

**NISTIR 7574**

# **Fast Iterative Solver for Convection-Diffusion Systems with Spectral Elements**

P. Aaron Lott  
Howard Elman

**NIST**  
National Institute of  
Standards and Technology  
U.S. Department of Commerce

NISTIR 7574

# Fast Iterative Solver for Convection-Diffusion Systems with Spectral Elements

P. Aaron Lott

*National Institute of Standards and Technology  
Information Technology Laboratory  
Mathematical and Computational Sciences Division*

Howard Elman

*University of Maryland  
College Park, MD*

March 2009



U.S. Department of Commerce  
*Gary Locke, Secretary*

National Institute of Standards and Technology  
*Patrick D. Gallagher, Deputy Director*

# FAST ITERATIVE SOLVER FOR CONVECTION-DIFFUSION SYSTEMS WITH SPECTRAL ELEMENTS

P. AARON LOTT\* AND HOWARD ELMAN<sup>†</sup>

**Abstract.** We introduce a solver and preconditioning technique based on Domain Decomposition and the Fast Diagonalization Method that can be applied to tensor product based discretizations of the steady convection-diffusion equation. The method is based on iterative substructuring where fast diagonalization is used to efficiently eliminate the interior degrees of freedom and subsidiary subdomain solves. We demonstrate the effectiveness of this method in numerical simulations using a spectral element discretization.

**Key words.** Convection-Diffusion, Domain Decomposition, Preconditioning, Spectral Element Method

**1. Introduction.** Numerical simulation of fluid flow allows for improved prediction and design of natural and engineered systems such as those involving water, oil, or blood. The interplay between inertial and viscous forces in a fluid flow dictates the length scale where energy is transferred, thus determining the resolution required to capture flow information accurately. This resolution requirement poses computational challenges in situations where the convective nature of the flow dominates diffusive effects. In such flows, convection and diffusion occur on disparate scales, causing sharp flow features that require fine numerical grid resolution. This leads to a large system of equations which is often solved using an iterative method. Exacerbating the challenge of solving a large linear system, the discrete convection-diffusion operator is non-symmetric and poorly conditioned. This leads to slow convergence of iterative solvers. In total, as convection dominates the flow the discrete fluid model becomes exceedingly challenging to solve.

In recent years, the spectral element method has gained popularity as a technique for numerical simulation of fluids [9],[18]. This is due in part to the method's high-order accuracy, which produces solutions with low dissipation and low dispersion with relatively few degrees of freedom. Also important is the inherent computational efficiency gained through the use of a hierarchical grid structure based on unstructured macro-elements with fine tensor-structured interiors. This structure has enabled the development of efficient multi-level solvers and preconditioners based on Fast Diagonalization and Domain Decomposition [8],[10],[21]. Application of these techniques, however, has been restricted to symmetric systems.

One way to apply such methods to non-symmetric systems is through use of time-splitting techniques, which split the system into symmetric and non-symmetric components. For convection-diffusion systems, the standard method for performing steady and unsteady flow simulations with spectral elements is operator integration factor splitting (OIFS) [12], which requires time integration even in steady flow simulations. Using this standard approach, convection and diffusion are treated separately; convection components are tackled explicitly using a sequence of small time steps that satisfy a CFL (Courant-Friedrichs-Lewy) condition, and diffusive components are treated implicitly with larger time steps via a backward differencing formula that couples the convection terms to the diffusion system. In this framework, the diffusive system is symmetric allowing fast solvers based on Fast Diagonalization and Domain Decomposition to be employed. However, to simulate fast moving flows the discretization must be refined to capture sharp flow features accurately; this in turn

---

\*Mathematical and Computational Sciences Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899; Aaron.Lott@nist.gov. This work was supported in part by a National Research Council Postdoctoral Fellowship.

<sup>†</sup>Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742; elman@cs.umd.edu. This work was supported in part by the U. S. Department of Energy under grant DEFG0204ER25619, and by the U. S. National Science Foundation under grant CCF0726017.

leads to severe constraints on the time-step size of the semi-implicit method. Such methods can become prohibitively expensive when simulating highly convective flows over long time periods.

An alternative approach to simulating convective flows is to perform implicit time integration, or to solve the steady state system directly. Such methods, however, require fast solvers that are able to resolve the disparate convective and diffusive scales efficiently. The steady convection-diffusion equation for the 2D transport of a scalar  $u(x, y)$  can be written as

$$-\varepsilon \nabla^2 u + (\vec{w} \cdot \nabla) u = f \text{ in } \Omega \quad (1.1)$$

where the vector field  $\vec{w} = (w_x(x, y), w_y(x, y))^T$  represents the given wind speed at each point in the domain  $\Omega$ ,  $f$  represents a source term, and  $\varepsilon$  represents the diffusivity. In addition to (1.1), we have the associated boundary conditions

$$u = u_D \text{ on } \partial\Omega_D, \quad \nabla u_i \cdot \vec{n} = 0 \text{ on } \partial\Omega_N, \quad (1.2)$$

where  $\vec{n}$  is the outward facing normal on the boundary, and subscripts  $D$  and  $N$  denote Dirichlet and Neumann boundary regions respectively. To measure the relative contributions of convection and diffusion, equation (1.1) can be non-dimensionalized by emphasizing the inertial terms through characteristic velocity and length scales,  $W$  and  $L$  respectively. That is, points in  $\Omega$  can be normalized by dividing by a characteristic length  $L$ , the source term is non-dimensionalized by taking  $f^* := fL/W^2$ , and the scalar  $u^*$  and wind  $w^*$  are non-dimensionalized by dividing by  $W$ . This leads to the equation in the normalized domain,

$$-\frac{1}{Pe} \nabla^{*2} u^* + (\vec{w}^* \cdot \nabla^*) u^* = f^*. \quad (1.3)$$

The quantity  $Pe := \frac{WL}{\varepsilon}$  is termed the *Peclet number*. This dimensionless number quantifies the contributions of convection and diffusion for a given flow. In diffusion-dominated flows,  $0 \leq Pe \leq O(1)$ , whereas in convection-dominated flows  $Pe \gg 1$ . We see in (1.3) that as  $Pe \rightarrow \infty$  the flow becomes dominated by convection and the diffusion term vanishes, leading to a hyperbolic system. Solutions to these nearly hyperbolic systems often exhibit sharp gradients in order to satisfy boundary conditions, thus requiring a fine numerical grid to resolve the flow accurately. We also observe that as  $Pe \rightarrow 0$ , (1.3) does not produce the Poisson equation, since the non-dimensionalization emphasizes the inertial terms. A similar non-dimensional form of the convection-diffusion equation can be written for slow moving flows by accentuating the viscous forces. In this form the source term is replaced with  $f^* = \frac{fL}{W\varepsilon}$ , whereas the other quantities are normalized as above, leading to the dimensionless form

$$-\nabla^{*2} u^* + Pe(\vec{w}^* \cdot \nabla^*) u^* = f^*. \quad (1.4)$$

Now, as  $Pe \rightarrow 0$  one is left with the Poisson equation to model the flow.

In this article, we introduce a new approach for simulating flows with spectral elements by developing efficient solvers for the steady convection-diffusion system (1.1). We emphasize the case where  $Pe$  is large, although these techniques apply to diffusion-dominated flows as well. Similar ideas are explored in [13] for stabilized spectral element discretizations via overlapping domain decomposition. Our method is based on non-overlapping domain decomposition and takes advantage of Fast Diagonalization to eliminate degrees of freedom in element interiors. The efficiency of our solution method is centered on two advancements: the first is the use of an accurate high-order matrix-free discretization to construct accurate discrete solutions while minimizing memory requirements; the second is the use of fast iterative solvers, which are accelerated by domain decomposition based preconditioners that take advantage of the local tensor product structure by exploiting Fast Diagonalization.

**2. Spectral Element Method applied to the Convection-Diffusion Equation.** The spectral element method is a numerical method for discretizing differential equations that uses a finite polynomial basis to represent the solution on a set of non-overlapping subdomains. The technique is a Galerkin method derived from the method of weighed residuals, in which a weak form equation is solved. The computational domain is tessellated by subdomains  $\Omega_e$  called elements, and the associated integrals are divided into a sum of integrals on individual elements which are then approximated by numerical quadrature. In one dimension, an orthogonal nodal spectral basis  $\pi^N$  is constructed in each element as a Lagrange interpolation polynomial based on a set of Gauss-Legendre-Lobatto (GLL) nodes,  $(\Xi_{N+1} := \xi_1, \xi_2, \dots, \xi_{N+1})$ . In higher dimensions, the basis is formed as a tensor product of these one-dimensional basis functions. This allows functions defined on  $\Omega_e$  to be written in terms of their spectral basis on each element. Using this nodal basis representation and applying numerical quadrature results in a system of linear matrix equations that numerically represents the original integral equation on each element. Solving this system allows one to obtain the solution at the GLL interpolation points. Inter-element coupling of these equations ensures continuity along elemental boundaries. These inter-element couplings can be enforced either by constructing a fully coupled sparse linear system of equations, or by performing a gather-scatter operation that sums the solution along element boundaries after element-based matrix-vector products are performed. As described in [4], this gather-scatter operation, called “direct-stiffness-summation,” coupled with the tensor product formulation of elemental operators, yields a matrix-free discretization, in which only local matrices associated with one-dimensional phenomena need to be stored.

