

A TAXONOMY OF CONSISTENTLY STABILIZED FINITE ELEMENT METHODS FOR THE STOKES PROBLEM

TERI BARTH^{1,4}, PAVEL BOCHEV^{1,4} MAX GUNZBURGER², AND JOHN SHADID^{3,4},

Abstract. Stabilized mixed methods can circumvent the restrictive inf-sup condition without introducing penalty errors. For properly chosen stabilization parameters these methods are well-posed for all conforming velocity-pressure pairs. However, their variational forms have widely varying properties. First, stabilization offers a choice between weakly or strongly coercive bilinear forms that give rise to linear systems with identical solutions but very different matrix properties. Second, coercivity may be conditional upon a proper choice of a stabilizing parameter. Here we focus on how these two aspects of stabilized methods affect their accuracy and efficient iterative solution. We present results that indicate a preference of Krylov subspace solvers for strongly coercive formulations. Stability criteria obtained by finite element and algebraic analyses are compared with numerical experiments. While for two popular classes of stabilized methods sufficient stability bounds correlate well with numerical stability, our experiments indicate the intriguing possibility that the pressure stabilized Galerkin method is unconditionally stable.

Key words. stabilized finite element methods, mixed methods, iterative solvers, coercive forms

AMS subject classifications. 76D05, 76D07, 65F10, 65F30

1. Introduction. It is well-known that stable and accurate mixed methods for the Stokes problem require finite element spaces that satisfy the LBB compatibility condition (see, e.g., [5, 13, 15]). This restrictive condition can be circumvented by regularization of the mixed problem. Regularizations that also avoid penalty errors are known as *stabilized Galerkin methods*; see [2, 6, 8, 10, 11, 19, 20, 21]. Stabilized methods achieve consistency by embedding the stabilizing terms into expressions that vanish on all sufficiently smooth solutions. This necessarily makes such methods dependent upon various stabilizing parameters whose purpose is to blend the regularizing terms with the mixed Galerkin form.

During the last two decades, stabilized methods have attracted significant attention from mathematicians and engineers and today they are widely used in practice. However, despite their popularity, very few studies have addressed the impact of formulation choice and regularization parameters upon performance of iterative solvers, finite element stability and convergence rates. For instance, finite element analyses give sufficient but not necessary stability conditions and, as a result, they can lead to pessimistic estimates of stability ranges for the parameters. The sharpness of these bounds is not known, nor is it well documented numerically. There's also an ambiguity in the choice of stabilized variational forms. Some authors insist on using strongly coercive formulations (see, e.g., [6]) while others emphasize symmetry (see [19].) The extant literature gives very few clues about these issues and more importantly, about the relative advantages and disadvantages of the different methods. To add to the

¹Computational Mathematics and Algorithms Department, Sandia National Laboratories, Albuquerque, NM 87185-1110 (`{tbarth,pboche}@sandia.gov`).

²School of Computational Science and Information Technology, Florida State University, Tallahassee FL 32306-4120 `gunzburg@csit.fsu.edu`. Supported in part by CSRI, Sandia National Laboratories under contract 18407.

³Computational Sciences Department, Sandia National Laboratories, Albuquerque, New Mexico 87185. `jshadi@sandia.gov`

⁴Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC-94AL85000.

confusion, in many cases conclusions are drawn from numerical results computed by implementations that do not correspond to the formulations they are supposed to represent. A typical example is the use of an iterative solver that scales a matrix row whenever a negative diagonal entry is encountered. This seemingly innocuous trick can effectively switch the formulation from one based on a weakly coercive form to one based on a strongly coercive form! Equally widespread is the tendency to confuse the original Galerkin Least-Squares (GLS) formulation of [19] with a method that can be derived from the Douglas-Wang stabilization introduced in [8]. While the two variational equations do indeed look a lot like each other and can be easily mistaken, their properties differ quite substantially. The GLS form is symmetric and weakly coercive. In contrast, the form derived from [8] is non-symmetric and strongly coercive.

With this paper we aim to rectify this situation by first giving a rigorous taxonomy of consistently stabilized methods and second, by conducting a study of issues related to their numerical performance. Our main focus is on the finite element stability and accuracy of different methods and the performance of Krylov subspace solvers for the solution of the discrete systems.

At this point, we should define what we mean by a *consistent* method; perhaps a more apt terminology would be *variationally consistent*. In standard usage, consistency of numerical schemes for partial differential equations requires that the *pointwise* truncation error vanishes as the grid size goes to zero, i.e., if one substitutes a smooth solution of the partial differential equation into the numerical scheme, then the residual is at least $o(h)$, where h denotes the grid size. Finite element schemes are not, in general, consistent in this sense. However, for standard finite element methods, sufficiently smooth exact solutions of the partial differential equations *exactly* satisfy the variational equation that defines the discrete finite element equations. This is what we mean by a consistent finite element scheme. This allows us to differentiate between the methods we consider in this paper and methods which are not consistent in this latter sense. For example, penalty methods for the Stokes problem are not consistent finite element methods since substitution of an exact solution into the discrete equations leaves a residual that is proportional to the penalty parameter. Thus, we only consider methods that do not suffer from this type of variational inconsistency.

We subdivide consistently stabilized methods into three classes. Each class contains two complementary methods, one that involves a weakly coercive, and possibly symmetric bilinear form, the other a strongly coercive but non-symmetric form. The complementary methods lead to linear systems with identical solutions but different matrix properties. As a result, matrix problems engendered by consistent stabilization can vary from symmetric indefinite systems that are conditionally stable to non-symmetric, unconditionally positive definite problems. This means that successful implementation of a stabilized method requires the proper selection of parameter values so as to choose one of the three classes of methods and, whenever iterative solvers are employed, the proper selection of one of the two complementary methods within the class. In this paper, we provide information to help make these choices.

We have organized the paper as follows. Section 2 develops the taxonomy of consistently stabilized methods, including their precise definition and nomenclature. It is shown that consistently stabilized methods originate from a single regularized bilinear form by virtue of a proper choice of parameters. Finite element stability of the methods is discussed in Section 3, where we also show how stability ranges can be estimated by algebraic methods. Section 4 summarizes our numerical experiments on

finite element convergence rates and the performance of unpreconditioned GMRES and QMR solvers. The paper concludes, in Section 5, with a brief summary of our experiences.

1.1. Notations. In what follows, we will make use of the usual notations $H^p(\Omega)$, $\|\cdot\|_p$, and $(\cdot, \cdot)_p$ with p a non-negative integer to denote the Sobolev spaces consisting of all functions having square integrable derivatives up to order p on Ω , the standard Sobolev norm, and inner product, respectively. When $p = 0$, we will write $L^2(\Omega)$ instead of $H^0(\Omega)$ and drop the index from the inner product designation. The symbol $|\cdot|_k$, $0 < k \leq p$ will denote the k th seminorm on $H^p(\Omega)$. As usual, $H_0^p(\Omega)$ will denote the closure of $C_0^\infty(\Omega)$ with respect to the norm $\|\cdot\|_p$ and $H^{-p}(\Omega)$ will denote the dual space of $H_0^p(\Omega)$. Vector-valued functions are denoted by bold-face font, e.g., \mathbf{u} and \mathbf{v} . Vectors in Euclidean spaces are denoted by vector notation, e.g., $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ and the Euclidean inner product and norm are denoted by $\langle \cdot, \cdot \rangle$ and $|\cdot|$, respectively, e.g., $\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \rangle = \vec{\mathbf{x}}^* \vec{\mathbf{y}}$ and $|\vec{\mathbf{x}}| = \langle \vec{\mathbf{x}}, \vec{\mathbf{x}} \rangle^{1/2}$. Matrices are denoted by block letters, e.g., \mathbb{A} and \mathbb{B} .

The symbol S_p^h denotes a space of continuous, piecewise polynomial functions defined with respect to a regular subdivision \mathcal{T}_h of the domain Ω into finite elements \mathcal{K} . For example, \mathcal{K} can be hexahedrons or tetrahedrons in three dimensions or triangles or quadrilaterals in two dimensions. It is assumed that for every $v \in H^{p+1}(\Omega)$ there exists $v^h \in S_p^h$ such that

$$(1.1) \quad \|v - v^h\|_0 + h\|v - v^h\|_1 \leq Ch^{p+1}\|v\|_{p+1}.$$

We will also need the inverse inequality

$$(1.2) \quad \|v^h\|_1 \leq C_I h^{-1} \|v^h\|_0$$

that holds for C^0 finite element spaces on regular subdivisions; see [7]. Lastly, we recall Poincaré's inequality (see, e.g., [13]) in the form

$$(1.3) \quad C_P \|u\|_1 \leq |u|_1$$

that holds for all functions belonging to $H^p(\Omega) \cap H_0^1(\Omega)$.

1.2. Galerkin mixed methods for the Stokes equations. We consider the incompressible Stokes equations

$$(1.4) \quad -\nu \Delta \mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega$$

$$(1.5) \quad \nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega$$

$$(1.6) \quad \mathbf{u} = \mathbf{0} \quad \text{on } \Gamma.$$

In (1.4)–(1.6), Ω denotes a bounded, open region in \mathcal{R}^d , $d = 2, 3$, with Lipschitz continuous boundary Γ . A weak formulation of the Stokes problem is to seek $(\mathbf{u}, p) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega)$ such that

$$(1.7) \quad A(\mathbf{u}, \mathbf{v}) + B(\mathbf{v}, p) = F(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega)$$

$$(1.8) \quad B(\mathbf{u}, q) = 0 \quad \forall q \in L_0^2(\Omega),$$

where $A(\cdot, \cdot)$, $B(\cdot, \cdot)$ and $F(\cdot)$ are defined by

$$A(\mathbf{u}, \mathbf{v}) = \int_{\Omega} \nu \nabla \mathbf{u} : \nabla \mathbf{v} \, d\Omega, \quad B(\mathbf{v}, p) = - \int_{\Omega} p \nabla \cdot \mathbf{v} \, d\Omega \quad \text{and} \quad F(\mathbf{v}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, d\Omega,$$

respectively. We recall that (1.7)–(1.8) is the optimality system for the saddle-point (\mathbf{u}, p) of the Lagrangian functional

$$(1.9) \quad L(\mathbf{v}, q) = \frac{1}{2}A(\mathbf{v}, \mathbf{v}) - F(\mathbf{v}) + B(\mathbf{v}, q).$$

Therefore, the pressure p is the Lagrange multiplier that is introduced into (1.9) to enforce the (weak) incompressibility constraint (1.8).

