Technical Report

TR-2009-002

Solution of Linear Systems from an Optimal Control Problem Arising in Wind Simulation

by

M. Benzi, L. Ferragut, M. Pennacchio, V. Simoncini

Mathematics and Computer Science

EMORY UNIVERSITY

NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2000; **00**:1–18 *Prepared using* nlaauth.cls *[Version: 2002/09/18 v1.02]*

Solution of Linear Systems from an Optimal Control Problem Arising in Wind Simulation

M. Benzi[1], L. Ferragut[2], M. Pennacchio[3,*] and V. Simoncini[4]

[1] *Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA*
[2] *Instituto Universitario de Física y Matemáticas y Departamento de Matemática Aplicada, Plaza de la Merced s/n, Universidad de Salamanca, Salamanca, 37008, Spain*
[3] *Istituto di Matematica Applicata e Tecnologie Informatiche, via Ferrata, 1, 27100 Pavia, Italy*
[4] *Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato, 5, 40127 Bologna, Italy*

SUMMARY

Several solution strategies for a class of large, sparse linear systems with a block 2-by-2 structure arising from the finite element discretization of an optimal control problem in wind simulation are introduced and analyzed. Block preconditioners and a sparse direct solver on the original coupled system are compared with a preconditioned GMRES iteration applied to a reduced system (Schur complement). Theoretical and experimental results demonstrate the effectiveness of the reduced system approach.

KEY WORDS: Block preconditioning, Schur complement, sparse direct solvers, AMG, eigenvalue bounds

1. INTRODUCTION

In this paper we develop an efficient solver for large, sparse systems of linear equations with a 2-by-2 block structure arising from finite element discretizations of control problems resulting from a mathematical model of wind field adjustment.

Wind models are important tools that allow the study of several problems related to the atmosphere, such as the effect of wind on structures, wind power, pollutant transport, fire spreading, and so forth. Our starting point is a mass consistent vertical diffusion wind field model. If the significant phenomena that we want to simulate occur in a zone where the

*Correspondence to: Micol Pennacchio, Istituto di Matematica Applicata e Tecnologie Informatiche, via Ferrata, 1, 27100 Pavia, Italy. E-mail: micol@imati.cnr.it

Contract/grant sponsor: U. S. National Science Foundation; contract/grant number: DMS-0511336.

Contract/grant sponsor: Ministerio de Ciencia e Innovación, Spain; contract/grant number: CGL2008-06003-C03-03/CLI

Contract/grant sponsor: Junta de Castilla y León, Spain; contract/grant number: SA124A08

Received
Revised

horizontal dimensions are much larger than the vertical one, then an asymptotic approximation of the primitive Navier–Stokes equations can be derived as in the model developed in [1]. The most salient feature of this asymptotic approach is that it provides a three-dimensional velocity wind field (which satisfies the incompressibility condition in the air layer) governed by a two–dimensional equation, so that it can be coupled with the temperature surface distribution in order to take into account thermal effects such as sea breezes. In addition, the terrain elevation information is also taken into account by the model.

The validity of this model has the following limits: the nonlinear terms are neglected, and it is assumed that the air temperature decreases linearly with the height. Also, the air compressibility is neglected. On the other hand, the model takes into account buoyancy forces, slope effects, and mass conservation. The wind model presented in this article is an adaptation of the wind model proposed in [1]. When the data are given by meteorological predictions, an optimal control problem is obtained [2] which can be solved using the adjoint equation-based method. We refer the reader to [3] and [4] for the details of this approach. The corresponding numerical approximation leads to linear algebraic systems of equations which are very ill-conditioned and quite challenging to solve. In practical applications, the number of equations can be high (roughly between 100,000 and 600,000) and many systems may have to be solved in the course of a simulation, thus justifying the search for efficient iterative methods.

The remainder of the paper is organized as follows. In section 2 we describe the mathematical model leading to the control problem that we are interested in solving. The variational (weak) formulation used to establish the well-posedness of the continuous problem and for its discretization by means of finite elements is given in section 3. Section 4 is devoted to a discussion of the discrete problem and some of its properties. Several solution methods, including a preconditioned iteration on the Schur complement system, are described in section 5. In section 6 we analyze the spectrum of the preconditioned Schur complement, and in section 7 we report on numerical experiments. Some conclusions are given in section 8.

2. THE CONTROL PROBLEM

In this section we present the wind model. An asymptotic analysis gives a three-dimensional convective model governed by a two-dimensional equation. This model adjusts a three-dimensional velocity wind field in a layer under the influence of the orography and temperature distribution.

2.1. Notation

Let us consider the three-dimensional domain $\Omega = \{(\mathbf{x}, z)|\mathbf{x} \in \omega, H(\mathbf{x}) < z < \delta\}$ representing the air layer under study. We assume that the height δ is small compared to the width, and that the surface height at point $\mathbf{x}$, $H(\mathbf{x})$, is smaller than δ. We decompose the boundary of Ω into $\partial\Omega = S \cup A \cup L$, where $S = \{(\mathbf{x}, z)|\mathbf{x} \in \omega,\ z = H(\mathbf{x})\}$ is the surface, $A = \{(\mathbf{x}, z)|\mathbf{x} \in \omega,\ z = \delta\}$ is the air upper boundary and $L = \{(\mathbf{x}, z)|\mathbf{x} \in \partial\omega,\ H(\mathbf{x}) < z < \delta\}$ is the air lateral boundary; $\omega \subset \mathbb{R}^2$ is a two–dimensional normalized bounded domain, representing the projection of the three-dimensional geographical surface S. We denote by $(\mathbf{x}, z)$ any point of the three-dimensional domain Ω, and by $\mathbf{x}$ any point of the two-dimensional domain ω.

2.2. Asymptotic equations

Consider an air velocity field $\mathbf{U} = (U, V, W)$ and a potential P satisfying the Navier–Stokes equations. Using the fact that the thickness δ of the considered air layer is small compared with its width, we obtain the following vertical diffusion model:

$$\begin{aligned} -\partial^2_{zz}\mathbf{V} + \nabla_{\mathbf{x}} P &= 0, & (1)\\ \partial_z P &= \mu T, & (2)\\ \nabla_{\mathbf{x}} \cdot \mathbf{V} + \partial_z W &= 0, & (3) \end{aligned}$$

where $\mathbf{V} = (U, V)$ denotes the horizontal velocity, $\mu = \phi Re$ (with ϕ representing buoyancy forces and Re the Reynolds number), and T is the temperature. We define the horizontal flux at a point $\mathbf{x} \in \omega$ by

$$\overline{\mathbf{V}} = \int_{H(\mathbf{x})}^{\delta} \mathbf{V}(\mathbf{x}, z) dz.$$

Denoting by $\mathbf{N}$ and $\mathbf{n}$ the inner unit normal vector field to $\partial\Omega$ and to $\partial\omega$, respectively, the boundary conditions can be written as

$$\begin{aligned} \partial_z \mathbf{V} = \zeta \mathbf{V}, \quad (\mathbf{V}, W) \cdot \mathbf{N} = 0, \qquad & \text{on } S, & (4)\\ \partial_z \mathbf{V} = 0, \quad W = 0, \qquad & \text{on } A, & (5)\\ \overline{\mathbf{V}} \cdot \mathbf{n} = (\delta - H)\mathbf{v}_m \cdot \mathbf{n}, \qquad & \text{on } \partial\omega. & (6) \end{aligned}$$

Here $\mathbf{v}_m$ denotes the meteorological wind, which is assumed to be known, horizontal, independent of z and with zero total flux through the lateral boundary, that is,

$$\partial_z \mathbf{v}_m = 0, \quad \int_{\partial\omega} (\delta - H)\, \mathbf{v}_m \cdot \mathbf{n}\, ds = 0.$$

Equations (1) to (6) are well posed: for given T and $\mathbf{v}_m$, there exists a unique solution $(\mathbf{V}, W, P)$ (up to an additive constant for P). For more details about this convection asymptotic model, see [1].

