Proceedings of ALGORITMY 2000
Conference on Scientific Computing, pp. 20–27

THE SQP-METHOD FOR TRACKING TYPE CONTROL OF THE INSTATIONARY NAVIER-STOKES EQUATIONS*

MICHAEL HINZE†

Abstract. The SQP-method is investigated for tracking-type optimal control of the instationary Navier-Stokes equations. It is argued that the a-priori formidable SQP-step can be decomposed into linear primal and linear adjoint systems, which is amenable for existing CFL-software. We report a numerical test which demonstrates the feasibility of the approach. In addition the functional analytic setting of the convergence analysis is presented.

Key words. Optimal control, SQP-method, Navier-Stokes equations

AMS subject classifications. 34H05, 49J20, 49K20, 65K10, 76D55

1. The optimal control problem. We consider the optimal control problem

$$(1)\qquad \begin{cases} \min_{(y,u)\in W\times U} \; J(y,u) := \frac{1}{2}\int\limits_{Q_o} |y-z|^2\,dxdt + \frac{\alpha}{2}\int\limits_{Q_c} |u|^2\,dxdt \\ \text{subject to} \\ \frac{\partial y}{\partial t} + (y\cdot\nabla)y - \nu\Delta y + \nabla p = Bu & \text{in } Q = (0,T)\times\Omega, \\ \text{div } y = 0 & \text{in } Q, \\ y(t,\cdot) = 0 & \text{on } \Sigma = (0,T)\times\partial\Omega, \\ y(0,\cdot) = y_0 & \text{in } \Omega. \end{cases}$$

where $Q_c := \Omega_c \times (0,T)$ and $Q_o := \Omega_o \times (0,T)$, with Ω_c and Ω_o subsets of $\Omega = (0,1)^2$ denoting control and observation volumes, respectively. The first term in the cost functional values the control gain which here is to track the state z, and the second term measures the control cost, where $\alpha > 0$ denotes a weighting factor. In this form solving (1) appears at first to be a standard task. However, the formidable size of (1) and the goal of analyzing second order methods necessitate an independent analysis. One of the few contributions focusing on second order methods for optimal control of fluids are given by Ghattas *et al* [3] and Heinkenschloss [4]. These works are restricted to stationary problems, however. Among other things analytical investigations on second order methods are given by the author in [5] and by Kunisch and the author in [6], where also further references can be found.

To define the spaces and operators required for the investigation of (1) we introduce the solenoidal spaces

$$H = \{v \in C_0^\infty(\Omega)^2 \colon \text{ div } v = 0\}^{-|\cdot|_{L^2}}, V = \{v \in C_0^\infty(\Omega)^2 \colon \text{ div } v = 0\}^{-|\cdot|_{H^1}},$$

with the superscripts denoting closures in the respective norms. Further we define

$$W = \{v \in L^2(V) \colon v_t \in L^2(V^*)\} \quad \text{and} \quad Z := L^2(V)\times H,$$

where W is endowed with the norm

$$|v|_W = (|v|^2_{L^2(V)} + |v_t|^2_{L^2(V^*)})^{1/2},$$

*This work was supported by the Special Research Project 557 of the Deutsche Forschungsgemeinschaft

†Fachbereich Mathematik, Technische Universität Berlin, Straße des 17. Juni 13, D-10623 Berlin, E-mail: hinze@math.tu-berlin.de, WWW: www.math.tu-berlin.de/~hinze

and set $\langle \cdot,\cdot\rangle := \langle \cdot,\cdot\rangle_{L^2(V^*),L^2(V)}$, with V^* denoting the dual space of V. Here $L^2(V)$ is an abbreviation for $L^2(0,T;V)$ and similarly $L^2(V^*) = L^2(0,T;V^*)$. Recall that up to a set of measure zero in $(0,T)$ elements $v \in W$ can be identified with elements in $C([0,T];H)$. In (1) further $U = L^2(Q_c)$ denotes the Hilbert space of controls which is identified with its dual U^*. The state z is assumed to an element of $L^2(H)$. It is not hard to show that the cost functional $J: L^2(H) \times U \to R$ is bounded from below, weakly lower semi-continuous, twice Fréchet differentiable with locally Lipschitzean second derivative, and radially unbounded in u, i.e. $J(y,u) \to \infty$ as $|u|_U \to \infty$, for every $y \in W$. We define the nonlinear mapping

