SIAM J. NUMER. ANAL.
Vol. 57, No. 3, pp. 1494–1523

MULTIGRID PRECONDITIONERS FOR THE NEWTON–KRYLOV METHOD IN THE OPTIMAL CONTROL OF THE STATIONARY NAVIER–STOKES EQUATIONS*

ANA MARIA SOANE† AND ANDREI DRĂGĂNESCU‡

Abstract. The focus of this work is on the construction and analysis of optimal-order multigrid preconditioners to be used in the Newton–Krylov method for a distributed optimal control problem constrained by the stationary Navier–Stokes equations. As in our earlier work [*Appl. Math. Comput.*, 219 (2013), pp. 5622–5634] on the optimal control of the stationary Stokes equations, the strategy is to eliminate the state and adjoint variables from the optimality system and solve the reduced nonlinear system in the control variables. While the construction of the preconditioners extends naturally the work in the aforementioned, the analysis shown in this paper presents a set of significant challenges that are rooted in the nonlinearity of the constraints. We also include numerical results that showcase the behavior of the proposed preconditioners and show that for low to moderate Reynolds numbers they can lead to significant drops in the number of iterations and wall-clock savings.

Key words. multigrid methods, PDE-constrained optimization, Navier–Stokes equations, finite elements

AMS subject classifications. 65F08, 65K15, 65N21, 65N55, 90C06

DOI. 10.1137/18M1175264

1. Introduction.

We consider the optimal control problem

$$\min_{y,p,u} J(y,p,u) = \frac{\gamma_y}{2}\|y-y_d\|^2_{\mathbf{L}^2(\Omega)} + \frac{\gamma_p}{2}\|p-p_d\|^2_{L^2(\Omega)} + \frac{\beta}{2}\|u\|^2_{\mathbf{L}^2(\Omega)}, \tag{1}$$

subject to the stationary Navier–Stokes equations

$$\begin{aligned} -\nu\Delta y + (y\cdot\nabla)y + \nabla p &= u \quad \text{in } \Omega,\\ \operatorname{div} y &= 0 \quad \text{in } \Omega,\\ y &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{2}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded convex polygonal domain. The goal of the control problem is to find a force u that gives rise to a velocity y and/or pressure p to match a known target velocity y_d, respectively, pressure p_d. Since this problem is ill-posed, we consider a standard Tikhonov regularization for the force, with the regularization parameter β being a fixed positive number. The constants γ_y, γ_p are nonnegative, not both zero.

Model problems like (1)–(2) are commonly encountered in the literature on optimal control of partial differential equations (PDEs), where boundary conditions, forcing terms, initial values, or coefficients are treated as controls in order for the

*Received by the editors March 13, 2018; accepted for publication (in revised form) March 26, 2019; published electronically June 27, 2019.
http://www.siam.org/journals/sinum/57-3/M117526.html
Funding: The work of the authors was supported by the U.S. Department of Energy Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program under award DE-SC0005455 and by the National Science Foundation under awards DMS-1016177 and DMS-0821311.
†Department of Mathematics, United States Naval Academy, MD 21402 (soane@usna.edu).
‡Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, MD 21250 (draga@umbc.edu).

solution they determine to be close to a target state. For a given PDE-constraint, the associated control problem that is most studied is the case of distributed body forcing as control (e.g., see [33]).

Several works are centered on optimal control problems constrained by the Navier–Stokes equations (see, e.g., [19, 21, 22, 11, 20, 15, 16, 12, 5] and the references therein), where both optimality conditions and numerical methods are addressed, for the unconstrained, control-constrained, or mixed control-state constrained problems. For a comprehensive overview of optimal flow control we refer the reader to [18]. In light of the potentially very large scale of the problems involved, a critical issue for all PDE-constrained optimization problems is to devise efficient solvers. These solvers largely fall into two categories: the first kind targets the sparse but indefinite Karush–Kuhn–Tucker (KKT) systems [29, 26], while the second kind is centered on reduced systems. Our strategy falls in the second category. More precisely we focus on the efficient solution of the linear systems arising in the solution process of (1)–(2), specifically on the design of multigrid preconditioners for the reduced Hessian in the Newton-CG method. To the best of our knowledge, this has not been addressed in the literature for the Navier–Stokes optimal control problem.

The multigrid preconditioning technique in this paper is rooted in the two- and multilevel methods for linear inverse problems proposed by Rieder [30], Hanke and Vogel [24], and Drăgănescu and Dupont [6], the latter being primarily concerned with regularized time-reversal of parabolic equations. The method has since been extended to distributed control of linear elliptic equations for problems with control constraints [9, 8, 10], distributed optimal control of semilinear elliptic equations [31], distributed optimal control of linear parabolic equations [23], as well as boundary control of elliptic equations [23].

The research in this article extends our earlier work on the distributed optimal control of the Stokes equations [7]; essentially, we show that for low to moderate Reynolds numbers the constructed preconditioners display the same optimal behavior as in the case of the Stokes-constrained problem. The fundamental departure from [7] resides, of course, in the nonlinearity of the constraints. Due to the linearity of the Stokes equations in [7], the cost functional of the reduced system is quadratic; thus the Hessian operator, and hence the preconditioner, is independent of the control. Instead, for the problem (1)–(2), the Hessian depends on the control, and the preconditioner changes at every Newton iteration accordingly. While the construction of the preconditioner is a natural extension of the one in [7], the main contribution in this paper lies in the analysis. The key element of the analysis is the estimation of the $\mathbf{L}^2$-operator norm of the difference between the discrete Hessian and the two-grid preconditioner, as expressed in the main results of this paper, Theorems 9 and 14. Due to the nonlinearity of the constraints, the discrete and continuous Hessians of the reduced cost functionals for (1)–(2) are more involved than in [7], and hence the necessary error estimates leading to the aforementioned results are more challenging. By comparison, the transition from linear elliptic to semilinear elliptic constraints [31] was facilitated by the existence of L^∞-estimates for the control-to-state operator. The merit of our analysis is that we were able to avoid $\mathbf{L}^\infty$ and $\mathbf{W}_1^\infty$-estimates for the Navier–Stokes equations and its linearization, which are more restrictive [13]. For convenience, the analysis is restricted to two-dimensional domains, though most of it can be extended to three dimensions with some restrictions on the discretization, at least for the case of velocity control ($\gamma_p = 0$). The preconditioning formulation can be used without change for three-dimensional problems, since it is based on a velocity-pressure formulation.

The paper is organized as follows. In section 2, we introduce the reduced optimal control problem and review a set of results that will be needed in what follows. In section 3, we introduce the discrete optimal control problem and discuss finite element estimates that will be needed for the multigrid analysis. Section 4 contains the analysis of the two-grid preconditioner and the main results of the paper. In section 5, we showcase numerical experiments that illustrate our theoretical results. Conclusions and a discussion of possible extensions are presented in section 6.

2. Problem formulation.

2.1. Preliminaries. In this section we introduce notation and review some classical existence, uniqueness, and regularity results regarding the Navier–Stokes equations which will allow us to formulate the reduced form (15) of (1)–(2), and will play an essential role in the analysis. We use standard notation for the Sobolev spaces $H^m(\Omega)$ and for their vector-valued counterparts we use the boldface notation. We denote by $\tilde{\mathbf{H}}^{-m}(\Omega)$ the dual (with respect to $\mathbf{L}^2$-inner product) of $\mathbf{H}^m(\Omega) \cap \mathbf{H}_0^1(\Omega)$ and define $Q = L_0^2(\Omega) = \{p \in L^2(\Omega) : \int_\Omega p\,dx = 0\}$, $X = \mathbf{H}_0^1(\Omega)$, and $V = \{v \in \mathbf{H}_0^1(\Omega) : (\operatorname{div} v, q) = 0\ \forall q \in Q\}$. Throughout this paper we write $(\cdot,\cdot)$ for the inner product in $L^2(\Omega)$ or $\mathbf{L}^2(\Omega)$, according to context, if there is no risk of misunderstanding. The $\mathbf{H}^m(\Omega)$ or $H^m(\Omega)$-norm will be denoted by $\|\cdot\|_m$, while $\|\cdot\|$ denotes the $\mathbf{L}^2(\Omega)$ or $L^2(\Omega)$-norm. Furthermore, define the norm in V' by

$$\|u\|_{V'} = \sup_{\phi \in V\setminus\{0\}} (u,\phi)/\|\nabla\phi\|.$$

To define the weak formulation of (2), we introduce the bilinear forms

$$a(y,\phi) = \nu(\nabla y, \nabla\phi) = \nu \sum_{i=1}^{2} \int_\Omega \nabla y_i \cdot \nabla \phi_i\,dx \quad \forall y,\phi \in X, \tag{3}$$

$$b(\phi,p) = -\int_\Omega p \operatorname{div}\phi\,dx \quad \forall \phi \in X, \forall p \in Q, \tag{4}$$

as well as the trilinear form

$$c(y;\phi,\psi) = ((y\cdot\nabla)\phi,\psi) \quad \forall y,\phi,\psi \in \mathbf{H}^1(\Omega). \tag{5}$$

A weak formulation of the Navier–Stokes equations is given by the following:

Given $u \in \mathbf{H}^{-1}(\Omega)$, find $(y,p) \in X \times Q$ satisfying

$$\begin{aligned} a(y,\phi) + c(y;y,\phi) + b(\phi,p) &= \langle u,\phi\rangle \quad && \forall \phi \in X,\\ b(y,q) &= 0 && \forall q \in Q, \end{aligned} \tag{6}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $\mathbf{H}_0^1(\Omega)$ and $\mathbf{H}^{-1}(\Omega)$. Following [27], the system (6) can be written equivalently as

Find $y \in V$ that satisfies

$$a(y,\phi) + c(y;y,\phi) = \langle u,\phi\rangle \quad \forall \phi \in V. \tag{7}$$

We recall here a standard result regarding the existence of a solution of (6) and uniqueness for small data (see, e.g., [14, 27]). For $\mathbf{H}^2$ regularity see [4].

THEOREM 1. *Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with Lipschitz continuous boundary. Then for any $\nu > 0$ and $u \in \mathbf{H}^{-1}(\Omega)$ there exists at least one solution $(y, p) \in V \times Q$ of the stationary Navier–Stokes problem* (6) *that satisfies the estimate*

$$\|\nabla y\| \leq \nu^{-1}\|u\|_{V'}. \tag{8}$$

Moreover, the solution is unique if the data satisfies the smallness condition

$$\mathcal{M}\nu^{-2}\|u\|_{V'} < 1 \text{ with } \mathcal{M} = \sup_{\phi,\psi,\chi \in X\setminus\{0\}} \frac{|c(\phi;\psi,\chi)|}{\|\nabla\phi\|\|\nabla\psi\|\|\nabla\chi\|}. \tag{9}$$

If Ω is convex and polygonal, and $u \in \mathbf{L}^2(\Omega)$, then $y \in \mathbf{H}^2(\Omega)$, $p \in H^1(\Omega)$ and

$$\|y\|_2 + \|p\|_1 \leq C(1 + \|u\|^3). \tag{10}$$

Recall that throughout this paper we will assume Ω to be a convex polygonal domain, so that the $\mathbf{H}^2$-regularity of the Navier–Stokes problem is ensured. We state here some well-known results concerning the trilinear form defined in (5) that will be needed in what follows [3, 14, 22].

LEMMA 2. *The trilinear form $c(y;\phi,\psi)$ defined in* (5) *has the following properties:*

$$\begin{aligned}
&c(y;\phi,\psi) = -c(y;\psi,\phi) \quad \forall y \in V, \forall \phi,\psi \in \mathbf{H}^1(\Omega),\\
&c(y;\phi,\phi) = 0 \quad \forall y \in V, \phi \in \mathbf{H}^1(\Omega),\\
&c(y;\phi,\psi) = ((\nabla\phi)^T\psi, y) \quad \forall y,\phi,\psi \in \mathbf{H}^1(\Omega),\\
&|c(y;\phi,\psi)| \leq \|y\|_1\|\phi\|_1\|\psi\|_1 \quad \forall y,\phi,\psi \in V,\\
&|c(y;\phi,\psi)| \leq \mathcal{M}\|\nabla y\|\|\nabla\phi\|\|\nabla\psi\| \quad \forall y,\phi,\psi \in X,\\
&|c(y;\phi,\psi)| \leq C\|u\|_1\|\phi\|_1\|\psi\|_1 \quad \forall y,\phi,\psi \in \mathbf{H}^1(\Omega),\\
&|c(y;\phi,\psi)| \leq C\|y\|\|\phi\|_2\|\psi\|_1 \quad \forall y,\psi \in X, \phi \in \mathbf{H}^2(\Omega),\\
&|c(y;\phi,\psi)| \leq C\|y\|_1\|\phi\|_2\|\psi\| \quad \forall y,\psi \in X, \phi \in \mathbf{H}^2(\Omega)
\end{aligned} \tag{11}$$

with $\mathcal{M}$ given in (9) *and C independent of y, ϕ, ψ.*

Proof. While the others are standard, we prove here only the last estimate. Using Hölder's inequality and the embedding $\mathbf{H}^1(\Omega) \hookrightarrow \mathbf{L}^4(\Omega)$, we have

$$|c(y;\phi,\psi)| = |((y\cdot\nabla)\phi,\psi)| \leq \|y\|_{L^4(\Omega)}\|\nabla\phi\|_{L^4(\Omega)}\|\psi\|_{L^2(\Omega)} \leq C\|y\|_1\|\phi\|_2\|\psi\|. \qquad \square$$

When discretizing (6) using finite elements, in order to preserve the antisymmetry in the last two arguments of the trilinear form c on the finite element spaces, it is standard to introduce a modified trilinear form [21, 27]

$$\tilde{c}(y;\phi,\psi) = \frac{1}{2}\{c(y;\phi,\psi) - c(y;\psi,\phi)\} \quad \forall y,\phi,\psi \in X \tag{12}$$

that has the following properties:

$$\begin{aligned}
&c(y;\phi,\psi) = \tilde{c}(y;\phi,\psi) \quad \forall y \in V, \phi,\psi \in X,\\
&\tilde{c}(y;\phi,\psi) = -\tilde{c}(y;\psi,\phi) \quad \forall y,\phi,\psi \in X,\\
&\tilde{c}(y;\psi,\psi) = 0 \quad \forall y,\psi \in X,\\
&|\tilde{c}(y;\phi,\psi)| \leq \mathcal{M}\|\nabla y\|\|\nabla\phi\|\|\nabla\psi\| \quad \forall y,\phi,\psi \in X
\end{aligned} \tag{13}$$

for the same $\mathcal{M} = \mathcal{M}(\Omega)$ as in (9). Thus, another variational formulation of (6) is the following:

Given $u \in \mathbf{H}^{-1}(\Omega)$, find $(y,p) \in X \times Q$ satisfying

$$\begin{aligned} a(y,\phi) + \tilde{c}(y;y,\phi) + b(\phi,p) &= \langle u,\phi\rangle \quad \forall \phi \in X, \\ b(y,q) &= 0 \qquad\quad \forall q \in Q. \end{aligned} \tag{14}$$

We define the set of admissible controls $U = \{u : \mathbf{L}^2(\Omega) : \|u\| < \nu^2/(\mathcal{M}\kappa)\}$ with $\mathcal{M}$ defined in (9) and κ the embedding constant of $\mathbf{L}^2(\Omega)$ into V'. By Theorem 1, the Navier–Stokes equations have a unique solution for each $u \in U$ on the right-hand side of (6). We introduce the control-to-state operators $Y : U \to V$, $P : U \to Q$ that assign to each $u \in U \subset \mathbf{L}^2(\Omega)$ the corresponding Navier–Stokes velocity $y = Y(u)$ and pressure $p = P(u)$, and rewrite problem (1) in reduced form as

$$\min_{u\in U} \hat{J}(u) = \frac{\gamma_y}{2}\|Y(u) - y_d\|^2 + \frac{\gamma_p}{2}\|P(u) - p_d\|^2 + \frac{\beta}{2}\|u\|^2. \tag{15}$$

Throughout this paper we will assume that the target velocity field y_d is from $\mathbf{H}^1(\Omega)$; the target pressure p_d is assumed for now to be in Q.