The *Galerkin mixed method* for the Stokes equations is obtained by restricting (1.7)–(1.8) to a pair of finite element subspaces $\mathbf{V}^h \subset \mathbf{H}_0^1(\Omega)$ and $S^h \subset L_0^2(\Omega)$. The corresponding discrete problem is to seek $(\mathbf{u}^h, p^h) \in \mathbf{V}^h \times S^h$ such that

$$(1.10) \quad A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) = F(\mathbf{v}^h) \quad \forall \mathbf{v}^h \in \mathbf{V}^h$$

$$(1.11) \quad B(\mathbf{u}^h, q^h) = 0 \quad \forall q^h \in S^h.$$

Evidently, (1.10)–(1.11) is the optimality system for the saddle-point (\mathbf{u}^h, p^h) of (1.9) out of $\mathbf{V}^h \times S^h$. Because (1.10)–(1.11) is a discrete saddle-point problem, it leads to stable and accurate approximations of (\mathbf{u}, p) if and only if the pair (\mathbf{V}^h, S^h) satisfies the following conditions. First, the inf-sup condition (see [5, 13, 15]): there exists $\gamma > 0$, independent of h , such that

$$(1.12) \quad \sup_{\mathbf{v}^h \in \mathbf{V}^h} \frac{B(\mathbf{v}^h, q^h)}{\|\mathbf{v}^h\|_1} \geq \gamma \|q^h\|_0 \quad \forall q^h \in S^h.$$

Second, A is coercive on $\mathbf{Z}^h \times \mathbf{Z}^h$, where $\mathbf{Z}^h = \{\mathbf{v}^h \in \mathbf{V}^h \mid B(q^h, \mathbf{v}^h) = 0 \quad \forall q^h \in S^h\}$ is the subspace of discretely solenoidal functions belonging to \mathbf{V}^h . Examples of unstable pairs include equal order finite elements on the same mesh and some other combinations such as the bilinear-constant pair; see [13, 15].

Problem (1.10)–(1.11) is equivalent to the symmetric, indefinite linear system

$$(1.13) \quad \begin{pmatrix} \mathbb{A} & \mathbb{B}^T \\ \mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{\mathbf{u}} \\ \vec{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \vec{\mathbf{f}} \\ \vec{\mathbf{0}} \end{pmatrix},$$

where the elements of $\vec{\mathbf{u}}$ and $\vec{\mathbf{p}}$ are the coefficients in the representation in terms of bases of the finite element pair (\mathbf{u}^h, p^h) ; the matrices \mathbb{A} and \mathbb{B} are deduced in the usual manner, using the bases for \mathbf{V}^h and S^h , from the bilinear forms $A(\cdot, \cdot)$ and $B(\cdot, \cdot)$, respectively.

If the second equation in (1.10)–(1.11) is multiplied by -1, we obtain the linear system

$$(1.14) \quad \begin{pmatrix} \mathbb{A} & \mathbb{B}^T \\ -\mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{\mathbf{u}} \\ \vec{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \vec{\mathbf{f}} \\ \vec{\mathbf{0}} \end{pmatrix}.$$

This system has the same solution as (1.13) but is non-symmetric. Thus, iterative methods may perform differently on (1.13) and (1.14). The dichotomy between the algebraically equivalent systems (1.13) and (1.14) will be relevant to our discussions about stabilized methods. For this reason, we will (using the acronym GMM for Galerkin mixed methods) refer to (1.13) and (1.14) as the GMM+ and GMM– methods, respectively.

For future reference, note that (1.7)–(1.8) can be rewritten in the form

$$(1.15) \quad Q_{gmm}(\mathbf{u}, p; \mathbf{v}, q) = F(\mathbf{v}) \quad \text{for all } (\mathbf{v}, q) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega),$$

where F is the same as before and

$$(1.16) \quad Q_{gmm}(\mathbf{u}, p; \mathbf{v}, q) = A(\mathbf{u}, \mathbf{v}) + B(\mathbf{v}, p) + B(\mathbf{u}, q).$$

2. Consistently stabilized methods for the Stokes equations. To define consistently stabilized mixed finite element methods for the Stokes equation, we consider families of bilinear forms

$$(2.1) \quad \begin{aligned} Q_{\alpha\beta}(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h) &= A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) + \beta B(\mathbf{u}^h, q^h) \\ &\quad - \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (-\Delta \mathbf{u}^h + \nabla p^h, -\alpha \Delta \mathbf{v}^h + \beta \nabla q^h)_{0, \mathcal{K}} \end{aligned}$$

and linear functionals

$$(2.2) \quad F_{\alpha\beta}(\mathbf{v}^h, q^h) = F(\mathbf{v}^h) - \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (\mathbf{f}, -\alpha \Delta \mathbf{v}^h + \beta \nabla q^h)_{0, \mathcal{K}}$$

parametrized by α , β , and δ ; α and β take on the values $\{-1, 0, 1\}$ and $\{-1, 1\}$, respectively, and δ is a positive, real valued parameter. The stabilized methods we consider are then as follows: find $(\mathbf{u}^h, p^h) \in \mathbf{V}^h \times S^h$ such that

$$(2.3) \quad Q_{\alpha\beta}(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h) = F_{\alpha\beta}(\mathbf{v}^h, q^h) \quad \text{for all } (\mathbf{v}^h, q^h) \in \mathbf{V}^h \times S^h.$$

Comparing (2.1)–(2.3) with (1.15)–(1.16), one can easily identify the terms added to the latter to effect stabilization.

Introducing the bilinear forms

$$D(\mathbf{u}^h, \mathbf{v}^h) = \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (-\Delta \mathbf{u}^h, -\Delta \mathbf{v}^h)_{0, \mathcal{K}}, \quad C(\mathbf{v}^h, q^h) = \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (\nabla q^h, \Delta \mathbf{v}^h)_{0, \mathcal{K}},$$

and

$$K(p^h, q^h) = \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (\nabla p^h, \nabla q^h)_{0, \mathcal{K}}$$

defined on $\mathbf{V}^h \times \mathbf{V}^h$, $\mathbf{V}^h \times S^h$, and $S^h \times S^h$, respectively, we can write (2.1) as

$$\begin{aligned} Q_{\alpha\beta}(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h) &= A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) + \beta B(\mathbf{u}^h, q^h) \\ &\quad - \alpha D(\mathbf{u}^h, \mathbf{v}^h) + \alpha C(\mathbf{v}^h, p^h) + \beta C(\mathbf{u}^h, q^h) - \beta K(p^h, q^h). \end{aligned}$$

It is then easy to see that the discrete problem (2.3) is equivalent to a family of linear algebraic systems of the form

$$(2.4) \quad \begin{pmatrix} \mathbb{A} - \alpha \mathbb{D} & (\mathbb{B} + \alpha \mathbb{C})^T \\ \beta(\mathbb{B} + \mathbb{C}) & -\beta \mathbb{K} \end{pmatrix} \begin{pmatrix} \vec{\mathbf{u}} \\ \vec{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \vec{\mathbf{f}}_1 \\ \vec{\mathbf{f}}_2 \end{pmatrix},$$

where the matrices \mathbb{C} , \mathbb{D} , and \mathbb{K} are respectively deduced in the usual manner from the bilinear forms $C(\cdot, \cdot)$, $D(\cdot, \cdot)$, and $K(\cdot, \cdot)$.

Choosing different α and β gives rise to different bilinear forms in (2.1) and to different matrices in (2.4). It is clear that changing β to $-\beta$ is the same as changing q^h to $-q^h$ in (2.3), or scaling the last row in (2.4) by -1 . Therefore, the problems corresponding to $\{\alpha, \beta\}$ and $\{\alpha, -\beta\}$ are equivalent in the sense that their solutions are identical. We will refer to these problems and the associated forms and matrices *complementary*.

REMARK 1. Writing (2.3) in the form

$$A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) + \beta B(\mathbf{u}^h, q^h) - \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (-\Delta \mathbf{u}^h + \nabla p^h - \mathbf{f}, -\alpha \Delta \mathbf{v}^h + \beta \nabla q^h)_{0, \mathcal{K}} = F(\mathbf{v}^h),$$

shows that (2.3) is satisfied by all sufficiently smooth solutions of (1.4). This type of *consistency*, which is a hallmark of stabilized methods we consider, is obtained by using residuals of the momentum equation to effect stabilization.

REMARK 2. For affine families of linear or bilinear finite element spaces, the second-order terms in (2.1) vanish and (2.3) reduces to a penalty method with penalty parameter proportional to h^2 ; see [6]. In this case, the consistency error can be reduced to within discretization error by employing a projection; see [21]. However, numerical studies of stability ranges require all residual terms to be represented in the discretization. Consequently, we restrict attention to quadratic and higher-order finite element spaces, even though on unstructured meshes second order terms will not vanish for bilinear elements.

REMARK 3. Stabilized methods lead to well-posed discrete problems for a wide range of velocity-pressure pairs, including equal order C^0 spaces and other combinations such as C^0 velocities and discontinuous pressures. The latter pairs however, have not gained much acceptance in stabilized methods because they require the use of nonstandard assembly process and face-based data structures; see [12]. Indeed, one of the main appeals of stabilized methods is that they allow for the use of *standard* C^0 , equal-order spaces for all variables. This greatly simplifies the code development and solution processes and so here we restrict attention to these type of spaces.

REMARK 4. Many of the methods defined by (2.3) do not represent an optimality system of some regularized Lagrangian and can only be derived as modifications of (1.10)–(1.11). Such methods cannot be related to optimization problems even though the underlying mixed method can be.

REMARK 5. It is tempting to study, instead of (2.1), the still consistent but even more general bilinear form

$$(2.5) \quad Q_{\alpha\beta_1\beta_2}(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h) = A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) + \beta_1 B(\mathbf{u}^h, q^h) - \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (-\Delta \mathbf{u}^h + \nabla p^h, -\alpha \Delta \mathbf{v}^h + \beta_2 \nabla q^h)_{0, \mathcal{K}},$$

where α is as before and β_1 and β_2 are allowed to take on the values ± 1 . However, whenever $\beta_2 = -\beta_1$, the form (2.5) leads to an unstable method so that we need only consider the case $\beta_1 = \beta_2$, i.e., the form (2.1).

2.1. The Galerkin least-squares class. For $\alpha = 1$, the complimentary forms are

$$(2.6) \quad Q_{gls}^{\pm}(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h) = A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) \pm B(\mathbf{u}^h, q^h) - \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (-\Delta \mathbf{u}^h + \nabla p^h, -\Delta \mathbf{v}^h \pm \nabla q^h)_{0, \mathcal{K}}.$$

For $\beta = 1$, the form (2.6) is symmetric and leads to a symmetric linear system. This case corresponds to the original Galerkin least-squares method of Franca, Hughes, and Balestra [19] denoted here by GLS+. The term “least-squares” can be justified

by noting that the resulting variational equation is the optimality system for the penalized Lagrangian functional

$$(2.7) \quad L_\delta(\mathbf{v}, q) = L(\mathbf{v}, q) - \delta \sum_{\mathcal{K} \in \mathcal{T}_h} \frac{1}{2} h_{\mathcal{K}}^2 \| -\Delta \mathbf{u}^h + \nabla p^h - \mathbf{f} \|_{0, \mathcal{K}}^2.$$

The regularizing contribution

$$(2.8) \quad \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (-\Delta \mathbf{u}^h + \nabla p^h, -\Delta \mathbf{v}^h + \nabla q^h)_{0, \mathcal{K}}$$

is the first variation of the least-squares penalty term

$$\delta \sum_{\mathcal{K} \in \mathcal{T}_h} \frac{1}{2} h_{\mathcal{K}}^2 \| -\Delta \mathbf{u}^h + \nabla p^h - \mathbf{f} \|_{0, \mathcal{K}}^2.$$

The GLS- form, obtained with $\beta = -1$, leads to a linear system with a non-symmetric matrix. This method has not been previously reported on in the literature. Its regularizing contribution

$$(2.9) \quad \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (-\Delta \mathbf{u}^h + \nabla p^h, -\Delta \mathbf{v}^h - \nabla q^h)_{0, \mathcal{K}}$$

cannot be viewed as originating from a least-squares term nor does the resulting variational equation correspond to an optimality system.