Equation (1), together with conditions in (4) and (5), yields

$$\mathbf{V}(\mathbf{x}, z) = m(\mathbf{x}, z)\nabla_{\mathbf{x}} p(\mathbf{x}) + k(\mathbf{x}, z)\nabla_{\mathbf{x}} \hat{t}(\mathbf{x}), \tag{7}$$

where

$$\begin{aligned} m(\mathbf{x}, z) &= \frac{1}{2}z^2 - \delta z - \frac{1}{2}H^2(\mathbf{x}) + (\delta + \xi)H(\mathbf{x}) - \xi\delta,\\ k(\mathbf{x}, z) &= -\frac{1}{24}z^4 + \frac{1}{6}\delta z^3 - \frac{1}{3}\delta^3 z + \frac{1}{24}H^4(\mathbf{x}) - \frac{1}{6}H^3(\mathbf{x})(\delta + \xi)\\ &\quad + \frac{1}{2}\xi\delta H^2(\mathbf{x}) + \frac{1}{3}\delta^3 H(\mathbf{x}) - \frac{1}{3}\xi\delta^3, \end{aligned}$$

being $\xi = \frac{1}{\zeta}$ the inverse of the friction coefficient ζ and $\hat{t}$ a re-scaled temperature related to the surface temperature $t = t(\mathbf{x})$ by $\hat{t}(\mathbf{x}) = \frac{\mu t(\mathbf{x})}{\delta - H(\mathbf{x})}$. We are assuming that the air temperature decreases linearly with the height, $T(\mathbf{x}, z) = t(\mathbf{x})\frac{\delta - z}{\delta - H(\mathbf{x})}$. The function $p(\mathbf{x})$ is a potential that satisfies the following boundary value problem:

$$-\nabla_{\mathbf{x}}(a\nabla_{\mathbf{x}}p) = \nabla_{\mathbf{x}}(r\,\nabla_{\mathbf{x}}\hat{t}) \quad \text{in } \omega, \tag{8}$$

$$a\frac{\partial p}{\partial \mathbf{n}} = -r\,\frac{\partial \hat{t}}{\partial \mathbf{n}} + (\delta - H)\mathbf{v}_m \cdot \mathbf{n} \quad \text{on} \quad \partial\omega, \tag{9}$$

where

$$a = a(\mathbf{x}) = \frac{1}{3}(\delta - H(\mathbf{x}))^2(3\xi + \delta - H(\mathbf{x})),$$

and

$$r = r(\mathbf{x}) = \frac{1}{30}(\delta - H(\mathbf{x}))^2\Big(2\delta^2(2\delta + 5\xi) - 2\delta(\delta - 5\xi)H(\mathbf{x}) - (3\delta + 5\xi)H^2(\mathbf{x}) + H^3(\mathbf{x})\Big).$$

2.3. Adjustment of point data by solution of an optimal control problem

To simplify the notation, and since in the following we are only concerned with the two-dimensional problem, we omit the subscript $(\ )_{\mathbf{x}}$ in the differential operators.

Let $v = (\delta - H)\mathbf{v}_m \cdot \mathbf{n}$, then $v \in L^2_0(\partial\omega)$, where $L^2_0(\partial\omega) = \{v \in L^2(\partial\omega) | \int_{\partial\omega} v\,ds = 0\}$. We are going to reformulate the original problem as an optimal control problem. Given N experimental measurements of the wind velocity $\mathbf{V}_i,\ i = 1, \ldots, N$, at N given points $P_i = (\mathbf{x}_i, z_i),\ i = 1, \ldots, N$, we look for the function $v \in L^2_0(\partial\omega)$ such that the values of $\mathbf{V}(\mathbf{x}_i, z_i)$ given by the expression in (7) are as close as possible to the experimental values of $\mathbf{V}_i$. Thus, in the optimal control framework we have:

i) $v \in L^2_0(\partial\omega)$ is the control;

ii) Equations (8),(9) are the state equations;

iii) The regularized cost functional to be minimized is given by:

$$J(v) = \frac{1}{2}\sum_{i=1}^{N}\int_{\omega} \rho_{\varepsilon,i}(\mathbf{x})|m(\mathbf{x}, z_i)\nabla p(\mathbf{x}) + k(\mathbf{x}, z_i)\nabla\hat{t}(\mathbf{x}) - \mathbf{V}_i|^2 d\mathbf{x} + \frac{\alpha}{2}\int_{\partial\omega} v^2 ds,$$

where α is a regularization parameter (usually $\alpha = 0.001$) and $\rho_{\varepsilon,i}$ is a suitable smoothing function given for example by

$$\rho_{\varepsilon,i}(\mathbf{x}) = \frac{1}{\varepsilon^2}\rho\left(\frac{\mathbf{x} - \mathbf{x}_i}{\varepsilon}\right), \qquad \rho(\mathbf{x}) = \begin{cases} G\,\mathrm{e}^{-\frac{1}{1-|\mathbf{x}|^2}} & \text{for } |\mathbf{x}| < 1 \\ 0 & \text{for } |\mathbf{x}| \geq 1, \end{cases}$$

for a small ε, where G is a constant such that $\int \rho_{\epsilon,i}(\mathbf{x})d\mathbf{x} = 1$. Here and in the rest of the paper, we denote the Euclidean length of a vector $\mathbf{x}$ by $|\mathbf{x}|$.

With these definitions, the optimal control problem to be solved may be posed as follows: Find $u \in L^2_0(\partial\omega)$ such that

$$J(u) = \inf_{v \in L^2_0(\partial\omega)} J(v). \tag{10}$$

The solution u is characterized by the vanishing of the first variation: $J'(u) = 0$.

3. WEAK FORMULATION

Let $V = H^1(\omega)$, then using general optimal control theory (see, e.g., [3]) and introducing the adjoint state q, problem (10) may be formulated as:

find $p \in V$, $q \in V$ *such that*

$$\int_\omega a\,\nabla p \cdot \nabla \varphi \; d\mathbf{x} + \frac{1}{\alpha}\int_{\partial\omega} q\varphi \; d\sigma = -\int_\omega r\,\nabla \hat{T} \cdot \nabla \varphi \; d\mathbf{x} \qquad \forall \varphi \in V, \tag{11}$$

$$\int_\omega a\,\nabla q \cdot \nabla \psi \; d\mathbf{x} - \sum_{i=1}^{N} \int_\omega \rho_{\varepsilon,i}\, m^2\, \nabla p \cdot \nabla \psi \; d\mathbf{x} = \sum_{i=1}^{N} \int_\omega g_i \nabla \psi \; d\mathbf{x} \qquad \forall \psi \in V, \tag{12}$$

together with the relation

$$u = -\frac{1}{\alpha} q \quad \text{on} \quad \partial\omega, \tag{13}$$

where $g_i(\mathbf{x}) = \rho_{\varepsilon,i}(\mathbf{x}) \left(k\,\nabla \hat{T} - V_i \right) m(\mathbf{x}, z_i)$.