We note that for all pairs $(y(u), u)$ with $u \in U$, we have

$$\nu > \mathcal{M}(y) \quad \text{with } \mathcal{M}(y) := \sup_{v\in X} \frac{|c(v;y,v)|}{\|\nabla v\|^2}, \tag{16}$$

which ensures the ellipticity of the linearized equations about y. The following estimate establishes the regularity of the solution of the linearized equations about y and will be used in section 4.

Lemma 3. *Let $u \in U$ and $y = Y(u) \in V$. Then for every $g \in V'$ there exists a unique weak solution $(w,r) \in X \times Q$ of the linearized Navier–Stokes system*

$$\begin{aligned} -\nu\Delta w + (w\cdot\nabla)y + (y\cdot\nabla)w + \nabla r &= g \quad \text{in } \Omega, \\ \operatorname{div} w &= 0 \quad \text{in } \Omega, \\ w &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{17}$$

and

$$\|\nabla w\| \le \frac{2}{\nu}\|g\|_{V'}. \tag{18}$$

If Ω is convex and polygonal, and $g \in \mathbf{L}^2(\Omega)$, then $w \in \mathbf{H}^2(\Omega)$, $r \in H^1(\Omega)$, and

$$\|w\|_2 \le C(y)\|g\|. \tag{19}$$

Proof. Existence and uniqueness follow from the Lax–Milgram lemma, using (16) to prove the ellipticity of the associated bilinear form. For the proof of (18) see [34, Corollary 3.7]. To prove (19), we note that for $g \in \mathbf{L}^2(\Omega)$, we have $(w\cdot\nabla)y$, $(y\cdot\nabla)w$ $\in \mathbf{L}^2(\Omega)$ (see estimates below); thus by rewriting (17) as

$$-\nu\Delta w + \nabla r = g - (w\cdot\nabla)y - (y\cdot\nabla)w,$$

we can use standard regularity results for the Stokes equations to obtain

$$\|\nabla\nabla w\| \le C_1(\Omega)(\|g\| + \|(w\cdot\nabla)y\| + \|(y\cdot\nabla)w\|). \tag{20}$$

We have

$$\|(w\cdot\nabla)y\|^2 = \int_\Omega |(w\cdot\nabla)y|^2 dx \le \int_\Omega |w|^2|\nabla y|^2 dx \le \|w\|^2_{\mathbf{L}^4(\Omega)}\|\nabla y\|^2_{\mathbf{L}^4(\Omega)},$$

which implies

$$\|(w\cdot\nabla)y\| \le C\|w\|_1\|\nabla y\|_1 \le C_1(y)\|g\|, \tag{21}$$

since $\mathbf{H}^1(\Omega)\hookrightarrow \mathbf{L}^4(\Omega)$. Similarly, it can be shown that

$$\|(y\cdot\nabla)w\| \le C\|y\|_1\|\nabla w\|_{\mathbf{L}^4(\Omega)} \le C_2(y)\|\nabla w\|^{1/2}\|\nabla\nabla w\|^{1/2},$$

where we used Ladyzhenskaya's inequality,

$$\|\nabla w\|_{\mathbf{L}^4(\Omega)} \le C\|\nabla w\|^{1/2}\|\nabla\nabla w\|^{1/2}.$$

Finally, using Young's inequality we obtain

$$\begin{aligned}\|(y\cdot\nabla)w\| &\le C_2(y)\Big(\frac{1}{2}C_2(y)C_1(\Omega)\|\nabla w\| + \frac{1}{2C_2(y)C_1(\Omega)}\|\nabla\nabla w\|\Big)\\ &= \frac{1}{2}C_2^2(y)C_1(\Omega)\|\nabla w\| + \frac{1}{2C_1(\Omega)}\|\nabla\nabla w\|.\end{aligned}$$

Substituting in (20) gives

$$\|\nabla\nabla w\| \le C_1(\Omega)\Big(\|g\| + C_1(y)\|g\| + \frac{1}{2}C_2^2(y)C_1(\Omega)\|\nabla w\| + \frac{1}{2C_1(\Omega)}\|\nabla\nabla w\|\Big),$$

from which (19) follows immediately. ∎

We recall here the following results from [5] regarding the differentiability of the solution operators Y, P.

THEOREM 4. *Let $u\in U$ and $y = Y(u)$. The control-to-state operators Y, P are twice Fréchet differentiable at u and their derivatives $w = Y'(u)g$, $r = P'(u)g$ and $\lambda = Y''(u)[g_1, g_2]$, $\mu = P''(u)[g_1, g_2]$ are given by the unique weak solutions of the systems*

$$\begin{aligned} -\nu\Delta w + (w\cdot\nabla)y + (y\cdot\nabla)w + \nabla r &= g \quad &&\textit{in } \Omega,\\ \operatorname{div} w &= 0 &&\textit{in } \Omega,\\ w &= 0 &&\textit{on } \partial\Omega \end{aligned} \tag{22}$$

and

$$\begin{aligned} -\nu\Delta\lambda + (y\cdot\nabla)\lambda + (\lambda\cdot\nabla)y + \nabla\mu &= -(Y'(u)g_1\cdot\nabla)Y'(u)g_2 \\ &\quad - (Y'(u)g_2\cdot\nabla)Y'(u)g_1 \quad &&\textit{in } \Omega,\\ \operatorname{div}\lambda &= 0 &&\textit{in } \Omega,\\ \lambda &= 0 &&\textit{on } \partial\Omega, \end{aligned} \tag{23}$$

respectively.

Lemma 5. *Let $u \in U$, $y = Y(u)$, and $Y'(u)^*$ be the adjoint of $Y'(u)$. Then $z = Y'(u)^* g$ is the first component of the unique weak solution (z, ρ) of the system*

$$\begin{aligned} -\nu\Delta z - (y\cdot\nabla)z + (\nabla y)^T z + \nabla\rho &= g \quad \text{in } \Omega, \\ \operatorname{div} z &= 0 \quad \text{in } \Omega, \\ z &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{24}$$

If $\Omega \subset \mathbb{R}^2$ is a convex polygonal domain, then $z \in \mathbf{H}^2(\Omega)$, $\rho \in H^1(\Omega)$, and

$$\|z\|_2 \le C(y)\|g\|. \tag{25}$$

Proof. See [34, Theorem 3.10] and Lemma 3. ∎

2.2. Optimality conditions. We derive next the first-order necessary optimality conditions associated with the optimal control problem (15). For $g \in \mathbf{L}^2(\Omega)$,

$$\hat{J}'(u)g = \gamma_y(Y(u) - y_d, Y'(u)g) + \gamma_p(P(u) - p_d, P'(u)g) + \beta(u, g),$$

and therefore

$$\nabla\hat{J}(u) = \gamma_y Y'(u)^*(Y(u) - y_d) + \gamma_p P'(u)^*(P(u) - p_d) + \beta u. \tag{26}$$

Thus, the optimal control u is the solution of the nonlinear equation

$$\gamma_y Y'(u)^*(Y(u) - y_d) + \gamma_p P'(u)^*(P(u) - p_d) + \beta u = 0. \tag{27}$$

The reduced Hessian is computed using the second variation of $\hat{J}$: if $g_1, g_2 \in \mathbf{L}^2(\Omega)$,

$$\begin{aligned} \hat{J}''(u)[g_1, g_2] = {} & \gamma_y(Y'(u)g_2, Y'(u)g_1) + \gamma_y(Y(u) - y_d, Y''(u)[g_2, g_1]) \\ & + \gamma_p(P'(u)g_2, P'(u)g_1) + \gamma_p(P(u) - p_d, P''(u)[g_2, g_1]) + \beta(g_1, g_2). \end{aligned} \tag{28}$$

We use different approaches in proving the main multigrid results, depending on whether the pressure term is present in the cost functional (1) or not, and therefore we will derive the reduced Hessian for the two cases separately.

2.2.1. Velocity control only. We consider first the case of velocity control only, i.e., $\gamma_y = 1, \gamma_p = 0$. In this case the second variation of $\hat{J}$ becomes

$$\hat{J}''(u)[g_1, g_2] = (Y'(u)g_2, Y'(u)g_1) + (Y(u) - y_d, Y''(u)[g_2, g_1]) + \beta(g_1, g_2). \tag{29}$$

We denote by L and M the solution operators of (22), such that $Lg = Y'(u)g$, $Mg = P'(u)g$. Although L, M depend on $y = y(u)$ in (22), we use the notation L, M instead of $L(u)$, $M(u)$, for simplicity, when there is no risk of misunderstanding. Following Theorem 4 $\lambda = Y''(u)[g_1, g_2]$ is the solution of

$$\begin{aligned} a(\lambda, \phi) + c(y; \lambda, \phi) + c(\lambda; y, \phi) + b(\phi, \mu) & \\ = -c(Lg_1; Lg_2; \phi) - c(Lg_2; Lg_1, \phi) & \quad \forall \phi \in X, \\ b(\lambda, q) = 0 & \quad \forall q \in Q. \end{aligned} \tag{30}$$

Similarly, we let $z = L^*(Y(u) - y_d)$. Note that z is the solution of

$$\begin{aligned} a(z, \phi) + c(y; \phi, z) + c(\phi; y, z) + b(\phi, \rho) &= (y - y_d, \phi) \quad \forall \phi \in X, \\ b(z, q) &= 0 \quad \forall q \in Q. \end{aligned} \tag{31}$$

By taking $\phi = z$ in (30) and $\phi = \lambda$ in (31) we obtain

$$-c(Lg_1; Lg_2; z) - c(Lg_2; Lg_1, z) = (Y(u) - y_d, \lambda). \tag{32}$$

Using this in (29) we get

$$\begin{aligned}\hat{J}''(u)[g_1, g_2] &= (Lg_1, Lg_2) - c(Lg_1; Lg_2, z) - c(Lg_2; Lg_1, z) + \beta(g_1, g_2)\\ &= (Lg_1, Lg_2) + ((Lg_1 \cdot \nabla)z, Lg_2) - ((\nabla Lg_1)^T z, Lg_2) + \beta(g_1, g_2).\end{aligned}$$

The Hessian operator associated with $\hat{J}$, defined by $(H_\beta(u)v, g) = \hat{J}''(u)[v, g]$, is

$$H_\beta(u)v = \beta v + L^* Lv + L^*((Lv \cdot \nabla) - (\nabla Lv)^T)L^*(y - y_d). \tag{33}$$

To simplify the presentation we introduce the notation

$$A(u)v = L^* Lv, \quad \mathcal{C}(u)v = L^*((Lv \cdot \nabla) - (\nabla Lv)^T)L^*(Y(u) - y_d),$$

which we will use throughout the paper. Note that

$$(\mathcal{C}(u)v, v) = -2c(Lv; Lv, L^*(Y(u) - y_d)). \tag{34}$$

2.2.2. Mixed/pressure control. Here we consider the general case of mixed velocity/pressure control or pressure control only, i.e, $\gamma_p \neq 0$.

Let $(\tilde{z}, \tilde{\rho})$ be the solution of the problem

$$\begin{aligned}a(\tilde{z}, \phi) + c(y; \phi, \tilde{z}) + c(\phi; y, \tilde{z}) + b(\phi, \tilde{\rho}) &= \gamma_y(y - y_d, \phi) \quad \forall \phi \in X,\\ b(\tilde{z}, q) &= \gamma_p(p - p_d, q) \quad \forall q \in Q,\end{aligned} \tag{35}$$

which is the weak form of the problem

$$\begin{aligned}-\nu\Delta\tilde{z} - (y \cdot \nabla)\tilde{z} + (\nabla y)^T \tilde{z} + \nabla\tilde{\rho} &= \gamma_y(y - y_d) \quad \text{in } \Omega,\\ \operatorname{div} \tilde{z} &= \gamma_p(p_d - p) \quad \text{in } \Omega,\\ \tilde{z} &= 0 \quad \text{on } \partial\Omega.\end{aligned} \tag{36}$$

By taking $\phi = \lambda$ in (35) and $\phi = \tilde{z}$ in (30) and using $b(\lambda, \tilde{\rho}) = 0$, $b(\tilde{z}, \mu) = \gamma_p(p - p_d, \mu)$ we obtain

$$\gamma_y(y - y_d, \lambda) + \gamma_p(p - p_d, \mu) = -c(Lg_1; Lg_2, \tilde{z}) - c(Lg_2; Lg_1, \tilde{z}).$$

Thus, the second variation of the reduced cost functional (28) becomes

$$\begin{aligned}\hat{J}''(u)[g_1, g_2] = \gamma_y(Y'(u)g_2, Y'(u)g_1) + \gamma_p(P'(u)g_2, P'(u)g_1)\\ - c(Lg_1; Lg_2, \tilde{z}) - c(Lg_2; Lg_1, \tilde{z}) + \beta(g_1, g_2)\end{aligned} \tag{37}$$

and the reduced Hessian is given by

$$H_\beta(u)v = \beta v + \gamma_y L^* Lv + \gamma_p M^* Mv + L^*((Lv \cdot \nabla)\tilde{z} - (\nabla Lv)^T \tilde{z}). \tag{38}$$

We introduce the notation

$$\tilde{\mathcal{C}}(u)v = L^*((Lv \cdot \nabla)\tilde{z} - (\nabla Lv)^T \tilde{z}) \tag{39}$$

and note that

$$(\tilde{\mathcal{C}}(u)v, v) = -2c(Lv; Lv, \tilde{z}). \tag{40}$$

Note that if we take $\gamma_y = 1$, $\gamma_p = 0$ in (35), then (35) is the adjoint linearized Navier–Stokes system and in this case (38) reduces to (33).