$$e: W \times U \to Z^*$$

by

$$e(y,u) = \left(\tfrac{\partial y}{\partial t} + (y\cdot\nabla)y - \nu\Delta y - Bu, y(0) - y_0\right),$$

where $Bu(t,\cdot)$ denotes the Leray-projection of the extension by zero of $u(t,\cdot)$ to the whole of Ω and $y_0 \in H$. In variational form the constraints in (1) can be equivalently expressed as: given $u \in U$ find $y \in W$ such that $y(0) = y_0$ in H and

$$\langle y_t, v\rangle + \langle (y\cdot\nabla)y, v\rangle + \nu(\nabla y, \nabla v)_{L^2(L^2)} = \langle Bu, v\rangle \;\; \forall\, v \in L^2(V). \tag{2}$$

It is well known, see Temam [8] that for every $u \in U$ (2) admits a unique solution $y(u) \in W$. Therefore, with respect to existence (1) can equivalently be rewritten as

$$\min\; \hat{J}(u) = J(y(u),u) \;\text{ subject to }\; u \in U, \tag{3}$$

where $y(u) \in W$ satisfies $e(y(u),u) = 0$. It is proved by Abergel *et al* [1] that (3) admits a solution $(y^*,u^*) \in W \times U =: X$.The Lagrangian $L: X \times Z \to R$ is given by

$$L(x,\lambda) = J(x) + \langle e(x), \lambda\rangle_{Z^*,Z}, \tag{4}$$

and we anticipate that the SQP-method can be interpreted as Newton's algorithm applied to the equation

$$L'(x,\lambda) \;=\; 0$$

with $L'(x,\lambda)$ denoting the gradient of the Lagrangian (4).

We shall frequently refer to the variational solution of the linearized Navier-Stokes system and the adjoint equations in the solenoidal setting: Given $f \in L^2(V^*)$ and $v_0 \in H$, find the variational solution $v \in W$ of

$$\left\{\begin{array}{l} v_t + (v\cdot\nabla)y + (y\cdot\nabla)v - \nu\Delta v = f \quad \text{in } L^2(V^*) \\ v(0) = v_0 \quad \text{in } H, \end{array}\right. \tag{5}$$

and, given $g \in W^* \cap L^\alpha(V^*) (\alpha \in [1,4/3])$, find the variational solution $w \in L^2(V)$ of

$$\left\{\begin{array}{l} -w_t + (\nabla y)^t w - (y\cdot\nabla)w - \nu\Delta v = g \quad \text{in } W^* \\ w(T) = 0 \quad \text{in } H. \end{array}\right. \tag{6}$$

The following proposition is proved in [5], compare also [6] for a similar analytic framework. It is essential for the analysis of SQP, Newton and quasi-Newton methods.

PROPOSITION 1.1. Let $x = (y,u) \in W \times U$. Then $e_y(x): W \to Z^*$ is a homeomorphism. Moreover, if the inverse of its adjoint $e_y^{-*}(x): W^* \to Z$ is applied to an element

$g \in W^* \cap L^\alpha(V^*)$, where $\alpha \in [1, 4/3]$ then setting $(w, w_0) := e_y^{-*}(x)g \in L^2(V) \times H$ we have $w_t \in L^\alpha(V^*)$, $w(0) = w_0$ and w is the variational solution to (6).

To apply the SQP-method to (1) we need second order information of the Lagrangian L. The basic ingredients are the derivatives of the operator e which were characterized in [5], compare also [6]. For the convenience of the reader we state

PROPOSITION 1.2. The operator $e = (e^1, e^2): X \to Z^*$ is twice continuously differentiable with Lipschitz continuous second derivative. The action of the first two derivatives of e^1 are given by

$$\begin{aligned}\langle e_x^1(x)(w,s), \phi\rangle = \langle w_t, \phi\rangle + \langle (w\cdot\nabla)y, \phi\rangle + \langle (y\cdot\nabla)w, \phi\rangle \\ +\nu(\nabla w, \nabla\phi)_{L^2(L^2)} - \langle Bs, \phi\rangle_{L^2(L^2)},\end{aligned}$$

where $x = (y, u) \in X, (w, s) \in X$ and $\phi \in L^2(V)$, and

$$\langle e_{xx}^1(x)(w,s)(v,r), \phi\rangle = \langle e_{yy}^1(x)(w,v), \phi\rangle = \tag{7}$$

$$\langle (w\cdot\nabla)v, \phi\rangle + \langle (v\cdot\nabla)w\phi, v\rangle =: \langle v, H(\phi)w\rangle_{W,W^*}, \tag{8}$$

where $(v, r) \in X$.