2.2. The simplified Galerkin least-squares class. For $\alpha = 0$, the complementary forms are

$$(2.10) \quad \begin{aligned} Q_{sgls}^\pm(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h) &= A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) \pm B(\mathbf{u}^h, q^h) \\ &- \delta \sum_{K \in \mathcal{T}_h} h_K^2 (-\Delta \mathbf{u}^h + \nabla p^h, \pm \nabla q^h)_{0, K}. \end{aligned}$$

The form in (2.10) can be viewed as being derived from (2.6) through a "simplification" of the weighting function. Thus, we will refer to these *simplified* Galerkin least-squares methods as the SGLS+ and SGLS- methods.

The SGLS- finite element method is the first example of a consistently stabilized formulation of the Stokes equations. It was introduced in [20] and is widely known as the *pressure-Poisson stabilized Galerkin method*. The SGLS+ method has not been reported on elsewhere. Neither one of the two complementary forms in (2.10) is symmetric; consequently both the SGLS+ and SGLS- linear systems are non-symmetric. Likewise, neither the SGLS+ nor the SGLS- variational problems can be derived as a first-order optimality condition of some Lagrangian functional.

2.3. The reflected Galerkin least-squares class. For $\alpha = -1$, the complementary forms are

$$(2.11) \quad \begin{aligned} Q_{rgls}^\pm(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h) &= A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) \pm B(\mathbf{u}^h, q^h) \\ &- \delta \sum_{K \in \mathcal{T}_h} h_K^2 (-\Delta \mathbf{u}^h + \nabla p^h, \Delta \mathbf{v}^h \pm \nabla q^h)_{0, K}. \end{aligned}$$

The form in (2.11) can be viewed as being obtained from (2.6) through "reflection" of the sign of $\Delta \mathbf{v}^h$. Accordingly, we will refer to these *reflected* Galerkin least-squares methods as the RGLS+ and RGLS- methods.

The RGLS+ method was introduced in [8] and is commonly known as the Douglas-Wang stabilized Galerkin method. Its RGLS- companion has not been formally introduced in the literature. The stabilizing term of RGLS- is given by

$$(2.12) \quad -\delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (-\Delta \mathbf{u}^h + \nabla p^h, -\Delta \mathbf{v}^h + \nabla q^h)_{0, \mathcal{K}}$$

and looks exactly like the stabilizing term (2.8) in the GLS+ formulation! As a result, RGLS- is often confused with the original GLS+ method. To see that RGLS- and GLS+ are in fact quite different, it helps to write their bilinear forms side by side. The GLS+ bilinear form is

$$(2.13) \quad \begin{aligned} Q_{gls}^+(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h) &= A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) + B(\mathbf{u}^h, q^h) \\ &\quad - \delta \sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 (-\Delta \mathbf{u}^h + \nabla p^h, -\Delta \mathbf{v}^h + \nabla q^h)_{0, \mathcal{K}} \end{aligned}$$

while the RGLS- form is

$$(2.14) \quad \begin{aligned} Q_{rgls}^-(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h) &= A(\mathbf{u}^h, \mathbf{v}^h) + B(\mathbf{v}^h, p^h) - B(\mathbf{u}^h, q^h) \\ &\quad + \delta \sum_{K \in \mathcal{T}_h} h_K^2 (-\Delta \mathbf{u}^h + \nabla p^h, -\Delta \mathbf{v}^h + \nabla q^h)_{0, K}. \end{aligned}$$

Despite their similarity, these two forms have strikingly different properties. We will see that the symmetric GLS+ form is weakly coercive only for sufficiently small δ . In contrast, the non-symmetric RGLS- form is strongly coercive for all positive values of δ . Another major difference is that (2.13) represents the optimality condition for the saddle points of (2.7), while (2.14) is not related to optimization.

It is also clear that the RGLS- and GLS+ forms are not complimentary, i.e., they cannot be obtained from each other by changing the sign of the pressure test function. As a result, the linear systems engendered by these two forms are not equivalent, cannot be obtained from each other, and have different solutions. Even though RGLS- has not been introduced in [8], or anywhere else in the literature, the misconception that (2.14) is the ‘‘Galerkin least-squares’’ method is quite widespread. However, as our discussion shows, (2.14) is a variant of Douglas-Wang stabilization and cannot be obtained from the true, original Galerkin least-squares formulation that uses (2.13).

3. Finite element stability analyses. We briefly review some relevant notions from functional analysis, specialized to our needs. For statements of general results, we refer to [1].

Let V be a Hilbert space. A bilinear form $Q(\cdot; \cdot) : V \times V \mapsto \mathcal{R}$ is called weakly coercive if there exists a positive constant C_2 such that

$$(3.1) \quad \sup_{v \in V, v \neq 0} \frac{Q(u; v)}{\|v\|_V} \geq C_2 \|u\|_V \quad \forall u \in V$$

and

$$(3.2) \quad \sup_{u \in V, u \neq 0} \frac{Q(u; v)}{\|u\|_V} > 0 \quad \forall v \in V.$$

An equivalent statement to (3.1) is that for any given $u \in V$, there exists $v \in V$ such that

$$(3.3) \quad Q(u; v) \geq C_2 \|u\|_V \|v\|_V.$$

A bilinear form $Q(\cdot; \cdot) : V \times V \mapsto \mathcal{R}$ is called strongly coercive (or V-elliptic) if

$$(3.4) \quad Q(u; u) \geq C_2 \|u\|_V^2 \quad \forall u \in V.$$

Finally, a form is called continuous if

$$(3.5) \quad |Q(u; v)| \leq C_1 \|v\|_V \|u\|_V.$$

The continuity and coercivity (either weak or strong) of a bilinear form implies that associated variational problems and the associated finite element discretizations are well posed [1].

THEOREM 3.1. *Assume that $F : V \mapsto \mathcal{R}$ is a continuous linear functional. For any weakly or strongly coercive bilinear form $Q(\cdot; \cdot)$, the variational problem: seek $u \in V$ such that*

$$(3.6) \quad Q(u; v) = F(v) \quad \forall v \in V$$

has a unique solution. Moreover, that solutions depends continuously on the data, i.e.,

$$(3.7) \quad \|u\|_V \leq \frac{1}{C_2} \|F\|.$$

THEOREM 3.2. *Assume all the hypotheses of Theorem 3.1 and let V^h be a closed subspace of V . Furthermore, assume that either $Q(\cdot; \cdot)$ is strongly coercive or that $Q(\cdot; \cdot)$ satisfies the discrete weak coercivity conditions*

$$(3.8) \quad \sup_{v^h \in V^h} \frac{Q(u^h; v^h)}{\|v^h\|_V} \geq C_2^h \|u^h\|_V \quad \forall u^h \in V^h;$$

and

$$(3.9) \quad \sup_{u^h \in V^h} \frac{Q(u^h; v^h)}{\|u^h\|_V} > 0 \quad \forall v^h \in V^h.$$

Then, the approximate problem: seek $u^h \in V^h$ such that

$$(3.10) \quad Q(u^h; v^h) = F(v^h) \quad \forall v^h \in V^h$$

has a unique solution u^h ; that solution satisfies the estimate

$$(3.11) \quad \|u^h\|_V \leq \frac{1}{C_1} \|F\|.$$

Moreover,

$$(3.12) \quad \|u - u^h\|_V \leq \left(1 + \frac{C_1}{C_2^h}\right) \inf_{w^h \in V^h} \|u - w^h\|_V.$$

REMARK 6. If $Q(\cdot, \cdot)$ is strongly coercive, then (3.8) trivially holds with $C_2^h = C_2$ for any conforming subspace $V^h \subset V$. Clearly, (3.9) holds too.

REMARK 7. For weakly coercive forms $Q(\cdot, \cdot)$, the inclusion $V^h \subset V$ may not be sufficient for (3.8)-(3.9) to hold. This inclusion only implies that for any $u^h \in V^h$, there exists $v \in V$ such that

$$Q(u^h, v) \geq C_2 \|u^h\|_V \|v\|_V.$$

However, existence of v is only guaranteed in the *larger* space V and not necessarily in the subspace V^h .

3.1. Stability of consistently modified Galerkin forms. If the variational problem (2.3) is to result in stable and accurate finite element approximations, then it must be well posed. Thus, we now turn our attention to the stability of the form $Q_{\alpha\beta}(\cdot; \cdot)$ defined in (2.1). For fixed α and β , this form depends on the parameter δ and so its stability, the ensuing well posedness of (2.3), and the validity of finite element error estimates will, in general, also depend upon the values of δ . Thus, given a particular form $Q_{\alpha\beta}(\cdot; \cdot)$ and a finite element subspace of $\mathbf{H}_0^1(\Omega) \times L_0^2(\Omega)$, we must determine whether or not there exists at least one positive value of δ and a positive constant $C_2^h(\delta)$ such that (3.8) holds.

We will say that $Q_{\alpha\beta}(\cdot; \cdot)$ is *unconditionally stable* if for any $\delta > 0$ one can find $C_2^h(\delta) \geq \gamma > 0$ such that either (3.4) or (3.8) holds for all $h > 0$. If weak or strong coercivity hold only for selected values of δ , we will call the form $Q_{\alpha\beta}(\cdot; \cdot)$ *conditionally stable*. Finally, we will call the set of all stable values of δ the *admissible range* for the form $Q_{\alpha\beta}(\cdot; \cdot)$.

In general, admissibility ranges cannot be determined exactly and must be estimated. We first consider an approach based on finite element techniques. For this purpose, we will need to employ the mesh dependent norm

$$\|p^h\|_{0,h} = \left(\sum_{\mathcal{K} \in \mathcal{T}_h} h_{\mathcal{K}}^2 \|\nabla p^h\|_{0,\mathcal{K}}^2 \right)^{1/2}.$$

The next theorem generalizes and extends results of [8, 6, 19, 20] to all three pairs of stabilized formulations.

THEOREM 3.3. *Assume that \mathbf{V}^h contains at least quadratic polynomials and that S^h is continuous. Then, for any $(\mathbf{u}^h, p^h) \in \mathbf{V}^h \times S^h$,*

$$(3.13) \quad \left\{ \begin{array}{l} \sup_{(\mathbf{v}^h, q^h) \in \mathbf{V}^h \times S^h} \frac{Q^+(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h)}{\|\mathbf{v}^h\|_1 + \|q^h\|_{0,h}} \\ \frac{Q^-(\mathbf{u}^h, p^h; \mathbf{u}^h, p^h)}{\|\mathbf{u}^h\|_1 + \|p^h\|_{0,h}} \end{array} \right\} \geq (C_P^2 - \delta C_I^2)^{1/2} \|\mathbf{u}^h\|_1 + \delta C_2 \|p^h\|_{0,h}$$

for the GLS and SGLS formulations and

$$(3.14) \quad \left\{ \begin{array}{l} \sup_{(\mathbf{v}^h, q^h) \in \mathbf{V}^h \times S^h} \frac{Q_{rgls}^+(\mathbf{u}^h, p^h; \mathbf{v}^h, q^h)}{\|\mathbf{v}^h\|_1 + \|q^h\|_{0,h}} \\ \frac{Q_{rgls}^-(\mathbf{u}^h, p^h; \mathbf{u}^h, p^h)}{\|\mathbf{u}^h\|_1 + \|p^h\|_{0,h}} \end{array} \right\} \geq C_1 \|\mathbf{u}\|_1 + \delta C_2 \|p\|_{0,h}$$

for the RGLS formulation. Furthermore, (3.9) holds for RGLS+ for all $\delta > 0$ and also for GLS+ and SGLS+ provided that $C_P^2 - \delta C_I^2 > 0$.