Let us define the following bilinear forms:

$$b(p, \varphi) := \int_\omega a\,\nabla p \cdot \nabla \varphi \; d\mathbf{x}, \tag{14}$$

$$c_2(p, \varphi) := \sum_{i=1}^{N} \int_\omega \rho_{\varepsilon,i}\, m^2\, a\,\nabla p \cdot \nabla \varphi \; d\mathbf{x}, \tag{15}$$

$$c_1(q, \psi) := \frac{1}{\alpha} \int_{\partial\omega} q\psi \, d\sigma. \tag{16}$$

From the definition of $c_2(\cdot,\cdot)$ and $b(\cdot,\cdot)$ we can see that there exists a constant γ related to the functions $\rho_{\varepsilon,i}$ and to m such that

$$c_2(p, \varphi) \le \gamma\, b(p, \varphi) \qquad \forall p, \varphi \in V. \tag{17}$$

To have uniqueness of the solution, the (singular) bilinear form $b(\cdot,\cdot)$ needs to be perturbed with a term of the form $\eta \int_\omega p\varphi \; d\mathbf{x}$. As a result, the perturbed bilinear form $\tilde{b}(\cdot,\cdot)$,

$$\tilde{b}(p, \varphi) := b(p, \varphi) + \eta \textstyle\int_\omega p\varphi \; d\mathbf{x}, \tag{18}$$

has an eigenvalue λ_{min} of order η. In the FreeFEM [5] code used below, $\eta = 0.001$.

It can be easily verified that there exist positive constants γ_1, γ_2 and k_η such that

$$|c_2(p, \varphi)| \le \gamma_1 \, \|p\|_{H^1(\omega)} \|\varphi\|_{H^1(\omega)}, \tag{19}$$

$$|\tilde{b}(p, \varphi)| \le \gamma_2 \, \|p\|_{H^1(\omega)} \|\varphi\|_{H^1(\omega)}, \tag{20}$$

$$|\tilde{b}(p, p)| \ge k_\eta \, \|p\|^2_{H^1(\omega)} \tag{21}$$

for all $p, \varphi \in V$ and with k_η dependent on the regularization parameter η.

Thus, the problem that we are dealing with can be written as:

find $p \in V$, $q \in V$ such that

$$\tilde{b}(p,\varphi) + c_1(q,\psi) = f(\varphi) \qquad \forall \varphi \in V, \tag{22}$$

$$\tilde{b}(q,\psi) - c_2(q,\varphi) = g(\psi) \qquad \forall \psi \in V, \tag{23}$$

with $f(\varphi) = -\int_\omega r\, \nabla \hat{T} \cdot \nabla \varphi \; d\mathbf{x}$ and $g(\psi) = \sum_{i=1}^{N} \int_\omega g_i \nabla \psi \; d\mathbf{x}$.

We now introduce linear continuous operators $\mathcal{B}, \mathcal{C}_1, \mathcal{C}_2 : V \longrightarrow V'$ associated to $\tilde{b}, c_1, c_2$ respectively, i.e.:

$$\langle \mathcal{B}p, \varphi \rangle := \tilde{b}(p,\varphi) \qquad \langle \mathcal{C}_1 p, \varphi \rangle := c_1(p,\varphi) \qquad \langle \mathcal{C}_2 q, \psi \rangle := c_2(q,\psi), \tag{24}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between V and its dual space V'.

The above weak formulation is the basis for the finite element discretization of the original continuous problem.

4. THE DISCRETE PROBLEM

Let $\mathcal{T}_h$ be a uniform triangulation of ω corresponding to a discretization parameter h, and let V_h be the associated space of P_1(or P_2) finite elements. Besides a better order of convergence, a reason in favor of P_2 against P_1 is that in practical applications, the variable of physical interest is the wind velocity $\mathbf{V}$ which is obtained from the potential p using expression (7), involving derivatives.

Choosing a finite element basis $\{\phi_i\}$ for V_h, we introduce the following matrices:

$$B = \left\{ b_{r,k} = \sum_{K \in \mathcal{T}_h} \int_K a\, \nabla \phi_r \cdot \nabla \phi_k \; d\mathbf{x} + \eta \int_K \phi_r \phi_k \; d\mathbf{x} \right\}, \; C_1 = \left\{ c^1_{r,k} = \frac{1}{\alpha} \int_{\partial\omega} \phi_r \phi_k \, d\mathbf{x} \right\},$$

$$C_2 = \left\{ c^2_{r,k} = \sum_{i=1}^{N} \sum_{K \in \mathcal{T}_h} \int_K \rho_{\varepsilon,i}\, m^2\, \nabla \phi_r \cdot \nabla \phi_k \; d\mathbf{x} \right\}.$$

Thus, the discrete problem can be written as the following linear algebraic system:

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \qquad \text{with} \quad \mathcal{A} = \begin{bmatrix} B & C_1 \\ -C_2 & B \end{bmatrix}. \tag{25}$$

Note that the system (25) is $2n \times 2n$, where $n = \dim V_h$.

4.1. Properties of C_1, C_2 and B

In order to develop efficient solvers for the linear system (25), it is important to have a good understanding of the properties of the matrices B, C_1 and C_2.

The entries of C_1 are all $O(h)$ since they are integrals on the boundary (which is 1-dimensional) not involving any derivatives. The rank of C_1 is the number of nodes (of the finite element mesh) on the boundary $\partial\omega$.

Since C_2 and B correspond to second-order elliptic operators discretized with P_1 (or P_2) finite elements, their entries are $O(1)$ (because we are in dimension two). Note that B has full

rank owing to the presence of regularization. The rank of C_2 is the number of nodes such that the associated basis function intersects the support of the functions $\rho_{\varepsilon,i}$, i.e., the number of nodes in a small neighborhood of the circle of radius ε with center $\mathbf{x}_i$. Hence, the rank of C_2 is usually much lower than the rank of B. Hence, C_1 and C_2 are both highly singular.

The following theorem provides some useful spectral information about the matrices B, C_1 and C_2.

Theorem 4.1. *Let us denote by $\lambda(C_1), \lambda(C_2), \lambda(B)$ the eigenvalues of C_1, C_2 and B respectively. Then it holds:*

1. C_1 is symmetric positive semidefinite with $\lambda_{\max}(C_1) \leq k_1 h$,
2. C_2 is symmetric positive semidefinite with $\lambda_{\max}(C_2) \leq k_2$,
3. B is symmetric positive definite with $\widetilde{k}_\eta\, h^2 \leq \lambda(B) \leq k_3$,

with $k_1, k_2, k_3, \widetilde{k}_\eta$ positive constants independent of h and $\widetilde{k}_\eta$ dependent on the regularization parameter η in (18).

Proof. Let v_h be the unique element in V_h such that its nodal values are the components v_i of the vector $\mathbf{v} \in \mathbb{R}^n$, i.e.,

$$v_h(x) = \sum_i v_i \phi_i(x)$$

and let us assume a regular triangulation. We have:

$$\exists\, \widetilde{c}, \widehat{c} > 0 : \forall v_h \in V_h \qquad \widetilde{c}\, h^2 |\mathbf{v}|^2 \leq ||v_h||^2_{L^2(\omega)} \leq \widehat{c}\, h^2 |\mathbf{v}|^2. \tag{26}$$

Moreover, the following inverse inequalities hold:

$$\exists \bar{c}_I > 0 : \forall v_h \in V_h \quad \|\nabla v_h\|_{L^2(\omega)} \leq \bar{c}_I h^{-1} \|v_h\|_{L^2(\omega)} \tag{27}$$

$$\exists c_I > 0 : \forall v_h \in V_h \quad \|v_h\|_{L^2(\partial\omega)} \leq c_I h^{-1/2} \|v_h\|_{L^2(\omega)}. \tag{28}$$