The weak formulation of (1.1) is: Find  $u \in X$  such that:

$$\varepsilon \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} (\vec{w} \cdot \nabla u) v = \int_{\Omega} f v \quad \forall v \in X_0, \quad (2.1)$$

where the continuous solution and test spaces are defined as

$$X := \{u \in H^1(\Omega)^2 \mid u = u_D \text{ on } \partial\Omega_D\} \quad (2.2)$$

$$X_0 := \{u \in H^1(\Omega)^2 \mid u = 0 \text{ on } \partial\Omega_D\}, \quad (2.3)$$

which are based on  $L^2(\Omega)$  - the space of all square integrable functions on  $\Omega$ .  $H^1(\Omega)$  is the space of all functions in  $L^2(\Omega)$  with first derivatives also in  $L^2(\Omega)$ . Existence and uniqueness

---

**Definition 1** Coercivity

A bilinear form  $a(\cdot, \cdot)$  is *coercive* with respect to the norm  $\|\cdot\|_V$  in a space  $V$  if there is a positive constant  $c$  such that  $a(u, u) \geq c\|u\|_V^2 \quad \forall u \in V$  (see [7] p. 121).

---



---

**Definition 2** Continuity

A bilinear form  $a(\cdot, \cdot)$  is *continuous* with respect to the norm  $\|\cdot\|_W$  in a space  $W$  if there is a positive constant  $C$  such that  $|a(u, v)| \leq c\|u\|_W\|v\|_W \quad \forall u, v \in W$  (see [7] p. 121).

---

of a corresponding weak solution can be demonstrated by establishing the coercivity and continuity of the bilinear form

$$a(u, v) := \varepsilon \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} (\vec{w} \cdot \nabla u) v. \quad (2.4)$$

For Dirichlet problems, coercivity can be established by observing that the convection term is skew-self-adjoint when  $\nabla \cdot \vec{w} = 0$  and  $\vec{w} \cdot \vec{n} = 0$  along the outflow, so  $a(u, u) = \varepsilon \|\nabla u\|^2$  for all admissible  $u$ , and continuity is verified over  $X_0$  using

$$|a(u, v)| \leq \varepsilon \left| \int_{\Omega} \nabla u \cdot \nabla v \right| + \left| \int_{\Omega} (\vec{w} \cdot \nabla u) v \right|, \quad (2.5)$$

and bounding the terms using the Cauchy-Schwarz inequality. We refer the reader to [3] for details.

Using the spectral element method, we construct a finite-dimensional basis for  $X$  and  $X_0$  by dividing the domain  $\Omega$  into  $E$  non-overlapping sub-domains (elements)  $\Omega = \cup_{e=1}^E \Omega_e$ ; each sub-domain is then discretized using tensor products of Lagrangian interpolants on degree  $N$  Legendre polynomials  $\pi^N$ . The corresponding discrete approximation space for the solution and test functions is defined as

$$X^N = X \cap \mathbb{P}_{N,E}^2(\Omega), \quad X_0^N = X_0 \cap \mathbb{P}_{N,E}^2(\Omega) \quad (2.6)$$

where

$$\mathbb{P}_{N,E}^2(\Omega) = \{v(x^e(r))|_{\Omega^e} \in \mathbb{P}_N(r_1) \otimes \mathbb{P}_N(r_2), e = 1, \dots, E\} \quad (2.7)$$

and  $\mathbb{P}_N(r)$  is the space of Lagrangian interpolants of degree less than or equal to  $N$  defined on the Gauss-Legendre-Lobatto points on the reference domain  $[-1, 1]$ .

The discrete weak form is: Find  $u \in X^N$  such that:

$$\varepsilon \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} (\vec{w} \cdot \nabla u) v = \int_{\Omega} f v \quad \forall v \in X_0^N. \quad (2.8)$$

The solution is then written as a sum of a known function  $u_D$ , that satisfies the Dirichlet boundary conditions at the nodes corresponding to points on the Dirichlet boundary and is zero otherwise, and an unknown function  $u_0$  that is zero on the Dirichlet boundaries,  $u = u_0 + u_D$ . Moving the known quantities to the right-hand side we have:

Find  $u \in X_0^N$  such that:

$$\varepsilon \int_{\Omega} \nabla u_0 \cdot \nabla v + \int_{\Omega} (\vec{w} \cdot \nabla u_0) v = \int_{\Omega} f v - \varepsilon \int_{\Omega} \nabla u_D \cdot \nabla v - \int_{\Omega} (\vec{w} \cdot \nabla u_D) v \quad \forall v \in X_0^N. \quad (2.9)$$

On each element the solution is of the form

$$u_0^e(x, y) = \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} u_{ij} \pi_i(x) \pi_j(y), \quad (2.10)$$

and after obtaining  $u_0^e$  on each element, this solution can be combined with  $u_D$  to form the discrete solution satisfying (1.1) and (1.2). The coefficients  $u_{ij}$  correspond to the nodal values of  $u$  on the tensored Gauss-Legendre-Lobatto (GLL) points. After inserting  $u$  and  $v$  represented via the bases indexed by  $ij$  and  $\hat{i}\hat{j}$  respectively into (2.9) we obtain the system of equations

$$\underbrace{\mathbb{M}^e \Sigma' F^e(\vec{w}^e)}_F u_0^e = \underbrace{\mathbb{M}^e \Sigma' (M^e f^e - F^e(\vec{w}) u_D^e)}_b. \quad (2.11)$$

Here, a mask matrix  $\mathbb{M}^e$  sets  $u_0^e$  to zero at all Dirichlet boundary nodes. The convection-diffusion operator  $F^e$  is represented on each element with dimensions  $h_x \times h_y$  through this

basis as

$$\begin{aligned}
F_x^e &= \varepsilon \frac{h_y}{h_x} \underbrace{\left( \int_{-1}^1 \pi_j \pi_{\hat{j}} \right)}_{\hat{M}_{j\hat{j}}} \underbrace{\left( \int_{-1}^1 \pi'_i \pi'_{\hat{i}} \right)}_{\hat{A}_{i\hat{i}}} + w_x^e \frac{h_y}{2} \underbrace{\left( \int_{-1}^1 \pi_j \pi_{\hat{j}} \right)}_{\hat{M}_{j\hat{j}}} \underbrace{\left( \int_{-1}^1 \pi'_i \pi'_{\hat{i}} \right)}_{\hat{C}_{i\hat{i}}} = \underbrace{\varepsilon (\hat{M} \otimes \frac{h_y}{h_x} \hat{A})}_{\text{Diffusion in x}} + \underbrace{W_x^e (\hat{M} \otimes \frac{h_y}{2} \hat{C})}_{\text{Convection in x}} \\
F_y^e &= \varepsilon \frac{h_x}{h_y} \underbrace{\left( \int_{-1}^1 \pi'_j \pi'_{\hat{j}} \right)}_{\hat{A}_{j\hat{j}}} \underbrace{\left( \int_{-1}^1 \pi_i \pi_{\hat{i}} \right)}_{\hat{M}_{i\hat{i}}} + w_y^e \frac{h_x}{2} \underbrace{\left( \int_{-1}^1 \pi'_j \pi'_{\hat{j}} \right)}_{\hat{C}_{j\hat{j}}} \underbrace{\left( \int_{-1}^1 \pi_i \pi_{\hat{i}} \right)}_{\hat{M}_{i\hat{i}}} = \underbrace{\varepsilon \left( \frac{h_x}{h_y} \hat{A} \otimes \hat{M} \right)}_{\text{Diffusion in y}} + \underbrace{W_y^e \left( \frac{h_x}{2} \hat{C} \otimes \hat{M} \right)}_{\text{Convection in y}} \\
F^e(w^e) &= F_x^e + F_y^e.
\end{aligned} \tag{2.12}$$

The discrete two-dimensional diffusion operator is formed via tensor products of the one-dimensional second derivative operator  $\hat{A}$  with the one-dimensional mass matrix  $\hat{M}$ . Similarly, the discrete convection operator is formed via a tensor product of the weak one-dimensional derivative operator  $\hat{C}$  with the mass matrix  $\hat{M}$ , then scaled by the wind speed at each node via the  $(N+1)^2 \times (N+1)^2$  diagonal matrices  $W_x^e = \text{diag}(w_x(\xi_i, \xi_j))$  and  $W_y^e = \text{diag}(w_y(\xi_i, \xi_j))$ . The mass matrix,  $M = \text{diag}(M_e)$  is composed of local mass matrices,  $M^e$ , represented via a tensor product of one-dimensional operators,  $\hat{M}$ , namely:

$$M^e = \frac{h_x h_y}{4} \underbrace{\left( \int_{-1}^1 \pi_i \pi_{\hat{i}} \right)}_{\hat{M}_{i\hat{i}}} \underbrace{\left( \int_{-1}^1 \pi_j \pi_{\hat{j}} \right)}_{\hat{M}_{j\hat{j}}} = \frac{h_x h_y}{4} \hat{M} \otimes \hat{M}. \tag{2.13}$$

Due to the orthogonality of the basis functions  $\pi_i$  and  $\pi_j$ , the mass matrix is a diagonal matrix.

The global system matrices can be defined by assembling the elemental matrices  $M^e$  and  $F^e$  across multiple elements via a boolean connectivity matrix  $Q$  that maps global nodal values to local nodal values; that is  $F(w) = \mathbb{M} Q^T F^e Q \mathbb{M}$ , is the global non-symmetric discrete convection-diffusion operator, and  $M = \mathbb{M} Q^T M^e Q \mathbb{M}$  the global diagonal mass matrix. However, in practice, the local form given in (2.11) is more convenient computationally since tensor-product operations can be readily applied without a scatter operation first being applied. In this format the direct-stiffness-summation operator  $\Sigma' = Q Q^T$  is applied to sum all elemental contributions. Thus (2.11) is a matrix-free formulation in the sense that no global matrix is assembled, allowing the discrete system to be written in terms of element-defined operators such that on each element several discrete operators are applied. We refer the reader to [4] for a detailed description of this formulation.