The proof of this theorem uses the fact that strong coercivity of a minus form always implies weak coercivity of a plus form. This easily follows from the identity

$$Q^+(\mathbf{u}, p; \mathbf{u}, -p) = Q^-(\mathbf{u}, p; \mathbf{u}, p),$$

because, if $Q^-(\cdot; \cdot)$ is strongly coercive, then (3.3) holds with $(\mathbf{v}^h, q^h) = (\mathbf{u}^h, -p^h)$, i.e., $Q^+(\cdot; \cdot)$ satisfies (3.8). The other inf-sup condition (3.9) also easily follows. On the other hand, weak coercivity does not imply strong coercivity. However, since complementary forms define problems with identical solutions, they are either simultaneously stable or unstable. This opens up an interesting possibility that is not

represented in Theorem 3.3, namely, that both the plus and the minus form can be *weakly coercive*.

Theorem 3.3 provides a sufficiency estimate of the stability ranges for GLS and SGLS methods. From (3.13), it follows that (2.6) and (2.10) will respectively be strongly and weakly coercive if

$$(3.15) \quad 0 < \delta < \frac{C_P^2}{C_I^2}.$$

When using this bound as an estimate for the stability range, it is important to recognize that it only represents a sufficient condition for the strong coercivity of the GLS- and SGLS- forms and the weak coercivity of their plus counterparts. The estimate says nothing about what happens for values of δ that violate the inequality in (3.15). For example, it may be the case that the GLS- and SGLS- forms that are strongly coercive if (3.15) is satisfied are weakly coercive when it is violated and thus can still result in a stable and accurate method.

On the other hand, Theorem 3.3 allows us unambiguously to conclude that RGLS forms are unconditionally stable. From (3.14), we see that (2.11) is coercive whenever $\delta > 0$.

Table 3.1 provides a summary of the stability and other properties for the three classes of stabilized methods that can be inferred from Theorem 3.3. For the sake of completeness, we also include information about the Galerkin mixed method.

TABLE 3.1
Summary of properties of mixed and stabilized variational forms.

| method | α | β | symmetry property | type of coercivity | stability property |
|--------|----------|---------|-------------------|--------------------|--------------------|
| GMM+ | - | 1 | symmetric | weak | inf-sup |
| GMM- | - | -1 | non-symmetric | weak | inf-sup |
| SGLS+ | 0 | 1 | non-symmetric | weak | conditional(?) |
| SGLS- | 0 | -1 | non-symmetric | strong | conditional(?) |
| GLS+ | 1 | 1 | symmetric | weak | conditional(?) |
| GLS- | 1 | -1 | non-symmetric | strong | conditional(?) |
| RGLS+ | -1 | 1 | non-symmetric | weak | unconditional |
| RGLS- | -1 | -1 | non-symmetric | strong | unconditional |

REMARK 8. The constants C_P and C_I are not known explicitly except in some special cases such as rectangular regions and uniform meshes. As a result, the upper bound in (3.15) is not in general easily computable.

REMARK 9. In practical computations, we must also impose a lower bound on δ in terms of the mesh size employed in the discretization, e.g., $0 < \delta_0 h < \delta$.

3.2. Computable bounds through algebraic stability analyses. We now turn our attention to computable estimates of the admissible range for δ . Such estimates can be derived from the equivalent algebraic representations of (3.1) and (3.4). Let $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ denote coefficient vectors of the finite element functions (\mathbf{u}^h, p^h) and (\mathbf{v}^h, q^h) , respectively. Let $\mathbb{Q}_{\alpha\beta}(\delta)$ denote the matrix in (2.4). For simplicity, we will write \mathbb{Q} instead of $\mathbb{Q}_{\alpha\beta}(\delta)$ with the understanding that \mathbb{Q} depends on α , β and δ . For a given subdivision \mathcal{T}_h , $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ are in \mathcal{R}^n , $n = O(h^{-d})$, where d is the space

dimension. In terms of the matrix \mathbb{Q} , the first weak coercivity condition (3.8) can be written as

$$(3.16) \quad \max_{\vec{y} \in \mathcal{R}^n} \frac{\vec{y}^T \mathbb{Q} \vec{x}}{(\vec{y}^T \mathbb{S} \vec{y})^{1/2}} \geq C_2^h(\delta) (\vec{x}^T \mathbb{S} \vec{x})^{1/2} \quad \forall \vec{x} \in \mathcal{R}^n; \quad \forall h > 0.$$

The matrix analogue of strong coercivity (3.4) is given by

$$(3.17) \quad \vec{x}^T \mathbb{Q} \vec{x} \geq C_2^h(\delta) \vec{x}^T \mathbb{S} \vec{x}, \quad \forall \vec{x} \in \mathcal{R}^n; \quad \forall h > 0.$$

Here, \mathbb{S} is the symmetric, positive definite Gram matrix of the finite element basis with respect to the norm $\|\mathbf{u}^h\|_1 + \|p\|_{0,h}$ on $V^h \times S^h$.

The algebraic conditions (3.16) and (3.17) have interesting interpretations. An equivalent statement of (3.16) is given by

$$(3.18) \quad \min_{\vec{x} \in \mathcal{R}^n} \max_{\vec{y} \in \mathcal{R}^n} \frac{\vec{y}^T \mathbb{Q} \vec{x}}{(\vec{y}^T \mathbb{S} \vec{y})^{1/2} (\vec{x}^T \mathbb{S} \vec{x})^{1/2}} \geq C_2^h(\delta) \quad \forall h > 0.$$

The left hand side in (3.18) defines the smallest generalized singular value of \mathbb{Q} , denoted here by $\sigma_1(\mathbb{Q}, \mathbb{S})$. Therefore, weak coercivity of the bilinear form amounts to having the smallest generalized singular value of the coefficient matrix in the discrete system bounded away from zero, independently of the mesh size h , i.e.,

$$\sigma_1(\mathbb{Q}, \mathbb{S}) > C_2^h(\delta) \geq \gamma > 0 \quad \forall h > 0$$

for some fixed γ .

Before interpreting (3.17), we first review some background information about the generalized field of values of a matrix. Let \mathbb{W} be a Hermitian positive definite matrix and let $\mathbb{G} \in \mathcal{C}^{n \times n}$. Then, the W -field of values of \mathbb{G} is defined as

$$(3.19) \quad \mathcal{F}_W(\mathbb{G}) := \left\{ \lambda : \lambda = \frac{\langle \mathbb{W} \mathbb{G} \vec{z}, \vec{z} \rangle}{\langle \mathbb{W} \vec{z}, \vec{z} \rangle} \text{ for some } \vec{z} \neq \vec{0} \in \mathcal{C}^n \right\}.$$

Recall that the W -adjoint of \mathbb{G} is the unique matrix \mathbb{G}^\dagger that satisfies

$$\langle \mathbb{W} \mathbb{G} \vec{z}, \vec{w} \rangle = \langle \mathbb{W} \vec{z}, \mathbb{G}^\dagger \vec{w} \rangle, \quad \forall \vec{z}, \vec{w} \in \mathcal{C}^n$$

which yields $\mathbb{G}^\dagger = \mathbb{W}^{-1} \mathbb{G}^* \mathbb{W}$. A matrix is W -self adjoint if $\mathbb{G}^\dagger = \mathbb{G}$ and W -skew adjoint if $\mathbb{G}^\dagger = -\mathbb{G}$. Next, by decomposing \mathbb{G} into its W -self adjoint and W -skew adjoint parts,

$$\mathbb{G} = \mathbb{M} + \mathbb{N},$$

where

$$\mathbb{M} = \frac{1}{2}(\mathbb{G} + \mathbb{G}^\dagger) \quad \text{and} \quad \mathbb{N} = \frac{1}{2}(\mathbb{G} - \mathbb{G}^\dagger),$$

we can write

$$\frac{\langle \mathbb{W} \mathbb{G} \vec{z}, \vec{z} \rangle}{\langle \mathbb{W} \vec{z}, \vec{z} \rangle} = \frac{\langle \mathbb{W} \mathbb{M} \vec{z}, \vec{z} \rangle}{\langle \mathbb{W} \vec{z}, \vec{z} \rangle} + \frac{\langle \mathbb{W} \mathbb{N} \vec{z}, \vec{z} \rangle}{\langle \mathbb{W} \vec{z}, \vec{z} \rangle} \quad \forall \vec{z} \in \mathcal{C}^n,$$

where

$$\frac{\langle \mathbb{W} \mathbb{M} \vec{z}, \vec{z} \rangle}{\langle \mathbb{W} \vec{z}, \vec{z} \rangle} \in \mathcal{F}_W(\mathbb{M}) \quad \text{and} \quad \frac{\langle \mathbb{W} \mathbb{N} \vec{z}, \vec{z} \rangle}{\langle \mathbb{W} \vec{z}, \vec{z} \rangle} \in \mathcal{F}_W(\mathbb{N}).$$

Since \mathbb{M} is W -self adjoint, it is normal with respect to the W -inner product and $\mathcal{F}_W(\mathbb{M})$ is equivalent to the convex hull of the spectrum of \mathbb{M} which is contained in the reals. Similarly, since \mathbb{N} is W -skew adjoint, $\mathcal{F}_W(\mathbb{N})$ is pure imaginary. Thus, the real part of $\mathcal{F}_W(\mathbb{G})$ can be bounded by the minimum and the maximum eigenvalues of \mathbb{M} . For details regarding properties of the field of values of matrices, see [17, 18, 24].