2. The algorithm. This section contains a description of the SQP-method to solve (3). Throughout x^* denotes a (local) solution to (3). The basic SQP-algorithm consists in applying Newton's method to the first order optimality system

$$\begin{aligned} L_x(x,\lambda) = 0 \quad &\text{in } X^* \\ L_\lambda(x,\lambda) = 0 \quad &\text{in } Z^*, \end{aligned} \tag{9}$$

where the Lagrangian L is defined in (4). As a consequence of Proposition 1.1 $e_x(x^*)$ is surjective and there exists a unique Lagrange multiplier $\lambda^* \in Z$ such that (9) holds with $(x, \lambda) = (x^*, \lambda^*)$ [5, 6]. To formulate the algorithm we note that the second derivative of L with respect to x is given by

$$L_{xx}(x,\lambda) = \begin{pmatrix} J_{yy}(x) + \langle e_{yy}^1(x)(\cdot,\cdot), \lambda^1\rangle & 0 \\ 0 & J_{uu}(x)\end{pmatrix} \in \mathcal{L}(\mathcal{X}, \mathcal{X}^*).$$

Let $B((x^*, \lambda^*))$ denote a sufficiently small neighbourhood of (x^*, λ^*). In algorithmic form the method then can be formulated as follows.

ALGORITHM 2.1. (SQP–algorithm)

1. Choose $(x^0, \lambda^0) \in B(x^*, \lambda^*)$, set $k = 0$.
2. Do until convergence
   i) solve

$$\begin{bmatrix} L_{xx}(x^k,\lambda^k) & e_x^*(x^k) \\ e_x(x^k) & 0 \end{bmatrix}\begin{bmatrix}\delta x^k \\ \delta\lambda^k\end{bmatrix} = -\begin{bmatrix} J_x(x^k) & + e_x^*(x^k)\lambda^k \\ & e(x^k)\end{bmatrix} \tag{10}$$

   ii) update $(x^{k+1}, \lambda^{k+1}) = (x^k, \lambda^k) + (\delta x^k, \delta\lambda^k)$,
   iii) set $k = k + 1$.

Now let us introduce the matrix operators

$$T(x) = \begin{bmatrix} -e_y^{-1}(x)e_u(x) \\ Id_U \end{bmatrix}, \quad A(x) = \begin{bmatrix} e_y^{-1}(x) \\ 0 \end{bmatrix}$$

and the reduced Hessian

$$H(x,\lambda) = T^*(x)L_{xx}(x,\lambda)T(x).$$

Since by Proposition 1.1 $e_y(x)$ is a homeomorphism for every $x \in X$ the SQP-step (10) can be rewritten in Schur-complement form w.r.t. the control variable u.

ALGORITHM 2.2. (Modified SQP-step). Choose $(x^0, \lambda^0) \in B(x^*, \lambda^*)$, set $k = 0$. Do until convergence

i) solve $H(x^k, \lambda^k)\delta u^k = -T^*(x^k)J_x(x^k) + T^*(x^k)L_{xx}(x^k, \lambda^k)A(x^k)e(x^k)$
ii) set $\delta y^k = -e_y(x^k)^{-1}e(x^k) - e_y(x^k)^{-1}e_u(x^k)\delta u^k$,
iii) set $x^{k+1} = x^k + \delta x^k$,
iv) set $\lambda^{k+1} = A^*(x^k)\left\{J_x(x^k) + L_{xx}(x^k, \lambda^k)\left\{A(x^k)e(x^k) + T(x^k)\delta u^k\right\}\right\}$,
v) $k = k + 1$

Let us consider the linear system in i). Its dimension is that of the control space U. Since the computation of the reduced Hessian $H(x, \lambda)$ would involve the inversion of $e_y(x)$ together with its adjoint we conclude that this system for large dimensional practical applications itself has to be solved iteratively, e. g. by a conjugate gradient technique. This would correspond to an in-exact SQP-method. We shall then refer to i) as the "inner " loop as opposed to the do–loop in 2. of Algorithm 2.1 which is the "outer" loop of the SQP-algorithm. The inner loop at iteration level k of the outer loop requires to