3. Discretization and approximation results. In order to discretize the optimization problem (1)–(2) we adopt the strategy to first discretize the Navier–Stokes system, then optimize the cost functional J in (1) subject to the discrete constraints.

3.1. Finite element approximation. We consider a shape regular, quasi-uniform quadrilateral mesh $\mathscr{T}_h$ of $\bar{\Omega}$, and we assume that the mesh $\mathscr{T}_h$ results from a coarser regular mesh $\mathscr{T}_{2h}$ from one uniform refinement. We use the Taylor–Hood $\mathbf{Q}_2 - \mathbf{Q}_1$ finite elements to discretize the state equation. The velocity field y is approximated in the space $X_h^0 = X_h \cap \mathbf{H}_0^1(\Omega)$, where

$$X_h = \{v_h \in C(\bar{\Omega})^2 : v_h|_T \in \mathbf{Q}_2(T)^2 \text{ for } T \in \mathscr{T}_h\}$$

and the pressure p is approximated in the space

$$Q_h = \{q_h \in C(\Omega) \cap L_0^2(\Omega) : q_h|_T \in \mathbf{Q}_1(T) \text{ for } T \in \mathscr{T}_h\},$$

where $\mathbf{Q}_k(T)$ is the space of polynomials of degree less than or equal to k in each variable [2]. The control variable u is approximated by continuous piecewise biquadratic polynomial vector functions from X_h. We also introduce the space

$$V_h = \{v_h \in X_h^0 : (\operatorname{div} v_h, q_h) = 0 \ \forall q_h \in Q_h\} \tag{41}$$

and note that $V_h \not\subseteq V$.

Remark 1. The choice to work with quadrilateral $\mathbf{Q}_2 - \mathbf{Q}_1$ Taylor–Hood elements was made for convenience and clarity of exposition; our analysis can be extended to triangular $\mathbf{P}_2 - \mathbf{P}_1$ elements as well as other stable mixed finite elements.

For a given control $u_h \in X_h \cap U$, the solution (y_h, p_h) of the discrete state equation is given by

$$\begin{aligned} a(y_h, \phi_h) + \tilde{c}(y_h; y_h, \phi_h) + b(\phi_h, p_h) &= (u_h, \phi_h) \quad && \forall \phi_h \in X_h^0, \\ b(y_h, q_h) &= 0 && \forall q_h \in Q_h. \end{aligned} \tag{42}$$

Let Y_h and P_h be the solution mappings of the discretized state equation, defined analogously to their continuous counterparts. The discretized, reduced optimal control problem reads

$$\min_{u_h} \hat{J}_h(u_h) = \frac{\gamma_y}{2}\|Y_h(u_h) - y_d^h\|^2 + \frac{\gamma_p}{2}\|P_h(u_h) - p_d^h\|^2 + \frac{\beta}{2}\|u_h\|^2, \tag{43}$$

where y_d^h, p_d^h are the L^2-projections of the data onto X_h, respectively, Q_h.

We denote by L_h, M_h the solution operators of the discretized linearized Navier–Stokes equations (about y_h), i.e., $L_h g = w_h$, $M_h g = r_h$, where

$$\begin{aligned} a(w_h, \phi_h) + \tilde{c}(y_h; w_h, \phi_h) &+ \tilde{c}(w_h; y_h, \phi_h) + b(\phi_h, r_h) \\ &= (g, \phi_h) \quad \forall \phi_h \in X_h^0, \\ b(w_h, q_h) &= 0 \quad \forall q_h \in Q_h. \end{aligned} \tag{44}$$

We remark that, as in the continuous case, $z_h = L_h^* g$ satisfies

$$\begin{aligned} a(z_h, \phi_h) + \tilde{c}(y_h; \phi_h, z_h) &+ \tilde{c}(\phi_h; y_h, z_h) + b(\phi_h, \rho_h) \\ &= (g, \phi_h) \quad \forall \phi_h \in X_h^0, \\ b(z_h, q_h) &= 0 \quad \forall q_h \in Q_h. \end{aligned} \tag{45}$$

3.2. A priori estimates. In this section we collect several approximation results pertaining to the finite element approximation of the Navier–Stokes equations and the linearized/adjoint linearized Navier–Stokes equations that will be needed for the multigrid analysis.

LEMMA 6. *Let π_h be the L^2-orthogonal projection onto X_h. The following approximation properties hold:*

$$\|(I-\pi_h)v\|_{\tilde{\mathbf{H}}^{-k}(\Omega)} \le Ch^k\|v\| \quad \forall v \in \mathbf{L}^2(\Omega), \quad k=1,2, \tag{46}$$

$$\|(I-\pi_h)u\|_{\tilde{\mathbf{H}}^{-1}(\Omega)} \le Ch^2\|u\|_1 \quad \forall u \in \mathbf{H}^1(\Omega) \tag{47}$$

with C independent of h.

Proof. The estimate (46) is a standard result (e.g., see [6]). For (47), let I_h : $\mathbf{H}^1(\Omega) \to X_h$ be the interpolant introduced by Scott and Zhang in [32]. We have

$$\begin{aligned}\|u-\pi_h u\|_{\tilde{\mathbf{H}}^{-1}(\Omega)} &= \sup_{v\in\mathbf{H}^1_0(\Omega)\setminus\{0\}} \frac{(u-\pi_h u, v)}{\|v\|_1} = \sup_{v\in\mathbf{H}^1_0(\Omega)\setminus\{0\}} \frac{(u-\pi_h u, v-I_h v)}{\|v\|_1} \\ &\le \sup_{v\in\mathbf{H}^1_0(\Omega)\setminus\{0\}} \frac{\|u-\pi_h u\|\|v-I_h v\|}{\|v\|_1} \le Ch\|u-\pi_h u\|,\end{aligned}$$

where we have used $\|v-I_hv\| \le Ch\|v\|_1$ (see [32, equation (4.6)]). Moreover,

$$\|u-\pi_h u\| \le \|u-I_h u\| \le ch\|u\|_1,$$

which combined with the previous estimate leads to (47). ∎

THEOREM 7. *Let $u \in U$ and $y = Y(u) \in V \cap \mathbf{H}^2(\Omega)$ (so that $\nu > \mathcal{M}(y)$), and L, M be the velocity/pressure operators of the linearized Navier–Stokes equations about y, and L_h, M_h their discrete counterparts. There exists constants C, $C_1 = C_1(y)$, $C_2 = C_2(y)$, and $C_3 = C_3(y)$ such that the following hold:*

(a) *smoothing:*

$$\|Lv\| \le C_1\|v\|_{\tilde{\mathbf{H}}^{-2}(\Omega)} \quad \forall v \in \mathbf{L}^2(\Omega), \tag{48}$$

$$\|Mv\| \le C_2\|v\|_{\tilde{\mathbf{H}}^{-1}(\Omega)} \quad \forall v \in \mathbf{L}^2(\Omega). \tag{49}$$

(b) *approximation:*

$$\|Y(u)-Y_h(u)\| \le Ch^2\|u\| \quad \forall u \in U, \tag{50}$$

$$\|Lv-L_hv\|_1 \le C_1h\|v\| \quad \forall v \in \mathbf{L}^2(\Omega), \tag{51}$$

$$\|Lv-L_hv\| \le C_1h^2\|v\| \quad \forall v \in \mathbf{L}^2(\Omega), \tag{52}$$

$$\|Mv-M_hv\| \le C_2h\|v\| \quad \forall v \in \mathbf{L}^2(\Omega), \tag{53}$$

$$\|L^*v-L_h^*v\|_1 \le C_3h\|v\| \quad \forall v \in \mathbf{L}^2(\Omega), \tag{54}$$

$$\|L^*v-L_h^*v\| \le C_3h^2\|v\| \quad \forall v \in \mathbf{L}^2(\Omega), \tag{55}$$

(c) *stability:*

$$\|Y_h(u)\| \le C\|u\| \quad \forall u \in U, \tag{56}$$

$$\|L_hv\| \le C_1\|v\| \quad \forall v \in \mathbf{L}^2(\Omega), \tag{57}$$

$$\|M_hv\| \le C_2\|v\| \quad \forall v \in \mathbf{L}^2(\Omega), \tag{58}$$

$$\|L_h^*v\| \le C_3\|v\| \quad \forall v \in \mathbf{L}^2(\Omega). \tag{59}$$

Proof. The statement at (a) is similar to the case of the Stokes problem [7]. For (50) in (b) see [17, p. 32], and for (51)–(55) see [19]. The stability in (c) follows from (8), (a), and (b). ∎

Remark 2. Theorem 7 and Lemma 6 imply that there is a constant $C > 0$ independent of h such that

$$\|L(I-\pi_h)v\| \le Ch^2\|v\| \quad \forall v \in \mathbf{L}^2(\Omega) \tag{60}$$

and

$$\|M(I-\pi_h)v\| \le Ch\|v\| \quad \forall v \in \mathbf{L}^2(\Omega). \tag{61}$$

For a polygonal domain $\Omega \subset \mathbb{R}^2$, the weighted Sobolev space $W_0^{1,0}(\Omega)$ is defined to be the class of functions for which the following norm is finite:

$$\|w\|^2_{W_0^{1,0}(\Omega)} = \int_\Omega |\nabla w|^2 dx + \int_\Omega \delta(x)^{-2}|w|^2 dx,$$

where $\delta(x) = \min\{\mathrm{dist}(x,P) : P \text{ a vertex of } \Omega\}$. The following regularity and approximation result plays an important role in the analysis from section 4.

THEOREM 8. *Let $\Omega \subset \mathbb{R}^2$ be a convex polygonal domain, $u \in U$, $y = Y(u) \in V$, and $f \in \mathbf{L}^2(\Omega)$, $g \in W_0^{1,0}(\Omega)$, $\int_\Omega g dx = 0$. Furthermore, let $\tilde{z} = \tilde{L}(f,g)$, $\tilde{\rho} = \tilde{M}(f,g)$ be the weak solution of*

$$\begin{aligned} -\nu\Delta\tilde{z} - (y\cdot\nabla)\tilde{z} + (\nabla y)^T\tilde{z} + \nabla\tilde{\rho} &= f \quad \text{in } \Omega,\\ \operatorname{div}\tilde{z} &= g \quad \text{in } \Omega,\\ \tilde{z} &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{62}$$

Then $\tilde{z} \in \mathbf{H}^2(\Omega)$, $\tilde{\rho} \in H^1(\Omega)$ and there exists a constant $C = C(\Omega, y) > 0$ such that

$$\|\tilde{z}\|_{\mathbf{H}^2(\Omega)} + \|\nabla\tilde{\rho}\| \le C(\|f\|_{\mathbf{L}^2(\Omega)} + \|g\|_{W_0^{1,0}(\Omega)}). \tag{63}$$

Moreover, if $\tilde{z}_h$ is the velocity of the corresponding discrete problem, then

$$\|\tilde{z}-\tilde{z}_h\|_1 \le Ch(\|f\|_{\mathbf{L}_2(\Omega)} + \|g\|_{W_0^{1,0}(\Omega)}), \quad \|\tilde{z}_h\|_1 \le C(\|f\|_{\mathbf{L}_2(\Omega)} + \|g\|_{W_0^{1,0}(\Omega)}). \tag{64}$$

Proof. The existence of a unique solution $(\tilde{z},\tilde{\rho}) \in X \times Q$ of (62) and the estimate

$$\|\tilde{z}\|_{\mathbf{H}^1(\Omega)} + \|\tilde{\rho}\| \le C(\|f\|_{-1} + \|g\|) \tag{65}$$

follow from standard results for saddle point problems [1]. In [25], it is shown that under the hypotheses of the theorem, the solution of the generalized Stokes system

$$\begin{aligned} -\nu\Delta z + \nabla\rho &= f \quad \text{in } \Omega,\\ \operatorname{div} z &= g \quad \text{in } \Omega,\\ z &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

satisfies $z \in \mathbf{H}^2(\Omega)$, $\rho \in H^1(\Omega)$, and

$$\|z\|_{\mathbf{H}^2(\Omega)} + \|\nabla\rho\| \le C(\|f\| + \|g\|_{W_0^{1,0}(\Omega)}).$$

Using this result together with (65), it is straightforward to show (63) using the same approach as in Lemma 3. For finite element spaces X_h, Q_h that satisfy the inf-sup condition, we have

$$\|\tilde{z}-\tilde{z}_h\|_{\mathbf{H}^1(\Omega)} + \|\tilde{\rho}-\rho_h\| \le C\Big(\inf_{\phi_h\in X_h}\|\tilde{z}-\phi_h\|_{\mathbf{H}^1(\Omega)} + \inf_{q_h\in Q_h}\|\tilde{\rho}-q_h\|\Big),$$

which combined with interpolation estimates yields (64). ∎

4. Two-grid preconditioner. In this section we present the construction of the two-grid preconditioners for the velocity control and mixed/pressure control problems and their analysis. The main results of this paper are Theorems 9 and 14 and their Corollaries 1 and 2. We begin with the description of the discrete Hessians for the two problems in section 4.1, followed by the construction and analysis of the two-grid preconditioners in section 4.2. The velocity control and mixed/pressure control are treated separately, as the form of the Hessian differs significantly in the two cases.

