CALCOLO 41, 203 –225 (2004)
DOI: 10.1007/s10092-004-0092-7

CALCOLO

A comparison of algorithms for control constrained optimal control of the Burgers equation

J.C. de los Reyes, K. Kunisch[⋆]

Institut für Mathematik, Karl-Franzens-Universität Graz, Graz, Austria

Received: December 2003 / Revised version: September 2004 – © Springer-Verlag 2004

Abstract. A comparison of three different but related numerical methods for control constrained optimal control of the Burgers equation is carried out. We develop the principal ideas of the different strategies considered and present detailed numerical examples. Lastly, conclusions about the behavior of the approaches are obtained.

1 Introduction

Introduced in the 1940s as a simplified model for turbulence, the Burgers equation was used thereafter in diverse physical problems such as shock waves, supersonic flow and acoustic transmission. Its study usually serves as a first approximation to more complex convection diffusion phenomena, and in many cases this model deals efficiently with the basic features of the problem involved.

Development over the past decades has yielded important results on existence and uniqueness for this equation in the stationary as well as in the evolutionary cases. In the 1950s, Hopf [12] gave pioneering results by studying an explicit solution of the time-dependent equation and observed its behavior as time tends to infinity. Today the functional analytic setting of the problem is well understood; see, e.g., [18].

The optimal control of this equation has a more recent history. We cite one of the earlier papers [6] as well as [4,5,13]. A rather complete analysis of optimal control problems and their approximation is given in [20]. This

⋆ Research partially supported by the Fonds zur Förderung der wissenschaftlichen Forschung under SFB 03 “Optimierung und Kontrolle”.

work contains an augmented Lagrangian-SQP method for the solution of the control problem and its local convergence analysis.

In the present paper we consider the same class of optimal control problems as in [20], but in the case where the control is subject to pointwise constraints. This introduces an additional nonlinearity to the optimal control problem. Its numerical treatment by super-linearly convergent methods is the focus of the present work. Detailed numerical examples are given, which allow us to draw conclusions about the behavior of the methods.

The paper is organized as follows. In Sect. 2 the optimal control problem is stated and an initial optimality system is obtained. In Sect. 3, based on the generalized equation approach, an SQP (sequential quadratic programming) approach, with a primal-dual active set method for the inner loop, is stated and analyzed for the solution of our problem. In Sect. 4 a nonlinear primal-dual active set method is explained and the scheme is applied, together with an SQP algorithm for the auxiliary nonlinear problems, to the solution of the constrained optimal control problem. In Sect. 5 a semi-smooth Newton method is introduced and its relationship with each of the previous methods is shown. In Sect. 6 numerical experiments are discussed and basic features of the implementations are explained. Lastly, in Sect. 7 conclusions, based on experiments for the three methods, are obtained.

2 Optimal control problem

Considering the domain $\Omega = (0, 1)$, the optimal control problem for the stationary Burgers equation with a tracking-type cost functional can be formulated as:

$$\begin{cases} \min & J(y,u) = \frac{1}{2}\int_0^1 |y - z_d|^2 dx + \frac{\alpha}{2}\int_{\tilde{\Omega}} |u|^2 dx \\ \text{s.t.} & -\nu y'' + yy' = \mathcal{B}u \\ & y(0) = y(1) = 0 \\ & u \le b \quad \text{a.e.}, \end{cases} \tag{2.1}$$

where $b \in L^2(\tilde{\Omega})$, $\tilde{\Omega} \subset \Omega$, $z_d \in H^1(\Omega)$, $\alpha > 0$ and

$$\mathcal{B}u = \begin{cases} u & \text{in} \quad \tilde{\Omega} \\ 0 & \text{in} \quad \Omega \backslash \tilde{\Omega}. \end{cases}$$

The parameter $\nu > 0$ stands for the viscosity coefficient of the fluid, which is equal to $\frac{1}{\mathrm{Re}}$, where Re denotes the Reynolds number.

Next the constraints in (2.1) are expressed by means of an appropriately defined operator. The differential equation is realized in variational form.

By introducing

$$\begin{aligned} G : H_0^1(\Omega) \times L^2(\tilde{\Omega}) &\rightarrow H^{-1}(\Omega) \times L^2(\tilde{\Omega}) \\ (y, u) &\rightarrow \begin{pmatrix} \nu(y', \cdot') + (yy', \cdot) - (\mathcal{B}u, \cdot) \\ b - u \end{pmatrix}, \end{aligned}$$

the constraints in (2.1) can be expressed as $G(y, u) \in K$, where K is the closed convex cone defined by $K = \{(z, t) \in H^{-1}(\Omega) \times L^2(\tilde{\Omega}) : z = 0,\ t \geq 0 \text{ a.e.}\}$. The operator G is twice Fréchet differentiable (see [20], p. 81) and its first derivative is given by

$$G'(y, u)(w, h) = \begin{pmatrix} \nu(w', \cdot') + ((yw)', \cdot) - (\mathcal{B}h, \cdot) \\ -h \end{pmatrix}.$$

The Lagrangian associated to (2.1) can be written as

$$\mathcal{L}(y, u, p, \lambda) = J(y, u) + \nu(y', p') + (yy', p) - (\mathcal{B}u, p) - (\lambda, b - u). \tag{2.2}$$

It is simple to argue the existence of a solution for (2.1). Next we verify the regular point condition [16], which allows us to argue the existence of Lagrange multipliers.