(i) evaluate $-T^*(x^k)J_x(x^k) + T^*(x^k)L_{xx}(x^k, \lambda^k)A(x^k)e(x^k)$. Given (x^k, λ^k) this amounts to solving system (5) with $f = e^1(x^k)$, $v_0 = e^2(x^k)$, and system (6) twice with $g = J_y(x^k)$ and $g = L_{yy}(x^k, \lambda^k)e_y(x^k)^{-1}e(x^k)$.
(ii) iteratively evaluate the action of $H(x^k, \lambda^k)$ on δ_j^k, the j–th iterate of the inner loop on the k–th level of the outer loop.

The iterate $q = H(x^k, \lambda^k)\delta_j^k$ can be evaluated by successively applying the steps

a) solve (5) with $v_0 = 0$ and $f = B\delta_j^k$ in $L^2(V^*)$ for $v \in W$,
b) evaluate $J_{yy}(x^k)v + \langle e^1_{yy}(x^k)(v, \cdot), \lambda^{1^k}\rangle$,
c) solve (6) with $g = J_{yy}(x^k)v + \langle e^1_{yy}(x^k)(v, \cdot), \lambda^{1^k}\rangle$ in W^* for $w \in L^2(V)$,
d) and finally set $q := J_{uu}\delta u + B^*w$.

It can be shown that the functional appearing in b) is an element of $W^* \cap L^{4/3}(V^*)$, and hence (6) is meaningful in the sense of admitting a solution w in $L^2(V)$ with $w_t \in L^{4/3}(V^*)$. Summarizing, for the outer iteration of the SQP-method three linearized Navier–Stokes solves are required. For the inner loop one forward (–in time) as well as one backwards linearized Navier–Stokes solve per iteration is necessary.

REMARK 2.1. The performance of the conjugate gradient method depends on the spectral condition number $\kappa(H)$ of the reduced Hessian H. Since H can be decomposed as $H = J_{uu} + K$ with a compact, selfadjoint and positive operator K the spectral condition number of H satisfies $\kappa(H) \leq (\lambda_{\max}(J_{uu}) + \|K\|)/\lambda_{\min}(J_{uu})$. In our application $J_{uu} = \alpha Id$ holds, i.e. $\kappa(H) = \mathcal{O}(1/\alpha)$ $(\alpha \to 0)$ which is worse for small $\alpha > 0$.

To prove local quadratic convergence of the method (with exact updates) we need to impose a smallness assumption on the difference $|y^* - z|_{L^2(H)}$. Due to the structure of the cost functional J it guarantees second order sufficient optimality of the local solution on x^*. The proof of the following theorem can be found in [5], compare also [6].

THEOREM 2.1. Let $x^* = (y^*, u^*)$ denote a solution to problem (1). Let $|y^* - z|_{L^2(H)}$ be sufficiently small and let λ^* denote the Lagrange-multiplier associated to x^*. Then there exist a neighbourhood $B(x^*, \lambda^*) \subset X \times Z$ such that for all $(x^0, \lambda^0) \in B(x^*, \lambda^*)$ the SQP-algorithm 2.1 is well defined and its iterates $\{(x^n, \lambda^n)\}_{n \in R}$ converge quadratically to (x^*, λ^*).

3. Numerical Example. In the numerical example we did not restrict the state y to satisfy homogeneous boundary conditons, thus extending the computations beyond the theoretical treatment of problem (1). In what follows we present numerical computations for the control of driven cavity flow in $\Omega = (0,1)^2$. The uncontrolled flow is driven by a non-zero tangential velocity component, i.e. $y^1 = 1, y^2 = 0$ at the top wall, and satisfies homogeneous boundary conditions on the remaining part of the boundary, compare Fig. 1. We note that with this boundary condition the state not even is an element of $H^1(\Omega)^2$.

The discretization of the Navier-Stokes equations, its linearization and adjoint was carried out by using parts of the code developed by Bänsch [2], which is based on Taylor-Hood finite elements for the spatial discretization. As time step size we took $\delta t = .00625$, which resulted in 160 grid points for the time grid, and 545 pressure and 2113 velocity nodes for the spatial discretization. This results in a total number of unknowns (primal-, adjoint-, and control variables) of the optimization problem of order $2.2{*}10^6$. The time horizon could still be increased or the mesh size decreased by utilizing reduced storage techniques at the expense of additional cpu-time, but we shall not pursue this aspect here. All computations were performed on a ORIGIN^{TM} 500.