The system in (2.11) can be solved using an iterative scheme, such as preconditioned GMRES (Generalized Minimal Residual Method) [16]. In the case of constant wind, fast direct solvers in conjunction with domain decomposition with iteration for subsidiary problems are available as a more efficient solver. In the next section, we explain how to solve constant coefficient problems by this technique, and provide empirical results in section 4. Then, in section 5 we discuss how this domain decomposition scheme can be used as a preconditioner to accelerate convergence convection-diffusion problems with non-constant wind.

**3. Domain Decomposition via Iterative Substructuring.** In this section we outline the use of a matrix-free domain decomposition method based on iterative substructuring to solve the discrete convection-diffusion problem with constant coefficients. The solution method obtains the discrete solution on a set of non-overlapping subdomains by first solving for unknowns on inter-element interfaces, and then performing back substitution to compute the interior degrees of freedom. For large systems, the solution on elemental interfaces is obtained by an iterative method. The system that governs the elemental interfaces may be poorly

conditioned, and thus requires preconditioning. We describe here a domain decomposition method that we use in conjunction with a Robin-Robin preconditioner for the interface solve. In addition we present a generalization of the Fast Diagonalization Method for obtaining interior degrees of freedom.

**3.1. Algorithm Overview.** As discussed in the previous section, the domain is subdivided into  $E$  spectral elements. Let  $\Gamma$  denote the degrees of freedom contained on the union of elemental interfaces, and let the sets  $I^e$  denote the degrees of freedom belonging to the interior of the  $e^{\text{th}}$  element. The discrete system (2.11) can be written in the form

$$\begin{bmatrix} \bar{F}_{II}^1 & 0 & \dots & 0 & \bar{F}_{I\Gamma}^1 \\ 0 & \bar{F}_{II}^2 & 0 & \dots & \bar{F}_{I\Gamma}^2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{F}_{II}^E & \bar{F}_{I\Gamma}^E \\ \bar{F}_{\Gamma I}^1 & \bar{F}_{\Gamma I}^2 & \dots & \bar{F}_{\Gamma I}^E & \bar{F}_{\Gamma\Gamma} \end{bmatrix} \begin{pmatrix} u_{I^1} \\ u_{I^2} \\ \vdots \\ u_{I^E} \\ u_{\Gamma} \end{pmatrix} = \begin{pmatrix} b_{I^1} - \bar{F}u_D|_{I^1} \\ b_{I^2} - \bar{F}u_D|_{I^2} \\ \vdots \\ b_{I^E} - \bar{F}u_D|_{I^E} \\ b_{\Gamma} - \bar{F}u_D|_{\Gamma} \end{pmatrix} = \begin{pmatrix} \hat{b}_{I^1} \\ \hat{b}_{I^2} \\ \vdots \\ \hat{b}_{I^E} \\ \hat{b}_{\Gamma} \end{pmatrix}. \quad (3.1)$$

Note that the unknowns here are ordered by block in a specific way: those associated with the interiors of an individual element are listed first, followed in the end by those lying on element interfaces. The term  $\bar{F}_{\Gamma\Gamma}$  is the sum of the individual element operators  $\bar{F}_{I\Gamma}^e$  defined on each element. We denote the discrete constant coefficient convection-diffusion operator by  $\bar{F}$  because in section 5 we will use this operator as a preconditioner for variable wind convection-diffusion problems.

As we mentioned in the last section, Dirichlet boundary conditions are implemented outside of the system operator in (3.1), by subtracting  $\bar{F}u_D$  from the right hand side vector.  $\bar{F}u_D$  denotes the full convection-diffusion system matrix applied to the Dirichlet boundary vector  $u_D$ .

The goal is to solve for interface values  $u_{\Gamma}$  and then to perform back substitution and solve for  $u_{I^e}$  on each sub-domain interior. The solution algorithm is based on a formal factorization of (3.1) into the product of lower and upper block matrices

$$\begin{bmatrix} I & 0 & \dots & 0 & 0 \\ 0 & I & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & I & 0 \\ \bar{F}_{\Gamma I}^1 \bar{F}_{II}^{1-1} & \bar{F}_{\Gamma I}^2 \bar{F}_{II}^{2-1} & \dots & \bar{F}_{\Gamma I}^E \bar{F}_{II}^{E-1} & I \end{bmatrix} \begin{bmatrix} \bar{F}_{II}^1 & 0 & \dots & 0 & \bar{F}_{I\Gamma}^1 \\ 0 & \bar{F}_{II}^2 & 0 & \dots & \bar{F}_{I\Gamma}^2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{F}_{II}^E & \bar{F}_{I\Gamma}^E \\ 0 & 0 & \dots & 0 & \bar{F}_S \end{bmatrix} \quad (3.2)$$

where

$$\bar{F}_S = \sum_{e=1}^E \bar{F}_S^e = \sum_{e=1}^E (\bar{F}_{\Gamma\Gamma}^e - \bar{F}_{\Gamma I}^e \bar{F}_{II}^{e-1} \bar{F}_{I\Gamma}^e) \quad (3.3)$$

represents the Schur complement of the system. By multiplying both sides of (3.1) with the inverse of the lower triangular matrix, one obtains the system

$$\begin{bmatrix} \bar{F}_{II}^1 & 0 & \dots & 0 & \bar{F}_{I\Gamma}^1 \\ 0 & \bar{F}_{II}^2 & 0 & \dots & \bar{F}_{I\Gamma}^2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{F}_{II}^E & \bar{F}_{I\Gamma}^E \\ 0 & 0 & \dots & 0 & \bar{F}_S \end{bmatrix} \begin{pmatrix} u_{I^1} \\ u_{I^2} \\ \vdots \\ u_{I^E} \\ u_{\Gamma} \end{pmatrix} = \begin{pmatrix} \hat{b}_{I^1} \\ \hat{b}_{I^2} \\ \vdots \\ \hat{b}_{I^E} \\ g_{\Gamma} \end{pmatrix} \quad (3.4)$$



with

$$g_\Gamma = \sum_{e=1}^E (\hat{b}_{\Gamma^e} - \bar{F}_{\Gamma I}^e \bar{F}_{II}^{e-1} \hat{b}_{I^e}). \quad (3.5)$$

The interface variables are then obtained by solving

$$\bar{F}_S u_\Gamma = g_\Gamma, \quad (3.6)$$

and interior variables are obtained by solving

$$\bar{F}_{II}^e u_{I^e} = \hat{b}_{I^e} - \bar{F}_{\Gamma I}^e u_\Gamma \quad (3.7)$$

on each element. This involves a three-step procedure:

- Perform E subdomain solves to apply the action of  $\bar{F}_{II}^e$  giving  $g_\Gamma$ , in (3.5)
- Perform interface solve for  $\bar{F}_S u_\Gamma = g_\Gamma$ , as in (3.6)
- Perform E subdomain solves to apply the action of  $\bar{F}_{II}^e$  yielding  $u_{I^e}$ , as in (3.7).

The Schur complement operator,  $\bar{F}_S$ , can be constructed element-by-element by writing  $\bar{F}_S = \sum \bar{F}_S^e$ . This allows for efficient tensor-product based (see appendix (7)) computation of the elemental matrix-vector products, which can be used to apply the matrix on each element inside an iterative solver. Once  $u_\Gamma$  is obtained it is substituted back into (3.4) to provide elemental boundary conditions for the interior solves. Note that the subdomain solves required for step 1 (to compute  $g_\Gamma$ ) and step 3 (to compute  $u_{I^e}$ ) entail the same operations, a set of independent solves for applying the action of  $\bar{F}_{II}^{e-1}$  on each element  $e$ . Moreover, the same operation is required to apply the elemental matrix-vector products in the iteration used for step 2. Each of these operations is performed using the element-wise Fast Diagonalization Method (FDM), which we discuss in section 3.3.

**3.2. Interface Solve.** To compute the interface variables in (3.6) we use preconditioned GMRES. Our choice of preconditioner is the Robin-Robin preconditioner developed in [1], which extends the popular Neumann-Neumann preconditioner used to accelerate the convergence of interface solvers corresponding to the Poisson equation, [14], [17], [20] to non-symmetric systems. This technique uses a pseudo-inverse of the locally defined Schur complement operator, with Robin and Neumann boundary conditions applied to elemental interfaces. The preconditioning matrix is given by

$$\sum_{e=1}^E D^{(e)} R_e^T (\bar{F}_S^e)^{-1} R_e D^{(e)}, \quad (3.8)$$

where  $\bar{F}_S^e$  from equation (3.3) is a modified Schur complement operator on the  $e^{th}$  element, the matrices  $D^e$  are chosen to provide an appropriate inverse scaling factor, and  $R_e$  restricts a vector to the  $e^{th}$  element. The number of iterations required for the preconditioned system to converge is bounded by  $\frac{C}{H}(1 + \log(N))^2$  ([20], p. 321) where  $N$  is the order of the spectral element basis functions,  $H$  is the diameter of a typical element  $\Omega_e$ , and the constant  $C$  depends on the Peclet number. One may gain insight into this preconditioner by considering a simple two-domain case. In the case of two subdomains, the Schur complement of equation (3.3) is  $\bar{F}_S = \bar{F}_S^1 + \bar{F}_S^2$ . If the two domains are of equal size, and the wind is constant on the domain such that the matrices  $\bar{F}_S^1 \approx \bar{F}_S^2$  are of similar character, then the preconditioned system is approximately a scaled identity

$$\begin{aligned} (\bar{F}_S^{1-1} + \bar{F}_S^{2-1})(\bar{F}_S^1 + \bar{F}_S^2) &= 2I + \bar{F}_S^{2-1} \bar{F}_S^1 + \bar{F}_S^{1-1} \bar{F}_S^2 \\ &= 2I + \bar{F}_S^{2-1} \bar{F}_S^1 + (\bar{F}_S^{2-1} \bar{F}_S^1)^{-1} \end{aligned} \quad (3.9)$$