Theorem 2.1 *The operator $G'(y, u)$ is surjective for each $(y, u) \in H_0^1(\Omega) \times L^2(\tilde{\Omega})$.*

Proof For $(f, g) \in H^{-1}(\Omega) \times L^2(\tilde{\Omega})$, it must shown that there exists a pair $(w, h) \in H_0^1(\Omega) \times L^2(\tilde{\Omega})$ such that $G'(y, u)(w, h) = (f, g)$.

Reformulating the equation

$$-\nu w'' + (yw)' = f \quad \text{in} \quad H^{-1}(\Omega) \tag{2.3}$$

as

$$((\nu\Delta)^{-1}\mathcal{K} - I)w = (\nu\Delta)^{-1} f,$$

with $\mathcal{K} : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$ defined by $\mathcal{K}w = (yw)'$, it can be shown, by applying Fredholm's alternative (cf. [20]), that (2.3) has a solution w for each $f \in H^{-1}(\Omega)$.

By setting $h = -g$ and considering $f + \mathcal{B}h$ as right-hand side of (2.3), the existence of a solution w, and hence the surjectivity of $G'(y, u)$, follow. □

With the surjectivity of $G'(y, u)$ established, we can insure the existence of Lagrange multipliers $(p, \lambda) \in H_0^1(\Omega) \times L^2(\tilde{\Omega})$ such that, if (y, u) is the optimal solution of the problem, then

$$\begin{aligned} &\mathcal{L}_{(y,u)}(y, u, p, \lambda)(w, h) = (y - z_d, w) + \alpha(u, h) + \nu(p', w') \\ &-(yp', w) - (\mathcal{B}^* p, h) + (\lambda, h) = 0 \quad \text{for all} \quad (w, h) \in H_0^1(\Omega) \times L^2(\tilde{\Omega}). \end{aligned} \tag{2.4}$$

Taking $h = 0$ implies that

$$\nu(p', w') - (yp', w) = (z_d - y, w) \quad \text{for all} \quad w \in H_0^1(\Omega). \tag{2.5}$$

Also setting $w = 0$ the following relation must hold:

$$\alpha(u, h) - (\mathcal{B}^* p, h) + (\lambda, h) = 0 \quad \text{for all} \quad h \in L^2(\tilde{\Omega}). \tag{2.6}$$

In addition, the complementarity condition $(\lambda, u - b) = 0,\ (\lambda, v - b) \leq 0$, for all $v \leq b$ and the constraint $u \leq b$ must hold.

Consequently, if (y, u) is a local optimal solution of our problem, then it must satisfy the following optimality system in the variational sense:

$$\begin{cases} -\nu y'' + yy' = \mathcal{B}u \\ y(0) = y(1) = 0 \\ -\nu p'' - yp' = z_d - y \\ p(0) = p(1) = 0 \\ \mathcal{B}^* p - \alpha u = \lambda \\ (\lambda, u - b) = 0 \\ u \leq b \\ \lambda \geq 0. \end{cases} \tag{2.7}$$

The last system has two kinds of nonlinearities, the first appearing in the primal equation and the second due to the presence of the inequality constraint resulting in the complementarity condition, which is expressed by the last three equations of (2.7).

Subsequently we discuss three related but different strategies for solving (2.7) numerically.

3 SQP-primal dual method

A first approach to solving (2.7) consists in applying the SQP method and in insuring in each iteration of this method that the solution satisfies the control constraints and the complementarity condition. Such an idea can be formalized through the general equation approach given in [1] and successfully applied to the constrained optimal control of the instationary Burgers equation in [7, 19].

Under appropriate second-order conditions, the equivalence between the SQP and the Lagrange-Newton methods holds locally and the convergence of the algorithm can be proved for sufficiently close starting points (cf. [1]).