We present the results for $\Omega_c = \Omega_o = (0,1)^2$, $\alpha = 10^{-2}$ and $1/\nu =$ Re=400. As initial guess for the SQP-algorithm we utilized the tuple (x^0, λ^0) where $x^0 = (y^0, u^0)$ is obtained from applying one step of Newton's method with initial control equal to zero to the numerical solution of problem (3), i.e. with u^0 computed by Newton's method y^0 denotes the solution of the Navier-Stokes equations in (1) with $u = u^0$. Finally, the initial guess for the Lagrange multiplier is set to $\lambda^0 = -e_y(x^0)J_y(x^0)$. The termination criterion for the outer iteration is chosen as $\frac{|L'(x^k,\lambda^k)|}{|L'(x^0,\lambda^0)|} \le 10^{-3}$. For the iterative solution of i) in Algorithm 2.2 we utilize the conjugate gradient algorithm whose termination criterion for the j-th iterate δu_j^k is chosen as

$$\frac{|H(x^k,\lambda^k)\delta u_j^k - r(x^k,\lambda^k)|}{|L'(x^0,\lambda^0)|} \le \min\left\{ \left(\frac{|L'(x^k,\lambda^k)|}{|L'(x^0,\lambda^0)|}\right)^{\frac{3}{2}}, 10^{-2}\frac{|L'(x^k,\lambda^k)|}{|L'(x^0,\lambda^0)|} \right\},$$

where $r(x^k,\lambda^k)$ denotes the right-hand-side in i) of Algorithm 2.2. This termination criterion is motivated by stopping rules that are utilized for inexact Newton methods, say, in the finite dimensional setting in order to guarantee superlinear convergence, compare [7]. We therefore can only expect super-linear convergence of the SQP-method which in fact is confirmed in Table 1 for the last five iterates. The numerical computation of the optimal solution took about 90 minutes cpu-time on the workstation environment mentioned above. It is worth noting that Newton's method applied to (3) with initial control $u = u^0$ computes the numerical approximations to the optimal control and the optimal state as twice as fast [6].

In Figure 2 the evolution of the cost functional and the control cost as a function of time are documented. It can be observed that graphically there is no significant change after the third iteration of the method. These comments hold for quite a wide range of values for α. In Fig. 1 the target flow together with the uncontrolled flow, the controlled flow and the control action at the end of the time interval are presented.

Note that different observation and control volumes result in smaller control and observation volumes as in our example, and thus the primal and adjoint equations are numerically simpler to solve. We refer to [5, 6] for a more detailed discussion of this topic.

TABLE 1
Performance of the SQP- method for Example 1

Iteration	CG-steps	$\frac{\lvert L'(x^k,\lambda^k)\rvert}{\lvert L'(x^0,\lambda^0)\rvert}$	$\frac{\lvert\delta(x^k,\lambda^k)\rvert_{L^2(Q)}}{\lvert\delta(x^{k-1},\lambda^{k-1})\rvert_{L^2(Q)}}$	$J(x^k)$
0	-	-	-	1.188772e-2
1	11	1.e0	1.	3.216904e-3
2	3	2.342777e-1	0.456	1.661840e-3
3	16	5.846246e-1	1.110	2.041436e-3
4	5	1.574504e-1	0.959	1.548152e-3
5	14	2.718657e-2	0.554	1.683024e-3
6	23	6.744024e-2	0.914	1.485434e-3
7	18	8.005254e-2	0.874	1.521882e-3
8	18	1.852064e-2	0.197	1.480751e-3
9	23	1.343532e-3	0.146	1.479234e-3
10	26	1.698641e-4	0.127	1.479219e-3

REFERENCES

[1] F. ABERGEL AND R. TEMAM, *On some Control Problems in Fluid Mechanics*, Theoret. Comput. Fluid Dynamics **1** (1990), 303–325.

[2] E. BÄNSCH, *An adaptive Finite-Element-Strategy for the three-dimensional time-dependent Navier-Stokes-Equations*, J. Comp. Math. **36** (1991)

[3] O. GHATTAS AND J.J. BARK, *Optimal control of two- and three-dimensional incompressible Navier-Stokes flows*, Journal of Computational Physics **136** (1997), 231–244.