4.1. The discrete Hessian. The discrete Hessian operator at $u \in U \cap X_h$ is defined by the equality

$$(H_\beta^h(u)v, g) = \hat{J}_h''(u)[v, g] \quad \forall v, g \in X_h. \tag{66}$$

4.1.1. Velocity control. As in the continuous case, when $\gamma_p = 0$ we have

$$\nabla \hat{J}_h(u) = Y_h'(u)^*(Y_h(u) - y_d^h) + \beta u, \quad u \in U \cap X_h, \tag{67}$$

with the second variation of the discrete cost functional being given by

$$\hat{J}_h''(u)[g_1, g_2] = (Y_h'(u)g_2, Y_h'(u)g_1) + (Y_h(u) - y_d^h, Y_h''(u)[g_2, g_1]) + \beta(g_1, g_2). \tag{68}$$

The second variation $\lambda_h = Y_h''(u)[g_1, g_2] \in X_h^0$ is the solution of

$$\begin{aligned} &a(\lambda_h, \phi_h) + \tilde{c}(y_h; \lambda_h, \phi_h) + \tilde{c}(\lambda_h; y_h, \phi_h) + b(\phi_h, \mu_h) \\ &\qquad = -\tilde{c}(Y_h'(u)g_1; Y_h'(u)g_2, \phi_h) - \tilde{c}(Y_h'(u)g_2; Y_h'(u)g_1, \phi_h) \quad \forall \phi_h \in X_h^0, \\ &b(\lambda_h, q_h) = 0 \quad \forall q_h \in Q_h. \end{aligned} \tag{69}$$

The discrete adjoint $z_h = Y_h'(u)^*(y_h - y_d^h) = L_h^*(Y_h(u) - y_d^h)$ is the solution of

$$\begin{aligned} &a(z_h, \phi_h) + \tilde{c}(y_h; \phi_h, z_h) + \tilde{c}(\phi_h; y_h, z_h) + b(\phi_h, \rho_h) \\ &\qquad = (y_h - y_d^h, \phi_h) \quad \forall \phi_h \in X_h^0, \\ &b(z_h, q_h) = 0 \quad \forall q_h \in Q_h. \end{aligned} \tag{70}$$

Using the same approach as in the continuous case, we obtain

$$-\tilde{c}(L_h g_1; L_h g_2, z_h) - \tilde{c}(L_h g_2; L_h g_1, z_h) = (y_h - y_d^h, \lambda_h)$$

and

$$\hat{J}_h''(u)[g_1, g_2] = (L_h g_1, L_h g_2) - \tilde{c}(L_h g_1; L_h g_2, z_h) - \tilde{c}(L_h g_2; L_h g_1, z_h) + \beta(g_1, g_2).$$

Hence, the discrete Hessian is given by

$$H_\beta^h(u)v = \beta v + L_h^* L_h v + \mathcal{C}_h(u)v = \beta v + A_h(u)v + \mathcal{C}_h(u)v, \tag{71}$$

where

$$(\mathcal{C}_h(u)v, v) = -2\tilde{c}(L_h v; L_h v, z_h). \tag{72}$$

4.1.2. Mixed/pressure control. Similarly with the derivation in section 2.2.2, in the case of mixed/pressure control, the discrete Hessian takes the form

$$H_\beta^h(u)v = \beta v + \gamma_y L_h^* L_h + \gamma_p M_h^* M_h + \tilde{\mathcal{C}}_h(u)v, \tag{73}$$

where

$$(\tilde{\mathcal{C}}_h(u)v, v) = -2\tilde{c}(L_h v; L_h v, \tilde{z}_h) \tag{74}$$

and $\tilde{z}_h$ is the solution of the discrete problem (35).

4.2. Two-grid preconditioner for discrete Hessian. In this section, we construct and analyze a two-grid preconditioner for the discrete Hessian $H^h_\beta(u)$ defined in (71) and (73). The construction is a natural extension of the technique used for the optimal control of the Stokes equations in [7] and is the same for both velocity- and mixed/pressure control. Let $X_h = X_{2h} \oplus W_{2h}$ be the L^2-orthogonal decomposition, where we consider on X_h the Hilbert-space structure inherited from $\mathbf{L}^2(\Omega)$. Let π_{2h} be the L^2-projector onto X_{2h}. For $u \in U \cap X_h$ we define the two-grid preconditioner

$$T^h_\beta(u) = H^{2h}_\beta(\pi_{2h}u)\pi_{2h} + \beta(I - \pi_{2h}). \tag{75}$$

It is worth noting that

$$(T^h_\beta(u))^{-1} = (H^{2h}_\beta(\pi_{2h}u))^{-1}\pi_{2h} + \beta^{-1}(I - \pi_{2h}). \tag{76}$$

We should remark that the difference between the preconditioner in (75) and the one in [7] is given by the dependence of the Hessian on the control u, which forces us to choose a coarse-level control $u_c \in X_{2h}$ at which the coarse Hessian $H^{2h}_\beta(u_c)$ in (75) is computed. The natural choice is $u_c = \pi_{2h}u$.

4.2.1. Analysis for the case of velocity control. To assess the quality of the preconditioner we use the spectral distance between $H^h_\beta(u)$ and $T^h_\beta(u)$ defined in [6] for two symmetric positive definite operators $T_1, T_2 \in \mathcal{L}(V_h)$ as

$$d_h(T_1, T_2) = \sup_{w \in V_h \setminus \{0\}} \left| \ln \frac{(T_1 w, w)}{(T_2 w, w)} \right|. \tag{77}$$

From (71) and (75) we have

$$T^h_\beta(u) = (\beta I + A_{2h}(\pi_{2h}u) + \mathcal{C}_{2h}(\pi_{2h}u))\pi_{2h} + \beta(I - \pi_{2h}). \tag{78}$$

The key result is the following.

THEOREM 9. *Given $u \in U \cap X_h$, there exists a constant $C = C(\Omega, u, y_d)$ such that*

$$\|(H^h_\beta(u) - T^h_\beta(u))v\| \le Ch^2\|v\| \quad \forall v \in X_h. \tag{79}$$

It is noteworthy that the estimate in Theorem 9 is symmetric in the sense that the same norm (namely, the $\mathbf{L}^2$-norm) appears on both sides of (79) and that the estimate is of optimal order with respect to h. This enables us to prove the following result.

COROLLARY 1. *Let $u \in U \cap X_h$. If $\mathcal{C}_h(u)$ is symmetric positive definite, then*

$$d(H^h_\beta(u), T^h_\beta(u)) \le \frac{C}{\beta}h^2 \tag{80}$$

for $h < h_0(\beta, \Omega, L)$.

Proof. By Theorem 9,

$$\left| \frac{(T^h_\beta(u)v, v)}{(H^h_\beta(u)v, v)} - 1 \right| \le \frac{C}{\beta} \frac{h^2\|v\|^2}{\|v\|^2 + \beta^{-1}(\|L_h v\|^2 + (\mathcal{C}_h(u)v, v))} \le \frac{C}{\beta}h^2.$$

Assume $C\beta^{-1}h_0^2 = \alpha < 1$ and $0 < h \le h_0$. Hence $T_\beta^h(u)$ is positive definite and

$$\sup_{v\in X_h\setminus\{0\}}\left|\ln\frac{(T_\beta^h(u)v,v)}{(H_\beta^h(u)v,v)}\right| \le \frac{|\ln(1-\alpha)|}{\alpha}\sup_{v\in X_h\setminus\{0\}}\left|\frac{(T_\beta^h(u)v,v)}{(H_\beta^h(u)v,v)}\right|$$
$$\le \frac{|\ln(1-\alpha)|}{\alpha}\frac{C}{\beta}h^2 \le \frac{C}{\beta}h^2,$$

where we also used that for $\alpha \in (0,1)$, $x \in [1-\alpha, 1+\alpha]$ we have

$$\frac{\ln(1+\alpha)}{\alpha}|1-x| \le |\ln x| \le \frac{|\ln(1-\alpha)|}{\alpha}|1-x|. \qquad \square$$

Prior to presenting the proof of Theorem 9 we prove some preliminary lemmas.

LEMMA 10. *Let $u \in U \cap X_h$ and $y = Y(u)$, $p = P(u)$, $\bar p = P(\pi_{2h}u)$, $\bar y = Y(\pi_{2h}u)$. Also, let $v \in X_h$ and $w = L(u)v$, $q = M(u)v$, $\bar w = L(\pi_{2h}u)v$, $\bar q = M(\pi_{2h}u)v$. Then there exists a constant $K = K(u,\nu,\Omega) > 0$ such that*

$$|y-\bar y|_1 \le K\|u-\pi_{2h}u\|_{\tilde{\mathbf{H}}^{-1}(\Omega)}, \tag{81}$$
$$\|p-\bar p\| \le Kh^2\|u\|_1, \tag{82}$$

and a constant C independent of h such that

$$\|w-\bar w\|_1 \le Ch^2\|u\|_1\|v\|, \tag{83}$$
$$\|q-\bar q\| \le Ch^2\|u\|_1\|v\|. \tag{84}$$

Proof. Since y and $\bar y$ are the solutions of the Navier–Stokes equations with forcing u, $\pi_{2h}u$, respectively, we have

$$a(y,\phi) + c(y;y,\phi) = (u,\phi) \quad \forall\phi \in V,$$
$$a(\bar y,\phi) + c(\bar y;\bar y,\phi) = (\pi_{2h}u,\phi) \quad \forall\phi \in V.$$

By taking $\phi = y - \bar y$ and subtracting the equations we obtain

$$a(y-\bar y, y-\bar y) + c(y-\bar y; y, y-\bar y) + c(\bar y; y-\bar y, y-\bar y) = (u-\pi_{2h}u, y-\bar y).$$

Given that $c(\bar y; y-\bar y, y-\bar y) = 0$, we obtain

$$\nu|y-\bar y|_1^2 = (u-\pi_{2h}u, y-\bar y) - c(y-\bar y; y, y-\bar y)$$
$$\le \|u-\pi_{2h}u\|_{\tilde{\mathbf{H}}^{-1}}\|y-\bar y\|_1 + \mathcal{M}(y)|y-\bar y|_1^2.$$

Since $\mathcal{M}(y) < \nu$ and $y, \bar y \in X = \mathbf{H}_0^1(\Omega)$, we get

$$(\nu - \mathcal{M}(y))|y-\bar y|_1^2 \le \|u-\pi_{2h}u\|_{\tilde{\mathbf{H}}^{-1}}\|y-\bar y\|_1 \le C\|u-\pi_{2h}u\|_{\tilde{\mathbf{H}}^{-1}}|y-\bar y|_1,$$

which implies (81). From the weak formulations of the Navier–Stokes equations in X, with forcing u, $\pi_{2h}u$, respectively, we have

$$b(\phi, p-\bar p) = (u-\pi_{2h}u,\phi) - a(y-\bar y,\phi) + c(\bar y;\bar y,\phi) - c(y;y,\phi).$$

Thus for $\phi \in X$

$$|b(\phi,p-\bar p)| \le \|u-\pi_{2h}u\|_{-1}\|\phi\|_1 + \nu\|y-\bar y\|_1\|\phi\|_1 + |c(\bar y; y-\bar y,\phi) + c(\bar y - y; y,\phi)|$$
$$\le \|u-\pi_{2h}u\|_{-1}\|\phi\|_1 + \nu\|y-\bar y\|_1\|\phi\|_1 + \|y-\bar y\|_1\|\phi\|_1(\|\bar y\|_1 + \|y\|_1).$$

Then, from the inf-sup condition

$$\beta^* \|q - \bar{q}\| \leq \sup_{0 \neq \phi \in X} \frac{|b(q - \bar{q}, \phi)|}{\|\nabla \phi\|}, \tag{85}$$

combined with (81), (47), we obtain

$$\|p - \bar{p}\| \leq C(\nu, u, \beta^*) h^2 \|u\|_1.$$

Recall that (w, q) (resp., $(\bar{w}, \bar{q})$) satisfy the linearized Navier–Stokes equations (22) about y (resp., $\bar{y}$) with forcing v, whose weak form in V reads

$$a(w, \phi) + c(w; y, \phi) + c(y; w, \phi) = (v, \phi) \quad \forall \phi \in V, \tag{86}$$

$$a(\bar{w}, \phi) + c(\bar{w}; \bar{y}, \phi) + c(\bar{y}; \bar{w}, \phi) = (v, \phi) \quad \forall \phi \in V. \tag{87}$$

By taking $\phi = w - \bar{w}$ in the equations above and subtracting we obtain

$$\begin{aligned} -a(w - \bar{w}, w - \bar{w}) =& c(w; y; w - \bar{w}) + c(y; w, w - \bar{w}) \\ & - c(\bar{w}; \bar{y}, w - \bar{w}) - c(\bar{y}; \bar{w}, w - \bar{w}). \end{aligned} \tag{88}$$

We have

$$\begin{aligned} &c(w; y, w - \bar{w}) - c(\bar{w}; \bar{y}, w - \bar{w}) = c(w; y - \bar{y}, w - \bar{w}) + c(w - \bar{w}; \bar{y}, w - \bar{w}), \\ &c(y; w, w - \bar{w}) - c(\bar{y}; \bar{w}; w - \bar{w}) = c(y - \bar{y}; w, w - \bar{w}), \end{aligned}$$

where we used $c(\bar{y}; w - \bar{w}, w - \bar{w}) = 0$ (see Lemma 2). Using these in (88), we obtain

$$\nu |w - \bar{w}|_1^2 = |c(y - \bar{y}; w; w - \bar{w}) + c(w; y - \bar{y}, w - \bar{w}) + c(w - \bar{w}; \bar{y}; w - \bar{w})|.$$

From the continuity of the trilinear form c and (16) we get

$$\nu |w - \bar{w}|_1^2 \leq \mathcal{M} \left(|y - \bar{y}|_1 |w|_1 |w - \bar{w}|_1 + |w|_1 |y - \bar{y}|_1 |w - \bar{w}|_1 \right) + \mathcal{M}(\bar{y}) |w - \bar{w}|_1^2,$$

which leads to

$$(\nu - \mathcal{M}(\bar{y})) |w - \bar{w}|_1^2 \leq 2\mathcal{M} |w|_1 |y - \bar{y}|_1 |w - \bar{w}|_1.$$

Since $\|\pi_{2h} u\| \leq \|u\|$, $\pi_{2h} u \in U$, and so $\nu - \mathcal{M}(\bar{y}) > 0$; hence we obtain

$$|w - \bar{w}|_1 \leq C |y - \bar{y}|_1 |w|_1 \overset{(18),(47),(81)}{\leq} C h^2 \|u\|_1 \|v\| \tag{89}$$

with C depending on ν, y, κ, $\mathcal{M}$, but not on h. To prove (84), we consider the weak formulations of (86) and (87) in X

$$\begin{aligned} &a(w, \phi) + c(w; y, \phi) + c(y; w, \phi) + b(q, \phi) = (v, \phi) \quad \forall \phi \in X, \\ &a(\bar{w}, \phi) + c(\bar{w}; \bar{y}, \phi) + c(\bar{y}; \bar{w}, \phi) + b(\bar{q}, \phi) = (v, \phi) \quad \forall \phi \in X, \end{aligned}$$

from which we obtain

$$\begin{aligned} b(q - \bar{q}, \phi) &= -a(w - \bar{w}, \phi) - c(w; y, \phi) - c(y; w, \phi) + c(\bar{w}; \bar{y}, \phi) + c(\bar{y}; \bar{w}, \phi) \\ &= -a(w - \bar{w}, \phi) - c(w; y - \bar{y}, \phi) - c(w - \bar{w}; \bar{y}, \phi) - c(y; w - \bar{w}, \phi) \\ &\qquad - c(y - \bar{y}; \bar{w}, \phi) \quad \forall \phi \in X. \end{aligned}$$