3.1 Algorithm

In order to use the theory developed in [1] we define a function $F : H_0^1(\Omega) \times L^2(\tilde{\Omega}) \times H_0^1(\Omega) \times L^2(\tilde{\Omega}) \to (H_0^1(\Omega) \times L^2(\tilde{\Omega}))' \times L^2(\tilde{\Omega}) \times H^{-1}(\Omega)$ in the form:

$$F(y, u, p, \lambda) = \begin{pmatrix} \mathcal{L}_{(y,u)}(y, u, p, \lambda) \\ u - b \\ -\nu(y', \cdot') - (yy', \cdot) + (\mathcal{B}u, \cdot) \end{pmatrix}. \tag{3.1}$$

We also define a multivalued operator T as follows:

$$\begin{aligned} T : H_0^1(\Omega) \times L^2(\tilde{\Omega}) \times H_0^1(\Omega) \times L^2(\tilde{\Omega}) & \\ \to \quad (H_0^1(\Omega) \times L^2(\tilde{\Omega}))' \times L^2(\tilde{\Omega}) \times H^{-1}(\Omega) & \\ (y, u, p, \lambda) \quad \to \quad \{0\} \times \partial_{\Phi_{K^*}}(\lambda) \times \{0\}, & \end{aligned}$$

where $\partial_{\Phi_{K^*}}(\lambda) := \begin{cases} \{z \in L^2(\tilde{\Omega}) : (\beta - \lambda, z) \le 0, \forall \beta \in K^*\} & \text{if} \quad \lambda \in K^* \\ \emptyset \quad \text{if} \quad \lambda \notin K^* \end{cases}$

and $K^* = K = \{f \in L^2(\tilde{\Omega}) : f \ge 0 \quad \text{a.e.}\}$.

The necessary condition can thus be expressed as

$$0 \in F(y, u, p, \lambda) + T(y, u, p, \lambda). \tag{3.2}$$

Applying Newton's method for generalized equations (cf. [1]) to (3.2) yields an algorithm which in the kth iteration, for $x_k := (y_k, u_k, p_k, \lambda_k)$, computes x_{k+1} as solution of

$$0 \in F(x_k) + F'(x_k) \begin{pmatrix} y - y_k \\ u - u_k \\ p - p_k \\ \lambda - \lambda_k \end{pmatrix} + T(x). \tag{3.3}$$

Remark 3.1 Equation (3.3) corresponds to the first-order necessary condition of the optimization problem:

$$\begin{cases} \min \; J_{(y,u)}(y_k, u_k)((y, u) - (y_k, u_k)) + \\ \qquad \dfrac{1}{2}\mathcal{L}''(x_k)((y, u) - (y_k, u_k), (y, u) - (y_k, u_k)) \\ \text{s.t. } \nu(y_k', \cdot') + (y_k y_k', \cdot) - (\mathcal{B}u_k, \cdot) \\ \qquad + \nu(y' - y_k', \cdot') + ((y_k(y - y_k))', \cdot) - (\mathcal{B}(u - u_k), \cdot) = 0 \\ \quad b - u \in K. \end{cases}$$

With $\delta_y := y - y_k,\ \delta_u := u - u_k,\ \delta_p := p - p_k$ and $\delta_\lambda := \lambda - \lambda_k$, (3.3) results in the equations:

$$\begin{aligned}&\nu(\delta_y', w') + (y_k\delta_y', w) + (y_k'\delta_y, w) - (\mathcal{B}\delta_u, w)\\ &\quad = -\nu(y_k', w') - (y_k y_k', w) + (\mathcal{B}u_k, w) \quad \text{for all} \quad w \in H_0^1(\Omega), \end{aligned} \tag{3.4}$$

$$\begin{aligned}&(\delta_y, w) - (\delta_y p_k', w) + \nu(\delta_p', w') - (y_k\delta_p', w)\\ &= (z_d - y_k, w) - \nu(p_k', w') + (y_k p_k', w) \quad \text{for all} \quad w \in H_0^1(\Omega), \end{aligned} \tag{3.5}$$

$$\begin{aligned}&(\alpha\delta_u, h) - (\mathcal{B}^*\delta_p, h) + (\delta_\lambda, h) = (\mathcal{B}^* p_k, h)\\ &\qquad - (\lambda_k, h) - (\alpha u_k, h) \quad \text{for all} \quad h \in L^2(\tilde{\Omega}), \end{aligned} \tag{3.6}$$

$$b - \delta_u - u_k \in \partial_{\Phi_{K^*}}(\lambda). \tag{3.7}$$

By applying Green's formula, the system of equations to be solved can be written as:

$$\begin{cases} -\nu\delta_y'' + y_k\delta_y' + y_k'\delta_y - \mathcal{B}\delta_u = \nu y_k'' - y_k y_k' + \mathcal{B}u_k \\ \delta_y(0) = \delta_y(1) = 0 \\ -\nu\delta_p'' + \delta_y - p_k'\delta_y - y_k\delta_p' = z_d - y_k + \nu p_k'' + y_k p_k' \\ \delta_p(0) = \delta_p(1) = 0 \\ \alpha\delta_u - \mathcal{B}^*\delta_p + \delta_\lambda = \mathcal{B}^* p_k - \lambda_k - \alpha u_k \\ u_k + \delta_u - b \le 0 \\ \lambda_k + \delta_\lambda \ge 0 \\ (\lambda_k + \delta_\lambda, u_k + \delta_u - b) = 0. \end{cases} \tag{3.8}$$