[4] M. HEINKENSCHLOSS, *Formulation and Analysis of a Sequential Quadratic Programming method for the Optimal Dirichlet Boundary Control of Navier-Stokes Flow*, in Optimal Control, Theory, Algorithms, and Applications. Kluwer Academic Publishers B.V. (1998)

[5] M. HINZE, *Optimal and instantaneous control of the instationary Navier-Stokes equations*, Habilitationsschrift, Technische Universität Berlin (1999)

[6] M. HINZE AND K. KUNISCH, *Second order methods for optimal control of time-dependent fluid flow*, Karl-Franzens Universität Graz, Institut für Mathematik, Spezialforschungsbereich F003, Bericht 165 (1999)

[7] C.T. KELLEY, *Iterative Methods for Linear and Nonlinear Equations*, SIAM (1995)

[8] R. TEMAM, *Navier-Stokes Equations*, North Holland, Amsterdam, 1979.

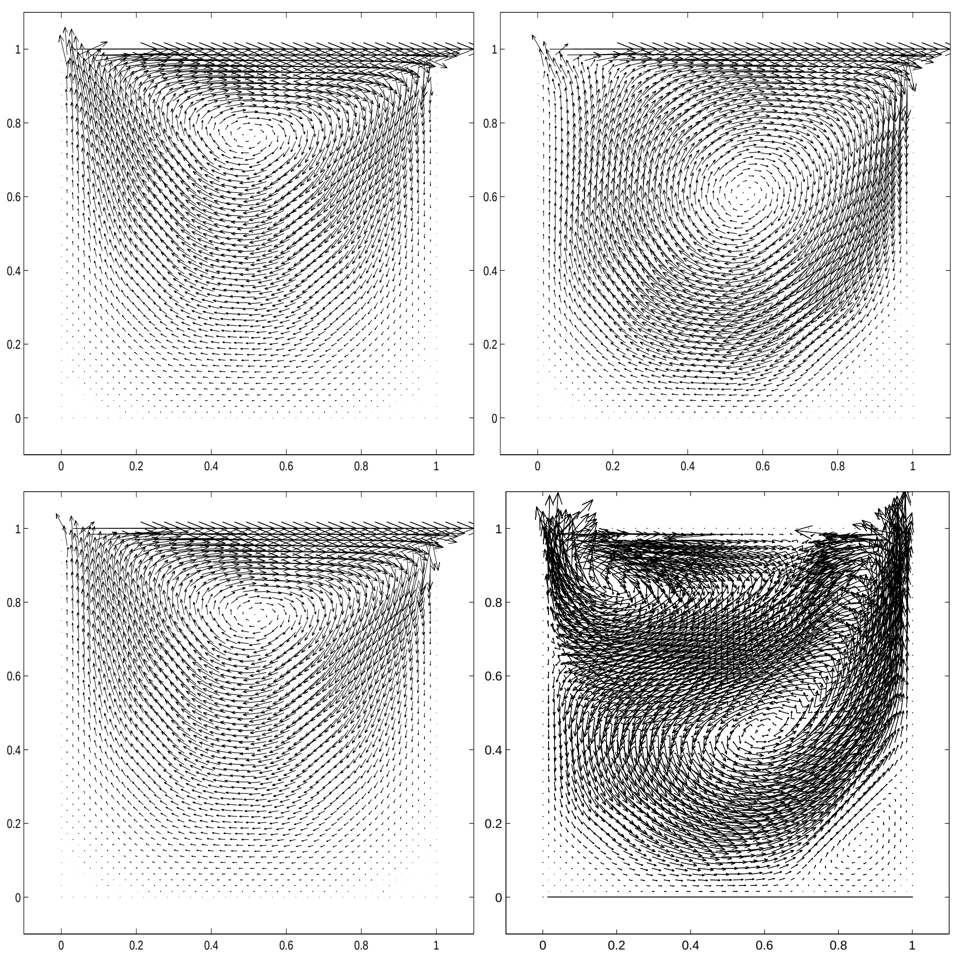
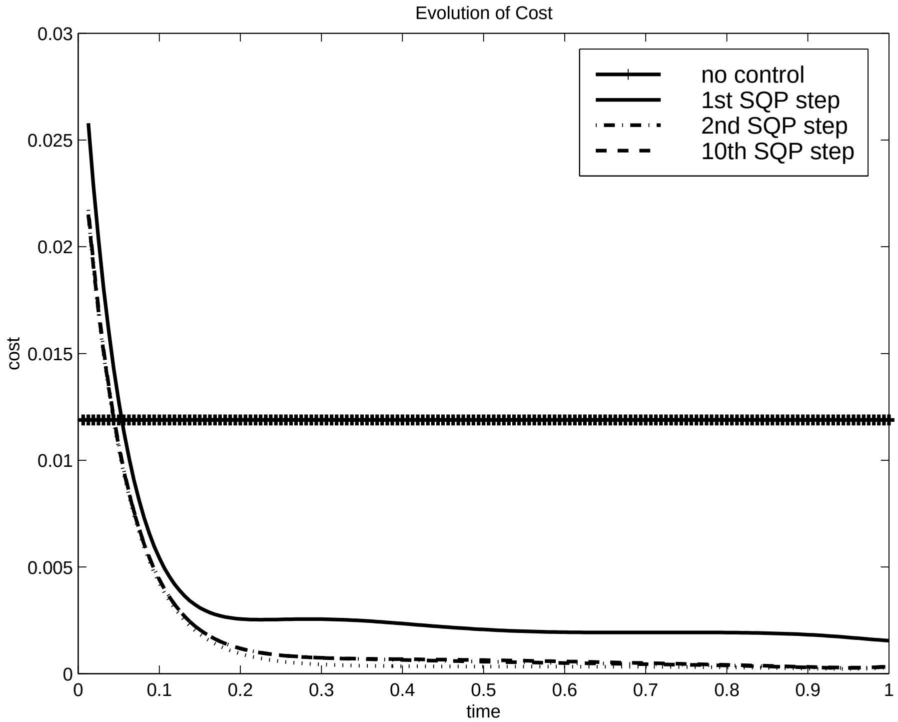