Thus, $\forall \phi \in X$

$$|b(q-\bar{q},\phi)| \le C|\phi|_1\left(|w-\bar{w}|_1 + |w-\bar{w}|_1(|y|_1+|\bar{y}|_1) + |y-\bar{y}|_1(|w|_1+|\bar{w}|_1)\right).$$

Using the inf-sup condition (85) we obtain

$$\begin{aligned}\|q-\bar{q}\| &\le C(|w-\bar{w}|_1 + |y-\bar{y}|_1(|w|_1+|\bar{w}|_1) + |w-\bar{w}|_1(|y|_1+|\bar{y}|_1)) \\ &\overset{(89)}{\le} C|y-\bar{y}|_1(|\bar{w}|_1 + |w|_1(1+|y|_1+|\bar{y}|_1)) \\ &\le C|y-\bar{y}|_1\|v\| \overset{(47),(81)}{\le} Ch^2\|u\|_1\|v\|. \qquad \square\end{aligned}$$

LEMMA 11. *Let $u \in U \cap X_h$ and $y = Y(u)$, $\bar{y} = Y(\pi_{2h}u)$. Also, let $v \in X_h$ and $z = L^*(y-y_d)$, $\bar{z} = L^*(\bar{y}-y_d)$. Then there exists a constant $C = C(u, y_d)$ independent of h such that*

$$\|z-\bar{z}\| \le Ch^2\|u\|_1. \tag{90}$$

Proof. Recall that z and $\bar{z}$ are solutions of

$$\begin{aligned} a(z,\phi) + c(y;\phi,z) + c(\phi;y,z) &= (y-y_d,\phi) \quad \forall \phi \in V \\ a(\bar{z},\phi) + c(\bar{y};\phi,\bar{z}) + c(\phi;\bar{y},\bar{z}) &= (\bar{y}-y_d,\phi) \quad \forall \phi \in V.\end{aligned}$$

By taking $\phi = z - \bar{z}$ in the previous equations and subtracting them we obtain

$$\begin{aligned}&\nu|z-\bar{z}|_1^2 \\ &\le |(y-\bar{y},z-\bar{z})| + |c(y-\bar{y};z-\bar{z},\bar{z})| + |c(z-\bar{z};y-\bar{y},z)| + |c(z-\bar{z};\bar{y},z-\bar{z})| \\ &\le C_1\|y-\bar{y}\|_{\mathbf{H}^{-1}}|z-\bar{z}|_1 + \|y-\bar{y}\|_1\|z-\bar{z}\|_1(\|\bar{z}\|_1+\|z\|_1) + \mathcal{M}(\bar{y})|z-\bar{z}|_1^2,\end{aligned}$$

which gives

$$(\nu - \mathcal{M}(\bar{y}))|z-\bar{z}|_1^2 \le |z-\bar{z}|_1(C_1\|y-\bar{y}\| + \|y-\bar{y}\|_1(\|z\|_1+\|\bar{z}\|_1)).$$

Hence,

$$|z-\bar{z}|_1 \le \|y-\bar{y}\|_1(C_1 + C_2\|y-y_d\| + C_3\|\bar{y}-y_d\|) \overset{(81),(47)}{\le} C(u,y_d)h^2\|u\|_1$$

from which (90) follows immediately. $\square$

LEMMA 12. *Let $u \in U \cap X_h$, $y = Y(u)$, $y_h = Y_h(u)$. Also, let $z = L^*(y-y_d)$ and $z_h = L_h^*(y_h - y_d^h)$. Then there exists $C = C(u, y_d)$ independent of h so that*

$$\|y_h - y_d^h\| \le C(\|u\| + \|y_d\|_1), \tag{91}$$

$$\|z - z_h\|_k \le Ch^{2-k}\|u\|, \quad k = 0, 1. \tag{92}$$

Proof. We have

$$\begin{aligned}\|y_h - y_d^h\| \le \|y_h\| + \|y_d\| + \|y_d - y_d^h\| &\overset{(56)}{\le} C\|u\| + \|y_d\| + \|y_d - y_d^h\| \\ &\le C\|u\| + \|y_d\| + \|y_d - I_h y_d\| \le C(\|u\| + \|y_d\| + h\|y_d\|_1).\end{aligned}$$

For $h < 1$ this leads to (91). To prove (92), recall that z and z_h satisfy (31) and (70), respectively. Let $(\bar{z}_h, \bar{\rho}_h)$ be the solution of

$$\begin{aligned} a(\bar{z}_h, \phi_h) + \tilde{c}(y_h; \phi_h, \bar{z}_h) + \tilde{c}(\phi_h; y_h, \bar{z}_h) + b(\phi_h, \bar{\rho}_h) & \\ = (y - y_d, \phi_h) \quad \forall \phi_h \in X_h^0, & \\ b(\bar{z}_h, q_h) = 0 \quad \forall q_h \in Q_h. & \end{aligned} \tag{93}$$

From (54)–(55), we have

$$\|z - \bar{z}_h\|_k \le Ch^{2-k}\|y - y_d\|, \quad k = 0, 1. \tag{94}$$

By taking $\phi_h = z_h - \bar{z}_h$ in (70) and (93) and subtracting the equations we obtain

$$\begin{aligned} \nu|z_h - \bar{z}_h|_1^2 + \tilde{c}(y_h; z_h - \bar{z}_h, z_h - \bar{z}_h) + \tilde{c}(z_h - \bar{z}_h; y_h, z_h - \bar{z}_h) + b(z_h - \bar{z}_h, \rho_h - \bar{\rho}_h) \\ = (y - y_h, z_h - \bar{z}_h) - (y_d - y_d^h, z_h - \bar{z}_h), \end{aligned}$$

which, by using (12) and $(y_d - y_d^h, z_h - \bar{z}_h) = 0$, simplifies to

$$\nu|z_h - \bar{z}_h|_1^2 + \tilde{c}(z_h - \bar{z}_h; y_h, z_h - \bar{z}_h) = (y - y_h, z_h - \bar{z}_h).$$

Thus,

$$\nu|z_h - \bar{z}_h|_1^2 \le \|y - y_h\|\|z_h - \bar{z}_h\| + \mathcal{M}(y_h)|z_h - \bar{z}_h|_1^2.$$

Since $\nu - \mathcal{M}(y_h) > 0$, we obtain

$$\|z_h - \bar{z}_h\|_1 \le C\|y - y_h\| \overset{(50)}{\le} Ch^2\|u\|,$$

which combined with (94) proves the lemma. ∎

We now present the proof of Theorem 9.

Proof. Using (71) and (78),

$$T_\beta^h(u) - H_\beta^h(u) = A_{2h}(\pi_{2h}u)\pi_{2h} - A_h(u) + \mathcal{C}_{2h}(\pi_{2h}u)\pi_{2h} - \mathcal{C}_h(u).$$

We first estimate

$$\begin{aligned} A_{2h}(\pi_{2h}u)\pi_{2h} - A_h(u) = {} & [A_{2h}(\pi_{2h}u) - A(\pi_{2h}u)]\pi_{2h} + A(\pi_{2h}u)(\pi_{2h} - I) \\ & + A(\pi_{2h}u) - A(u) + A(u) - A_h(u). \end{aligned} \tag{95}$$

For any $v \in X_h$ we have

$$\begin{aligned} |(A(u) - A_h(u))v, v)| &= |(L^*Lv - L_h^*L_hv, v)| = |\|Lv\|^2 - \|L_hv\|^2| \\ &\le \|(L - L_h)v\|(\|Lv\| + \|L_hv\|) \le Ch^2\|v\|^2, \end{aligned}$$

which implies

$$\|(A(u) - A_h(u))v\| \le Ch^2\|v\|,$$

since $A(u) - A_h(u)$ is symmetric on X_h. Similarly, it can be shown that

$$\|(A_{2h}(\pi_{2h}u) - A(\pi_{2h}u))\pi_{2h}v\| \le Ch^2\|v\|.$$

For the second term in (95) we have

$$\|A(\pi_{2h}u)(\pi_{2h}-I)v\| = \|L^*(\pi_{2h}u)L(\pi_{2h}u)(\pi_{2h}-I)v\| \overset{(60)}{\leq} Ch^2\|v\|.$$

Finally, we have

$$\begin{aligned}
&|(A(\pi_{2h}u)v - A(u)v, v)| = |(L^*(\pi_{2h}u)L(\pi_{2h}u)v - L^*(u)L(u)v, v)| \\
&\quad = |\|L(\pi_{2h}u)v\|^2 - \|L(u)v\|^2| \leq \|(L(\pi_{2h}u)v - L(u)v\|(\|L(\pi_{2h}u)v\| + \|L(u)v\|) \\
&\quad \overset{(83)}{\leq} Ch^2\|u\|_1\|v\|,
\end{aligned}$$

which implies

$$\|(A(\pi_{2h}u) - A(u))v\| \leq Ch^2\|v\|.$$

Combining this with the previous estimates we obtain

$$\|(A_{2h}(\pi_{2h}u) - A_h(u))v\| \leq Ch^2\|v\|. \tag{96}$$

Next, we estimate

$$\begin{aligned}
\mathcal{C}_{2h}(\pi_{2h}u)\pi_{2h} - \mathcal{C}_h(u) = (\mathcal{C}_{2h}(\pi_{2h}u) - \mathcal{C}(\pi_{2h}u))\pi_{2h} + \mathcal{C}(\pi_{2h}u)(\pi_{2h} - I) \\
+ \mathcal{C}(\pi_{2h}u) - \mathcal{C}(u) + \mathcal{C}(u) - \mathcal{C}_h(u).
\end{aligned} \tag{97}$$

We begin by estimating the term $\|\mathcal{C}(u)v - \mathcal{C}_h(u)v\|$. Let $y = Y(u)$, $y_h = Y_h(u)$, $z = L^*(y - y_d)$, $z_h = L_h^*(y_h - y_d^h)$. Using (34) and (72),

$$(\mathcal{C}(u)v, v) = -2c(Lv; Lv, z) \text{ and } (\mathcal{C}_h(u)v, v) = -2\tilde{c}(L_hv; L_hv, z_h).$$

We have $c(Lv; Lv, z) = \tilde{c}(Lv; Lv, z)$ since $Lv \in V$. Therefore,

$$\begin{aligned}
|(\mathcal{C}(u)v - \mathcal{C}_h(u)v, v)| &= |2\tilde{c}(Lv; Lv, z) - 2\tilde{c}(L_hv; L_hv, z_h)| \\
&\leq |c(Lv; Lv, z) - c(L_hv; L_hv, z_h)| \\
&\quad + |c(Lv; z, Lv) - c(L_hv; z_h, L_hv)|.
\end{aligned}$$

The first term in the inequality above can be bounded by

$$\begin{aligned}
&|c(L_hv; L_hv, z_h) - c(Lv; Lv, z)| \\
&\qquad \leq |c((L_h - L)v; L_hv, z_h)| + |c(Lv; L_hv, z_h - z)| + |c(Lv; z, (L_h - L)v|,
\end{aligned}$$

where we used $c(Lv; (L_h - L)v, z) = -c(Lv; z, (L_h - L)v)$, since $Lv \in V$.
We have

$$\begin{aligned}
&|c((L_h - L)v; L_hv, z_h)| \\
&\leq |c((L_h - L)v; (L_h - L)v, z_h)| + |c((L_h - L)v; Lv, z_h)| \\
&\leq \|(L_h - L)v||_1\|(L_h - L)v\|_1\|z_h\|_1 + \|(L_h - L)v\|\|Lv\|_2\|z_h\|_1 \\
&\overset{(51),(52)}{\leq} Ch^2\|v\|^2\|z_h\|_1 \overset{(59)}{\leq} Ch^2\|v\|^2\|y_h - y_d^h\| \overset{(91)}{\leq} C(u, y_d)h^2\|v\|^2
\end{aligned}$$

and

$$\begin{aligned}
|c(Lv; L_hv, z_h - z)| &\leq |c(Lv; (L_h - L)v, z_h - z)| + |c(Lv; Lv, z_h - z)| \\
&\leq \|Lv\|_1\|(L_h - L)v\|_1\|z_h - z\|_1 + C\|Lv\|_1\|Lv\|_2\|z_h - z\| \\
&\overset{(51),(19),(92)}{\leq} Ch^2\|v\|^2\|y_h - y_d^h\| \overset{(91)}{\leq} C(u, y_d)h^2\|v\|^2.
\end{aligned}$$

Combining these estimates with

$$|c(Lv; z, (L_h - L)v)| \leq C\|Lv\|_1\|z\|_2\|(L_h - L)v\| \overset{(25),(52)}{\leq} C(u, y_d)h^2\|v\|^2,$$

we obtain

$$|c(L_hv; L_hv, z_h) - c(Lv; Lv, z)| \leq C(u, y_d)h^2\|v\|^2.$$

Similarly,

$$\begin{aligned}
&|c(L_hv; z_h, L_hv) - c(Lv; z, Lv)| \\
&\quad \leq |c((L_h - L)v; z_h, L_hv)| + |c(Lv; z_h, (L_h - L)v)| + |c(Lv; Lv, z - z_h)| \\
&\quad \leq |c((L_h - L)v; z_h - z, L_hv)| + |c((L_h - L)v; z, L_hv)| \\
&\qquad + |c(Lv; z_h - z, (L_h - L)v)| + |c(Lv; z, (L_h - L)v)| + |c(Lv; Lv, z - z_h)| \\
&\quad \leq \|(L_h - L)v\|_1\|z_h - z\|_1\|L_hv\|_1 + C\|(L_h - L)v\|\|z\|_2\|L_hv\|_1 \\
&\qquad + \|Lv\|_1\|z_h - z\|_1\|(L_h - L)v\|_1 + C\|Lv\|_1\|z\|_2\|(L_h - L)v\| \\
&\qquad + C\|Lv\|_1\|Lv\|_2\|z - z_h\| \overset{(25),(51),(52),(92)}{\leq} C(u, y_d)h^2\|v\|^2.
\end{aligned}$$