For the solution of (3.8), the main difficulty is given by the complementarity problem expressed by the last three equations of the system. For its solution we apply the primal-dual active set strategy [2], after reformulating the complementarity problem in the following equivalent way:

$$\lambda_k + \delta_\lambda = c\max\{0, u_k + \delta_u - b + \frac{\lambda_k + \delta_\lambda}{c}\} \quad \text{for any} \quad c > 0. \tag{3.9}$$

The primal-dual active set strategy is consequently applied in each SQP iteration to solve the linearized constraint problem and the proposed active and inactive sets in the nth iteration are

$$\mathcal{A}_n^k = \{x : \delta_u + u_k - b + \frac{\lambda_k + \delta_\lambda}{c} > 0\}, \tag{3.10}$$

$$\mathcal{I}_n^k = \{x : \delta_u + u_k - b + \frac{\lambda_k + \delta_\lambda}{c} \le 0\}. \tag{3.11}$$

The algorithm in the kth iteration is then given in the following steps.

Algorithm (PD)

1. Given y_k, p_k, u_k, λ_k. Set $\delta_u = 0$ and solve
$$-\nu\delta_y'' + y_k\delta_y' + y_k'\delta_y = \nu y_k'' - y_k y_k' + \mathcal{B}u_k$$
$$-\nu\delta_p'' + \delta_y - p_k'\delta_y - y_k\delta_p' = z_d - y_k + \nu p_k'' + y_k p_k'.$$
Set $\delta_\lambda = -\alpha u_k + \mathcal{B}^* p_k + \mathcal{B}^*\delta_p - \lambda_k$, $n = 1$.
2. Determine the sets
$$\mathcal{A}_n^k = \{x : u_k + \delta_u - b + \lambda_k + \delta_\lambda > 0\},$$
$$\mathcal{I}_n^k = \{x : u_k + \delta_u - b + \lambda_k + \delta_\lambda \le 0\}.$$
3. If $n \ge 2$ and $\mathcal{A}_n^k = \mathcal{A}_{n-1}^k$, stop.
4. Else, find (δ_y, δ_p) such that
$$-\nu\delta_y'' + y_k\delta_y' + y_k'\delta_y = \nu y_k'' - y_k y_k' + \begin{cases} 0 & \text{if} \quad x \in \Omega\backslash\tilde{\Omega} \\ b & \text{if} \quad x \in \mathcal{A}_n^k \\ \frac{\mathcal{B}^*(p_k+\delta_p)}{\alpha} & \text{if} \quad x \in \mathcal{I}_n^k \end{cases}$$
$$-\nu\delta_p'' + \delta_y - p_k'\delta_y - y_k\delta_p' = z_d - y_k + \nu p_k'' + y_k p_k'.$$
Set $\delta_u = \begin{cases} b - u_k & \text{if} \quad x \in \mathcal{A}_n^k \\ \frac{1}{\alpha}(p_k + \delta_p - \alpha u_k) & \text{if} \quad x \in \mathcal{I}_n^k \end{cases}$.
5. Set $\delta_\lambda = \mathcal{B}^* p_k + \mathcal{B}^*\delta_p - \alpha u_k - \alpha\delta_u - \lambda_k$, update $n := n + 1$ and goto 2.

In the case of high Reynolds numbers, our test examples are linked to positive velocities and, consequently, the use of backward finite differences in the primal equation and forward finite differences in the adjoint equation is justified to insure stability of the finite difference discretization in those cases (cf. [8,9]).

The finite difference discretization of Algorithm (PD) is now standard ([17], p. 77). By introducing the matrix notation:

$$A := \begin{pmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \end{pmatrix}, \quad D := \begin{pmatrix} 1 & 0 & & \\ -1 & 1 & 0 & \\ & \ddots & \ddots & \ddots \end{pmatrix},$$

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_{n-1} \end{pmatrix}, \ \mathbf{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_{n-1} \end{pmatrix}, \ \boldsymbol{\delta}_\mathbf{p} = \begin{pmatrix} \delta_p^1 \\ \vdots \\ \delta_p^{n-1} \end{pmatrix}, \ \boldsymbol{\delta}_\mathbf{y} = \begin{pmatrix} \delta_y^1 \\ \vdots \\ \delta_y^{n-1} \end{pmatrix},$$

$$Y := \mathrm{diag}(\mathbf{y}) = \begin{pmatrix} y_1 & & \\ & y_2 & \\ & & \ddots \end{pmatrix}, \ P := \mathrm{diag}(\mathbf{p}),$$

$$Z := \mathrm{diag}(\mathbf{z}), \ \mathrm{Act} := \mathrm{diag}(\mathrm{act}), \ \mathrm{Inact} := \mathrm{diag}(\mathrm{inact}),$$

where act is the index vector of active points, i.e.,

$$\mathrm{act}_i = \begin{cases} 1 & \text{if} \quad x_i \quad \text{is active,} \\ 0 & \text{if not,} \end{cases}$$

