\nabla\Phi||$ and $M_\infty = q_\infty / a_\infty$. Here, a is the sound speed, q the flow speed, and ∞ refers to the freestream. When the flow is everywhere subsonic the full potential formulation fits within the monotone nonlinear elliptic framework of additive Schwarz

methods [4]. For a simple non-lifting model problem of an airfoil lying along the symmetry axis $y = 0$, we choose boundary conditions as follows:

- Upstream and Freestream: $\Phi = q_\infty x$ (zero angle of attack),
- Downstream: $\Phi_{,n} = q_\infty$,
- Symmetry: $\Phi_{,n} = 0$,
- On the parameterized airfoil with shape $y = f(x)$: $\Phi_{,n} = -q_\infty f'(x)$.

The farfield boundary conditions lead to inaccuracies if applied too near the airfoil, but our interest is in algebraic convergence rates.

TABLE 3

Average number of GMRES steps per Newton step for full potential Newton-Krylov-Schwarz solver with varying coarse grid size.

Coarse Grid	0×0	4×5	8×9	12×13	16×17	20×21
Analytical	177	35	28	27	24	21
Matrix-free	183	41	28	27	25	23

A uniform fine grid of mesh size h and a uniform coarse grid of mesh size H_c are defined over the entire domain. The coarse grid is not necessarily nested within the fine grid, and its elements are not necessarily vertices of the subdomains, which have sides of size H_s, [3]. Bilinear rectilinear elements are used for both coarse and fine grids, and bilinear interpolation for intergrid transfers. An overlap $2h$ is employed on each of the subdomains, and problems posed on the subdomains as part of the Schwarz preconditioning are solved inexactly by ILU(0) with a drop tolerance 0.01. M_∞ is 0.1 and the airfoil is the scaled upper surface of a NACA0012. Nonlinear convergence is declared following a 10^{-3} relative reduction in the steady-state residual, which requires only three Newton steps independent of inner linear method. Inner iteration convergence is a relative residual reduction of 10^{-4}. We restart GMRES every 20 iterations. Table 3 shows convergence performance for a fixed-size problem of 128×128 uniform cells with a fixed number of subdomains in an 8×8 array as the density of the unnested uniform coarse grid varies. Key observations from this example are: (1) even a modest coarse grid makes a significant improvement in an additive Schwarz preconditioner; (2) a law of diminishing returns sets in at roughly one point per subdomain; and (3) matrix-free "matvecs" degrade convergence as much as 15-20% in the less well-conditioned cases.

6. An Euler Problem

Our Euler example is a two-dimensional transonic airfoil flow modeled using an EAGLE-derivative code [14] that employs a finite volume discretization over a body-fitted coordinate grid. The Euler equations for dependent variable vector $Q \equiv [\rho, \rho u, \rho v, e]^T$ are expressed in strong conservation curvilinear coordinate form as

$$\tilde{Q}_\tau + (\tilde{F})_\xi + (\tilde{G})_\eta = 0, \tag{1}$$

where $\tilde{Q}$ and the contravariant flux vectors, $\tilde{F}$ and $\tilde{G}$, are defined in terms of the Cartesian fluxes and the Jacobian determinant of the coordinate system transformation, $J = x_\xi y_\eta -$

$y_\xi x_\eta$, through

$$\begin{aligned}\tilde{Q} &= J^{-1}Q \\ \tilde{F} &= J^{-1}\left(\xi_t Q + \xi_x F + \xi_y G\right) \\ \tilde{G} &= J^{-1}\left(\eta_t Q + \eta_x F + \eta_y G\right).\end{aligned}$$

C-grids of 128×16 or 128×32 cells (from [10]) around a NACA0012 airfoil at an angle of attack of 1.25° and an M_∞ of 0.8 are considered. To obtain a representative matrix/RHS pair on which to test the behavior of Euler Jacobians under Krylov-Schwarz, we first ran a demonstration case from [14] partway to convergence and linearized about the resulting flow state.