As shown in [17], when applying the preconditioner, it is not necessary to form  $\bar{F}_S^{e-1}$  on each element, since the matrix

$$\bar{F}^e = \begin{bmatrix} \bar{F}_{II}^e & \bar{F}_{I\Gamma}^e \\ \bar{F}_{\Gamma I}^e & \bar{F}_{\Gamma\Gamma}^e \end{bmatrix}. \quad (3.10)$$

can be factored as

$$\bar{F}^e = \begin{pmatrix} I & 0 \\ \bar{F}_{\Gamma I}^e & \bar{F}_{II}^{e-1} \end{pmatrix} \begin{pmatrix} \bar{F}_{II}^e & 0 \\ 0 & \bar{F}_S^e \end{pmatrix} \begin{pmatrix} I & \bar{F}_{II}^{e-1} \bar{F}_{I\Gamma}^e \\ 0 & I \end{pmatrix}, \quad (3.11)$$

which means that

$$\bar{F}^{e-1} = \begin{pmatrix} I & -\bar{F}_{II}^{e-1} \bar{F}_{I\Gamma}^e \\ 0 & I \end{pmatrix} \begin{pmatrix} \bar{F}_{II}^{e-1} & 0 \\ 0 & \bar{F}_S^{e-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -\bar{F}_{\Gamma I}^e & \bar{F}_{II}^{e-1} \end{pmatrix}. \quad (3.12)$$

Thus  $\bar{F}_S^e$  can be applied to a vector by applying the inverse of  $\bar{F}^e$  to the vector restricted to the elemental boundary, and then restricting the resulting vector to the elemental boundary, namely

$$\bar{F}_S^{e-1} x_\Gamma = (0 \quad I) \bar{F}^{e-1} \begin{pmatrix} 0 \\ x_\Gamma \end{pmatrix}. \quad (3.13)$$

Note that with this methodology, the matrices  $\bar{F}^e$  correspond to discretizations of the operator with Neumann boundary conditions applied on elemental boundaries. The Neumann boundary conditions applied at elemental boundaries produces a preconditioning operator that corresponds to the bilinear form

$$a_e(u, v) = \int_{\Omega_e} (\varepsilon \nabla u \cdot \nabla v + (\vec{w} \cdot \nabla u) v), \quad (3.14)$$

which can be derived from the element-based Neumann problem

$$-\varepsilon \nabla^2 u + (\vec{w} \cdot \nabla) u = f \quad \text{in } \Omega_e, \quad (3.15)$$

$$-\varepsilon \frac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma_e. \quad (3.16)$$

Using  $\bar{F}_S^e$  as a preconditioner corresponds to the Neumann-Neumann domain decomposition method which is commonly used in solving Poisson's equation. Because of the Neumann boundary conditions in the elemental operators,  $\bar{F}_{II}^e$  contains a zero eigenvalue and a pseudo-inverse operation can be performed via Fast Diagonalization.

However, as convection becomes dominant, these Neumann element interface conditions (3.16) cause the bilinear form (3.14) to lose coercivity [14], thus rendering the Neumann-Neumann method ineffective. Achdou et al. [1] adapted the Neumann-Neumann preconditioner to non-symmetric convection-diffusion systems by choosing interface boundary conditions that reflect the movement of the flow across element boundaries. In particular, Robin boundary conditions are used instead of Neumann conditions at inflow boundaries where  $\vec{w} \cdot n < 0$ ; Neumann boundary conditions are still imposed at outflows where  $\vec{w} \cdot n > 0$ . This strategy is known as a Robin-Robin preconditioner. It ensures that the symmetric part of the preconditioning operator is positive-definite (see [1] & [20]), since the underlying bilinear form,

$$a_e(u, v) = \int_{\Omega_e} (\varepsilon \nabla u \cdot \nabla v + (\vec{w} \cdot \nabla u) v) - \int_{\Gamma_i} (\vec{w} \cdot \vec{n}) uv, \quad (3.17)$$

derived from the element-based Robin problem

$$-\varepsilon \nabla^2 u + (\vec{w} \cdot \nabla) u = f \quad \text{in } \Omega_e, \quad (3.18)$$

$$-\varepsilon \frac{\partial u}{\partial n} + (\vec{w} \cdot \vec{n}) u = 0 \quad \text{on } \Gamma_e, \quad (3.19)$$

is coercive.

The difference between Neumann-Neumann and Robin-Robin preconditioners is the additional  $(\vec{w} \cdot \vec{n})u$  term at the element interfaces; thus the two preconditioners are equivalent when  $\vec{w} = 0$ . This added term allows the flow to move between elements, and ensures that the preconditioned system matrix is positive definite, as described in [1] and [20]. In practice, the elemental boundary conditions are determined by considering the sign of the convection term at each element boundary. This involves modifying the one-dimensional convection-diffusion operators  $\hat{F}_x$  and  $\hat{F}_y$  from equation (2.12) by a single entry corresponding to the inflow condition

$$\begin{aligned} \hat{F}_{xRR}(1,1) &= \hat{F}_x(1,1) + w_x \hat{M}(1,1) & \text{if } w_x > 0 \\ \hat{F}_{xRR}(N+1,N+1) &= \hat{F}_x(N+1,N+1) - w_x \hat{M}(N+1,N+1) & \text{if } w_x < 0 \\ \hat{F}_{yRR}(1,1) &= \hat{F}_y(1,1) + w_y \hat{M}(1,1) & \text{if } w_y > 0 \\ \hat{F}_{yRR}(N+1,N+1) &= \hat{F}_y(N+1,N+1) - w_y \hat{M}(N+1,N+1) & \text{if } w_y < 0. \end{aligned}$$

In section 4 we provide empirical results comparing the Neumann-Neumann and Robin-Robin preconditioning schemes using our solution method. We note that in both cases the operation is performed using Fast Diagonalization, which we introduce in the following subsection.

**3.3. Fast Diagonalization Method (FDM).** Throughout this section we have seen that the domain decomposition strategy requires interior solves during calculations involving  $\bar{F}_S$  and  $g_\Gamma$  as well as in solving (3.4), and computing the pseudo-inverse (3.13) in the application of the interface preconditioner. In this section we describe an efficient technique known as the Fast Diagonalization Method (FDM) for performing these solves. The Fast Diagonalization Method was originally constructed to solve problems arising from tensor-product based finite difference discretizations of constant coefficient partial differential equations. This method only depends on the inverses of diagonal matrices, and of small matrices corresponding to one-dimensional phenomena. We exploit this in the application of our solvers and preconditioners as discussed in this section. In particular, we show how to generalize the FDM described in [11], to non-symmetric discrete convection-diffusion systems with constant wind  $\vec{w}$ . We then examine the use of FDM applied to (2.11) on a single element.

The spectral element discretization enables the convection-diffusion equation to be written as sums of tensor products on each element. This form is particularly useful when performing matrix-vector products, and when solving certain elemental systems of equations. For systems in which the coefficient matrix is of order  $n^d$  and has a tensor product structure, where  $d$  represents the number of spatial dimensions and  $n$  represents the number of grid points used along each dimension on a single element, the FDM enables the solution to be computed in  $O(n^{d+1})$  operations.

Consider equation (2.12) in the special case where  $W_x^e = c_x$  and  $W_y^e = c_y$  are both constant on each element. In this special case, the convection-diffusion operator on each element can be written as

$$F^e(c_x, c_y) = \hat{M} \otimes \hat{F}_x + \hat{F}_y \otimes \hat{M} =: \bar{F}^e. \quad (3.20)$$

We use the fact that  $\hat{M}$  is diagonal to apply a transformation to  $\bar{F}^e$  that will allow for Fast Diagonalization. That is, we can write  $\bar{F}^e = M^{1/2} \tilde{F}^e M^{1/2}$  where  $M = \hat{M} \otimes \hat{M}$ , and

$$\begin{aligned} \tilde{F}^e &= M^{-1/2} \bar{F}^e M^{-1/2} \\ &= (\hat{M}^{-1/2} \otimes \hat{M}^{-1/2})(\hat{M} \otimes \hat{F}_x + \hat{F}_y \otimes \hat{M})(\hat{M}^{-1/2} \otimes \hat{M}^{-1/2}) \\ &= (I \otimes \hat{M}^{-1/2} \hat{F}_x \hat{M}^{-1/2}) + (\hat{M}^{-1/2} \hat{F}_y \hat{M}^{-1/2} \otimes I) \\ &= (I \otimes B) + (A \otimes I). \end{aligned} \quad (3.21)$$

Assuming both  $A$  and  $B$  are diagonalizable, we have  $A = S\Lambda_y S^{-1}$ ,  $B = T\Lambda_x T^{-1}$ . This gives

$$\tilde{F}^e = (S \otimes T)(\Lambda_y \otimes I + I \otimes \Lambda_x)(S^{-1} \otimes T^{-1}) \quad (3.22)$$

so that

$$\tilde{F}^{e-1} = (S \otimes T)(\Lambda_y \otimes I + I \otimes \Lambda_x)^{-1}(S^{-1} \otimes T^{-1}). \quad (3.23)$$

That is, the transformed matrix  $\tilde{F}$  can be diagonalized cheaply and the action of the inverse of  $\bar{F}^e$  can also be inexpensively applied as

$$\bar{F}^{e-1} = (\hat{M}^{-1/2} \otimes \hat{M}^{-1/2})(S \otimes T)(\Lambda_y \otimes I + I \otimes \Lambda_x)^{-1}(S^{-1} \otimes T^{-1})(\hat{M}^{-1/2} \otimes \hat{M}^{-1/2}). \quad (3.24)$$

We use this method to apply the action of  $\bar{F}_H^{e-1}$  for each element in the domain decomposition method described earlier in this section. In flows where the wind coefficient  $\vec{w}$  is constant, the resulting algorithm defined by (3.5)-(3.7) can be viewed as a direct solver for equation (2.11), although it should be noted that in practice, an iteration is needed to obtain the solution to the Schur complement system (3.6) on the union of element interfaces. We demonstrate the use of this methodology for solving constant coefficient problems in two examples in the next section. We then describe in section 5 how this method can be used to accelerate convergence of GMRES for computing solutions of more general flows. Included in this discussion is the impact and convergence properties of the inner iteration for the Schur complement system (3.6).