and inact is the index vector of inactive points, the first equation of (4) of Algorithm (PD), by taking $\tilde{\Omega} = \Omega$ and dropping the iteration index, can be written as

$$\left(-\frac{\nu}{h^2}A + \frac{1}{h}YD + \frac{1}{h}\mathrm{diag}(D\,\mathbf{y}),\ -\frac{1}{\alpha}\mathrm{Inact}\right)\begin{pmatrix}\delta_y \\ \delta_p\end{pmatrix} = \left(\frac{\nu}{h^2}A\,\mathbf{y} - \frac{1}{h}(YD)\,\mathbf{y} + \mathrm{Act}\,\mathbf{b} + \frac{1}{\alpha}\mathrm{Inact}\,\mathbf{p}\right).$$

Proceeding in the same way for the second equation, we get

$$\delta_y^i - \left(\frac{p_{i+1}-p_i}{h}\right)\delta_y^i - \nu\left(\frac{\delta_p^{i-1} - 2\delta_p^i + \delta_p^{i+1}}{h^2}\right) - y_i\left(\frac{\delta_p^{i+1}-\delta_p^i}{h}\right) = z_i - y_i + \nu\left(\frac{p_{i+1}-2p_i+p_{i+1}}{h^2}\right) + y_i\left(\frac{p_i - p_{i-1}}{h}\right), \quad (3.12)$$

which in matrix notation is equivalent to

$$\left(I + \frac{1}{h}\mathrm{diag}(D^T p)\ ,\ -\frac{\nu}{h^2}A + \frac{1}{h}YD^T\right)\begin{pmatrix}\delta_y \\ \delta_p\end{pmatrix} = \left(Z - Y + \frac{\nu}{h^2}A\,\mathbf{p} + \frac{1}{h}(YD)\mathbf{p}\right).$$

The complete algorithm can be stated as follows.

Algorithm (SQP-PD)

1. Initialization: give y_0, p_0, u_0, λ_0 and set $k = 0$.
2. While $\|y_k - y_{k-1}\| + \|p_k - p_{k-1}\| + \|u_k - u_{k-1}\| + \|\lambda_k - \lambda_{k-1}\| \geq \sqrt{\varepsilon}$ solve PD).
   Update $y_{k+1} := y_k + \delta_y$, $p_{k+1} := p_k + \delta_p$, $u_{k+1} := u_k + \delta_u$, $\lambda_{k+1} := \lambda_k + \delta_\lambda$.

4 Primal dual-SQP method

An alternative approach for the solution of (2.7) consists in realizing the control constraint and the complementarity condition first, and then solving the resulting nonlinear problem. In this context, based on an alternative formulation of the complementarity condition, iteratively a primal-dual active set strategy is applied to (2.7) and, once the active and inactive sets are defined, the control is determined for the nonlinear problem over the inactive set only.

4.1 Algorithm

As in the previous section, we observe that the three last equations in (2.7) are equivalent to $\lambda = c\max(0, u - b + \frac{\lambda}{c})$ for any $c > 0$.

With the active and inactive sets defined as

$$\mathcal{A} = \{x : u - b + \frac{\lambda}{c} > 0\} \quad \text{for all} \quad c > 0, \tag{4.1}$$

$$\mathcal{I} = \{x : u - b + \frac{\lambda}{c} \leq 0\} \quad \text{for all} \quad c > 0, \tag{4.2}$$

and as $\tilde{\Omega} = \Omega$, the following system of equations must solved:

$$\begin{cases} -\nu y'' + yy' = \begin{cases} b & \text{in} \quad \mathcal{A} \\ \frac{p}{\alpha} & \text{in} \quad \mathcal{I} \end{cases} \\ y(0) = y(1) = 0 \\ -\nu p'' - yp' = z_d - y \\ p(0) = p(1) = 0. \end{cases} \tag{4.3}$$

Thus, the nonlinear P-D method can be stated in the following steps.

Algorithm (Nonlinear P-D)

1. Initialization: given u_0, solve the state and adjoint equations to obtain $y_0,\ p_0,\ \lambda_0 = p_0 - \alpha u_0, n = 1$.
2. Put $\mathcal{A}_n = \{x : u_{n-1} + \lambda_{n-1} > b\},\ \mathcal{I}_n = \{x : u_{n-1} + \lambda_{n-1} \leq b\}$.
3. If $n \geq 2$ and $\mathcal{A}_n = \mathcal{A}_{n-1}$, stop.
4. Otherwise, solve:

$$\begin{cases} -\nu y_n'' + y_n y_n' = \begin{cases} b & \text{in} \quad \mathcal{A}_n \\ \frac{p_n}{\alpha} & \text{in} \quad \mathcal{I}_n \end{cases} \\ y_n(0) = y_n(1) = 0 \\ -\nu p_n'' - y_n p_n' = z_d - y_n \\ p_n(0) = p_n(1) = 0. \end{cases} \tag{4.4}$$

5. Put $u_n = \begin{cases} b & \text{in} \quad \mathcal{A}_n \\ \frac{p_n}{\alpha} & \text{in} \quad \mathcal{I}_n \end{cases}$, $\lambda_n = p_n - \alpha u_n$ and goto 2.