The discrete equations take the form

$$\left[I + \Delta\tau(\delta_\xi A^+_\bullet + \delta_\eta B^+_\bullet + \delta_\xi A^-_\bullet + \delta_\eta B^-_\bullet)\right]\Delta Q^l = -\Delta\tau f^l,$$

where

$$A = \frac{\partial \tilde{F}^{\text{impl}}}{\partial \tilde{Q}} \quad \text{and} \quad B = \frac{\partial \tilde{G}^{\text{impl}}}{\partial \tilde{Q}},$$

the eigenvalues of A and B, respectively, are the components of the characteristic velocities in the ξ and η directions, δ is the first-order spatial difference operator, superscripts $\pm$ denote the characteristic (upwind) direction in which the differencing occurs, and the bullets signify that each spatial differencing is carried out on the entire product to the right, for example, δ_ξ on $(A^+\Delta Q^l)$. Following the defect correction practice of [16], a flux vector split scheme of Van Leer type is employed for the implicit operators, A and B, and the vector of steady-state residuals, $f(u)$, is discretized by a Roe-type flux difference split scheme. Characteristic variable boundary conditions are employed at farfield boundaries using an explicit, first-order accurate formulation.

TABLE 4

Iteration counts for transonic Euler flow Jacobian with 8192 degrees of freedom discretized at local CFL numbers of 1 and 100, for various preconditioners and decomposition into 4, 16, or 64 subdomains.

Precond.	Block Jacobi		Add. Schwarz		Mult. Schwarz	
CFL No.	1	10^2	1	10^2	1	10^2
1×1	1	1	1	1	1	1
2×2	4	14	7	14	2	7
4×4	4	18	7	17	3	8
8×8	5	28	10	23	3	8

For a given granularity of decomposition, curvilinear "box" decompositions are generally better than curvilinear "strip" decompositions for this problem [10]. Table 4 shows that the zero-overlap results are only slightly less convergent than the corresponding h-overlapped additive Schwarz results at high Courant-Friedrichs-Lewy (CFL) number, and that h-overlapped multiplicative Schwarz is significantly better, though the latter is a less

TABLE 5

Iteration counts and serial execution time for a transonic Euler flow Jacobian with 16384 degrees of freedom discretized at a local Courant number of 100, for Block Jacobi run in serial on a Sparc10.

Decomp.	Iterations	Serial Time
1×1	1	38.1
2×2	12	37.9
4×4	14	21.0
8×8	22	20.3

parallel algorithm. Though we have not yet experimented with a coarse grid in the Euler context, [20] shows that even a piecewise constant coarse grid operator substantially improves Krylov-Schwarz convergence rates in unstructured problems.

The serial benefit of the Schwarz preconditioning is illustrated in Table 5. Here, a direct solve with a nested dissection ordering is found inferior to an additive Schwarz-GMRES iteration with nested dissection applied to successively smaller subdomains.

The Euler code has been executed on an ethernet network of workstations using a package of distributed sparse linear system routines developed at Argonne National Laboratory by Gropp and Smith [11], with p4 [9] as the data exchange layer.

Table 6 shows CPU times for each of three Schwarz preconditioners with different parallel granularity, for a fixed number of iterations on a fixed size problem. When exact solvers are used on each subdomain, speedups on a per iteration basis are seen on up to 16 processors. This advantage may be artificial in the sense that global incomplete LU is superior to a Schwarz method using exact subdomain solvers in serial, and ILU should replace nested dissection as the subdomain solver. In recent experiments [1] on Jacobians from the TRANAIR code, a threshhold drop tolerance form of ILU was employed as the subdomain solver in a Schwarz preconditioner, and the drop tolerance was varied from zero (exact nested dissection factorization) to a value that filtered out all but the largest entries in the approximate factors. The optimal drop tolerance in terms of number of operations per gridpoint varied with problem size and required converged precision, but the optimal average number of nonzeros per row hovered around 20 to 40, which is much sparser than a full nested dissection.

To test the nonlinear matrix-free approach in a situation with four differently scaled components per gridpoint, we approached the steady solution via a pseudo-transient continuation with a local adaptation of CFL number. Starting from a small initial CFL

TABLE 6

Execution time, maximized over the processor gang, for an equal number (ten) of Block Jacobi preconditioned iterations run in parallel over a network of Sparc10's and SparcELC's, for a transonic Euler flow Jacobian with 16384 degrees of freedom discretized at a local Courant number of 100.