Using the same approach, it can be shown that

$$\|(\mathcal{C}_{2h}(\pi_{2h}u) - \mathcal{C}(\pi_{2h}u)\pi_{2h}v\| \leq Ch^2\|v\|.$$

Let $\bar{z} = L^*(Y(\pi_{2h}u) - y_d)$. The third term in (97) can be bounded as

$$\begin{aligned}
|(\mathcal{C}(\pi_{2h}u)v - \mathcal{C}(u)v, v)| &= 2|c(Lv; Lv, \bar{z}) - c(Lv; Lv, z)| = 2|c(Lv; Lv, \bar{z} - z)| \\
&\leq C\|Lv\|_1\|Lv\|_2\|\bar{z} - z\| \overset{(19)}{\leq} C\|v\|^2\|\bar{z} - z\| \\
&\overset{(90)}{\leq} C(u, y_d)h^2\|u\|_1\|v\|^2.
\end{aligned}$$

Finally let $w = (\pi_{2h} - I)v$. With $L = L(\pi_{2h}u)$, we have

$$\begin{aligned}
&|(\mathcal{C}(\pi_{2h}u)(\pi_{2h} - I)v, v)| \\
&= |((Lw \cdot \nabla)\bar{z} - (\nabla Lw)^T\bar{z}, Lv)| \leq |((Lw \cdot \nabla)\bar{z}), Lv)| + |(\nabla Lw)^T\bar{z}, Lv)| \\
&= |c(Lw; \bar{z}, Lv)| + |c(Lv; Lw, \bar{z})| \overset{(11)}{=} |c(Lw; \bar{z}, Lv)| + |c(Lv; \bar{z}, Lw)| \\
&\overset{(11)}{\leq} C\|Lw\|\|\bar{z}\|_2\|Lv\|_1 \overset{(60)}{\leq} Ch^2\|v\|\|\bar{z}\|_2\|v\| \overset{(25)}{\leq} C(u, y_d)h^2\|v\|^2. \qquad \square
\end{aligned}$$

4.2.2. Analysis for the case of mixed/pressure control. Recall from (73) that in the case of mixed/pressure control, the discrete Hessian takes the form

$$H_\beta^h(u)v = \beta v + \gamma_y A_hv + \gamma_p B_hv + \tilde{\mathcal{C}}_h(u)v,$$

where $B_h = M_h^* M_h$ and $A_h = L_h^* L_h$ as in (71). Following the definition in (75), the two-grid preconditioner takes the form

$$T_\beta^h(u) = (\beta I + \gamma_y A_{2h}(\pi_{2h}u) + \gamma_p B_{2h}(\pi_{2h}u) + \tilde{\mathcal{C}}_{2h}(\pi_{2h}u))\pi_{2h} + \beta(I - \pi_{2h}). \tag{98}$$

LEMMA 13. *Let $u \in U \cap X_h$ and $y = Y(u)$, $p = P(u)$, $\bar{y} = Y(\pi_{2h}u)$, $\bar{p} = P(\pi_{2h}u)$. Also, let $v \in X_h$ and $\tilde{z} = \tilde{L}(\gamma_y(y - y_d), \gamma_p(p_d - p))$, $\hat{z} = \tilde{L}(\gamma_y(\bar{y} - y_d), \gamma_p(p_d - \bar{p}))$ with $\tilde{L}$ defined in Theorem 8. Then there exists a constant $C = C(u, y_d, p_d, \gamma_y, \gamma_p)$ independent of h such that*

$$\|\tilde{z} - \hat{z}\|_1 \leq Ch\|u\|_1^{1/2}. \tag{99}$$

Proof. Recall that $(\tilde{z}, \tilde{\rho})$ is the solution of (35), and $(\hat{z}, \hat{\rho})$ satisfies

$$\begin{aligned} a(\hat{z}, \phi) + c(\bar{y}; \phi, \hat{z}) + c(\phi; \bar{y}, \hat{z}) + b(\phi, \hat{\rho}) &= \gamma_y(\bar{y} - y_d, \phi) \quad \forall \phi \in X, \\ b(\hat{z}, q) &= \gamma_p(\bar{p} - p_d, q) \quad \forall q \in Q. \end{aligned} \tag{100}$$

By subtracting (100) from (35) we obtain

$$\begin{aligned} a(\tilde{z} - \hat{z}, \phi) + c(y; \phi, \tilde{z} - \hat{z}) + c(\phi; y, \tilde{z} - \hat{z}) + b(\phi, \tilde{\rho} - \bar{\rho}) &= \\ \gamma_y(y - \bar{y}) + c(\bar{y} - y; \phi, \hat{z}) &+ c(\phi; \bar{y} - y, \hat{z}) \\ b(\tilde{z} - \bar{z}, q) = \gamma_p(\bar{p} - p) \ \ \forall \phi \in X, q \in Q&, \end{aligned}$$

which represents the weak form of

$$\begin{aligned} -\nu\Delta(\tilde{z} - \hat{z}) - (y \cdot \nabla)(\tilde{z} - \hat{z}) + (\nabla y)^T(\tilde{z} - \hat{z}) + \nabla(\tilde{\rho} - \hat{\rho}) &= \gamma_y(y - \bar{y}) + ((y - \bar{y}) \cdot \nabla)\hat{z} \\ &\quad - (\nabla(y - \bar{y}))^T\hat{z} \quad \text{in } \Omega, \\ \operatorname{div}(\tilde{\rho} - \hat{\rho}) &= \gamma_p(\bar{p} - p) \quad \text{in } \Omega \\ \tilde{z} - \hat{z} &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Using (65), we get

$$\|\tilde{z} - \hat{z}\|_1 \leq C(\gamma_y\|y - \bar{y}\| + \gamma_p\|p - \bar{p}\| + \|((y - \bar{y}) \cdot \nabla)\hat{z}\| + \|(\nabla(y - \bar{y}))^T\hat{z}\|). \tag{101}$$

From (21), combined with (81), (47), (65), we have

$$\|((y - \bar{y}) \cdot \nabla)\hat{z}\| \leq C\|y - \bar{y}\|_1\|\nabla\hat{z}\| \leq Ch^2(\gamma_y\|\bar{y} - y_d\| + \gamma_p\|\bar{p} - p_d\|).$$

Of the four terms in the right-hand side of (101), only the last is of order one in h:

$$\begin{aligned} \|(\nabla(y - \bar{y}))^T\hat{z}\| &\leq \|\nabla(y - \bar{y})\|_{\mathbf{L}^4(\Omega)}\|\hat{z}\|_{\mathbf{L}^4(\Omega)} \leq C\|\nabla(y - \bar{y})\|^{1/2}\|\nabla\nabla(y - \bar{y})\|^{1/2}\|\hat{z}\|_1 \\ &\overset{(10),(81),(47),(65)}{\leq} C(u)h\|u\|_1^{1/2}(\|\bar{y} - y_d\| + \|\bar{p} - p_d\|). \end{aligned}$$

Using these estimates together with (81), (82), (47) in (101) we obtain

$$\|\tilde{z} - \hat{z}\|_1 \leq C(u)h\|u\|_1^{1/2}(\gamma_y\|\bar{y} - y_d\| + \gamma_p\|\bar{p} - p_d\|). \qquad \square$$

We are now in a position to prove the main result for mixed/pressure control.

THEOREM 14. *Let $u \in U \cap X_h$ be so that $p = P(u) \in W_0^{1,0}(\Omega)$. If $p_d \in W_0^{1,0}(\Omega) \cap Q$, then there exists a constant $C = C(\Omega, u, y_d, p_d, \gamma_y, \gamma_p)$ such that*

$$\|(H_\beta^h(u) - T_\beta^h(u))v\| \leq Ch\|v\| \quad \forall v \in X_h.$$

Proof. For any $u \in U \cap X_h$, we have

$$T^h_\beta(u) - H^h_\beta(u) = \gamma_y(A_{2h}(\pi_{2h}u)\pi_{2h} - A_h(u)) + \gamma_p(B_{2h}(\pi_{2h}u)\pi_{2h} - B_h(u)) \\ + \tilde{\mathcal{C}}_{2h}(\pi_{2h}u)\pi_{2h} - \tilde{\mathcal{C}}_h(u).$$

We use the same approach as in the case of velocity control only. We have already shown in (96) that the first term is $O(h^2\|v\|)$. The second term is estimated similarly:

$$B_{2h}(\pi_{2h}u)\pi_{2h} - B_h(u) = [B_{2h}(\pi_{2h}u) - B(\pi_{2h}u)]\pi_{2h} + B(\pi_{2h}u)(\pi_{2h} - I) \\ + B(\pi_{2h}u) - B(u) + B(u) - B_h(u). \tag{102}$$

For any $v \in X_h$ we have

$$|((B(u) - B_h(u))v, v)| = |(M^*Mv - M_h^*M_hv, v)| = |\|Mv\|^2 - \|M_hv\|^2| \\ \leq \|Mv - M_hv\|(\|Mv\| + \|M_hv\|) \leq Ch\|v\|^2,$$

where we used (53) and (58). Similarly, it can be shown that

$$\|(B_{2h}(\pi_{2h}u) - B(\pi_{2h}u))v\| \leq Ch\|v\|^2 \quad \forall v \in X_h.$$

The second term in (102) can be bounded as

$$\|B(\pi_{2h}u)(\pi_{2h} - I)v\| = \|M^*(\pi_{2h}u)M(\pi_{2h}u)(\pi_{2h} - I)v\| \overset{(61)}{\leq} Ch\|v\|. \tag{103}$$

Finally, we have

$$|(B(\pi_{2h}u)v - B(u)v, v)| = |(M^*(\pi_{2h}u)M(\pi_{2h}u)v - M^*(u)M(u)v, v)| \\ = |\|M(\pi_{2h}u)v\|^2 - \|M(u)v\|^2| \\ \leq \|(M(\pi_{2h}u)v - M(u)v\|(\|M(\pi_{2h}u)v\| + \|M(u)v\|) \overset{(84)}{\leq} Ch^2\|u\|_1\|v\|,$$

which gives

$$\|(B(\pi_{2h}u) - B(u))v\| \leq Ch^2\|u\|_1\|v\|. \tag{104}$$

Next, we estimate

$$\tilde{\mathcal{C}}_{2h}(\pi_{2h}u)\pi_{2h} - \tilde{\mathcal{C}}_h(u) = [\tilde{\mathcal{C}}_{2h}(\pi_{2h}u) - \tilde{\mathcal{C}}(\pi_{2h}u)]\pi_{2h} + \tilde{\mathcal{C}}(\pi_{2h}u)(\pi_{2h} - I) \\ + \tilde{\mathcal{C}}(\pi_{2h}u) - \tilde{\mathcal{C}}(u) + \tilde{\mathcal{C}}(u) - \tilde{\mathcal{C}}_h(u). \tag{105}$$

We first estimate the term $\|\tilde{\mathcal{C}}(u)v - \tilde{\mathcal{C}}_h(u)v\|$ and recall that

$$(\tilde{\mathcal{C}}(u)v, v) = -2c(Lv; Lv, \tilde{z}) \text{ and } (\tilde{\mathcal{C}}_h(u)v, v) = -2\tilde{c}(L_hv; L_hv, \tilde{z}_h)$$

with $\tilde{z} = \tilde{L}(\gamma_y(y - y_d), \gamma_p(p_d - p))$, $\tilde{z}_h = \tilde{L}_h(\gamma_y(y_h - y_d^h), \gamma_p(p_d^h - p_h))$, with the operators $\tilde{L}$ and $\tilde{L}_h$ as defined in Theorem 8. Thus,

$$|(\tilde{\mathcal{C}}(u)v - \tilde{\mathcal{C}}_h(u)v, v)| = |2\tilde{c}(Lv; Lv, \tilde{z}) - 2\tilde{c}(L_hv; L_hv, \tilde{z}_h)| \\ \leq |c(Lv; Lv, \tilde{z}) - c(L_hv; L_hv, \tilde{z}_h)| \\ + |c(Lv; \tilde{z}, Lv) - c(L_hv; \tilde{z}_h, L_hv)|.$$

The first term in the inequality above can be bounded by

$$\begin{aligned}|c(L_hv; L_hv, \tilde{z}_h) - c(Lv; Lv, \tilde{z})| \le |c((L_h - L)v; L_hv, \tilde{z}_h)| + |c(Lv; L_hv, \tilde{z}_h - \tilde{z})|\\ + |c(Lv; \tilde{z}, (L_h - L)v|,\end{aligned}$$

where we used $c(Lv; L_hv - Lv, \tilde{z}) = -c(Lv; \tilde{z}, L_hv - Lv)$ since $Lv \in V$. Thus, we have

$$\begin{aligned}&|c(L_hv; L_hv, \tilde{z}_h) - c(Lv; Lv, \tilde{z})|\\ &\le C(\|Lv - L_hv\|_1\|L_hv\|_1\|\tilde{z}_h\|_1 + \|Lv\|_1\|L_hv\|_1\|\tilde{z}_h - \tilde{z}\|_1 + \|Lv\|_1\|\tilde{z}\|_1\|L_hv - Lv\|_1)\\ &\overset{(51),(64)}{\le} Ch\|v\|^2(\gamma_y\|y - y_d\| + \gamma_p\|p - p_d\|_{W_0^{1,0}(\Omega)}).\end{aligned}$$

Note that we have used (64) also for $\|L_hv\|_1 \le C\|v\|$, since $L_hv = \tilde{L}_h(v, 0)$. Also,

$$\begin{aligned}|c(L_hv; \tilde{z}_h, L_hv) - c(Lv; \tilde{z}, Lv)| &\le |c((L_h - L)v; \tilde{z}_h, L_hv)| + |c(Lv; \tilde{z}_h, (L_h - L)v)|\\ &\quad + |c(Lv; Lv, \tilde{z} - \tilde{z}_h)|\\ &\le \|L_hv - Lv\|_1\|\tilde{z}_h\|_1\|L_hv\|_1 + \|Lv\|_1\|\tilde{z}_h\|_1\|L_hv - Lv\|_1\\ &\quad + \|Lv\|_1\|Lv\|_1\|\tilde{z} - \tilde{z}_h\|_1 \le Ch\|v\|(\gamma_y\|y - y_d\| + \gamma_p\|p - p_d\|_{W_0^{1,0}(\Omega)}).\end{aligned}$$