Theorem 4.1 *If there exists $n \geq 1$ such that $\mathcal{A}_n = \mathcal{A}_{n+1}$, then the algorithm stops and the last iterate satisfies:*

$$\begin{cases} -\nu y_n'' + y_n y_n' = u_n \\ y_n(0) = y_n(1) = 0 \\ -\nu p_n'' - y_n p_n' = z_d - y_n \\ p_n(0) = p_n(1) = 0, \end{cases} \tag{4.5}$$

with $\lambda_n = 0$ on $\mathcal{I}_n$ and $\lambda_n > 0$ on $\mathcal{A}_n$.

Proof By construction the system (4.5) is satisfied. It remains only to verify that $u_n \leq b$ and $\lambda_n = 0$ in $\mathcal{I}_n$, and $\lambda_n > 0$ on $\mathcal{A}_n$.

On $\mathcal{I}_n$ we know by construction that $\lambda_n = 0$. Since $\mathcal{A}_n = \mathcal{A}_{n+1}$,

$$u_n + \frac{\lambda_n}{c} = u_n \leq b.$$

On $\mathcal{A}_n$ we have $u_n = b$ and $u_n + \frac{\lambda_n}{c} > b$. Hence $\lambda_n > 0$.

Hence, if the algorithm stops, the last iterate satisfies the optimality system (2.7). □

Applying Newton's method to (4.3) yields the equations:

$$\begin{cases} -\nu \delta_y'' + y' \delta_y + y \delta_y' = \nu y'' - y y' + \begin{cases} b & \text{in} \quad \mathcal{A} \\ \frac{p + \delta_p}{\alpha} & \text{in} \quad \mathcal{I} \end{cases} \\ \delta_y(0) = \delta_y(1) = 0 \\ -\nu \delta_p'' + \delta_y - p' \delta_y - \delta_p' y = z_d - y + \nu p'' + p' y \\ \delta_p(0) = \delta_p(1) = 0. \end{cases} \tag{4.6}$$

By considering finite differences and using the same matrix notation as in Sect. 2, the system can be expressed as:

$$\begin{pmatrix} -\frac{\nu}{h^2} A + \frac{1}{\alpha} Y D + \frac{1}{h} \mathrm{diag}(D\mathbf{y}) & -\frac{1}{\alpha} \mathrm{Inact} \\ I + \frac{1}{h} \mathrm{diag}(D^T p) & -\frac{\nu}{h^2} A + \frac{1}{h} Y D^T \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_p \end{pmatrix} = \begin{pmatrix} \frac{\nu}{h^2} A\, \mathbf{y} - \frac{1}{h} (YD)\mathbf{y} + \mathrm{Act}\, \mathbf{b} + \frac{1}{\alpha} \mathrm{Inact}\, \mathbf{p} \\ Z - Y + \frac{\nu}{h^2} A\, \mathbf{p} + \frac{1}{h} (YD)\mathbf{p} \end{pmatrix}. \tag{4.7}$$

Algorithm (PD-SQP)

1. Initialization: given u_0, solve the state and adjoint equations to obtain $y_0,\ p_0,\ \lambda_0 = p_0 - \alpha u_0, n = 1$.
2. Put $\mathcal{A}_n = \{x_i : u_{n-1} + \lambda_{n-1} > b\},\ \mathcal{I}_n = \{x_i : u_{n-1} + \lambda_{n-1} \leq b\}$.

3. If $n \geq 2$ and $\mathcal{A}_n = \mathcal{A}_{n-1}$, stop.
4. Otherwise, with initial values $\mathbf{y}_{n,0} = \mathbf{y}_{n-1}$, $\mathbf{p}_{n,0} = \mathbf{p}_{n-1}$, solve iteratively, until $\|\delta_y\| + \|\delta_p\| < \sqrt{\varepsilon}$ is satisfied,

$$\begin{pmatrix} -\frac{\nu}{h^2}A + \frac{1}{\alpha}YD + \frac{1}{h}\mathrm{diag}(D\mathbf{y}) & -\frac{1}{\alpha}\mathrm{Inact} \\ I + \frac{1}{h}\mathrm{diag}(D^T p) & -\frac{\nu}{h^2}A + \frac{1}{h}YD^T \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_p \end{pmatrix} = \begin{pmatrix} \frac{\nu}{h^2}A\,\mathbf{y} - \frac{1}{h}(YD)\mathbf{y} + \mathrm{Act}\,\mathbf{b} + \frac{1}{\alpha}\mathrm{Inact}\,\mathbf{p} \\ Z - Y + \frac{\nu}{h^2}A\,\mathbf{p} + \frac{1}{h}(YD)\mathbf{p} \end{pmatrix} \quad (4.8)$$

$\mathbf{y}_{n,k+1} = \mathbf{y}_{n,k} + \delta_y$
$\mathbf{p}_{n,k+1} = \mathbf{p}_{n,k} + \delta_p$.