Decomp.	Parallel Time
2×2	20.3
4×2	11.5
4×4	8.3

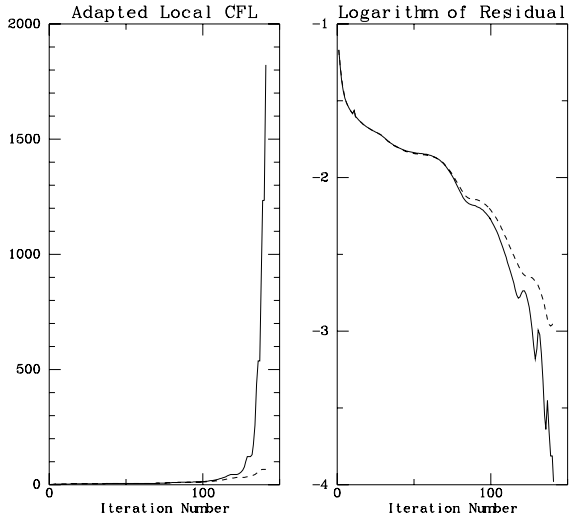

FIG. 2. *CFL and steady-state residual versus iteration count for defect correction and Newton-Krylov solvers.*

number (10), CFL may be adaptively advanced according to:

$$\mathrm{CFL}^{l+1} = \mathrm{CFL}^{l} \cdot \frac{||f(u)||^{l-1}}{||f(u)||^{l}}$$

This was found preferable to an alternative local strategy:

$$\mathrm{CFL}_{ij}^{l+1} = \mathrm{CFL}_{ij}^{l} \cdot \frac{|f_{ij}(u)|^{l-1}}{|f_{ij}(u)|^{l}}$$

and also to higher powers of the ratio of successive norms.

Use of the baseline approximate factorization defect correction algorithm produces the dashed curves in Fig. 2. To obtain the solid curves, the explicitly available (Van Leer) flux vector split Jacobian (J_{VL}) is used to precondition the implicitly defined (Roe) flux difference split Jacobian (J_R) at each implicit time step. In matrix terms, the corrections u are obtained as the approximate solutions of, respectively,

$$J_{VL}u = -f_R \quad \text{and} \quad (J_{VL})^{-1}J_R u = -(J_{VL})^{-1}f_R.$$

Unfortunately, in the retrofit of the existing code, transition to a full Newton method (CFL number approaching infinity) is precluded by explicit boundary conditions, but CFL number can be advanced, as shown in the figure, to $\mathcal{O}(10^3)$ with advantage.

The Schwarzian theory relies on the dominance of the elliptic operator. However, as in this example, multidimensional hyperbolic problems are often discretized in a defect-correction manner, with an artificially diffusive left-hand side operator. Some of the benefits of elliptic algorithms will be realized when these methods are accelerated with a Krylov method.

7. Conclusions

Several aspects of Newton-Krylov-Schwarz methods that are expected to be of value in a practical parallelized Navier-Stokes code are validated in model contexts. Simultaneous demands for higher-order discretizations in the Jacobian proper and diagonally dominant,

easily invertible preconditioners are satisfied by mixed (inconsistent) discretizations. As in defect correction methods, the Jacobian proper need not be computed explicitly. However, by accelerating the stationary defect correction process, arbitrarily good convergence to the true Newton direction can be obtained, providing an asymptotic advantage. It is shown to be possible to find a differencing parameter for the matrix-free Newton method that is simultaneously small enough for good truncation error in approximating the action of the Jacobian and large enough to avoid the pitfalls of double precision round-off, at least for problems of $\mathcal{O}(10^2)$ mesh cells in each dimension, and in the presence of significant grid stretching. This problem is expected to grow more acute for high-aspect ratio Navier-Stokes grids. The use of a non-nested coarse grid provides a continuously adjustable "knob" with significant leverage in reducing the condition number of an elliptically dominated problem. The coarse grid entails a high communication cost in parallel implementations, but in additive algorithms it can be solved simultaneously with the subdomain fine grids operators, permitting overlap of its excess communication with useful subdomain computation. Finally, preliminary experience with treating Unix workstations spread throughout a building and connected only by a single ethernet is enouraging in that even for a modest fixed-size problem, wall clock execution time per iteration improves on up to sixteen workstations. When parallelism is exploited in its more advantageous scaling of fixed subdomain size per processor, we expect yet more encouraging results.

REFERENCES

[1] D. P. Young, C. C. Ashcraft, R. G. Melvin, M. B. Bieterman, W. P. Huffman, T. F. Johnson, C. L. Hilmes, and J. E. Bussoletti, *Ordering and Incomplete Factorization Issues for Matrices Arising from the TRANAIR CFD Code*, Boeing Computer Services TR BCSTECH-93-025, Seattle, 1993.