In our case, \mathbb{Q} is real and \mathbb{S} is symmetric, positive definite. Notice that we can write

$$\frac{\langle \mathbb{Q}\vec{z}, \vec{z} \rangle}{\langle \mathbb{S}\vec{z}, \vec{z} \rangle} = \frac{\langle \mathbb{S}(\mathbb{S}^{-1}\mathbb{Q})\vec{z}, \vec{z} \rangle}{\langle \mathbb{S}\vec{z}, \vec{z} \rangle}.$$

This is a Rayleigh quotient generalized to the S -inner product. It follows from (3.19) that the set of all quotients of this form can be referred to as $\mathcal{F}_S(\mathbb{S}^{-1}\mathbb{Q})$, the S -field of values of $\mathbb{S}^{-1}\mathbb{Q}$. Since $\mathbb{S}^{-1}\mathbb{Q}$ is real,

$$\frac{\langle \mathbb{S}(\mathbb{S}^{-1}\mathbb{Q})\vec{x}, \vec{x} \rangle}{\langle \mathbb{S}\vec{x}, \vec{x} \rangle} = \frac{\langle \mathbb{S}(\mathbb{S}^{-1}\mathbb{Q})_{SA} \vec{x}, \vec{x} \rangle}{\langle \mathbb{S}\vec{x}, \vec{x} \rangle} \quad \forall \vec{x} \in \mathcal{R}^n,$$

where $(\mathbb{S}^{-1}\mathbb{Q})_{SA}$ is the S -self adjoint part of $\mathbb{S}^{-1}\mathbb{Q}$. With this information, (3.17) can be rewritten as

$$(3.20) \quad \frac{\vec{x}^T \mathbb{Q} \vec{x}}{\vec{x}^T \mathbb{S} \vec{x}} = \frac{\langle \mathbb{Q}\vec{x}, \vec{x} \rangle}{\langle \mathbb{S}\vec{x}, \vec{x} \rangle} = \frac{\langle \mathbb{S}(\mathbb{S}^{-1}\mathbb{Q})_{SA} \vec{x}, \vec{x} \rangle}{\langle \mathbb{S}\vec{x}, \vec{x} \rangle} \geq C_2^h(\delta), \quad \forall \vec{x} \in \mathcal{R}^n; \quad \forall h > 0.$$

Since the real part of $\mathcal{F}_S(\mathbb{S}^{-1}\mathbb{Q})$, denoted here by $Re(\mathcal{F}_S(\mathbb{S}^{-1}\mathbb{Q}))$, is bounded below by the minimum eigenvalue of $(\mathbb{S}^{-1}\mathbb{Q})_{SA}$ and the eigenvectors of $(\mathbb{S}^{-1}\mathbb{Q})_{SA}$ are real, the condition (3.17) for strong coercivity can be interpreted as requiring the S -field of values of $\mathbb{S}^{-1}\mathbb{Q}$ to be in the right half plane, independent of h , i.e.,

$$Re(\mathcal{F}_S(\mathbb{S}^{-1}\mathbb{Q})) > C_2^h(\delta) \geq \gamma > 0 \quad \forall h > 0.$$

In contrast to (3.15), both (3.18) and (3.20) are computable in the sense that $\sigma_1(\mathbb{Q}, \mathbb{S})$ and $Re(\mathcal{F}_S(\mathbb{S}^{-1}\mathbb{Q}))$ can be computed numerically for any given \mathcal{T}_h and δ . However, the type of estimates for the admissible range implied by (3.15) on the one hand and enabled by (3.18) or (3.20) on the other hand, are very different. The bound in (3.15) guarantees stability for any h and represents an inner estimate of the true admissibility range. In contrast, (3.18) and (3.20) can only provide outer estimates for this range. Indeed, given a subdivision \mathcal{T}_h and a desired value for $C_2^h(\delta)$, one can determine computationally whether or not there is a δ that verifies (3.16) or (3.17) and then estimate the largest such δ . The catch is that finite element stability requires more than just a nonsingular matrix (which could easily be the case even for unstable forms); we also need for $C_2^h(\delta)$ to be mesh independent. This fact may not be so easy to verify as we can only compute the field of values or the smallest singular value for a finite number of mesh sizes h and parameter values δ . Therefore, while computable, these bounds should be treated with caution.

3.3. Remarks about stability. We saw that weakly coercive formulations are associated with linear systems with indefinite matrices whose smallest generalized singular value is bounded away from zero independently of h . Strongly coercive formulations lead to linear systems with real, positive definite¹ matrices whose field of

¹A matrix \mathbb{A} is real positive definite if $\vec{u}^T \mathbb{A} \vec{u} > 0$ for any real-valued vector \vec{u} . Equivalently, \mathbb{A} is real positive definite if $(\mathbb{A} + \mathbb{A}^T)/2$ is positive definite.

values has its real part bounded away from zero independently of h . In addition, for conditionally stable forms, these properties may hold only for a limited range of values for δ .

From these wide and varied properties of stabilized methods and their associated algebraic problems, there arises a number of issues about their use and implementation. First, each class of stabilized forms, i.e., GLS, SGLS, and RGLS, offers a choice of a linear system with either a positive definite or an indefinite matrix. When the complimentary systems generated by the $\text{GLS}\pm$, $\text{SGLS}\pm$, or $\text{RGLS}\pm$ formulations are solved by a direct method, the answers and the solver performance will be identical. However, if these systems are solved iteratively, it is not clear which formulation should be preferred: a strongly coercive one, yielding a positive definite but nonsymmetric matrix, or a weakly coercive one which gives an indefinite but possibly symmetric (as in the case of $\text{GLS}+$ method) matrix. Moreover, the answer to this question may depend on the type of iterative solver and preconditioning used.

A second set of questions concerns the admissible range bound (3.15) and its algebraic estimates computed using (3.18) or (3.20). We recall that (3.15) is biased with respect to the minus forms, i.e., it is a sufficient condition for a strong coercivity of $\text{GLS}-$ and $\text{SGLS}-$. However, a minus form may fail to be strongly coercive and still lead to a stable method if it happens to be weakly coercive. Thus, the possibilities are that either loss of strong coercivity in a minus form will immediately lead to an unstable complementary pair, or that the pair will remain stable even if the minus form is not strongly coercive. In the former case, (3.15) and its algebraic estimates should be able to reliably predict loss of stability, while in the latter case these estimates will be overly pessimistic.

4. Computational studies. The main goal of this section is to document the numerical behavior of the three stabilized classes, including the performance of iterative solvers, and to compare them with one another. Specifically, for the GLS and SGLS classes, we want to assess the accuracy of the algebraic estimates of the admissibility range by comparing finite element stability inside and outside the estimated range. Then, we will compare and contrast Krylov subspace solvers for two typical scenarios that may arise in the use of stabilized methods. In the first case, the same stabilized formulation is solved for different values of δ . The goal is to compare solver convergence for stable vs. unstable values of δ . The second scenario is to alternate between a plus and minus formulation for the same δ . Here our objective is to determine whether or not iterative solvers have a markedly better performance for one of the two complementary formulations.

In all numerical experiments, Ω is taken to be the unit square in \mathcal{R}^2 . The stabilized methods were implemented using C^0 , piecewise quadratic finite element spaces on triangles for all dependent variables, i.e., using a $P^2 - P^2$ discretization. The reasons for using such elements were explained in Remarks 2 and 3. We recall that the $P^2 - P^2$ pair is unstable for the mixed Galerkin problem (1.10)–(1.11). To eliminate the hydrostatic pressure mode without changing the effective condition number, a rank-one update defined by $\bar{\mathbf{c}}^T \bar{\mathbf{c}}$, where $\bar{\mathbf{c}}$ is the coefficient vector of the constant pressure mode, is added to the lower right block of the matrices; see [3] for details about this approach.

4.1. Numerical estimates of the admissible region. To estimate the admissible range (3.15) for δ for the conditionally stable GLS and SGLS methods, we can use the approach based on the generalized field of values outlined in Section 3.2. First, we choose a positive \tilde{C} and a sequence of subdivisions \mathcal{T}_h^k . For each \mathcal{T}_h^k , we seek to

determine, if possible, the largest δ such that $Re(\mathcal{F}_S(\mathbb{S}^{-1}\mathbb{Q})) > \tilde{C}$, or equivalently, for which

$$(4.1) \quad \vec{\mathbf{x}}^T \mathbb{Q} \vec{\mathbf{x}} \geq \tilde{C} \vec{\mathbf{x}}^T \mathbb{S} \vec{\mathbf{x}} \quad \forall \vec{\mathbf{x}} \in \mathcal{R}^n.$$

Let δ_k denote this value. Then, the stable range of δ is estimated as

$$0 < \delta < \min_k \delta_k.$$

It is clear that the value of δ_k is itself an estimate that can be determined by solving for the eigenvalues of $(\mathbb{S}^{-1}\mathbb{Q})_{SA}$ with several values of δ .

We can simplify this process by using the field of values $\mathcal{F}(\mathbb{Q})$ instead of $\mathcal{F}_S(\mathbb{S}^{-1}\mathbb{Q})$. Since \mathbb{S} is symmetric, positive definite,

$$\vec{\mathbf{x}}^T \mathbb{S} \vec{\mathbf{x}} \geq \lambda_{min}(\mathbb{S}) \vec{\mathbf{x}}^T \vec{\mathbf{x}} \quad \forall \vec{\mathbf{x}} \in \mathcal{R}^n.$$

Therefore, if δ is such that

$$(4.2) \quad \vec{\mathbf{x}}^T \mathbb{Q} \vec{\mathbf{x}} \geq \hat{C} \vec{\mathbf{x}}^T \vec{\mathbf{x}}$$

for some given \hat{C} , then (4.1) holds with $\tilde{C} = \hat{C}/\lambda_{min}(\mathbb{S})$. Since the real part of $\mathcal{F}(\mathbb{Q})$ is bounded by the smallest and the largest eigenvalues of $(\mathbb{Q} + \mathbb{Q}^T)/2$, we can proceed as in the first case, except that now estimating δ_k for each \mathcal{T}_h^k requires solving just one symmetric eigenvalue problem. If \mathcal{T}_h^k are obtained through a uniform refinement process, then δ_k should be approximately the same for all values of k . In general, δ_k will vary with the geometry of the mesh and so this approach should only be used on sequences of meshes that are not too different.

Let us apply the simplified procedure to compute an estimate for the admissible range of the GLS- method. Ideally we would like to use only few, relatively coarse, grids to estimate the range for δ . Here, we employ two uniform triangulations of the unit square, denoted by \mathcal{T}_h^1 and \mathcal{T}_h^2 , whose respective mesh sizes are $h = 1/4$ and $h = 1/8$. Table 4.1 shows the smallest and the largest eigenvalues of $(\mathbb{Q} + \mathbb{Q}^T)/2$, i.e., the range of $Re(\mathcal{F}(\mathbb{Q}))$, on these meshes for 10 different values of δ .

From Table 4.1 it follows that $\delta_1 = \delta_2 \approx 0.04$. As a result, based on the two grids \mathcal{T}_h^1 and \mathcal{T}_h^2 , the admissible range can be estimated by $(0, 0.04)$.

REMARK 10. For uniform Q_2 spaces (piecewise continuous bi-quadratic polynomials on squares) on $[0, 1]^2$, the upper bound in (3.15) has been analytically shown to be equal to $11/270 \approx 0.0407$; see [16] and [10]. Note that this is essentially the same value as the one that we have obtained from the data in Table 4.1 for quadratic elements on uniform triangulations.

To compute an estimate for the SGLS- method, we use the same two grids but a slightly different range of values for δ . From Table 4.2, we see that $\delta_1 = 0.2$ and $\delta_2 = 0.1$ so the estimate for the admissibility range based on \mathcal{T}_h^1 and \mathcal{T}_h^2 is approximately $(0, 0.1)$.