5. Put $u_n = \begin{cases} b & \text{in} \quad \mathcal{A}_n \\ \frac{p_n}{\alpha} & \text{in} \quad \mathcal{I}_n \end{cases}$, $\lambda_n = p_n - \alpha u_n$ and goto 2.

5 Semi-smooth Newton method

In this section, a semi-smooth Newton method (SSN) is applied, as an alternative strategy, to solve the optimality system (2.7). The aim of this approach is the simultaneous linearization of the primal equation and the complementarity condition.

In the case of linear constrained optimal control problems, the application of the primal dual active set strategy and the semi-smooth Newton method coincide (cf. [11]). In the present case, however, the method coincides with the SQP-PD and PD-SQP methods, if only one primal-dual iteration or one SQP iteration are allowed, respectively.

5.1 Algorithm

For the application of the semi-smooth Newton method, the last three equations in (2.7) are reformulated by the equivalent operator equation

$$C(u, \lambda) := \lambda - \max(0, \lambda + c(u - b)) = 0.$$

By choosing, in particular, $c = \alpha$, the complete system (2.7) can be equivalently expressed as

$$\mathcal{C}(u) = 0,$$

where $\mathcal{C}(u) = -\alpha\, u + \mathcal{B}p - max(0, \mathcal{B}^* p - \alpha b)$ and $p = p(y)$, $y = y(u)$, satisfy the primal and adjoint equations, respectively. Recall from [11] that $\mathrm{Max} : v \to \max(0, v)$ from $L^p(\Omega) \to L^2(\Omega)$ admits a generalized derivative whenever $p > 2$. Since $u \to y(u) \to p(y)$ is smoothing, $u \to \mathcal{C}$

therefore admits a generalized derivative. A generalized derivative for Max is given by

$$D(v) = \begin{cases} 1 & \text{if} \quad v > 0 \\ 0 & \text{if} \quad v \leq 0. \end{cases}$$

With this choice of derivative for the max-operation a Newton step can be applied to $\mathcal{C} = 0$ or, equivalently, to (2.7). The resulting system can again be expressed in terms of an active/inactive set structure. We define these sets as

$$\mathcal{A} = \{x : \lambda + c(u - b) > 0\} \quad \text{and} \quad \mathcal{I} = \{x : \lambda + c(u - b) \leq 0\}. \quad (5.1)$$

Denoting by act and inact the indicator functions for the active and inactive sets, the system to be solved by applying a semi-smooth Newton step to (2.7) becomes:

$$\begin{cases} -\nu\delta_y'' + y'\delta_y + y\delta_y' = \nu y'' - yy' + u + \delta_u \\ \delta_y(0) = \delta_y(1) = 0 \\ -\nu\delta_p'' + \delta_y - p'\delta_y - \delta_p' y = z_d - y + \nu p'' + p'y \\ \delta_p(0) = \delta_p(1) = 0 \\ \delta_p - \delta_\lambda - \alpha\delta_u = -p + \alpha u + \lambda \\ -c\delta_u \cdot \text{act} + \delta_\lambda \cdot \text{inact} = -\lambda + \max\{0, \lambda + c(u - b)\} \end{cases} \quad (5.2)$$

This implies that, on $\mathcal{A}$, $u + \delta_u = b$ and, on $\mathcal{I}$, $\lambda + \delta_\lambda = 0$. Consequently, the last two equations in (5.2) imply that

$$u + \delta_u = \begin{cases} b & \text{in} \quad \mathcal{A} \\ \frac{p+\delta_p}{\alpha} & \text{in } \mathcal{I}. \end{cases} \quad (5.3)$$

Using the same matrix notation as in previous sections, the system to be solved, once the active and inactive sets are estimated, is given by

$$\begin{pmatrix} -\frac{\nu}{h^2}A + \frac{1}{\alpha}YD + \frac{1}{h}\text{diag}(D\mathbf{y}) & -\frac{1}{\alpha}\text{Inact} \\ I + \frac{1}{h}\text{diag}(D^T p) & -\frac{\nu}{h^2}A + \frac{1}{h}YD^T \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_p \end{pmatrix} = \begin{pmatrix} \frac{\nu}{h^2}A\,\mathbf{y} - \frac{1}{h}(YD)\mathbf{y} + \text{Act}\,\mathbf{b} + \frac{1}{\alpha}\text{Inact}\,\mathbf{p} \\ Z - Y + \frac{\nu}{h^2}A\,\mathbf{p} + \frac{1}{h}(YD)\mathbf{p} \end{pmatrix}. \quad (5.4)$$

The complete algorithm can consequently be formulated in the following way.