Similarly, it can be shown that $\|(\tilde{\mathcal{C}}_{2h}(\pi_{2h}u) - \tilde{\mathcal{C}}(\pi_{2h}u)\pi_{2h}v\| \le Ch\|v\|$. To estimate the third term in (105), let $\hat{z} = \tilde{L}(\gamma_y(Y(\pi_{2h}u) - y_d), \gamma_p(p_d - P(\pi_{2h}u)))$. Then

$$\begin{aligned}&|(\tilde{\mathcal{C}}(\pi_{2h}u)v - \tilde{\mathcal{C}}(u)v, v)| = 2|c(Lv; Lv, \tilde{z}) - c(Lv; Lv, \hat{z})| = 2|c(Lv; Lv, \tilde{z} - \hat{z})|\\ &\le C\|Lv\|_1^2\|\tilde{z} - \hat{z}\|_1 \overset{(64)}{\le} C\|v\|^2\|\tilde{z} - \hat{z}\|_1 \overset{(99)}{\le} C(u, y_d, p_d, \gamma_y, \gamma_p)h\|v\|^2.\end{aligned}$$

Finally, let $w = (\pi_{2h} - I)v$. With $L = L(\pi_{2h}u)$ we have (see (39))

$$\begin{aligned}&|(\tilde{\mathcal{C}}(\pi_{2h}u)(\pi_{2h} - I)v, v)| = |((Lw \cdot \nabla)\hat{z} - (\nabla Lw)^T\hat{z}, Lv)|\\ &\le |((Lw \cdot \nabla)\hat{z}), Lv)| + |(\nabla Lw)^T\hat{z}, Lv)| = |c(Lw; \hat{z}, Lv)| + |c(Lv; Lw, \hat{z})|\\ &= |c(Lw; \hat{z}, Lv)| + |c(Lv; \hat{z}, Lw)| \le C\|Lw\|\|\hat{z}\|_2\|Lv\|\\ &\overset{(60)}{\le} Ch^2\|v\|\|\hat{z}\|_2\|v\| \overset{(62)}{\le} C(u, y_d, p_d, \gamma_y, \gamma_p)h^2\|v\|^2,\end{aligned}$$

which combined with the other estimates yields the conclusion. ∎

The following result follows from Theorem 14 using arguments similar to the ones in the proof of Corollary 1. Essentially it shows a decline by one unit in the approximation order of the two-grid preconditioner for the case of mixed/pressure control compared to velocity control. This is consistent with the case of the Stokes-constrained problem studied in [7].

COROLLARY 2. *Under the conditions of Theorem* 14*, if* $\tilde{\mathcal{C}}_h(u)$ *is symmetric positive definite, then*

$$d(H_\beta^h(u), T_\beta^h(u)) \le \frac{C}{\beta}h \tag{106}$$

for $h < h_0(\beta, \Omega, L, M, \tilde{L})$.

Remark 3. We should note that the hypotheses of Theorem 14 are quite restrictive, in that we cannot expect p_d and $P(u) \in W_0^{1,0}(\Omega)$ in practice; smooth functions in

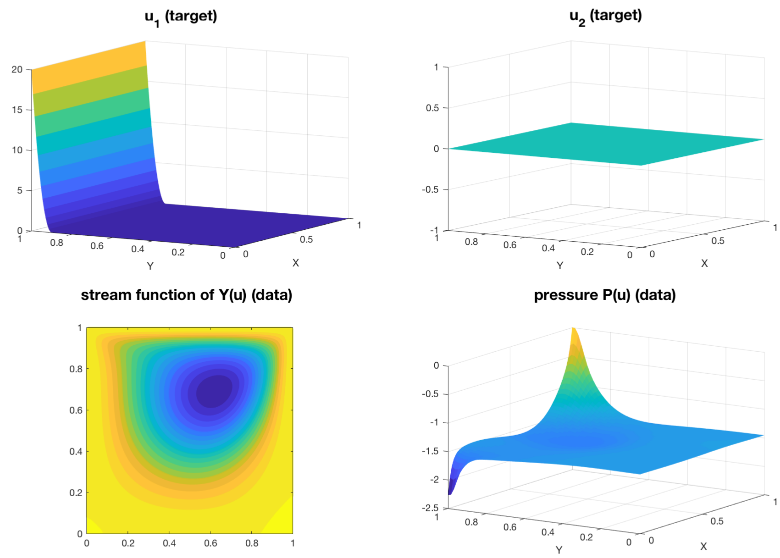

FIG. 1. *Top images: components of target control. Bottom images: velocity (stream function) and pressure data. Viscosity is* $\nu = 0.01$ *with* $\mathrm{Re} = \nu^{-1}\|Y(u)\|_\infty \approx 105$*, and* $h = 1/64$.

$W_0^{1,0}(\Omega)$ tend to zero near the corners of the domain and there is no a priori reason for a flow to have a zero-pressure near the corners of the domain. Nevertheless, this fact does not prevent the preconditioner from working quite well for the mixed/pressure control, as shown in section 5.

Remark 4. The two-grid preconditioner can be extended to a multigrid preconditioner following essentially the same strategy as in [7], and the analysis is extended in a similar fashion to show that the multigrid preconditioner satisfies the estimates (80) and (106). Suffice it to say that the correct multigrid preconditioner has a W-cycle structure, while the associated V-cycle gives suboptimal results; furthermore, the coarsest level has to be sufficiently fine in order for the optimal quality to be preserved.

5. Numerical results.

We present a set of numerical results to showcase the behavior of our multigrid preconditioner in the Newton iteration of (43) on $\Omega = (0,1)^2$. We consider uniform rectangular grids with mesh sizes $h = 1/32, 1/64, 1/128, 1/256$, and we use Taylor–Hood $\mathbf{Q}_2$–$\mathbf{Q}_1$ elements for velocity-pressure and $\mathbf{Q}_2$ elements for the controls. The data is given by $y_d^h = Y_h(u_h)$, $p_d^h = P_h(u_h)$ with u_h being the interpolant of the target control $u(x,y) = [10^3(\mathrm{sign}(y-0.9)+1)(y-0.9)^2, 0]$ (see Figure 1); the velocity field resembles one obtained from a lid-driven cavity flow. In Figure 2 we show the optimal control and the recovered velocity and pressure profile for the velocity control problem. As can be seen in the picture, if $\gamma_p = 0$, the pressure is not recovered. The Newton iteration is stopped when $\|\nabla \hat{J}_h\|_\infty \le 10^{-10}$. On the coarsest grid at $h = 1/32$ we use a zero-initial guess for the Newton solve, while for subsequent grids we start the iteration using the solution from the coarser problem. The linear systems at each iteration are solved in two ways: first we use CG preconditioned by the multigrid preconditioner (MGCG) (see Remark 4), with base cases $h_0 = 1/32$ or $1/64$, depending on necessity. Second, we solve the same systems using unpreconditioned CG. The reduced Hessian is applied matrix-free using (71)–

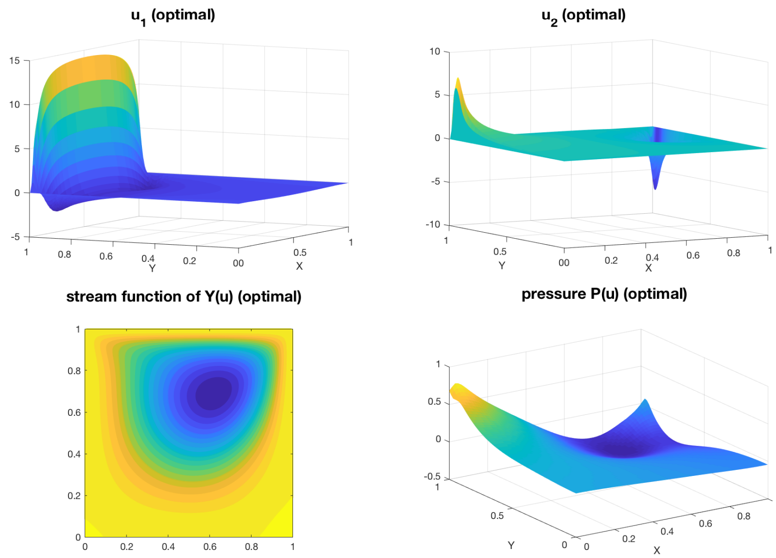

FIG. 2. *Top images: components of optimal control corresponding to data from Figure* 1. *Bottom images: optimal velocity (stream function) and pressure. Parameters values are* $\gamma_y = 1$, $\gamma_p = 10^{-3}$, *and* $\beta = 10^{-5}$.

(73). Obviously, the Hessian-vector multiplication (matvec) is the most expensive operation, as it essentially requires solving the linearized Navier–Stokes system twice. The goal is to show that, as a result of multigrid preconditioning, the number of matvecs at the highest resolution is relatively low compared to the unpreconditioned case.

We present in Table 1 results for low and in Table 2 for moderate Reynolds numbers, and we compare velocity control ($\gamma_p = 0$) with mixed velocity-pressure control with varying ratios of the two terms in $\hat{J}_h$ ($\gamma_y = 1$, $\gamma_p = 10^{-4}, 10^{-3}$). As for the regularization parameter we let $\beta = 10^{-4}, 10^{-5}$. For each of the twelve parameter choices and for each $h = 1/64, 1/128, 1/256$ we report the number of iterations of the MGCG/CG-based solves for each Newton iteration as well as the total (added) wall-clock time of the linear solves (since the overhead due to the gradient computation and Hessian setup is the same for both solves). For example, in the top left compartment of Table 2, we show the case $\nu = 0.01$, $\gamma_y = 1, \gamma_p = 0$ (velocity control only), $\beta = 10^{-4}$ with resolutions $1/64, 1/128, 1/256$. At resolution $h = 1/256$ two Newton iterations were required with CG necessitating 382 and 274 iterations, with a total time of linear solves of 11.4 hours, while the four-grid MGCG needed 6 and 4 iterations for a total of 0.58 hours, meaning almost twenty times faster. Note that at the coarsest level we actually build the Hessian at each Newton iteration and invert it using direct methods, the time of this operation being included in the reported wall-clock time. We should note that using direct methods at the coarsest level is critical for mixed/pressure control, but can be replaced with low-tolerance iterative solves for velocity control, a phenomenon that is still under scrutiny. For the experiments in Tables 1 and 2 the relative tolerance for both CG and MGCG is set at 10^{-8}. While this is quite low for the Newton-CG method, it forces a slightly higher number of linear iterations, thus allowing one to better observe the desired behavior of the MGCG preconditioner, namely, that for fixed β the number of linear iterations will decrease with $h \downarrow 0$. For

Table 1

Iteration counts and runtimes for MGCG vs. CG; $\nu = 0.1$, $\mathrm{Re} = \nu^{-1}\|Y(u)\|_\infty \approx 1.1$. *Tolerance is set at* 10^{-8}.

	$\beta = 10^{-4}$			$\beta = 10^{-5}$		
h_j^{-1}	64	128	256	64	128	256
$\gamma_p = 0$ (velocity control only)						
# cg	44	42	41	107	103	103
t_{cg} (sec.)	113	497	3662	285	1284	9015
# mg ($h_0 = 1/32$)	3	3	2	5	4	3
t_{mg} (sec.)	56	149	637	49	166	735
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$	2.02	3.34	5.75	5.82	7.73	12.27
$\gamma_p = 10^{-4}$ (some pressure control)						
# cg	46, 39	46, 40	46, 41	116, 102	116, 103	117, 106
t_{cg} (sec.)	227	1016	7610	579	2754	19483
# mg ($h_0 = 1/32$)	5, 5	5, 5	nc	8, 6	8, 6	nc
t_{mg} (sec.)	122	348	–	141	389	–
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$	1.86	2.92	–	4.11	7.08	–
# mg ($h_0 = 1/64$)	n/a	4, 4	4, 4	n/a	5, 5	5, 5
t_{mg} (sec.)		4216	4981		4073	5792
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$		0.24	1.53		0.68	3.36
$\gamma_p = 10^{-3}$ (more pressure control)						
# cg	43, 46	44, 48	46, 48	104, 124	106, 120	109, 123
t_{cg} (sec.)	239	1087	8329	605	2679	20676
# mg ($h_0 = 1/32$)	5, 7	5, 7	nc	8, 9	8, 9	nc
t_{mg} (sec.)	129	359	–	140	419	–
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$	1.85	3.03	–	4.32	6.39	–
# mg ($h_0 = 1/64$)	n/a	5, 7	5, 7	n/a	7, 8	6, 8
t_{mg} (sec.)		4360	5573		4427	5768
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$		0.25	1.49		0.61	3.58

two of the twelve cases we also vary the tolerance for the linear solves (10^{-2} and 10^{-4}) and show the results in Table 3. Here we also report the total wall-clock time for the computation (including the gradient computation and Hessian setup at each Newton iteration), since the number of Newton iterations begins to vary between CG to MGCG when using a large tolerance. This is due mainly to two factors related to MGCG: either it converges very fast and the gradients are better approximated than expected even if the tolerance is set at 10^{-2}, or the MG preconditioner is not positive definite and so MGCG fails to converge, in which case we still report 1 MGCG iteration. We allow a maximum of 10 Newton iterations.

Tables 1 and 2 indicate a behavior that is standard for the multigrid preconditioner presented in this work, and which is consistent with the analysis. First we notice that unpreconditioned CG is scalable, in the sense that for each case the number of CG iterations is bounded with respect to mesh-size. (The wall-clock times suffer due to the fact that we used direct solvers for the linearized Navier–Stoles solves in the matvec.) The MGCG instead shows an efficiency that increases over CG with decreasing h, measured both in terms of number of iterations and wall-clock time, and this can be seen for all the velocity control cases, and for the mixed control cases with base case $h_0 = 1/64$ at $\nu = 0.1$ (see Table 1). As usual with these types of algorithms, the lower order of approximation for the mixed/pressure control case leads not only

TABLE 2

Iteration counts and runtimes for MGCG vs. CG; $\nu = 0.01$, $\mathrm{Re} = \nu^{-1}\|Y(u)\|_\infty \approx 104$. *Tolerance is set at* 10^{-8}.