4.2. Finite element convergence rates. We will now test how well the computable bounds for δ predict the finite element stability of GLS and SGLS methods and whether or not the unconditional theoretical stability of RGLS formulations also translates into numerical stability over a wide range of parameter values. To assess stability, we will compute numerical convergence rates for the smooth exact solution

$$\begin{cases} \mathbf{u} = (\sin(\pi x - 0.7) \sin(\pi y + 0.2), \cos(\pi x - 0.7) \cos(\pi y + 0.2)) \\ p = \sin(x) \cos(y) + (\cos(1) - 1) \sin(1). \end{cases}$$

TABLE 4.1
Estimate of the admissible range for GLS-.

| h | δ | $\lambda_{\min}((\mathbb{Q} + \mathbb{Q}^T)/2)$ | $\lambda_{\max}((\mathbb{Q} + \mathbb{Q}^T)/2)$ |
|------|---------------|---|---|
| 1/4 | 0.1 | -0.8089484E+01 | 0.7241451E+01 |
| | 0.09 | -0.6520340E+01 | 0.7286919E+01 |
| | 0.08 | -0.4962731E+01 | 0.7344019E+01 |
| | 0.07 | -0.3421514E+01 | 0.7417573E+01 |
| | 0.06 | -0.1904588E+01 | 0.7515238E+01 |
| | 0.05 | -0.4267026E+00 | 0.7649627E+01 |
| | 0.04 | 0.1321793E-03 | 0.7842181E+01 |
| | 0.03 | 0.9913447E-04 | 0.8129579E+01 |
| | 0.02 | 0.6608964E-04 | 0.8570449E+01 |
| 0.01 | 0.3304482E-04 | 0.9236781E+01 | |
| 1/8 | 0.1 | -0.9977763E+01 | 0.7408766E+01 |
| | 0.09 | -0.8181848E+01 | 0.7451614E+01 |
| | 0.08 | -0.6396991E+01 | 0.7505553E+01 |
| | 0.07 | -0.4627906E+01 | 0.7575311E+01 |
| | 0.06 | -0.2882238E+01 | 0.7668542E+01 |
| | 0.05 | -0.1173041E+01 | 0.7798242E+01 |
| | 0.04 | 0.9987927E-05 | 0.7987464E+01 |
| | 0.03 | 0.7490945E-05 | 0.8278022E+01 |
| | 0.02 | 0.4993963E-05 | 0.8741929E+01 |
| 0.01 | 0.2496982E-05 | 0.9476134E+01 | |

TABLE 4.2
Estimate of the admissible range for SGLS-.

| h | δ | $\lambda_{\min}((\mathbb{Q} + \mathbb{Q}^T)/2)$ | $\lambda_{\max}((\mathbb{Q} + \mathbb{Q}^T)/2)$ |
|-----|----------|---|---|
| 1/4 | 0.5 | -0.2921313E-01 | 0.1017679E+02 |
| | 0.4 | -0.1095955E-01 | 0.1017352E+02 |
| | 0.3 | -0.2242186E-02 | 0.1017099E+02 |
| | 0.2 | 0.1567617E-05 | 0.1016918E+02 |
| | 0.1 | 0.1802104E-03 | 0.1016810E+02 |
| | 0.05 | 0.1282501E-03 | 0.1016783E+02 |
| | 0.01 | 0.3157644E-04 | 0.1016775E+02 |
| 1/8 | 0.5 | -0.9540696E-02 | 0.1053977E+02 |
| | 0.4 | -0.4307011E-02 | 0.1053954E+02 |
| | 0.3 | -0.1199751E-02 | 0.1053935E+02 |
| | 0.2 | -0.3833118E-04 | 0.1053922E+02 |
| | 0.1 | 0.1148637E-04 | 0.1053915E+02 |
| | 0.05 | 0.9140608E-05 | 0.1053913E+02 |
| | 0.01 | 0.2363616E-05 | 0.1053912E+02 |

As usual, the inhomogeneous boundary condition is handled by using a boundary interpolant that becomes an additional source term in the discrete problem.

The goal of our experiments is not as much to obtain sharp estimates of the theoretical convergence rates but to investigate whether or not the changes of sign in Tables 4.1 and 4.2 that were used to determine the stable values of δ also correlate with changes in the behavior of the finite element errors and the convergence rates.

Numerical rates of convergence are computed by solving SGLS, GLS and RGLS problems on uniform triangular meshes obtained by subdividing Ω into squares using 17×17 and 33×33 uniformly spaced grid lines and then drawing the diagonal in each square. Since we use P^2 elements, the first mesh has 33×33 nodes and the second mesh has 65×65 nodes. The linear systems are solved by direct solvers. We denote the errors in the approximations to \mathbf{u} and p by $E_{\mathbf{u}}$ and E_p , respectively.

Table 4.3 contains a sample of the computed numerical convergence rates. We see that the convergence rates for the SGLS and RGLS methods are essentially the same even for values of δ that are *outside* the estimated stability range for SGLS. In contrast, the convergence rates of GLS solutions become erratic once δ exceeds the *numerically estimated* threshold stability value of 0.04.

TABLE 4.3

Convergence rates for the GLS, SGLS and RGLS methods for different values of δ . Bold face numbers denote unstable values of δ for GLS. Columns (i) are for $\|E_{\mathbf{u}}\|_0$, columns (ii) are for $|E_{\mathbf{u}}|_1$, and columns (iii) are for $\|E_p\|_0$.

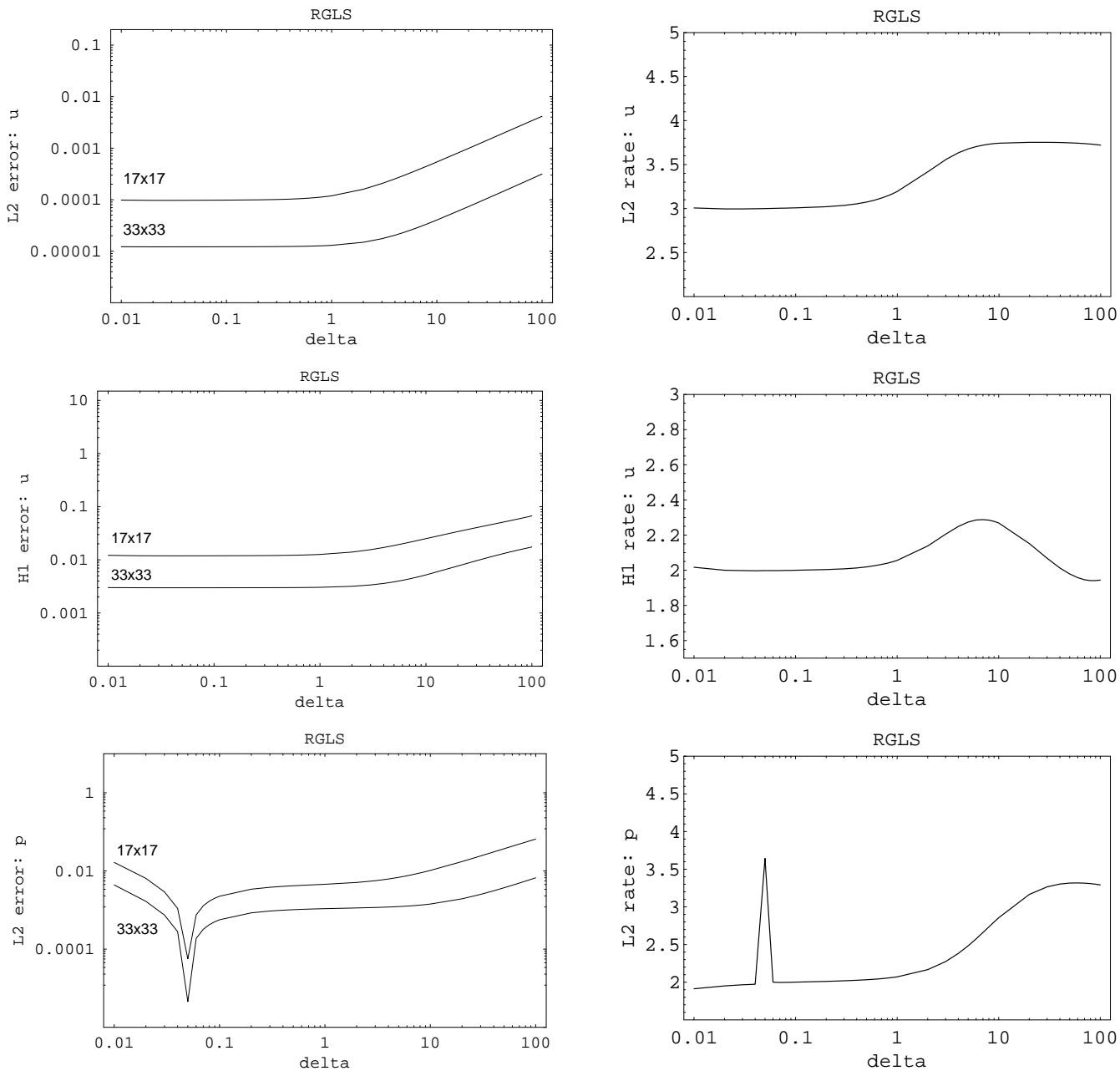
| GLS | | | | SGLS | | | RGLS | | |
|-------------|-----|------|-------|------|------|-------|------|------|-------|
| δ | (i) | (ii) | (iii) | (i) | (ii) | (iii) | (i) | (ii) | (iii) |
| 0.02 | 3.0 | 2.0 | 1.9 | 3.0 | 2.0 | 1.9 | 3.0 | 2.0 | 1.9 |
| 0.03 | 3.0 | 2.0 | 1.9 | 3.0 | 2.0 | 1.9 | 3.0 | 2.0 | 1.9 |
| 0.04 | 3.3 | 2.2 | 1.8 | 3.0 | 2.0 | 1.9 | 3.0 | 2.0 | 1.9 |
| 0.05 | 2.3 | 1.3 | 1.6 | 3.0 | 2.0 | 3.6 | 3.0 | 2.0 | 3.6 |
| 1.00 | 1.7 | 0.8 | 0.9 | 3.5 | 2.0 | 2.1 | 3.2 | 2.1 | 2.1 |
| 20.0 | 0.6 | -0.5 | -1.0 | 3.8 | 2.8 | 3.4 | 3.8 | 2.2 | 3.2 |

To further compare and contrast the behavior of the errors in the finite element approximations, we provide, in Figures 4.3 and 4.2 plots of the error norms on 17×17 and 33×33 meshes, and the associated convergence rates for values of δ ranging from 0.01 to 100. Figure 4.1 confirms the expected unconditional stability of RGLS methods.

Figure 4.2 is quite intriguing as it shows that errors and convergence rates of SGLS methods are almost indistinguishable from those of RGLS formulations. These results indicate that SGLS formulation remains numerically stable even after the minus SGLS form loses its strong coercivity. Therefore, it is very likely that the SGLS– form switches from strong to weak coercivity and the method retains its stability. This also means that well-posedness of the SGLS class of methods is not governed by the strong coercivity of the minus form. The curious “superconvergence” spike that occurs for the error in the pressure approximation is unexplained at this moment.

Figure 4.3 shows that exactly the opposite holds true for GLS methods. We see that for values of δ that exceed the threshold value of 0.04, both the errors and the convergence rates became erratic. The failure of finite element solutions to converge to the smooth exact solution clearly indicates the instability of the variational problem. Therefore, the algebraic estimate of the stability region turns out to be very sharp. This also means that well-posedness of the GLS class of methods is equivalent, at least numerically, to the strong coercivity of the GLS– form.