[2] X.-C. Cai, *An Optimal Two-level Overlapping Domain Decomposition Method for Elliptic Problems in Two and Three Dimensions*, SIAM J. Sci. Comp. (1993) 239–247.

[3] X.-C. Cai, *A Non-nested Coarse Space for Schwarz Type Domain Decomposition Methods*, Tech. Rep. CU-CS-705-94, Dept. of Comp. Sci., Univ. of Colorado at Boulder, 1994.

[4] X.-C. Cai and M. Dryja, *Domain Decomposition Methods for Monotone Nonlinear Elliptic Problems*, in D. E. Keyes and J. Xu, eds., Seventh International Conference on Domain Decomposition Methods in Scientific Computing (Penn. State Univ.), AMS, Providence, 1994, to appear.

[5] X.- C. Cai, W. D. Gropp, and D. E. Keyes, *A Comparison of Some Domain Decomposition and ILU Preconditioned Iterative Methods for Nonsymmetric Elliptic Problems*, J. Numer. Lin. Alg. Applics. (1994), to appear.

[6] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[7] M. Dryja and O. B. Widlund, *An Additive Variant of the Alternating Method for the Case of Many Subregions*, Courant Institute, NYU, TR 339, 1987.

[8] A. Ern, V. Giovangigli, D. E. Keyes and M. D. Smooke, *Towards Polyalgorithmic Linear System Solvers for Nonlinear Elliptic Problems*, SIAM J. Sci. Comp. **15**(1994), to appear.

[9] Ralph Butler and Ewing Lusk, *Monitors, Messages, and Clusters: The p4 Parallel Programming System*, Argonne National Laboratory MCS Div. preprint P362-0493, and J. Parallel Comput., to appear.

[10] W. D. Gropp, D. E. Keyes and J. S. Mounts, *Implicit Domain Decomposition Algorithms for Steady, Compressible Aerodynamics*, in A. Quarteroni et al., eds., Sixth International Symposium on Domain Decomposition Methods for Partial Differential Equations (Como, Italy), AMS, Providence, 1994.

[11] W. D. Gropp and B. F. Smith, *Simplified Linear Equations Solvers Users Manual*, ANL-93/8, Argonne National Laboratory, 1993.

[12] D. E. Keyes and W. D. Gropp, *Domain-Decomposable Preconditioners for Second-order Upwind*

Discretizations of Multicomponent Systems, in R. Glowinski et al., eds., Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations (Moscow), SIAM, Philadelphia, 1991, pp. 129–139.

[13] P. R. McHugh and D. A. Knoll, *Inexact Newton's Method Solutions to the Incompressible Navier-Stokes and Energy Equations Using Standard and Matrix-Free Implementations*, AIAA Paper, 1993.

[14] J. S. Mounts, D. M. Belk and D. L. Whitfield, *Program EAGLE User's Manual, Vol. IV – Multi-block Implicit, Steady-state Euler Code*, Air Force Armament Laboratory TR-88-117, Vol. IV, September 1988.

[15] Y. Saad and M. H. Schultz, *GMRES: A Generalized Minimal Residual Algorithm for Solving Non-symmetric Linear Systems*, SIAM J. Sci. Stat. Comp. **7**(1986), 865–869.

[16] J. L. Steger and R. F. Warming, *Flux Vector Splitting of the Inviscid Gasdynamics Equations with Applications to Finite-Difference Methods*, J. Comp. Phys. **40**(1981), 263–293.

[17] M. D. Tidriri, *PhD thesis*, Univ. of Paris XI, May 1992.

[18] L. N. Trefethen, *Approximation Theory and Numerical Linear Algebra*, in J. C. Mason and M. G. Cox, Algorithms for Approximation, Chapman (1990), pp. 336–360.

[19] V. Venkatakrishnan, *Convergence to Steady State Solutions of the Euler Equations on Unstructured Grids with Limiters*, J. Comp. Phys. (1994), submitted.

[20] V. Venkatakrishnan, *Parallel Implicit Unstructured Grid Euler Solvers*, AIAA Paper 94-0759, Reno, Nevada, January 1994.

[21] P. N. Brown and Y. Saad, *Hybrid Krylov Methods for Nonlinear Systems of Equations*, SIAM J. Sci. Stat. Comp. **11**(1990), 450–481.

[22] B. Van Leer, *Towards the Ultimate Conservative Difference Scheme, II. Monotonicity and conservation combined in a Second Order Scheme*, J. Comp. Phys. **14**(1974), 361–370.