**4. Solution of Problems with Constant Wind.** In this section we demonstrate the effectiveness of our solvers applied to two test problems with constant coefficients.

**4.1. Closed-Form solution with outflow boundary layer.** We consider a problem with a closed-form solution; this problem has a zero source term and a constant wind in the vertical direction. The solution is

$$u(x, y) = x \left( \frac{1 - e^{(y-1)/\varepsilon}}{1 - e^{-2/\varepsilon}} \right). \quad (4.1)$$

Plots of the computed solution and its contours are displayed in Figure 4.1. These were obtained using a spectral element discretization with 2 elements in each dimension, and polynomial degree 16 on each element. The solution exhibits dramatic change near the outflow boundary  $y = 1$ ; the width of this exponential boundary layer is proportional to  $\varepsilon$  (see [5] & [15]), so as the Peclet number increases the width of this boundary layer narrows.

We test the accuracy of our numerical scheme by comparing with the analytic solution and record the results in Tables 4.1 and 4.2. In the first table, a  $2 \times 2$  element grid is fixed as the polynomial degree on each element ranges from 4 to 32. We see that as the polynomial degree is doubled, the error decays exponentially, as expected from p-refinement. In the second table the polynomial degree on each element is fixed at  $N = 2$  as the number of elements range from 16 to 1024. Here we see that the solution converges algebraically as  $E$  is increased. These results are in agreement with the cubic convergence rate for quadratic elements [2].

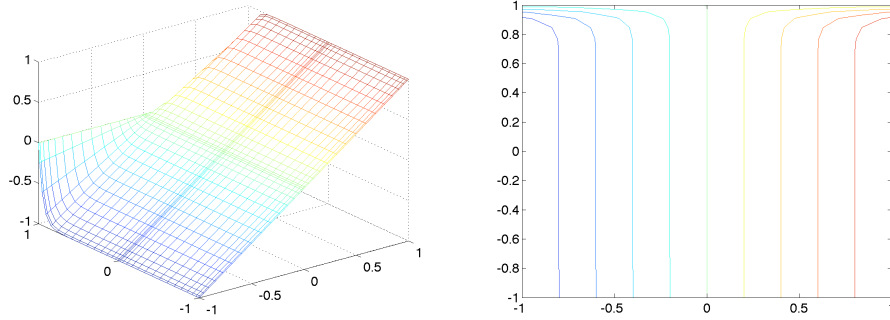


Fig. 4.1: Computed solution (left) and contours (right) of steady convection diffusion flow corresponding to example 4.1 with constant wind  $\vec{w} = (0, 1)$  and moderate convection  $Pe = 40$ .

| N  | $\ u - u_N\ _2$         |
|----|-------------------------|
| 4  | $5.535 \times 10^{-2}$  |
| 8  | $2.505 \times 10^{-3}$  |
| 16 | $2.423 \times 10^{-7}$  |
| 32 | $7.931 \times 10^{-13}$ |

Table 4.1: Exponential convergence for example 4.1 with moderate convection ( $Pe = 40$ ), as polynomial degree is varied on a fixed  $2 \times 2$  element grid.

| E    | $\ u - u_N\ _2$        | $13h^3$                |
|------|------------------------|------------------------|
| 16   | $8.594 \times 10^{-2}$ | $2.031 \times 10^{-1}$ |
| 64   | $2.593 \times 10^{-2}$ | $2.523 \times 10^{-2}$ |
| 256  | $3.558 \times 10^{-3}$ | $3.174 \times 10^{-3}$ |
| 1024 | $3.610 \times 10^{-4}$ | $3.967 \times 10^{-4}$ |

Table 4.2: Algebraic convergence for example 4.1 with moderate convection ( $Pe = 40$ ), as the number of quadratic elements are varied.

**4.2. Oblique wind with internal and outflow boundary layers.** In our second example we use a flow that exhibits two layers, one characteristic boundary layer along the top wall proportional to  $\epsilon$  and a second internal layer of width proportional to  $\sqrt{\epsilon}$  that results from a jump discontinuity in the boundary at  $(0, -1)$  (see Figure 4.2). Dirichlet boundary conditions are imposed along each boundary. The wind field in this test case is constant  $\vec{w} = (-\sin(\pi/6), \cos(\pi/6))$ , but unlike the previous example, here the wind is not aligned with the grid.

**4.3. Constant wind results.** We now focus on the iterative solution of the interface problem (3.6). We examine the influence of changes in the discretization and the Peclet number on the iterations by performing three sets of experiments for each example problem. First we modify the number of elements in the discretization for a fixed polynomial basis and Peclet number. Next we refine the polynomial degree while keeping the number of elements and the Peclet number fixed. Finally, we fix the number of elements and the polynomial basis, while modifying the Peclet number.

GMRES is used to perform the interface solve. In each experiment, we test three choices of preconditioner for the interface solve: no preconditioner (None), Neumann-Neumann preconditioning (N-N), and Robin-Robin preconditioning (R-R). The results are summarized in Tables 4.3, 4.4, and 4.5. There we record the number of iterations needed for GMRES to converge within a tolerance of  $10^{-12}$ .

For the first experiment we set the Peclet number to 40 and fix the polynomial degree at

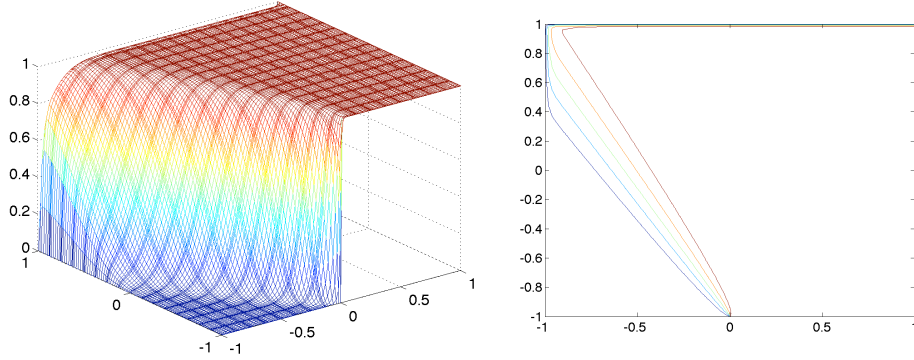


Fig. 4.2: Computed solution (left) and contours (right) of a convection-dominated steady convection-diffusion flow,  $Pe = 250$ , corresponding to example 4.2.

$N = 2$ . We then increase the number of elements from  $E = 16$  (as a  $4 \times 4$  grid) to  $E = 1024$  ( $32 \times 32$  grid). We record the iterations in Table 4.3. Each method depends on the increased grid resolution. The table shows that performance of the Robin-Robin method compares well with theory and is superior to the other two methods. We see that the number of (R-R) iterations for both test cases are essentially of the form  $\frac{C}{h}(1 + \log(N))^2$ , for  $C \approx 1$ . Where  $h$  is the diameter of a single element.

|                | E    | None | N-N   | R-R | $\frac{1}{h}(1 + \log(N))^2$ |
|----------------|------|------|-------|-----|------------------------------|
| Example<br>4.1 | 16   | 13   | 13    | 12  | 11.5                         |
|                | 64   | 49   | 47    | 25  | 22.9                         |
|                | 256  | 108  | 88    | 45  | 45.9                         |
|                | 1024 | 312  | 180   | 85  | 91.7                         |
| Example<br>4.2 | 16   | 29   | 33    | 21  | 11.5                         |
|                | 64   | 40   | 63    | 26  | 22.9                         |
|                | 256  | 69   | 117   | 46  | 45.9                         |
|                | 1024 | 132  | > 200 | 87  | 91.7                         |

Table 4.3: Convergence results for examples 4.1 (top) and 4.2 (bottom) with moderate convection  $Pe=40$ , as the number of quadratic ( $N=2$ ) elements are varied.

In our next experiment, we again set the Peclet number to 40, but now we fix the number of elements to 4 as a  $2 \times 2$  element grid, and view the influence of the polynomial degree on the number of interface solve iterations. The iterations are recorded in Table 4.4. We note that for this test problem with moderate convection and few subdomains, (N-N) and (R-R) converge in roughly the same number of steps, and here the iterations in both test cases are  $\frac{C}{h}(1 + \log(N))^2$ , for  $C \approx 1$ .