1. The only nonzero entries of C_1 are those corresponding to nodes belonging to the boundary $\partial\omega$. Moreover, from the definition and the symmetry of $c_1(\cdot,\cdot)$ we get that C_1 is symmetric and positive semidefinite. The eigenvalues λ of C_1 are such that:

$$\lambda \;=\; \frac{\mathbf{v}^T C_1 \mathbf{v}}{|\mathbf{v}|^2} = \frac{c_1(v_h, v_h)}{|\mathbf{v}|^2} \leq \frac{\|v_h\|^2_{L^2(\partial\omega)}}{|\mathbf{v}|^2} \leq c_I^2\, h^{-1} \frac{\|v_h\|^2_{L^2(\omega)}}{|\mathbf{v}|^2} = c_I^2 \widehat{c}\, h^{-1} h^2 \frac{|\mathbf{v}|^2}{|\mathbf{v}|^2} = k_1\, h$$

   hence $\lambda_{max}(C_1) \leq k_1\, h$ with k_1 positive constant independent of h.
2. Concerning C_2, we can apply the standard theory for matrices associated to the discretization of a second-order elliptic problem by P_1 (or P_2) finite elements in dimension two. The only nonzero entries of C_2 are those corresponding to nodes such that the associated basis function intersects the support of the functions $\rho_{\varepsilon,i}(\mathbf{x})$. Thanks to (19),(27), and (26) we get that the eigenvalues λ of C_2 satisfy

$$\lambda \;=\; \frac{\mathbf{v}^T C_2 \mathbf{v}}{|\mathbf{v}|^2} = \frac{c_2(v_h, v_h)}{|\mathbf{v}|^2} \leq \gamma_1 \frac{\|v_h\|^2_{H^1(\omega)}}{|\mathbf{v}|^2} \leq k_2\, h^{-2} h^2 \frac{|\mathbf{v}|^2}{|\mathbf{v}|^2} = k_2$$

   with k_2 positive constant independent of h. Hence $\lambda_{\max}(C_2) \leq k_2$, whereas $\lambda_{\min}(C_2) = 0$ since for any constant vector $\mathbf{x}$ we have $C_2 \mathbf{x} = \mathbf{0}$.

3. Finally, we note that B is the matrix associated to the discretization of a second-order elliptic operator using P_1 (or P_2) finite elements in dimension two; it is symmetric positive definite since B is associated to the regularized bilinear form $\tilde{b}(\cdot,\cdot)$ defined in (18). The eigenvalues of B then satisfy:

$$\lambda \;=\; \frac{\mathbf{v}^T B\mathbf{v}}{|\mathbf{v}|^2} = \frac{\tilde{b}(v_h, v_h)}{|\mathbf{v}|^2}$$

and by using (20) and (21) we have

$$k_\eta \frac{\|v_h\|^2_{H^1(\omega)}}{|\mathbf{v}|^2} \le \lambda \le \gamma_2 \frac{\|v_h\|^2_{H^1(\omega)}}{|\mathbf{v}|^2}.$$

Thanks to (27) and (26) we get

$$\widetilde{k}_\eta\, h^2 \le \lambda(B) \le k_3$$

with $k_3, \widetilde{k}_\eta$ positive constants independent of h and $\widetilde{k}_\eta$ dependent on the regularization parameter η.

The proof is complete. □

System (25) is nonsymmetric. Interchanging the first and second block columns of $\mathcal{A}$ leads to the symmetric indefinite system

$$\mathcal{A}\mathcal{Q}(\mathcal{Q}\mathbf{x}) = \mathbf{b}, \quad \text{with } \mathcal{A}\mathcal{Q} = \begin{bmatrix} C_1 & B \\ B & -C_2 \end{bmatrix}, \qquad \text{where} \quad \mathcal{Q} = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}. \tag{29}$$

Corollary 4.2. *The coefficient matrix $\mathcal{A}$ in (25) is nonsingular.*

Proof. The matrix $\mathcal{A}\mathcal{Q}$ in (29) is nonsingular. Indeed, since B is nonsingular it follows that $\mathrm{Ker}(C_1) \cap \mathrm{Ker}(B) = \{\mathbf{0}\}$. The nonsingularity of $\mathcal{A}\mathcal{Q}$ is then a consequence of Lemma 1.1 in [6]. Since $\mathcal{Q}$ is obviously nonsingular, $\mathcal{A}$ must be nonsingular as well. □

5. SOLUTION METHODS

In this section we describe some solution methods for the linear system (25). First we note that we have a choice between working with the original system (25), with the symmetric system (29), or with the nonsymmetric system

$$\mathcal{Q}\mathcal{A}\mathcal{Q}(\mathcal{Q}\mathbf{x}) = \mathcal{Q}\mathbf{b}, \qquad \text{where} \quad \mathcal{Q}\mathcal{A}\mathcal{Q} = \begin{bmatrix} B & -C_2 \\ C_1 & B \end{bmatrix}. \tag{30}$$

Although the symmetric formulation (29) would seem at first to be the most attractive, the singularity of the diagonal blocks C_1 and C_2 causes significant difficulties for both direct and preconditioned iterative methods applied to (29). Indeed, it is very difficult to find effective symmetric preconditioners for (29), which are necessary if one is to use symmetric solvers like SQMR or MINRES (the latter actually requires the preconditioner to be symmetric positive

definite, which is even more problematic); see, e.g., [7]. We mention that all these equivalent formulations lead to ill-conditioned linear systems.

It turns out that a highly effective solution method is obtained by means of a preconditioned Schur complement approach, leading to a nonsymmetric system that can be solved by GMRES [8] in a constant number of iterations. For the description of this approach it is convenient to work with the nonsymmetric formulation (30), which we rewrite more explicitly as

$$\begin{bmatrix} B & -C_2 \\ C_1 & B \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_q \\ \mathbf{b}_p \end{bmatrix}. \tag{31}$$

In the following, we denote with $\mathcal{M}$ the coefficient matrix in (31). This system can be solved using GMRES with a suitable block preconditioner. Using the 'ideal' preconditioner

$$\mathcal{P}_{ideal} = \begin{bmatrix} B & -C_2 \\ 0 & S \end{bmatrix},$$

where $S = B + C_1 B^{-1} C_2$ is the Schur complement, results in a preconditioned matrix $\mathcal{M}\mathcal{P}_{ideal}^{-1}$ with minimum polynomial of degree 2 (see [9]). This implies that GMRES preconditioned with $\mathcal{P}_{ideal}$ converges in at most two steps. It is interesting to observe that although B^{-1} is completely dense, the Schur complement S, while not as sparse as B, is still quite sparse. However, explicitly forming $S = B + C_1 B^{-1} C_2$ is not recommended. Besides efficiency considerations, matrix S is very ill-conditioned and badly scaled, with entries that vary over many orders of magnitude, owing to the presence of many large entries in B^{-1} and in C_1. Even with a state-of-the-art sparse LU factorization [10], the LU factors of S contain a large number of nonzeros originating from the need to perform pivoting to maintain numerical stability. This suggests that one should avoid forming the Schur complement explicitly. Instead, we proceed as follows. Consider the block triangular preconditioner

$$\mathcal{P}_{tr} = \begin{bmatrix} B & -C_2 \\ 0 & B \end{bmatrix}. \tag{32}$$

We have

$$\mathcal{M}\mathcal{P}_{tr}^{-1} = \begin{bmatrix} B & -C_2 \\ C_1 & B \end{bmatrix} \begin{bmatrix} B^{-1} & B^{-1}C_2B^{-1} \\ 0 & B^{-1} \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ C_1B^{-1} & I_n + C_1B^{-1}C_2B^{-1} \end{bmatrix}, \tag{33}$$

hence the spectrum of the preconditioned matrix consists of the eigenvalue $\lambda = 1$ (counted n times) plus the eigenvalues of the matrix $SB^{-1} = I_n + C_1B^{-1}C_2B^{-1}$; a more detailed description of the spectrum of $\mathcal{M}\mathcal{P}_{tr}^{-1}$ is given in the next section. Note that using $\mathcal{P}_{tr}$ as a preconditioner for GMRES applied to the system (31) necessitates the solution of two linear systems with coefficient matrix B at each iteration. Moreover, such an approach requires working with vectors of length $2n$ in each GMRES iteration. The following implementation instead only requires vectors of length n within GMRES: first we find the solution of the block lower triangular system

$$\begin{bmatrix} I_n & 0 \\ C_1B^{-1} & I_n + C_1B^{-1}C_2B^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{y}_q \\ \mathbf{y}_p \end{bmatrix} = \begin{bmatrix} \mathbf{b}_q \\ \mathbf{b}_p \end{bmatrix}, \tag{34}$$

and then we recover the solution of (31) by solving the block upper triangular system

$$\begin{bmatrix} B & -C_2 \\ 0 & B \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_q \\ \mathbf{y}_p \end{bmatrix}.$$

If a sparse Cholesky factorization of B is available, the latter system can be easily solved. Since B represents a discrete elliptic operator in 2D, it can be factored very efficiently and with relatively low fill-in by a sparse Cholesky factorization like the one described in [10].