FIG. 1. *top left: target flow, top right: uncontrolled flow, bottom left: controlled flow, bottom right: control force*



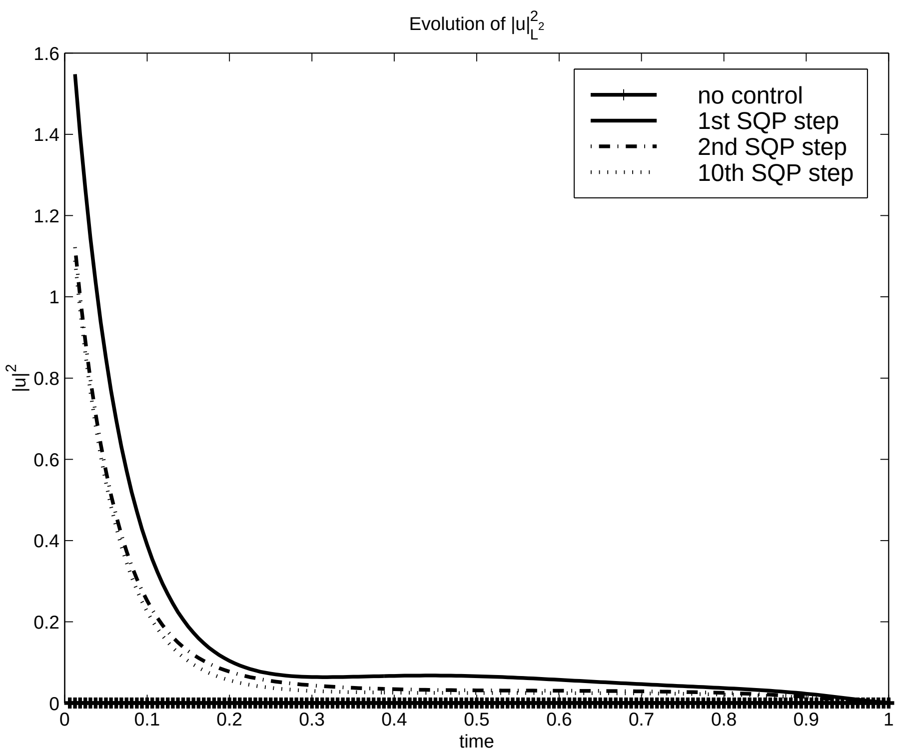


FIG. 2. *Re=400, $\alpha = 10^{-2}$: Evolution of cost functional (top) and control cost (bottom) for different iteration levels and relative accuracy = 1.d-3*