Finally, we examine how changes in the Peclet number affect convergence. For this study, we compare the number of iterations needed to obtain the interface solution as the Peclet number is increased. We use a fixed grid with polynomial degree  $N=8$ , and the number of elements at  $E = 1024$  ( $32 \times 32$  element grid). In Table 4.5, we compare convergence results when different preconditioners are used to obtain the interface solution. The Neumann-Neumann preconditioner (N-N) is no more effective than when no preconditioning is used (None), whereas the Robin-Robin preconditioner (R-R) shows little dependence on changes in the Peclet number. Indeed, for moderate values, as the Peclet number is increased

|                | N  | None | N-N | R-R | $\frac{1}{h}(1 + \log(N))^2$ |
|----------------|----|------|-----|-----|------------------------------|
| Example<br>4.1 | 4  | 3    | 3   | 3   | 5.6                          |
|                | 8  | 7    | 7   | 7   | 9.4                          |
|                | 16 | 15   | 11  | 14  | 14.2                         |
|                | 32 | 30   | 16  | 18  | 19.9                         |
| Example<br>4.2 | 4  | 13   | 13  | 13  | 5.6                          |
|                | 8  | 25   | 25  | 18  | 9.4                          |
|                | 16 | 36   | 28  | 20  | 14.2                         |
|                | 32 | 50   | 29  | 21  | 19.9                         |

Table 4.4: Convergence results for examples 4.1 (top) and 4.2 (bottom) with a moderate convection  $Pe = 40$ , as polynomial degree is varied on a fixed  $2 \times 2$  element grid.

the Robin-Robin preconditioner actually improves. For higher Peclet numbers, the iteration counts with Robin-Robin preconditioning increase only mildly with the Peclet number.

|                | Pe   | None  | N-N   | R-R |
|----------------|------|-------|-------|-----|
| Example<br>4.1 | 125  | 161   | 165   | 64  |
|                | 250  | 126   | 144   | 52  |
|                | 500  | 107   | 147   | 46  |
|                | 1000 | 109   | 164   | 43  |
|                | 2000 | 135   | > 200 | 42  |
|                | 5000 | > 200 | > 200 | 50  |
| Example<br>4.2 | 125  | 184   | 186   | 70  |
|                | 250  | 140   | 158   | 61  |
|                | 500  | 107   | 148   | 52  |
|                | 1000 | 88    | 166   | 46  |
|                | 2000 | 96    | > 200 | 52  |
|                | 5000 | > 200 | > 200 | 70  |

Table 4.5: Comparison of iteration counts for examples 4.1 (top) and 4.2 (bottom) with increasingly convection-dominated flows.  $N=8$ ,  $E=1024$  using  $32 \times 32$  element grid.

We conclude that Robin-Robin preconditioner performs well for both grid-aligned and non-grid aligned flows. In cases where the Peclet number is small, this technique behaves like the Neumann-Neumann preconditioner, which is often used to precondition the discrete Poisson equation [17]. We observed that when using the Robin-Robin preconditioner, the number of iterations required for  $\bar{F}_S^{-1}$  had little dependence on the mesh size as well as the Peclet number.

As a final note for this section we point out that for constant-wind problems this solution strategy based on iterative substructuring is significantly cheaper than naively applying GMRES directly to (2.11). The domain decomposition strategy allows one to solve highly convective systems, where in contrast without preconditioning, we have found that GMRES fails to converge. This is accomplished, in part, by eliminating interior degrees of freedom using FDM, thereby reducing memory overhead in high resolution cases by nearly an order of magnitude. For example, in the simulation above using  $N = 8$  and  $E = 32 \times 32$ , there are roughly five times fewer degrees of freedom along elemental interfaces than in the complete system. In three dimensions with large  $N$ , savings in memory would be even greater. For non-symmetric systems solved using GMRES this is very important since one must hold all computed iterates in memory to obtain an orthogonal search direction at each iterate. In the next section we attempt to take advantage of this inexpensive computational method as a



preconditioning technique for solving non-constant wind convection-diffusion systems.

**5. Non-Constant Wind Systems.** When the convection coefficient  $\vec{w}$  is not constant in each component, the domain decomposition solution technique described in the previous two sections does not directly apply. In this section we present a way to use our domain decomposition solver as a preconditioner to accelerate the convergence of GMRES or Flexible GMRES (FGMRES) for solving discrete convection-diffusion systems (2.11) arising from non-constant winds.

In the case of non-constant wind, the matrix  $F(\vec{w})$  from equation (2.12) cannot be written in the tensor product form (3.20) where Fast Diagonalization can be applied element-wise. However, if we approximate  $F(\vec{w})$  on each element in a certain way, we can construct an element-based matrix of this form to be used as a preconditioner for  $F(\vec{w})$ . In particular, we consider the approximation to  $F(\vec{w})$  where  $\vec{w}$  is approximated locally as a constant on each element. To determine this constant, we take the average of  $\vec{w}$  in each component to construct a new piecewise constant wind vector approximation  $\bar{w}$ . Using this vector we can construct a constant-wind approximation to  $F(\vec{w})$ , which we call  $\bar{F}$ .

In order to apply the preconditioner  $\bar{F}$  for multiple elements, we follow the domain decomposition strategy developed in section 3. The only difference is that  $\bar{F}$  is now an approximation to  $F$  based on average elemental wind speeds. Because of this, we note that it is not necessary to resolve the  $\bar{F}_S$  interface solution to high precision, since these values are not likely to reflect the values of the non-constant wind solution accurately. Instead, we seek a rough estimate of the interface values that we can obtain from a few iterations of the Schur complement solve for (3.6). That is, we apply  $\bar{F}_S^{-1}$  inexactly through a few steps of GMRES. It is also important to point out that the iteration used to solve the Schur complement problem in  $\bar{F}_S$  makes  $\bar{F}$  a non-linear operator. Since the preconditioning operator is non-linear, the outer GMRES iteration must be replaced with a variant such as Flexible GMRES [16], which allows for arbitrary changes in the preconditioner at each step of the iteration. Thus, FGMRES is used to solve the right-preconditioned system

$$\underbrace{[\mathbb{M}^e \Sigma' F^e(\vec{w}) \bar{F}^{-1}]}_F [\bar{F} u_0^e] = \underbrace{\mathbb{M}^e \Sigma' (M^e f^e - F^e(\vec{w}) u_D^e)}_b. \quad (5.1)$$

In summary, to solve convection-diffusion problems with non-constant wind we use Flexible GMRES. At each step of FGMRES, a preconditioner,  $\bar{F}$ , based on an element-wise average wind is used. In addition, during each FGMRES iteration we allow for the possibility of using an inexact inner iteration for the interface nodes associated with (3.6). This inner iteration is coupled with FDM to obtain interior degrees of freedom on each element.

**5.1. Double Glazing Problem.** We examine our non-constant wind solver by applying it to a flow that has a recirculating wind  $\vec{w} = (2y(1-x^2), -2x(1-y^2))$  and discontinuities in parts of the boundaries, which lead to boundary layers. This example is known as the *double-glazing problem*, and serves as a simple model of the spread of heat in a box with a hot wall.

The left pane of Figure 5.1 shows the computed solution for an example flow with  $Pe = 400$  using FGMRES, and the right pane shows an approximate solution obtained by replacing the coefficient matrix with the preconditioning matrix  $\bar{F}$ . It can be seen in the figure that in the approximate solution the constant wind approximation produces an error in regions where the convection field changes direction, such as at the corners and around the internal boundary layer.

**5.2. Variable wind with curved open streamlines.** For our second example, we use a test case from [13] with a variable flow field  $\vec{w} = \frac{1}{2}((1-x^2)(1+y), x((1+y)^2 - 4))$ . Boundary



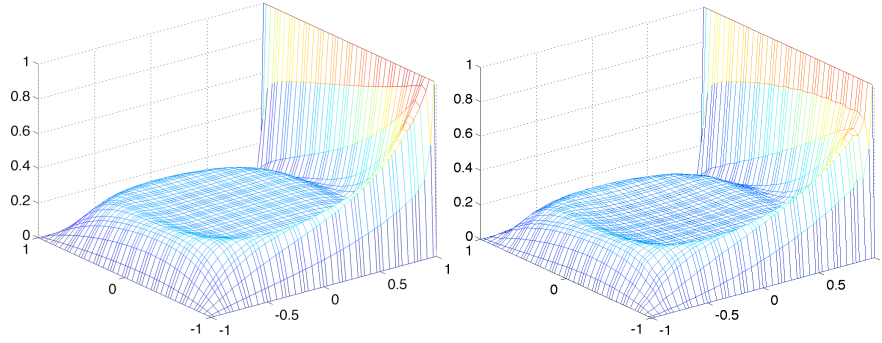


Fig. 5.1: Comparison of exact solution (left) and inexact solution (right) obtained by applying  $\bar{F}^{-1}$  as an inexact solver for example 5.1 with  $Pe = 400$ .

conditions are zero except along the inflow region  $(x, 1) : -1 < x < 0$ , where  $u = 1$ . The left pane of Figure 5.2 shows the computed solution for an example flow with  $Pe = 125$ , and the right pane shows a depiction of the wind field.

**5.3. Variable wind results.** We compare the use of  $\bar{F}$  as a preconditioner for FGMRES with unpreconditioned GMRES and block-Jacobi preconditioned GMRES. To construct the block Jacobi preconditioner, we use the block diagonal of  $\bar{F}$ . So on each element the interior nodes are obtained by solving a system using  $\bar{F}_{II}^e$  via FDM, and the boundary nodes are obtained by solving a system with  $\bar{F}_{BB}$  via GMRES.