The solution of the linear system (34) is given by $[\mathbf{b}_q; \mathbf{y}_p]$ where $\mathbf{y}_p$ solves the reduced system

$$(I_n + C_1 B^{-1} C_2 B^{-1})\mathbf{y}_p = \mathbf{b}_p - C_1 B^{-1}\mathbf{b}_q, \tag{35}$$

which can be written as

$$(B + C_1 B^{-1} C_2)B^{-1}\mathbf{y}_p = \mathbf{d}, \qquad \text{where} \quad \mathbf{d} = \mathbf{b}_p - C_1 B^{-1}\mathbf{b}_q. \tag{36}$$

Solving the reduced system (36) with GMRES is equivalent to applying right-preconditioned GMRES to the Schur complement system

$$S\,\mathbf{z}_p = \mathbf{d}, \qquad \mathbf{y}_p = B\mathbf{z}_p,$$

using B as the preconditioner. As shown below, this iteration converges at a rate independent of h. Clearly this requires solving two linear systems with coefficient matrix B at each step, just like GMRES preconditioned by $\mathcal{P}_{tr}$ applied to the unreduced system (31). The advantage of the reduced system approach is that it requires only vectors of length n (rather than $2n$) and this results in very substantial savings already for moderate n. Again, a sparse Cholesky factorization of B (computed once and for all at the outset) can be used to compute the action of B^{-1} on a vector.

Summarizing, the algorithm (which we call $\mathcal{P}_{tr}^S$) is the following:

$$\begin{aligned}
&R = \texttt{chol}(B)\\
&\mathbf{f} = R\backslash(R^T\backslash\mathbf{b}_q);\\
&\mathbf{d} = \mathbf{b}_p - C_1\,\mathbf{f}\\
\mathcal{P}_{tr}^S:\quad &\texttt{solve}\quad (B + C_1 B^{-1} C_2)B^{-1}\mathbf{y}_p = \mathbf{d} \quad \text{with} \quad \text{GMRES}\\
&\mathbf{p} = R\backslash(R^T\backslash\mathbf{y}_p);\\
&\mathbf{q} = \mathbf{f} + R\backslash(R^T\backslash(C_2\,\mathbf{p}))
\end{aligned} \tag{37}$$

where the Matlab-like 'backslash' notation $\mathbf{x} = A\backslash\mathbf{b}$ denotes the solution of $A\mathbf{x} = \mathbf{b}$. Furthermore, in GMRES the coefficient matrix $(B+C_1B^{-1}C_2)B^{-1}$ is not constructed explicitly. Instead its matrices are applied to a vector in sequence; B^{-1} is applied by using its Cholesky factors R and R^T, computed in $R = \texttt{chol}(B)$. In practice, the matrix B is first reordered using an approximate minimum degree (AMD) algorithm [11] before computing the Cholesky factor.

In addition to this 'reduced system' approach we also tested the use of GMRES on the whole system (31) with preconditioner $\mathcal{P}_{tr}$ as well as with the following block preconditioners:

$$\mathcal{P}_{tr}^{AMG} = \begin{bmatrix} \widehat{B} & -C_2 \\ 0 & \widehat{B} \end{bmatrix} \qquad \text{Block triang. with} \quad \widehat{B} \approx B \tag{38}$$

$$\mathcal{P}_d = \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} \qquad \text{Exact Block diagonal.} \tag{39}$$

In $\mathcal{P}_{tr}^{AMG}$, the approximation $\widehat{B}$ of B is implicitly defined by replacing the 'exact' inversion of B by an algebraic multigrid (AMG) V-cycle. This method has been implemented using the recently developed AMG code described in [12].

All these iterative methods have been compared with a sparse direct solver directly applied to the system (31). The sparse solver used is the one available as the 'backslash' operator in Matlab, see [10].

6. SPECTRAL ANALYSIS

In this section we study the spectral properties of the preconditioned Schur complement matrix $SB^{-1} = I_n + C_1B^{-1}C_2B^{-1}$, which determine to a large extent the convergence behavior of GMRES with the preconditioners $\mathcal{P}_{tr}^S$ and $\mathcal{P}_{tr}$. We begin with a simple lemma.

Lemma 6.1. *Let $A, B \in \mathbb{C}^n$ be Hermitian positive semidefinite. Then the eigenvalues of $C = AB$ are real and nonnegative.*

Proof. Since A is positive semidefinite, $\widehat{A} := A + \epsilon I_n$ is positive definite for any $\epsilon > 0$. Using the similarity transformation $\widehat{A}^{-1/2}(\widehat{A}B)\widehat{A}^{1/2} = \widehat{A}^{1/2}B\widehat{A}^{1/2}$ and Sylvester's Law of Inertia we observe that $\widehat{A}B$ has the same number of positive and zero eigenvalues as B, and no negative eigenvalues. The desired result is obtained letting $\epsilon \to 0$ and keeping in mind that the eigenvalues of a matrix are continuous functions of the matrix entries. □

The key result is the following.

Theorem 6.2. *The eigenvalues of $C_1B^{-1}C_2B^{-1}$ are real and nonnegative. Moreover, the maximum eigenvalue of $C_1B^{-1}C_2B^{-1}$ satisfies*

$$\lambda_{\max}(C_1B^{-1}C_2B^{-1}) \leq c(\eta) \tag{40}$$

with $c(\eta)$ positive constant dependent only on the regularization parameter η introduced in (18). In particular, $c(\eta)$ is independent of the mesh size h.

Proof. The matrix $C_1B^{-1}C_2B^{-1}$ is the product of two symmetric positive semidefinite matrices, C_1 and $B^{-1}C_2B^{-1}$. It follows from the previous lemma that the eigenvalues of $C_1B^{-1}C_2B^{-1}$ are real and nonnegative. Moreover, from $C_1B^{-1}C_2B^{-1}\mathbf{x} = \lambda\mathbf{x}$ and using the fact that B is symmetric and positive definite, we obtain

$$(B^{-1/2}C_1B^{-1/2})\,(B^{-1/2}C_2B^{-1/2})\mathbf{w} = \lambda\mathbf{w}, \qquad \mathbf{w} = B^{-1/2}\mathbf{x},$$

so that

$$\begin{aligned}
\lambda &= \frac{\mathbf{w}^T(B^{-1/2}C_1B^{-1/2})\,(B^{-1/2}C_2B^{-1/2})\mathbf{w}}{\mathbf{w}^T\mathbf{w}} \\
&\leq \frac{|B^{-1/2}C_1B^{-1/2}\mathbf{w}|}{|\mathbf{w}|}\,\frac{|B^{-1/2}C_2B^{-1/2}\mathbf{w}|}{|\mathbf{w}|} \\
&\leq \lambda_{\max}(B^{-1/2}C_1B^{-1/2})\lambda_{\max}(B^{-1/2}C_2B^{-1/2}).
\end{aligned}$$

We are thus left to show that the largest eigenvalues of $B^{-1/2}C_1B^{-1/2}$, $B^{-1/2}C_2B^{-1/2}$ only depend on the regularization parameter.