4.3. Performance of Krylov subspace solvers. This section gives preliminary results on the performance of various Krylov subspace solvers for the three classes of stabilized methods and the complementary plus and minus forms within each class.

FIG. 4.1. *Finite element errors and convergence rates: RGLS-.*

The linear systems that result from the stabilized methods studied in this paper vary from symmetric indefinite systems that are conditionally stable, to non-symmetric, unconditionally positive definite systems. Ideally, the Krylov subspace method and the preconditioner would be chosen specifically for each class and complementary form. Although preconditioning is necessary on real application problems, it would not be very meaningful to compare results obtained by using the same general pur-

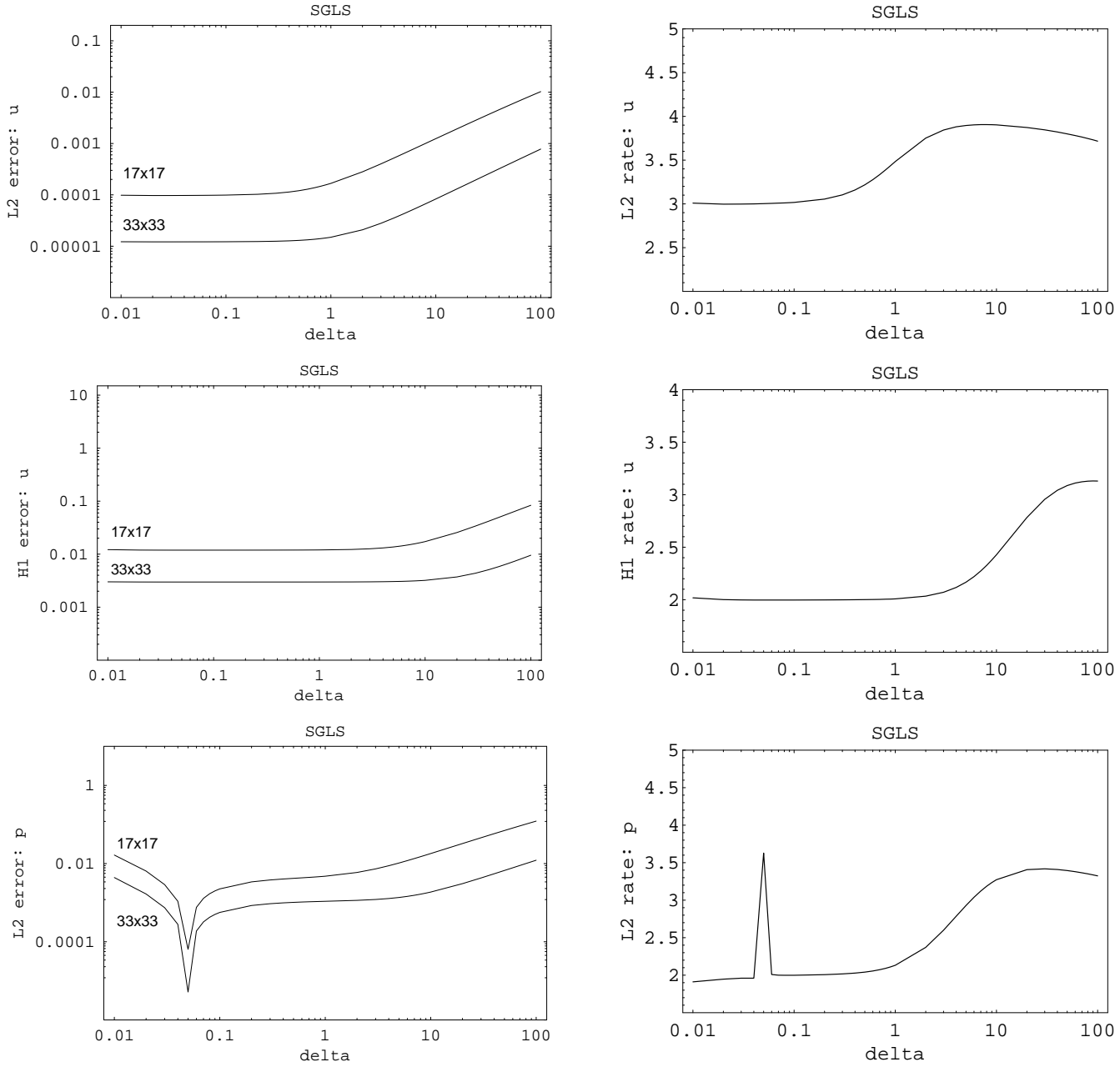
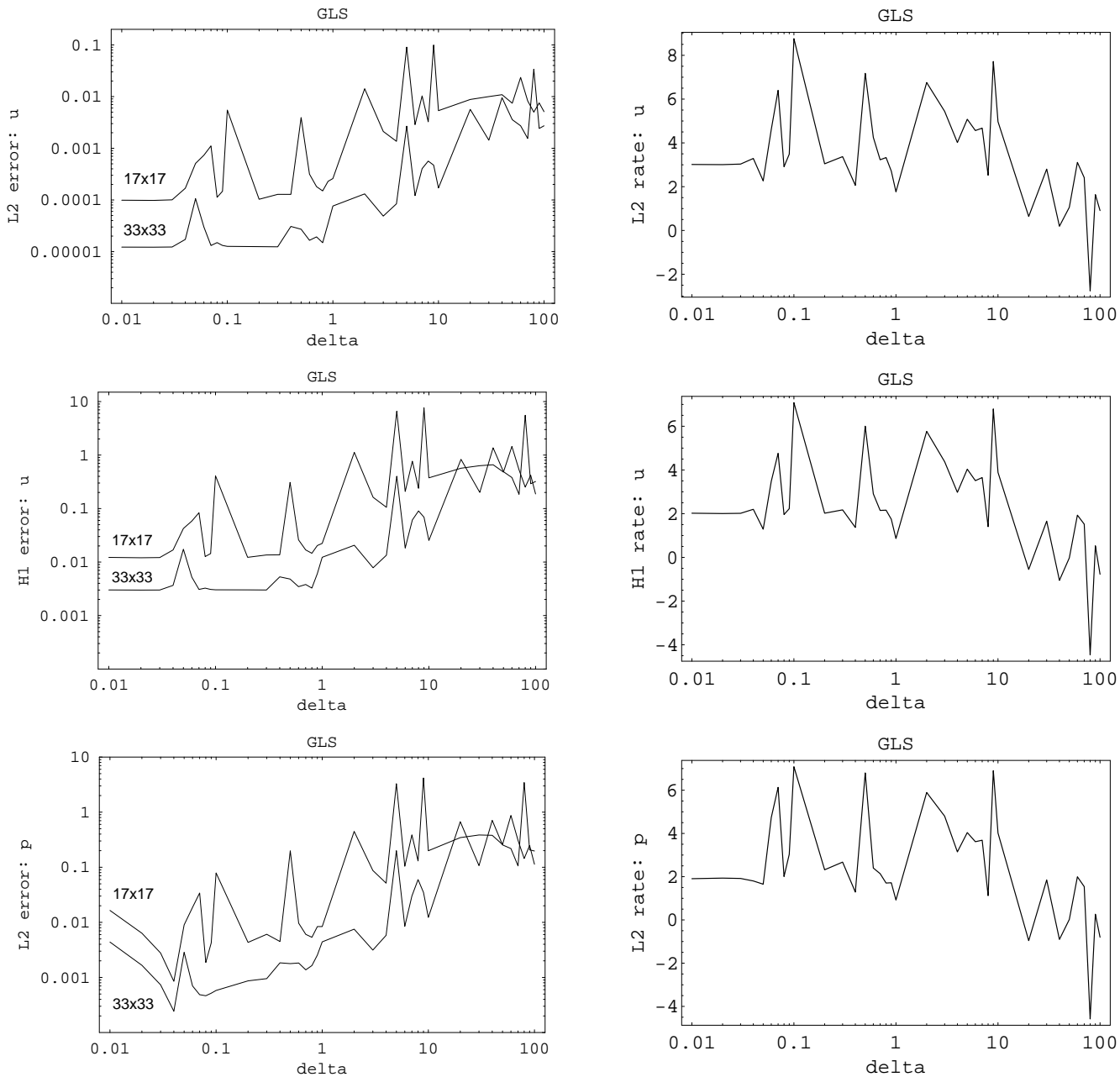


FIG. 4.2. Finite element errors and convergence rates: SGLS-.

pose preconditioners for all formulations. A more thorough investigation is needed to compare iterative solvers and preconditioning techniques that are designed specifically for each method and complementary form.

Here, we have tested unpreconditioned versions of restarted and full GMRES (see [14, 23, 24]) and QMR (see [9, 14, 24]) on the linear systems resulting from the various stabilized formulations on a small example problem. The purpose of this work is to

FIG. 4.3. *Finite element errors and convergence rates: GLS+.*

get a general idea how basic Krylov solvers perform on these stabilized formulations, and to use the results to help decide on the direction of future investigations.

Table 4.4 describes the convergence history for GMRES(10), full GMRES, and QMR on the various stabilized formulations. To discuss results, we will refer to values of δ inside and outside the numerically determined stability interval as “stable” and “unstable,” respectively. The history is given for three values of the stabilizing

parameter δ for the minus form of each class, and one (stable) value of δ for the plus form of each class. We present only one value of δ in the stable range for the plus formulations since these possibly symmetric, indefinite forms are less well suited to our brief study of unpreconditioned, nonsymmetric Krylov solvers. In addition, the only suitable formulation, GLS+, that can take advantage of symmetry has a very limited range of finite element stability as described above. In the table, stable values of δ are indicated by bold type. The ** symbols in the table mean the relative residual norm has not been reduced to the specified level in less than or equal to 4000 iterations for restarted GMRES(10), or 800 iterations for full GMRES or QMR.

TABLE 4.4
Iterative solver performance

| | | Number of iterations required for GMRES(10)/GMRES/QMR to reduce the relative residual norm to | | | |
|--------|--------------|---|--------------|--------------|--------------|
| Method | δ | 1.0e-2 | 1.0e-3 | 1.0e-4 | 1.0e-5 |
| SGLS- | 0.10 | 300/95/110 | 710/135/140 | 1170/180/225 | 1550/210/280 |
| SGLS- | 5.00 | 190/75/135 | 580/155/185 | 720/225/280 | 960/270/370 |
| SGLS- | 15.00 | 190/70/130 | 600/135/185 | 1010/190/280 | 1450/240/320 |
| SGLS+ | 0.10 | 700/105/115 | 2950/150/165 | **/190/200 | **/220/270 |
| GLS- | 0.04 | 250/110/110 | 650/165/200 | 1000/215/240 | 1650/230/275 |
| GLS- | 0.05 | **/170/220 | **/215/350 | **/250/405 | **/265/500 |
| GLS- | 1.00 | **/235/** | **/285/** | **/330/** | **/345/** |
| GLS+ | 0.04 | **/110/115 | **/170/180 | **/220/245 | **/240/285 |
| RGLS- | 0.01 | 200/155/210 | 3500/230/300 | **/250/390 | **/275/500 |
| RGLS- | 1.00 | 50/75/150 | 1300/185/380 | **/230/700 | **/260/** |
| RGLS- | 10.00 | 1000/25/50 | 3600/125/130 | **/180/700 | **/230/** |
| RGLS+ | 1.00 | **/110/150 | **/230/600 | **/255/** | **/280/** |

We see that iterative solver performance for the SGLS method is less sensitive to violations of (3.15) than for the GLS method. Note that observed solver performance also correlates well with the finite element stability of these methods discussed in Section 4.2. Iterative solver performance for the unconditionally stable RGLS method is also sensitive to the value of δ . Here, the value of δ affects only the condition number but not the definiteness of the matrix.