The left pane of Figure 5.3 shows the number of FGMRES iterations required to solve (5.1) to a tolerance of  $10^{-12}$  for example 5.1. The top two curves show that unpreconditioned GMRES and block Jacobi preconditioned GMRES are ineffective at solving this system. Next, we study the effect of inexactly applying  $F_S^{-1}$  using an inner iteration with unpreconditioned GMRES. The latter four curves show the results of using various approximations of  $\bar{F}$  to precondition the system. These curves show the influence of accuracy of the Schur complement solve  $\bar{F}_S^{-1}$  on the performance of the outer FGMRES iteration. The  $\circ$  curve shows that applying  $\bar{F}_S^{-1}$  with 1 step of GMRES to obtain an approximate solution of the interface

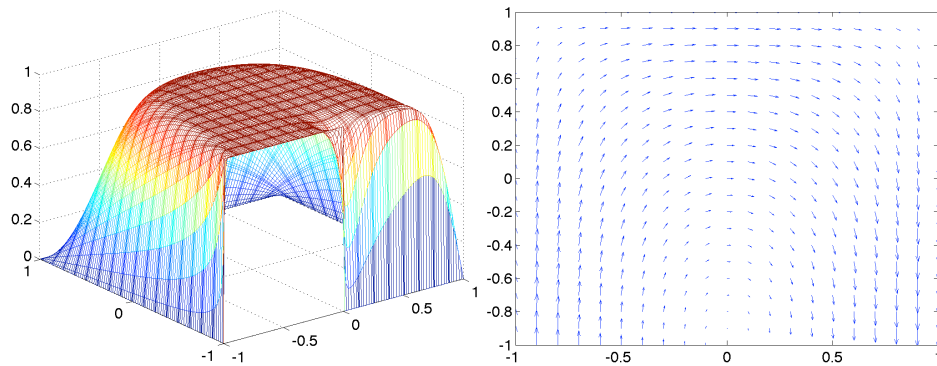
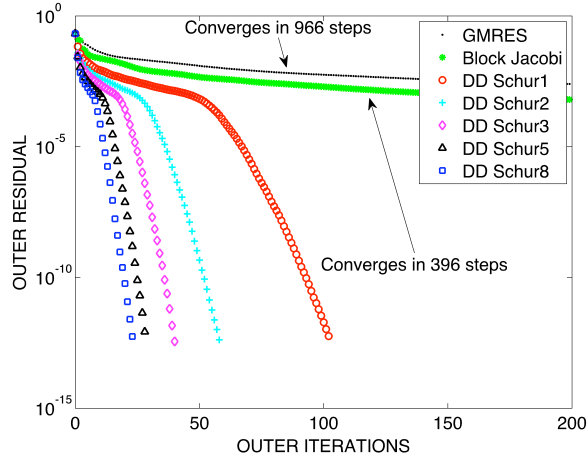


Fig. 5.2: Computed solution (left) and wind field (right) of flow corresponding to example 5.2 with  $Pe = 125$ .



| Preconditioner | Time (s) |
|----------------|----------|
| None           | 53       |
| Block Jacobi   | 67       |
| DD Schur1      | 32       |
| DD Schur3      | 20       |
| DD Schur5      | 17       |
| DD Schur8      | 19       |

Table 5.1: Timings corresponding to the experiments in Figure 5.3.

Fig. 5.3: Comparison of preconditioned Flexible GMRES iterations with varied inner iteration counts, Block Jacobi iterations and unpreconditioned GMRES iterations for solving equation (2.11) with  $Pe = 400$ .

allows the outer FGMRES to converge in 102 steps. In the opposite extreme, the  $\square$  curve shows that computing an inexact interface solution  $u_I$  by performing 8 GMRES iterations in the inner iteration allows the outer iterations via FGMRES to converge in 23 steps. It is evident from Figure 5.3 that only a few inner iterations for the preconditioner are required in order for the outer iteration to converge quickly. We provide a comparison of timings for this experiment in Table 5.1. Note that in this experiment, the fastest run is performed when using 5 GMRES inner iterations, even though this run takes a few more outer iterations than using 8 GMRES inner iterations.

Note that the results described in the previous paragraph come from using unpreconditioned GMRES for the Schur complement problem. We saw in the previous section that Robin-Robin preconditioning improves the performance for constant coefficient problems. However, here our conclusions are somewhat different. In Figure 5.3 we observed that using a few steps of unpreconditioned GMRES for the inner iteration allowed the outer iteration to converge rapidly. The left pane of Figure 5.4 shows the residual during the first 40 steps of the interface solve using GMRES without preconditioning (top) and GMRES with Robin-Robin preconditioning. We see that both solvers reduce the residual by nearly one order of magnitude in the first 10 steps before convergence slows significantly. The Robin-Robin preconditioned solver further reduces the residual around the 25<sup>th</sup> iteration, while the non-preconditioned system shows little improvement. Achdou et al. point out in [1] that in convection-dominated flows, the continuous Robin-Robin preconditioned system operator is close to an idempotent (or periodic) operator of order  $E/2$  where  $E$  is the number of elements. They argue that this causes GMRES to stagnate for  $E/2$  steps before converging asymptotically<sup>1</sup>. However, the right pane of Figure 5.4 shows that the residual of the outer FGMRES iteration is essentially unaffected by the difference in these two inner iteration residuals. The top curve represents the residual of the outer FGMRES iteration using unpreconditioned GM-

<sup>1</sup>Elman and Chernesky [6] report a similar convergence delay when applying block Gauss-Seidel line relaxation to node orderings that do not follow the flow in 1D convection-diffusion simulations.

RES for the inner iteration. In comparison, the bottom curve represents the residual of the outer FGMRES iteration when the Robin-Robin preconditioner is applied to the inner GMRES iteration. Although the Robin-Robin system (bottom curve) converges slightly faster (by 3 iterations), it requires an extra matrix-vector product at each inner iteration, thus in this example it is roughly 40 times more expensive to use Robin-Robin preconditioning instead of using no preconditioner to obtain an inexact interface solution.

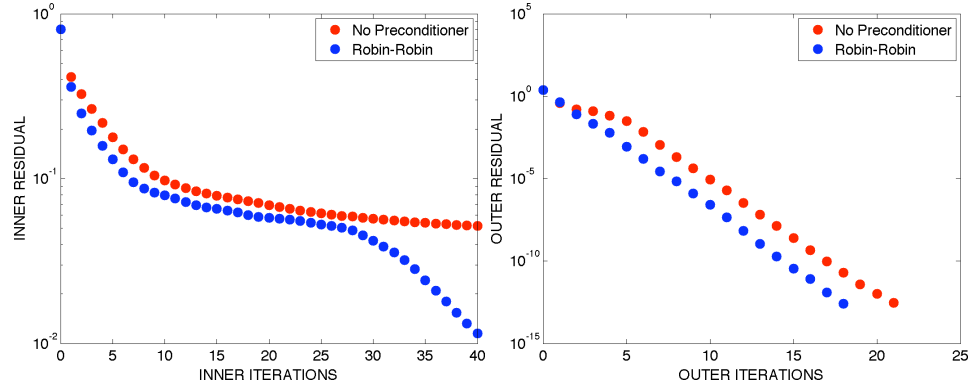


Fig. 5.4: Comparison of residuals for interface iterations obtained by GMRES without preconditioning and with Robin-Robin preconditioning (left) for example 5.1. Effect on FGMRES residuals with inexact  $\bar{F}^{-1}$  using no interface preconditioner and Robin-Robin (right).

Next we look at how the number of FGMRES iterations are influenced by mesh refinement. In Tables 5.2 and 5.3 we show the dependence of FGMRES iterations as the mesh is refined using both h-refinement and p-refinement. We use a stopping tolerance for FGMRES set at  $10^{-12}$ , and for the  $\bar{F}_S^{-1}$  interface solve we stop the iterations when the residual is less than  $10^{-1}$  or when a maximum of 20 steps are reached. We see that outer iterations for both p-refinement (Table 5.2 middle column) and h-refinement (Table 5.3 middle column) are roughly constant, whereas the inner iterations (right columns) require show a mild dependence on the mesh size.

|                | N  | Number of FGMRES<br>Outer Iterations | Number of<br>Inner Iterations |
|----------------|----|--------------------------------------|-------------------------------|
| Example<br>5.1 | 4  | 40                                   | 5                             |
|                | 8  | 51                                   | 5                             |
|                | 16 | 44                                   | 13                            |
|                | 32 | 48                                   | 20                            |
| Example<br>5.2 | 4  | 34                                   | 7                             |
|                | 8  | 35                                   | 8                             |
|                | 16 | 34                                   | 18                            |
|                | 32 | 34                                   | 20                            |

Table 5.2: Iteration counts for examples 5.1 (top) and 5.2 (bottom) as polynomial degree is increased with  $Pe = 400$  and  $E = 16$ .