	$\beta = 10^{-4}$			$\beta = 10^{-5}$		
h_j^{-1}	64	128	256	64	128	256
$\gamma_p = 0$ (velocity control only)						
# cg	281, 289	259, 247	382, 274	819, 837, 459	777, 716	1136, 816
t_{cg} (hours)	0.42	1.73	11.40	1.54	5.15	33.63
# mg ($h_0 = 1/32$)	10, 10	9, 9	6, 4	27, 31, 15	25, 27	16, 11
t_{mg} (hours)	0.05	0.15	0.58	0.11	0.29	0.93
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$	8.81	11.30	19.82	14.70	18.07	36.11
$\gamma_p = 10^{-4}$ (some pressure control)						
# cg	299, 318	309, 306	500, 493	834, 909, 507	891, 965	1644, 1354
t_{cg} (hours)	0.43	2.12	16.92	1.45	6.36	50.57
# mg ($h_0 = 1/32$)	12, 13	13, 13	nc	36, 38, 23	35, 38	nc
t_{mg} (hours)	0.05	0.19	–	0.11	0.38	–
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$	7.94	10.99	–	12.60	16.83	–
# mg ($h_0 = 1/64$)	n/a	7, 7	9, 8	n/a	12, 12	19, 14
t_{mg} (hours)		1.27	1.62		1.23	1.98
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$		1.67	10.45		5.16	25.48
$\gamma_p = 10^{-3}$ (more pressure control)						
# cg	294, 332, 213	296, 336, 161	481, 597	922, 995, 728	802, 1028, 561	1321, 1818, 629
t_{cg} (hours)	0.62	2.70	18.62	1.62	8.18	64.39
# mg ($h_0 = 1/32$)	14, 16 12	15, 18, 11	nc	42, 51, 39	41, 57, 32	nc
t_{mg} (hours)	0.08	0.31	–	0.13	0.59	–
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$	7.99	8.97	–	12.17	13.78	–
# mg ($h_0 = 1/64$)	n/a	8, 11, 6	9, 17, 9	n/a	14, 18, 13	16, 23, 16
t_{mg} (hours)		1.78	2.84		1.97	3.19
eff$=t_{\mathrm{cg}}/t_{\mathrm{mg}}$		1.52	6.56		4.16	20.22

to a slightly higher number of MGCG iterations, but also requires a finer base case; for all the mixed velocity-pressure control problems, the four-grid preconditioner at resolution $h = 1/256$ (base case $h_0 = 1/32$) led to an iteration that does not converge within 10 Newton iterations since the MG-preconditioner is not positive definite. However, the base case choice $h_0 = 1/64$ appears to be sufficient when $\nu = 0.1$. For the higher Reynolds number case, while we did not encounter divergence with base case $h_0 = 1/64$, it is conceivable that it may still be too coarse, that is, it will lead to divergence at higher resolutions. We should point out that we purposefully selected a set of parameters that exhibit a variety of behaviors expected from these types of algorithms. Yet we find it remarkable that whenever MGCG converges, it does so significantly faster than unpreconditioned CG, with significant wall-clock savings.

Table 3
Iteration counts and runtimes for MGCG vs. CG; $\nu = 0.1$, $\mathrm{Re} = \nu^{-1}\|Y(u)\|_\infty \approx 1.1$.

	$\beta = 10^{-4}$			$\beta = 10^{-5}$		
h_j^{-1}	64	128	256	64	128	256
$\gamma_p = 10^{-3}$, $tol = 10^{-2}$						
# cg	12, 17, 18, 17	10, 17 19, 17	10, 17 19, 17	16, 39 44, 46	16, 40, 45, 47	15, 37, 42, 44
t_{cg} (sec.)	141	765	5878	309	1817	12614
t_{total} (sec.)	370	1779	10323	536	2822	17175
# mg ($h_0 = 1/32$)	3, 4, 4	3, 4, 4	2, 1, 4, 1, 4		1, 4, 5, 5	nc
t_{mg} (sec.)	178	469	3277		624	–
t_{total} (sec.)	357	1250	8702		1625	–
# mg ($h_0 = 1/64$)		3, 4	3, 2, 4	1, 5, 3, 5	1, 4, 2	1, 4, 3, 4
t_{mg} (sec.)		4150	7771	197	6295	10787
t_{total} (sec.)		4722	11316	426	7070	15307
$\gamma_p = 10^{-3}$, $tol = 10^{-4}$						
# cg	22, 29	21, 28	21, 29	44, 79	45, 76	46, 78
t_{cg} (sec.)	110	586	4788	260	1440	11429
t_{total} (sec.)	239	1154	7375	392	2010	14025
# mg ($h_0 = 1/32$)	4, 5	4, 5	4, 1, 5, 5	6, 7	4, 7	nc
t_{mg} (sec.)	117	337	3114	124	344	–
t_{total} (sec.)	247	907	7605	257	907	–
# mg ($h_0 = 1/64$)	n/a	3, 5	3, 5	n/a	4, 6	4, 6
t_{mg} (sec.)		4160	5308		4497	5489
t_{total} (sec.)		4736	7894		5066	8064

The results in Table 3 are also instructive. First they show that for the two cases ($\gamma_p = 10^{-3}$ and $\beta = 10^{-4}, 10^{-5}$), a tolerance of 10^{-4} leads to the same number of Newton iterations as with 10^{-8}, as in the last group in Table 1. Second, it shows that a large tolerance may lead to an increase in the number of Newton iterations coupled with fewer CG/MGCG iterations per Newton iteration, as expected. However, in almost every case, this results in a higher wall-clock time for the entire computation due to the costly overhead at the beginning of each Newton iteration. To conclude, it is preferable to set the tolerance sufficiently low in order to minimize the number of Newton iterations; then one should use the coarsest base case for the MG preconditioner that preserves the good approximation properties of the two-grid preconditioner, as described in detail in [7]. Whether MGCG or unpreconditioned CG gives a faster wall-clock time certainly depends on the particular problem parameters, but what remains consistent is the increasing efficiency of MGCG over CG as $h \downarrow 0$.

6. Conclusions and extensions. We have developed and analyzed a two-grid preconditioner to be used in the Newton iteration for the optimal control of the stationary Navier–Stokes equations. Under the natural assumption that the iteration starts sufficiently close to the solution it is shown that the preconditioner has a behavior that is similar to the optimal control of the stationary Stokes equations [7]. While the extension to multigrid is not explicitly discussed due to the similarity with the Stokes-control case, numerical results confirm that the behavior is consistent with the analysis and can lead to significant savings over unpreconditioned CG-based solves.

The method described in this work extends naturally to boundary control and to space-time distributed control for time-dependent problems, although the optimal discretization of the controls and the analysis requires additional fine tuning. As shown in the case of optimal control of linear elliptic equations, the analysis of the analogous MGCG preconditioner for boundary control problems is challenging and is expected to behave differently than for distributed optimal control [23], with significant differences being observed between Dirichlet- and Neumann control. With respect to the space-time distributed optimal control for the time-dependent Navier–Stokes system, we expect to inherit the challenges encountered in the space-time distributed optimal control of parabolic equations, where we observed the optimal $O(h^2/\beta)$ two-grid approximation just for the case when coarsening is performed only in space [23]. Space-time coarsening, although more aggressive, leads to optimal two-grid approximation for the parabolic case only when using a piecewise constant discontinuous-in-time and continuous-in-space discretization as described in [28], which effectively lowers the approximation order of the preconditioner to $O((\Delta t + h^2)/\beta)$, where Δt is the time step. Extending those results to the optimal control of the Stokes and Navier–Stokes systems forms the subject of our current research. Last, but not least, adding control constraints to a semismooth Newton approach as in the elliptic control case presents a set of challenges. The technique developed in [10] for distributed optimal control of linear elliptic equations with control constraints produces a multigrid preconditioner which approximates the Hessian to $O(h/\beta)$, assuming a piecewise constant discretization of the controls. Thus, if applied to Navier–Stokes control, the method in [10] is expected to yield an optimal-order preconditioner for the mixed/pressure control, but not for velocity control. Improving that quality to the optimal order $O(h^2/\beta)$ also forms the subject of current research.

Acknowledgments. The authors thank the editor and the anonymous referees for the careful reading of the manuscript and useful suggestions.

REFERENCES

[1] D. Braess, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, Cambridge, 1997.

[2] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics Appl. Math. 40, SIAM, Philadelphia, 2002.

[3] J. C. De los Reyes, *A primal-dual active set method for bilaterally control constrained optimal control of the Navier-Stokes equations*, Numer. Funct. Anal. Optim., 25 (2004), pp. 657–683.

[4] J. C. De los Reyes and R. Griesse, *State-constrained optimal control of the three-dimensional stationary Navier-Stokes equations*, J. Math. Anal. Appl., 343 (2008), pp. 257–272, https://doi.org/10.1016/j.jmaa.2008.01.029.

[5] J. C. de los Reyes and F. Tröltzsch, *Optimal control of the stationary Navier-Stokes equations with mixed control-state constraints*, SIAM J. Control Optim., 46 (2007), pp. 604–629, https://doi.org/10.1137/050646949.

[6] A. Drăgănescu and T. F. Dupont, *Optimal order multilevel preconditioners for regularized ill-posed problems*, Math. Comp., 77 (2008), pp. 2001–2038.

[7] A. Drăgănescu and A. M. Soane, *Multigrid solution of a distributed optimal control problem constrained by the Stokes equations*, Appl. Math. Comput., 219 (2013), pp. 5622–5634.
[8] A. Drăgănescu, *Multigrid preconditioning of linear systems for semi-smooth Newton methods applied to optimization problems constrained by smoothing operators*, Optim. Methods Softw., 29 (2014), pp. 786–818, https://doi.org/10.1080/10556788.2013.854356.
[9] A. Drăgănescu and C. Petra, *Multigrid preconditioning of linear systems for interior point methods applied to a class of box-constrained optimal control problems*, SIAM J. Numer. Anal., 50 (2012), pp. 328–353, https://doi.org/10.1137/100786502.
[10] A. Drăgănescu and J. Saraswat, *Optimal-order preconditioners for linear systems arising in the semismooth Newton solution of a class of control-constrained problems*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 1038–1070, https://doi.org/10.1137/140997002.
[11] A. Fursikov, M. Gunzburger, L. S. Hou, and S. Manservisi, *Optimal control problems for the Navier-Stokes equations*, in Lectures on Applied Mathematics, Springer, Berlin, 2000, pp. 143–155.
[12] A. V. Fursikov, M. D. Gunzburger, and L. S. Hou, *Optimal boundary control for the evolutionary Navier-Stokes system: The three-dimensional case*, SIAM J. Control Optim., 43 (2005), pp. 2191–2232, https://doi.org/10.1137/S0363012904400805.
[13] V. Girault, R. H. Nochetto, and L. R. Scott, *Max-norm estimates for Stokes and Navier-Stokes approximations in convex polyhedra*, Numer. Math., 131 (2015), pp. 771–822, https://doi.org/10.1007/s00211-015-0707-8.
[14] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer Ser. Comput. Math. 5, Springer, Berlin, 1986.
[15] M. Gunzburger, *Adjoint equation-based methods for control problems in incompressible, viscous flows*, Flow Turbul. Combust., 65 (2000), pp. 249–272, https://doi.org/10.1023/A:1011455900396.
[16] M. Gunzburger and S. Manservisi, *Flow matching by shape design for the Navier-Stokes system*, in Optimal Control of Complex Structures, Internat. Ser. Numer. Math. 139, Birkhäuser, Basel, Switzerland, 2002, pp. 279–289.
[17] M. D. Gunzburger, *Finite Element Methods for Viscous Incompressible Flows: A Guide to Theory, Practice, and Algorithms*, Comput. Sci. Sci. Comput., Academic Press, Boston, MA, 1989.
[18] M. D. Gunzburger, *Perspectives in Flow Control and Optimization*, Adv. Des. Control 5, SIAM, Philadelphia, 2003.
[19] M. D. Gunzburger, L. Hou, and T. P. Svobodny, *Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with distributed and Neumann controls*, Math. Comp., 57 (1991), pp. 123–151.
[20] M. D. Gunzburger, H. Kim, and S. Manservisi, *On a shape control problem for the stationary Navier-Stokes equations*, ESAIM Math. Model. Numer. Anal., 34 (2000), pp. 1233–1258, https://doi.org/10.1051/m2an:2000125.
[21] M. D. Gunzburger and S. Manservisi, *Analysis and approximation of the velocity tracking problem for Navier-Stokes flows with distributed control*, SIAM J. Numer. Anal., 37 (2000), pp. 1481–1512, https://doi.org/10.1137/S0036142997329414.
[22] M. D. Gunzburger and S. Manservisi, *The velocity tracking problem for Navier-Stokes flows with boundary control*, SIAM J. Control Optim., 39 (2000), pp. 594–634, https://doi.org/10.1137/S0363012999353771.
[23] M. Hajghassem, *Efficient Multigrid Methods for Optimal Control of Partial Differential Equations*, PhD thesis, University of Maryland, Baltimore, MD, 2017.
[24] M. Hanke and C. R. Vogel, *Two-level preconditioners for regularized inverse problems.* I. *Theory*, Numer. Math., 83 (1999), pp. 385–402.
[25] R. B. Kellogg and J. E. Osborn, *A regularity result for the Stokes problem in a convex polygon*, J. Funct. Anal., 21 (1976), pp. 397–431.
[26] M. Kollmann and W. Zulehner, *A robust preconditioner for distributed optimal control for stokes flow with control constraints*, Numer. Math. Adv. Appl. 2011 (2013), pp. 771–779.
[27] W. Layton, *Introduction to the Numerical Analysis of Incompressible Viscous Flows*, Comput. Sci. Eng. 6, SIAM, Philadelphia, 2008.
[28] D. Leykekhman and B. Vexler, *Pointwise best approximation results for Galerkin finite element solutions of parabolic problems*, SIAM J. Numer. Anal., 54 (2016), pp. 1365–1384, https://doi.org/10.1137/15M103412X.
[29] T. Rees and A. Wathen, *Preconditioning iterative methods for the optimal control of the Stokes equations*, SIAM J. Sci. Comput., 33 (2011), pp. 2903–2926.
[30] A. Rieder, *A wavelet multilevel method for ill-posed problems stabilized by Tikhonov regularization*, Numer. Math., 75 (1997), pp. 501–522.

[31] J. Saraswat, *Multigrid Solution of Distributed Optimal Control Problems Constrained by Semilinear Elliptic PDEs*, PhD thesis, University of Maryland, Baltimore, MD, 2014.

[32] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[33] F. Tröltzsch, *Optimal Control of Partial Differential Equations*, Grad. Stud. Math. 112, AMS, Providence, RI, 2010, https://doi.org/10.1090/gsm/112.

[34] F. Tröltzsch and D. Wachsmuth, *Second-order sufficient optimality conditions for the optimal control of Navier-Stokes equations*, ESAIM Control Optim. Calc. Var., 12 (2006), pp. 93–119, https://doi.org/10.1051/cocv:2005029.