Recall that for the GLS method, the computed admissible range for δ is $(0, 0.04)$. For the SGLS method, the computed admissible range for δ is wider, $(0, 0.1)$. Figure 4.4 compares $\mathcal{F}(\mathbb{Q})$ for the GLS- method with $\delta = 0.05$, and the SGLS- method with $\delta = 15$. The solid lines enclose the field of values of the matrices, the vertical dashed lines bound $Re(\mathcal{F}(\mathbb{Q}))$, and the small green circles denote the eigenvalues of the matrices \mathbb{Q} . The dotted vertical line indicates the location of the line $x = 0$. Notice that $\delta = 0.05$ is slightly outside of the GLS admissible δ range, yet both $\mathcal{F}(\mathbb{Q})$ and the spectrum of \mathbb{Q} extend into the left half-plane. On the other hand, $\delta = 15$ is significantly outside of the computed admissible δ range for the SGLS method. Here, $\mathcal{F}(\mathbb{Q})$ extends well into the left half-plane, but the eigenvalues of \mathbb{Q} remain in the right half-plane. This may in part explain the greater sensitivity of the GLS- method.

5. Conclusions. The choice of the stabilizing parameter δ is critical to the success of the consistently stabilized finite element methods studied in this report. We

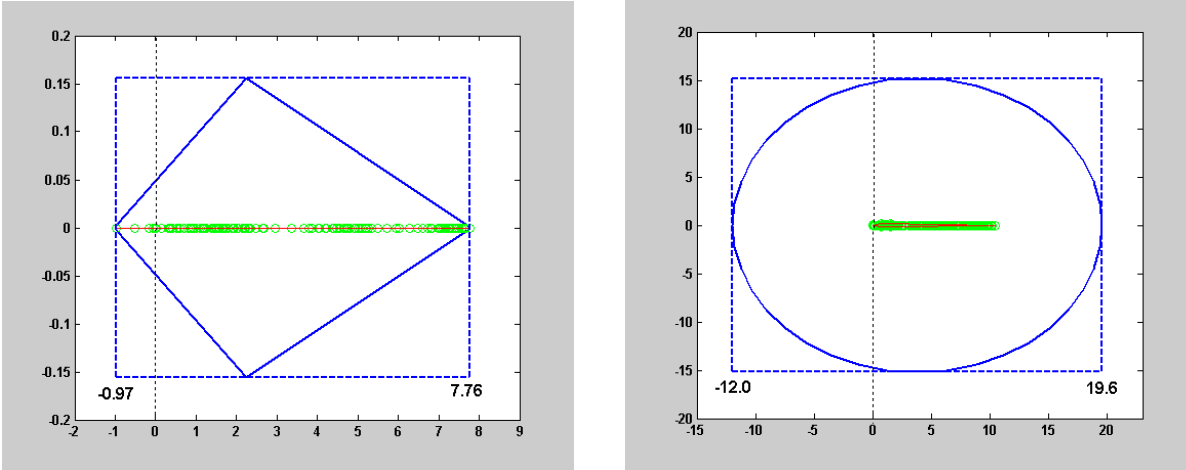


FIG. 4.4. *Left: $\mathcal{F}(\mathbb{Q})$ for GLS- with $\delta = 0.05$, Right: $\mathcal{F}(\mathbb{Q})$ for SGLS- with $\delta = 15$.*

saw that this could even be important for the unconditionally stable RGLS class. Here, the range of δ mattered most for the solver convergence and much less for the finite element convergence rates. We found that smaller values of δ yield faster convergence of the iterative solvers. Finite element error convergence essentially remained the same over all practical values of δ . These results are consistent with the unconditional stability of RGLS formulations established in Theorem 3.3.

According to Theorem 3.3, stability of GLS methods is possibly conditional. Indeed, we found that for this class of methods the choice of δ proved to be critical for the accuracy and stability of finite element solutions. Outside of the estimated admissible range for δ , we observed a sharp deterioration of finite element convergence rates. This strongly suggests that for GLS formulations, the field of values estimate has a very good predictive capability for the admissible range. Most notably, our numerical estimate of the admissible range for GLS coincided with the analytic result of [16] for a slightly different element type.

The sharpness of these bounds also means that the practical use of GLS must be accompanied by a careful determination of the admissible range for δ . Unfortunately, the inverse constant C_I in (3.15) depends upon the subdivision. As a result, one may have to recompute the admissible range estimate every time \mathcal{T}_h is refined or the region is remeshed. An alternative is to use iterative solvers as an indicator for the finite element stability of GLS. In practice, the wide variation in their convergence behavior makes such an indicator rather unreliable. For example, we saw that QMR converged for $\delta = 0.05$ and stalled for $\delta = 1$. Both are unstable values for GLS, but based on solver performance one may be led to believe that the method behaves reasonably well. Again, this highlights the distinction between having a matrix problem that defines a stable finite element approximation vs. one that is merely non-singular.

The most unexpected conclusion from our study is the unconditional stability of SGLS formulations. We found that strong coercivity of the SGLS- form is not necessary for the stability of this method. As a result, estimates of the stability range based on Theorem 3.3 turned out to be very pessimistic for SGLS methods. We saw that errors and convergence rates of SGLS are very close to those of RGLS methods over a wide range of values for δ . In addition, solver performance remained

robust for virtually all values of δ , even though the associated matrices were real positive definite only for some of the tested values. These numerical results led us to develop and analyze an SGLS like formulation for which we were able to prove unconditional stability [4]. The rigorous proof of unconditional stability for the specific SGLS formulation considered in this paper still remains an open question.

Nevertheless, we can conclude that SGLS is in actuality a strong contender among stabilized methods, first, because of its excellent finite element stability, and second, because our preliminary results for performance of unpreconditioned iterative solvers indicate better performance for this formulation. While encouraging, more careful studies that includes appropriate preconditioning methods need to be carried out. Further, of the three classes of stabilized methods, the structure of the SGLS matrices is closest to that of the unperturbed Galerkin mixed method matrices, as well as to that of the mixed method penalty matrices studied in [22], where an optimal preconditioner was designed that was independent of the discretization and penalty parameters. For these reasons, we plan a future, more thorough study of preconditioned Krylov solvers for SGLS matrices.

As a final note, the reader should be aware that the term GLS is often used in a generic sense to denote *any* consistently stabilized formulation. Our taxonomy clearly shows the distinctions between the three principal classes of stabilized methods and should be helpful to avoid the confusion between these methods that often results from such generic labeling. The information provided about the different methods should also be useful when making choices regarding their use and iterative methods for their solution.

6. Acknowledgements. The authors wish to thank the anonymous referees for the careful reading of the paper and their suggestions that helped to improve the manuscript.

REFERENCES

- [1] I. BABUSKA AND A. AZIZ, *Survey lectures on the mathematical foundations of the finite element method*, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, Academic, New York, 1972.
- [2] M. BEHR, L. FRANCA, AND T. TEZDUYAR, *Stabilized finite element methods for the velocity-pressure-stress formulation of incompressible flows*, Computer Methods in Appl. Mech. Engrg., 104 (1993) pp. 31–48.
- [3] P. BOCHEV AND R. B. LEHOUCQ *On finite element methods for the pure Neumann problem*, submitted to SIAM Review.
- [4] P. BOCHEV, M. GUNZBURGER *An absolutely stable pressure-Poisson stabilized method for the Stokes equations* submitted to SIAM J. Num. Anal.
- [5] F. BREZZI, *On existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers*, RAIRO Model. Math. Anal. Numer., 21 (1974) pp. 129–151.
- [6] F. BREZZI AND J. DOUGLAS, *Stabilized mixed methods for the Stokes problem*, Numer. Math., 53 (1988) pp.225–235.
- [7] P. CIARLET, *Finite Element Methods for Elliptic Problems*, North Holland, Amsterdam (1978).
- [8] J. DOUGLAS AND J. WANG, *An absolutely stabilized finite element method for the Stokes problem*, Math. Comp., 52 (1989) pp. 495–508.
- [9] R. FREUND AND N. NACHTIGAL, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991) pp. 315-339.
- [10] L. FRANCA, S. FREY, AND T. HUGHES, *Stabilized finite element methods: I. Application to the advective-diffusive model*, Comput. Meth. Appl. Mech. Engrg., 95 (1992) pp. 253–276.
- [11] L. FRANCA AND R. STENBERG, *Error analysis of some Galerkin least-squares methods for the elasticity equations*, SIAM J. Numer. Anal., 28 (1991) pp. 1680–1697.

- [12] L. FRANCA, T.J.R. HUGHES AND R. STENBERG, *Stabilized finite element methods*, in: *Incompressible Computational Fluid Dynamics*, edited by M. Gunzburger and R. Nicolaides, Cambridge University Press, 1993
- [13] V. GIRAULT AND P. RAVIART, *Finite Element Methods for Navier-Stokes Equations*, Springer, Berlin (1986).
- [14] A. GREENBAUM, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia (1997).
- [15] M. GUNZBURGER, *Finite Element Methods for Viscous Incompressible Flows*, Academic, Boston, (1989).
- [16] I. HARARI, *Computational methods for problems of acoustics with particular reference to exterior domains*. PhD Thesis, Division of Applied Mathematics, Stanford University, Stanford, California, 1991.
- [17] R. HORN, AND C. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, London, U.K., 1991.
- [18] A. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, (1964).
- [19] T. HUGHES AND L. FRANCA, *A new finite element formulation for computational fluid dynamics: VII. The Stokes problem with various well-posed boundary conditions: symmetric formulations that converge for all velocity pressure spaces*, *Comput. Meth. Appl. Mech. Engrg.*, 65 (1987) pp. 85–96.
- [20] T. HUGHES, L. FRANCA, AND M. BALESTRA, *A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuska-Brezzi condition: A stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations*, *Comput. Meth. Appl. Mech. Engrg.*, 59 (1986) pp. 85–99.
- [21] K. JANSEN, S. COLLIS, C. WHITING, AND F. SHAKIB, *A better consistency for low-order stabilized finite element methods*, *Computer Methods in Applied Mechanics and Engineering*, 174 (1999) pp. 153–170.
- [22] A. KLAWONN, *An Optimal Preconditioner For A Class Of Saddle Point Problems With A Penalty Term*, *SIAM J. Sci. Comput.*, vol. 19, no. 2, (1998), pp. 540-552.
- [23] Y. SAAD, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, *SIAM J. Sci. Stat. Comput.*, 7 (1986) pp. 856-869.
- [24] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, PWS , Boston (1996).