An eigenvalue λ of the former matrix satisfies

$$\lambda = \frac{\mathbf{z}^TB^{-1/2}C_1B^{-1/2}\mathbf{z}}{\mathbf{z}^T\mathbf{z}} = \frac{\mathbf{w}^TC_1\mathbf{w}}{\mathbf{w}^TB\mathbf{w}},$$

$2n$	$C_1B^{-1}C_2B^{-1}$	C_1B^{-1}	C_2B^{-1}	C_1	C_2
282	12.6076	67.4967	67.6947	0.0696	26.4185
2266	5.2753	67.7020	85.7526	0.0232	55.4314
8698	5.4098	67.7024	167.2627	0.0116	118.6824
23878	5.4035	67.7024	187.1614	0.0070	170.1218
60906	5.4048	67.7024	192.5934	0.0043	166.1686
116882	5.4042	67.7024	194.8736	0.0032	177.3061

Table I. Maximum eigenvalue of various matrices, scaled by 10^{-4}.

with $\mathbf{w} = B^{-1/2}\mathbf{z}$. Let w_h be the unique element in V_h such that its nodal values are the components w_i of the vector $\mathbf{w}$, i.e. $w_h(x) = \sum_i w_i\phi_i(x)$, then

$$\lambda = \frac{\mathbf{w}^T C_1 \mathbf{w}}{\mathbf{w}^T B \mathbf{w}} = \frac{{}_{V'}\langle \mathcal{C}_1 w_h, w_h\rangle_V}{{}_{V'}\langle \mathcal{B} w_h, w_h\rangle_V} \leq \frac{\bar{k}\,\|\mathcal{C}_1 w_h\|_{(H^1(\omega))'}\|w_h\|_{H^1(\omega)}}{k_\eta \|w_h\|^2_{H^1(\omega)}} \tag{41}$$

with $\bar{k}$ positive constant and k_η coercivity constant defined in (21). By using the definition of $\mathcal{C}_1$ and standard trace inequalities we get

$$\|\mathcal{C}_1 w\|_{(H^1(\omega))'} := \sup_{v\in H^1(\omega), \|v\|\neq 0} \frac{\frac{1}{\alpha}\int_{\partial\omega} wv\, d\sigma}{\|v\|_{H^1(\omega)}} \leq k_1 \|w\|_{H^{1/2}(\partial\omega)} \leq k_1 \|w\|_{H^1(\omega)},$$

hence

$$\lambda = \frac{\mathbf{w}^T C_1 \mathbf{w}}{\mathbf{w}^T B \mathbf{w}} \leq \frac{k_1\,\bar{k}\,\|w_h\|^2_{H^1(\omega)}}{k_\eta \|w_h\|^2_{H^1(\omega)}} = k(\eta), \tag{42}$$

that is

$$\lambda_{\max}\left(C_1 B^{-1}\right) \leq k(\eta) \tag{43}$$

with $k(\eta)$ positive constant dependent only on η.

We proceed in a similar manner for the eigenvalues λ of $B^{-1/2}C_2B^{-1/2}$. Setting $\mathbf{w} = B^{-1/2}\mathbf{z}$ and thanks to (17) we can write

$$\lambda = \frac{\mathbf{z}^T B^{-1/2} C_2 B^{-1/2}\mathbf{z}}{\mathbf{z}^T\mathbf{z}} = \frac{\mathbf{w}^T C_2 \mathbf{w}}{\mathbf{w}^T B \mathbf{w}} = \frac{c_2(w_h, w_h)}{\tilde{b}(w_h, w_h)} \leq \kappa \frac{\tilde{b}(w_h, w_h)}{\tilde{b}(w_h, w_h)} = \kappa \tag{44}$$

that is

$$\lambda_{\max}\left(C_2 B^{-1}\right) \leq \kappa$$

with κ constant independent of η and h. This completes the proof of the result. □

In Table I we report the maximum eigenvalue of the matrices $C_1B^{-1}C_2B^{-1}$, C_1B^{-1}, C_2B^{-1}, C_1 and C_2 for a sequence of problems of increasing size. The matrix C_1 was first rescaled by dividing it by 10^4. These results are clearly in keeping with our theory.

Let $r_1 := \mathrm{rank}(C_1) = \mathrm{rank}(C_1B^{-1})$ and $r_2 := \mathrm{rank}(C_2) = \mathrm{rank}(C_2B^{-1})$. Furthermore, let $r := \min\{r_1, r_2\}$. From the inequality

$$\mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\},$$

n	C_1	C_2	$C_1B^{-1}C_2B^{-1}$
141	40	10	10
1133	120	44	41
4349	240	136	67

Table II. Rank of C_1, C_2 and of $C_1B^{-1}C_2B^{-1}$.

which holds for any two matrices A and B for which the product AB is well-defined, we obtain the following corollary.

Corollary 6.3. *The preconditioned Schur complement matrix $SB^{-1} = I_n + C_1B^{-1}C_2B^{-1}$ has the eigenvalue $\lambda = 1$ with multiplicity at least $n - r$ with the remaining eigenvalues satisfying $1 < \lambda \leq 1 + c(\eta)$, where the constant $c(\eta)$ is independent of the mesh size h.*

The sparsity structure of C_1 and C_2 is such that $C_1C_2 = 0$. This is because C_1 is nonzero only on the boundary of ω whereas C_2 is nonzero only on the nodes corresponding to the experimental measurements of the wind velocity V_i. These nodes are inside the domain ω and and far from the boundary, thus yielding $C_1C_2 = 0$. This leads to significant cancellation in the product $C_1B^{-1}C_2B^{-1}$, and it turns out that the rank of $C_1B^{-1}C_2B^{-1}$ is actually much less than $r = \min\{r_1, r_2\}$ for n sufficiently large. Hence, almost all the eigenvalues of SB^{-1} are equal to 1. The remaining ones are confined in a finite interval (independent of h and bounded below by 1), and their number grows slowly with the number n of degrees of freedom; see Table II.

Remark 6.4. *Additional results have been obtained for the block diagonal preconditioner $\mathcal{P}_d$ and for inexact variants of the block triangular preconditioner $\mathcal{P}_{tr}$. However, the performance of these preconditioners was found to be inferior to that of $\mathcal{P}_{tr}^S$. For this reason we do not report the results of our analysis.*

Remark 6.5. *The block triangular preconditioner $\mathcal{P}_{tr}$, and therefore $\mathcal{P}_{tr}^S$, can be interpreted as a* constraint preconditioner *applied to the symmetric indefinite system (29). This follows from the identity*

$$\begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \begin{bmatrix} B & -C_2 \\ C_1 & B \end{bmatrix} \begin{bmatrix} B & -C_2 \\ 0 & B \end{bmatrix}^{-1} \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} = \begin{bmatrix} C_1 & B \\ B & -C_2 \end{bmatrix} \begin{bmatrix} 0 & B \\ B & -C_2 \end{bmatrix}^{-1}.$$

Note that system (29) can be regarded as a (regularized) saddle point problem. The constraint preconditioner is obtained approximating the low-rank matrix C_1 with the zero matrix; see, e.g., [13, Section 10.2].

7. NUMERICAL RESULTS

In this section we report on a few numerical experiments aimed at assessing the performance of the solvers discussed in the previous sections. All the numerical experiments have been performed in Matlab 7.5.0 (R2007b) on an iMAC 2.66GHz Intel Core 2 Duo, 2 GB 800MHz DDR2 SDRAM - 2×1GB.

We deal with a square domain $\omega = [0,6] \times [0,6]$ (Kms.) and four experimental values of the wind velocity at points $(1.,1.),(5.,1.),(5.,5.),(1.,5.)$. By using the code FreeFEM, unstructured meshes $\mathcal{T}_h$ have been built through a Delaunay-Voronoi triangulation algorithm with V_h the associated space of P_2 finite elements (see Fig. 1).