Finally, we consider the dependence of the Peclet number on this solution method. In Table 5.4 we show how the number of FGMRES iterations are affected as the Peclet number is increased from 125 up to 5000. In this study we use a grid where  $N = 8$  and  $E = 1024$

|                | E    | Number of FGMRES<br>Outer Iterations | Number of<br>Inner Iterations |
|----------------|------|--------------------------------------|-------------------------------|
| Example<br>5.1 | 16   | 40                                   | 5                             |
|                | 64   | 25                                   | 12                            |
|                | 256  | 17                                   | 19                            |
|                | 1024 | 28                                   | 20                            |
| Example<br>5.2 | 16   | 34                                   | 7                             |
|                | 64   | 18                                   | 8                             |
|                | 256  | 11                                   | 20                            |
|                | 1024 | 16                                   | 20                            |

Table 5.3: Iteration counts for examples 5.1 (top) and 5.2 (bottom) as the number of elements are increased with  $Pe = 400$  and  $N = 4$ .

|                | Pe   | Number of FGMRES<br>Outer Iterations | Number of<br>Inner Iterations |
|----------------|------|--------------------------------------|-------------------------------|
| Example<br>5.1 | 125  | 27                                   | 20                            |
|                | 250  | 28                                   | 20                            |
|                | 500  | 30                                   | 20                            |
|                | 1000 | 32                                   | 20                            |
|                | 2000 | 37                                   | 20                            |
|                | 5000 | 48                                   | 20                            |
| Example<br>5.2 | 125  | 19                                   | 20                            |
|                | 250  | 16                                   | 20                            |
|                | 500  | 16                                   | 20                            |
|                | 1000 | 16                                   | 20                            |
|                | 2000 | 17                                   | 20                            |
|                | 5000 | 21                                   | 20                            |

Table 5.4: Iteration counts for examples 5.1 (top) and 5.2 (bottom) as the Peclet number is increased on a fixed grid with  $N = 8$  and  $E = 1024$ .

on a  $32 \times 32$  element grid. We use the same stopping criteria as in the mesh refinement study above. We see that outer iterations (middle column) are mildly dependent on the Peclet number whereas the inner iterations (right columns) reach the maximum number of iterations (20) for each Peclet number.

**6. Summary.** We have introduced two solution strategies for convection-diffusion systems. The first strategy applies to problems with constant wind coefficients. This method uses an extension to the Fast Diagonalization Method that we developed in order to solve convection-diffusion problems with constant wind coefficients on single domains. We coupled this result with Robin-Robin preconditioned Domain Decomposition to develop a matrix-free solution method for tensor-product based discretizations of the steady convection-diffusion equation with constant wind on each element.

This method uses iterative substructuring to resolve elemental interface values, together with Fast Diagonalization to eliminate interior degrees of freedom on each element via a direct solve. We demonstrated that this solution method has a weak dependence on Peclet number, and mild dependence on mesh refinement for both h-refinement and p-refinement. Similar conclusions have been reported in [1], [13] and [19] for overlapping and non-overlapping domain decomposition strategies. The use of FDM in our method significantly reduces the computational cost of the subsidiary interior solves throughout the calculation.

We then developed a solver for variable winds by demonstrating how the domain decom-

position method we developed for constant-winds can be used as a preconditioner for general convection-diffusion systems when combined with Flexible GMRES. This variable wind solution strategy showed significant improvement over non-preconditioned GMRES and Block-Jacobi preconditioning techniques. We showed that only an inexact interface solve is needed to significantly accelerate convergence of the outer iteration. Using  $\bar{F}$  as a preconditioner allowed FGMRES to obtain convergence rates independent of the mesh size, and mildly dependent of changes in convection strength. Our results compare favorably with those reported in [1], [7], [13] and [19].

**7. Appendix: Tensor-Based Nodal Ordering for Domain Decomposition.** In section 3.3 we explained how the tensor product basis of the spectral element method allows for efficient matrix-vector products and Fast Diagonalization. In Domain Decomposition Methods, however, it is common to use node orderings that enumerate interior degrees of freedom and then boundary degrees of freedom. In this section we provide the one-dimensional building blocks needed to formulate the two-dimensional operators in terms of their interior and boundary couplings within a lexicographically ordered tensor product framework. We let  $N$  be the degree of the polynomial basis for a given discretization.

We write  $\hat{F}_{(N+1) \times (N+1)}$  as the full 1D convection-diffusion (or diffusion) matrix, and  $\hat{M}_{(N+1) \times (N+1)}$  as the diagonal 1D mass matrix.  $F_{(N+1)^2 \times (N+1)^2} = \hat{F} \otimes \hat{M} + \hat{M} \otimes \hat{F}$  is the sparse 2D convection-diffusion matrix on a single element. We can decompose  $\hat{F}$  and  $\hat{M}$  into their interior and boundary couplings.

$$\begin{aligned}\hat{F}_{ii} &= \hat{F}(2 : N, 2 : N) \text{ Interior-Interior} \\ \hat{F}_{ib} &= \hat{F}(2 : N, 1 : N+1) \text{ Interior-Boundary} \\ \hat{F}_{bi} &= \hat{F}(1 : N+1, 2 : N) \text{ Boundary-Interior} \\ \hat{F}_{bb} &= \hat{F}(1, 1) + \hat{F}(1, N) + \hat{F}(N, 1) + \hat{F}(N, N) \text{ Boundary-Boundary} \\ \hat{M}_{ii} &= \hat{M}(2 : N, 2 : N) \text{ Interior-Interior} \\ \hat{M}_{bb} &= \hat{M}(1 : 1, \vec{0}, N+1 : N+1) \text{ Boundary-Boundary}\end{aligned}$$

This decomposition allows  $F$  to be written as  $F = F_{II} + F_{IT} + F_{TI} + F_{TT}$  with

$$\begin{aligned}F_{II} &= \hat{F}_{ii} \otimes \hat{M}_{ii} + \hat{M}_{ii} \otimes \hat{F}_{ii} \\ F_{IT} &= \hat{F} \otimes \hat{M}_{bb} + \hat{M}_{bb} \otimes \hat{F} + \hat{F}_{bb} \otimes \hat{M}_{ii} + \hat{M}_{ii} \otimes \hat{F}_{bb} \\ F_{TI} &= \hat{F}_{ib} \otimes \hat{M}_{ii} + \hat{M}_{ii} \otimes \hat{F}_{ib} \\ F_{TT} &= \hat{F}_{bi} \otimes \hat{M}_{ii} + \hat{M}_{ii} \otimes \hat{F}_{bi}.\end{aligned}$$

## REFERENCES

- [1] Y. ACHDOU, P. L. TALLEC, F. NATAF, AND M. VIDRASCU, *A domain decomposition preconditioner for an advection-diffusion problem*, Computer Methods in Applied Mechanics and Engineering, 184 (2000), pp. 145–170.
- [2] D. BRAESS, *Finite Elements Theory, fast solvers, and applications in solid mechanics*, Cambridge University Press, 2nd ed., 2001.
- [3] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.
- [4] M. O. DEVILLE, P. F. FISCHER, AND E. H. MUND, *High-Order Methods for Incompressible Fluid Flows*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2002.
- [5] W. ECKHAUS, *Asymptotic Analysis of Singular Perturbations*, Elsevier Science, Amsterdam, 1979.
- [6] H. C. ELMAN AND M. P. CHERNESKY, *Ordering effects on relaxation methods applied to the discrete one-dimensional convection-diffusion equation*, SIAM Journal on Numerical Analysis, 30 (1993), pp. 1268–1290.
- [7] H. C. ELMAN, D. SILVESTER, AND A. WATHEN, *Finite Elements and Fast Iterative Solvers with applications in incompressible fluid dynamics*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2005.

- [8] P. FISCHER, *An overlapping Schwarz method for spectral element solution of the incompressible Navier-Stokes equations*, Journal of Computational Physics, 133 (1997), pp. 84–101.
- [9] P. F. FISCHER, F. LOTH, S. E. LEE, S. W. LEE, D. S. SMITH, AND H. S. BASSIOUNY, *Simulation of high Reynolds number vascular flows*, Computer methods in Applied Mechanics and Engineering, 196 (2007), pp. 3049–3060.
- [10] J. LOTTES AND P. FISCHER, *Hybrid multigrid/Schwarz algorithms for the spectral element method*, Journal of Scientific Computing, 24 (2005), pp. 613–646.
- [11] R. LYNCH, J. RICE, AND D. THOMAS, *Direct solution of partial difference equations by tensor product methods*, Numerische Mathematik, 6 (1964), pp. 185–199.
- [12] Y. MADAY, A. PATERA, AND E. RÖNQUIST, *An operator-integration-factor splitting method for time dependent problems: Application to incompressible fluid flow*, Journal of Scientific Computation, 5 (1990), pp. 263–292.
- [13] L. PAVARINO, *Overlapping schwarz preconditioners for spectral element discretizations of convection-diffusion problems*, International Journal for Numerical Methods in Engineering, 53 (2002), pp. 1005–1023.
- [14] A. QUARTERONI AND A. VALLI, *Domain Decomposition Methods for Partial Differential Equations*, Numerical Mathematics and Scientific Computation, Oxford University Press, 1999.
- [15] H. G. ROOS, M. STYNES, AND L. TOBISKA, *Numerical Methods for Singularly Perturbed Differential Equations*, Springer-Verlag, Berlin, 1996.
- [16] Y. SAAD, *A flexible inner-outer preconditioned gmres algorithm*, SIAM Journal on Scientific Computing, 14 (1993), pp. 461–469.
- [17] B. F. SMITH, P. BJØRSTAD, AND W. D. GROPP, *Domain Decomposition Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, 1996.
- [18] S. J. THOMAS AND R. D. LOFT, *The NCAR spectral element climate dynamical core: Semi-implicit Eulerian formulation*, Journal of Scientific Computing, 25 (2005), pp. 307–322.
- [19] A. TOSELLI, *Feti domain decomposition methods for scalar advection-diffusion problems*, Computer Methods in Applied Mechanics and Engineering, 190 (2001), pp. 5759–5776.
- [20] A. TOSELLI AND O. WIDLUND, *Domain Decomposition Methods - Algorithms and Theory*, Springer Series in Computational Mathematics, Springer, 2005.
- [21] H. TUFO AND P. FISCHER, *Terascale spectral element algorithms and implementations*, in Supercomputing '99: Proceedings of the 1999 ACM/IEEE conference on Supercomputing (CDROM), New York, NY, USA, 1999, ACM Press, p. 68.