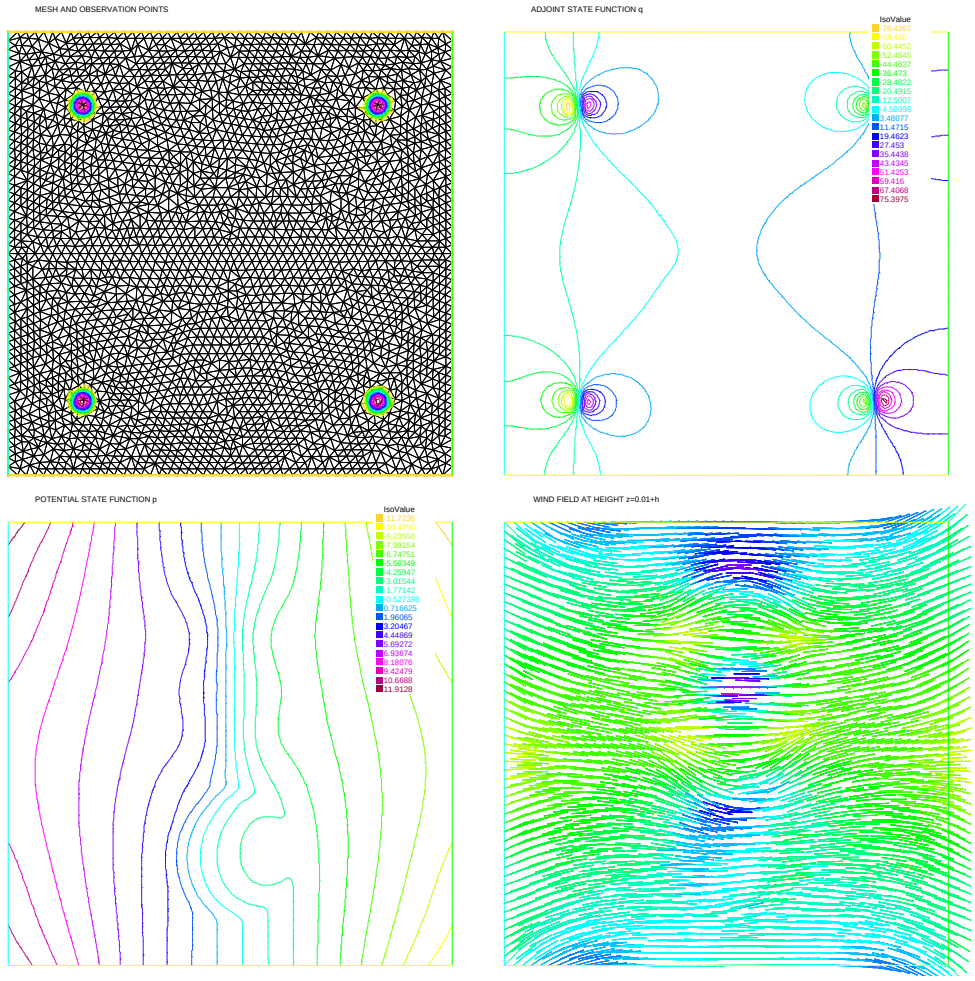

Figure 1. Mesh $\mathcal{T}_h$ and observation points, adjoint state function q, potential state function p and wind field at height $z = 0.01 + H$.

We have compared the preconditioners $\mathcal{P}_{tr}$, $\mathcal{P}_{tr}^S$, $\mathcal{P}_{tr}^{AMG}$ and $\mathcal{P}_d$ with the sparse direct solver ('backslash') in Matlab. For the latter method we apply a symmetric AMD reordering to the diagonal blocks B prior to solving the system, as this was found to improve performance relative to other orderings (or to no reordering at all).

$2n$	$\mathcal{P}_{tr}$	$\mathcal{P}_{tr}^{S}$	$\mathcal{P}_{tr}^{AMG}$	$\mathcal{A}\backslash\mathbf{b}$
2266	(23) 0.329	(25) 0.060	(31) 0.614	0.047
8698	(20) 0.780	(22) 0.342	(28) 1.599	0.180
23878	(20) 1.675	(21) 0.689	(29) 4.315	0.542
60906	(20) 4.130	(21) 1.532	(29) 10.07	1.780
116882	(21) 8.746	(21) 2.851	(29) 19.62	3.843
380010	(21) 34.05	(21) 10.32	(30) 71.40	26.47
592490	(21) 56.94	(21) 17.13	(30) 120.1	33.98

Table III. Iteration count (in parenthesis) and CPU time for $\mathcal{P}_{tr}$, $\mathcal{P}_{tr}^{S}$, $\mathcal{P}_{tr}^{AMG}$ and CPU time for $\mathcal{A}\backslash\mathbf{b}$.

The stopping criterion used was as follows:

$$\frac{|\mathbf{b} - \mathcal{A}\mathbf{x}_k|}{|\mathbf{b}| + \|\mathcal{A}\|_\infty |\mathbf{x}_k|} < tol$$

where $\|\cdot\|_\infty$ denotes the infinity norm of a matrix, and *tol* is chosen so that the relative error in the computed approximation $\mathbf{x}_k$ satisfies

$$\frac{|\mathbf{x}_* - \mathbf{x}_k|}{|\mathbf{x}_*|} \approx 10^{-5}.$$

Here the 'exact' solution $\mathbf{x}_*$ is the one computed by the direct method. The resulting values of *tol* where $tol = 10^{-10}$ for $\mathcal{P}_{tr}$, $tol = 10^{-9}$ for $\mathcal{P}_{tr}^{S}$, and $tol = 10^{-12}$ for $\mathcal{P}_{tr}^{AMG}$. These small values of *tol* reflect the ill-conditioning of the linear systems under consideration.

In Table III we report GMRES iteration counts (in parentheses) and CPU times for a sequence of problems of increasing size. We do not include the results for the block diagonal preconditioner $\mathcal{P}_d$ because it was found to require about twice the number of iterations (and CPU time) as the block triangular one.

It is clear from the results in Table III that the iterative solvers all exhibit h-independent convergence rates. The preconditioners $\mathcal{P}_{tr}$ and $\mathcal{P}_{tr}^{S}$ are mathematically equivalent; the small difference in the iteration counts for the four smallest problems is likely due to implementation details and round-off effects. Note, however, the striking difference in CPU time due to the use of vectors of half the size with the Schur complement reduction approach. In terms of CPU time, the direct solver is best only for problem sizes up to 23878, whereas $\mathcal{P}_{tr}^{S}$ is the winner for all the remaining problems. For the underlying application, $2n = 592490$ is a realistic problem size, therefore $\mathcal{P}_{tr}^{S}$ is the method of choice, requiring only half the time as its closest competitor.

We also note that the 'optimal' AMG-based preconditioner $\mathcal{P}_{tr}^{AMG}$ is actually not competitive due to the high cost of each preconditioned iteration compared with the (sub-optimal!) Cholesky-based preconditioners. We tried using different variants of this approach with different convergence tolerances but the performance of this preconditioner was also significantly inferior, in terms of CPU time, to that of $\mathcal{P}_{tr}^{S}$. An attempt was also made to replace the Cholesky factorization with an AMG inner iteration in the solution of the Schur complement system, but the results were not good.

Set-up times were found to be very small both for the Cholesky-based preconditioners and for the AMG-based one, accounting in all cases for a negligible fraction of total solution time.

η	$\lambda_{\min}(B)$	$\lambda_{\max}(C_1B^{-1}C_2B^{-1})$	$\lambda_{\max}(C_1B^{-1})$	$\lambda_{\max}(C_2B^{-1})$	# its $\mathcal{P}_{tr}^S$
10	0.0419	7.05e-04	1.41e+03	149.0	2
1	0.0075	135.7	4.46e+03	164.6	14
1.e-1	8.19e-04	9.77e+03	1.46e+04	167.0	23
1.e-2	8.27e-05	3.78e+04	7.67e+04	167.2	22
1.e-3	8.28e-06	5.41e+04	6.77e+05	167.3	22
1.e-4	8.27e-07	5.66e+04	6.68e+06	167.3	22
1.e-5	8.28e-08	5.69e+04	6.67e+07	167.3	22

Table IV. Minimum eigenvalue of B, maximum eigenvalues of $C_1B^{-1}C_2B^{-1}$, C_1B^{-1}, C_2B^{-1} and number of iterations required by $\mathcal{P}_{tr}^S$ for $2n = 8968$ and different values of η defined in (18).

Next, we assess the robustness of the $\mathcal{P}_{tr}^S$ solver with respect to the regularization parameter η. In Table IV we display the minimum eigenvalue of B, the maximum eigenvalue of C_1B^{-1}, C_2B^{-1} and $C_1B^{-1}C_2B^{-1}$, and the number of iterations required by GMRES preconditioned with $\mathcal{P}_{tr}^S$ for different values of η. The number of degrees of freedom is fixed ($2n = 8968$). The results show that $\mathcal{P}_{tr}^S$ is essentially insensitive to the value of η.

A further advantage of the Schur complement approach over the direct solver applied to the unreduced system is that it only requires the factorization of B. The matrices C_1, C_2 enter the computation only in the form of matrix-vector products. Hence, if in the course of the simulation some of the entries of either C_1 or C_2 change, no additional computations are needed for the Schur complement preconditioner. In contrast, with the direct solver the entire LU factorization of $\mathcal{A}$ must be computed anew, at a significant cost. We note that changes in C_1 occur whenever the locations $\mathbf{x}_i$ of the wind measurements change.

Finally, we have performed a few experiments with an augmented variant of GMRES. Augmented GMRES [14] is known to significantly improve the performance of GMRES when convergence delay is due to the presence of a separated group of eigenvalues. If available, good approximate eigenvectors may be injected in the approximation space so as to spare GMRES their dynamic approximation, resulting in the mitigation of a possible stagnation phase. The process can also be applied in a restarting setting, although we have not exploited this possibility here. Our spectral analysis in Corollary 6.3 shows that the coefficient matrix of the reduced system (35) has indeed a group of eigenvalues well separated from the rest of the spectrum, which GMRES takes a few iterations to 'spot' (see the near-stagnation phase in the solid curve of Figure 2). Whenever a group of (approximate) relevant eigenvectors is available from the problem, the use of the Augmented method may be highly beneficial. This is reported in the case $2n = 2266$ in Figure 2, where the dash-dotted line shows the convergence curve of the Augmented strategy when the exact eigenvectors corresponding to the 10 largest eigenvalues are included in the approximation space. The dashed line shows the convergence curve when the eigenvectors corresponding to 10 out of the 20 largest eigenvalues are included (those with odd index). Clearly, the inclusion of some of the largest 20 eigenvectors practically eliminates the plateau phase (dashed curve), whereas the inclusion of only the largest ten (which span a much smaller real interval) requires GMRES to capture more eigenvectors before the asymptotic linear behavior takes place. Note however that in this latter case, few eigenvectors are sufficient to initially drastically decrease the residual norm.

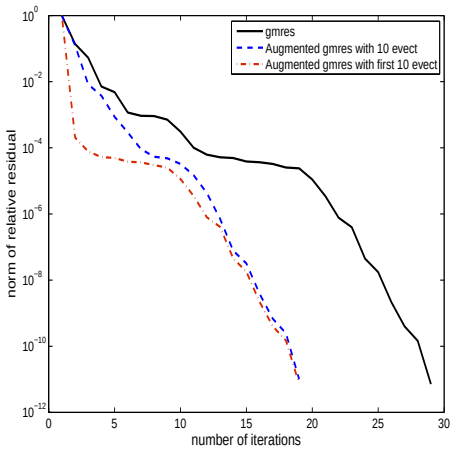

Figure 2. Number of iterations to converge for GMRES (solid line) and for Augmented GMRES. Dash-dotted line: eigenvectors corresponding to the largest 10 eigenvalues are included. Dashed line: eigenvectors corresponding to the largest 10 odd indexed eigenvalues are included.

8. CONCLUSIONS

We have investigated the solution of large linear systems with block 2-by-2 structure resulting from finite element discretizations of coupled systems of elliptic PDEs arising from a class of optimal control problems. We have shown that a combination of Schur complement reduction, GMRES and suitable preconditioning leads to h-independent convergence rates and is superior, in terms of CPU time, to several other approaches including a state-of-the-art sparse direct solver. The convergence of the iterative solver was also found to be independent of the choice of the regularization parameter. In particular, the fact that the reduced system approach works on vectors of length n (rather than $2n$) was found to result in very substantial savings in terms of CPU times over the other methods tested. While the original control problem considered in this paper is rather special, it is possible that the methods and results of this paper will find application to similarly structured problems arising in other areas of scientific computing. Indeed, in many optimal control problems governed by partial differential equations the use of the adjoint state method leads to large, sparse systems of linear equations with the same 2-by-2 block structure as (31); see, e.g., [4]. Moreover, if the control is on the boundary some of the matrices involved will have similar properties to those considered here.

ACKNOWLEDGEMENTS

This work was performed while the forth author was visiting the TU Berlin. Prof. Volker Mehrmann's warm hospitality is gratefully acknowledged. The research was partially supported by *Deutsche Forschungsgemeinschaft*, via the DFG Research Center MATHEON, Mathematics for Key Technologies, in Berlin and by Berlin Mathematical School.

REFERENCES

1. Asensio MI, Ferragut L, Simon J. A convection model for fire spread simulation. *Applied Mathematics Letters* 2005; **18**:673–677.
2. Ferragut L, Asensio MI, Simon J. 3D Wind field adjustment performing only 2D computations including thermal effects. *Preprint submitted to Comm. Numer. Meth. Engng.*, 2009.
3. Lions JL. *Contrôle Optimal de Systèmes Gouvernés par des Équations aux Dérivées Partielles*. Dunod, Paris, 1968.
4. Gunzburger M. *Perspectives in Flow Control and Optimization*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2003.
5. Pironneau O, Hecht F, Hyaric, AL. FreeFEM. `http://www.freefem.org/`.
6. Benzi M, Golub GH. A preconditioner for generalized saddle point problems. *SIAM Journal on Matrix Analysis and Applications* 2004; **26**:20–41.
7. Simoncini V., Szyld DB. Recent computational developments in Krylov Subspace Methods for linear systems. *Numerical Linear Algebra with Applications* 2007; **14**:1–59.
8. Saad Y, Schultz MH. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* 1986; **7**:856–869.
9. Ipsen ICF. A note on preconditioning nonsymmetric matrices. *SIAM Journal on Scientific Computing* 2001; **23**:1050–1051.
10. Davis TA. *Direct Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, 2006.
11. Amestoy PR, Davis TA, Duff IS. An approximate minimum degree ordering algorithm. *SIAM Journal on Matrix Analysis and Applications* 1996; **17**: 886–905.
12. Boyle J, Mihajlovic MD, Scott JA. *HSL-MI20: An Efficient AMG Preconditioner*. Technical Report RAL-TR-2007-021, Rutherford Appleton Laboratory, December 2007.
13. Benzi M, Golub GH, Liesen J. Numerical solution of saddle point problems. *Acta Numerica* 2005; **14**: 1–137.
14. Morgan RB. GMRES with deflated restarting. *SIAM Journal on Scientific Computing*, 2002; **24**: 20–37.