Algorithm (SSN)

1. Initialization: give $u_0,\ y_0,\ p_0,\ \lambda_0$ and set $k = 1$.
2. Until $\|y_k - y_{k-1}\| + \|p_k - p_{k-1}\| + \|u_k - u_{k-1}\| + \|\lambda_k - \lambda_{k-1}\| < \sqrt{\varepsilon}$ set $\mathcal{A}_k = \{x : u_{k-1} + \lambda_{k-1} > b\},\ \mathcal{I}_k = \{x : u_{k-1} + \lambda_{k-1} \leq b\}$, solve (5.4) and update $y_{k+1} := y_k + \delta_y, p_{k+1} := p_k + \delta_p$, $u_{k+1} = \begin{cases} b & \text{in } \mathcal{A}_k \\ \frac{p_{k+1}}{\alpha} & \text{in } \mathcal{I}_k \end{cases}$, $\lambda_{k+1} = p_{k+1} - \alpha u_{k+1},\ k = k + 1$.

The algorithm obtained shows that, besides the prediction strategy of active and inactive sets, the linear system is solved only once in order to obtain the increments and update the prediction of the sets. Consequently, the semi-smooth Newton method is equivalent to the PD-SQP algorithm if we allow the SQP to do only one iteration, and it is also equivalent to the SQP-PD algorithm if only one primal dual step is allowed.

6 Numerical results

In this section we explain some details concerning the implementation of the methods presented previously as well as their behavior with respect to four examples. For each test case the graph of the solution is given and a table, which depicts the performance of the methods, is displayed.

In the case of the SQP-PD algorithm, the chosen value of the stopping parameter is $\sqrt{\text{eps}}$, where eps denotes the precision of the machine. The algorithm is initialized with a subroutine, which starts with the parameter values equal to zero and continues with a run of the semi-smooth Newton algorithm until the same active sets are reached in two consecutive iterations. The values of the variables (y, u, p, λ) and the active and inactive sets obtained are used as initialization for the SQP-PD algorithm.

The norm used for the stopping criteria of the discretized problems is the discrete L^2 norm. We also remark that δ_{u_k} is initialized as 0 in each auxiliary problem.

In the case of the PD-SQP method, the process begins with the same initialization subroutine as for the SQP-PD. The stopping parameter value is equal to $\sqrt{\text{eps}}$. The times tabulated correspond to the total time needed to get the solution: initialization time + algorithm time.

As regards the semi-smooth Newton method, the stopping parameter value chosen is also $\sqrt{\text{eps}}$ and the initialization values are set to zero. As before, the norm used for the discretized problems is the Euclidean one.

The controls, states and multipliers obtained by the three methods are similar and the numerical l^∞ difference between them is on the order of the machine precision.

The tables presented in this section illustrate the performance of the methods explained. The number of iterations, total running time and convergence rates are recorded.

In the first type of table, the first column corresponds to values of the grid size and the second column shows the number of SSN iterations used for the initialization subroutine of the SQP-PD and the PD-SQP methods. The following three columns show the number of outer and inner iterations, as well as the total time needed by the SQP-PD method. Analogous data are present for the PD-SQP method in the next three columns. Lastly, in the last two columns, the number of iterations and total time for the SSN method are shown.

In the second type of table the convergence order is given. The tables show the values of the convergence factors for each method. For SQP-PD, quadratic convergence is illustrated, while for the SSN method superlinear convergence holds. In the case of PD-SQP, the behavior of the data suggests superlinear convergence. To obtain the convergence factors, we used as exact solution the one obtained by a SSN run with a precision of $\text{eps}^{0.9}$. Due to this approximation, the final iteration of the SSN method does not decrease. Also, due to this fact and the influence of the machine precision in the square of the difference norm, the SQP-PD method has a large factor in the last iterate. For these results on the convergence rates, the initialization for SQP-PD was done by setting all variables equal to zero. PD-SQP was initialized by taking $u = b$ and solving the primal and adjoint equations.

The numerical calculations were done on a DEC-ALPHA 500 machine with a precision of $\text{eps} = 2.22 \times 10^{-16}$. The resulting linear systems were solved exactly, using MATLAB's sparse solver.

6.1 Example 1

The first example to be shown has very standard parameter values. In this case with a constant value for b it is straightforward to inspect the satisfaction of the complementarity condition from the graphical representation.

The parameter values used in this case are: $\text{Re} = 12$, $\alpha = 0.1$, $b = 0.3$, $z_d = 0.3$. Figure 1 shows the graph of the optimal control and its multiplier. Figure 2 shows the final state reached with the application of the optimal control.

The example was solved for an increasing number of grid points and the corresponding results are shown in Table 1. The data for the convergence order with 100 grid points are depicted in Table 2.

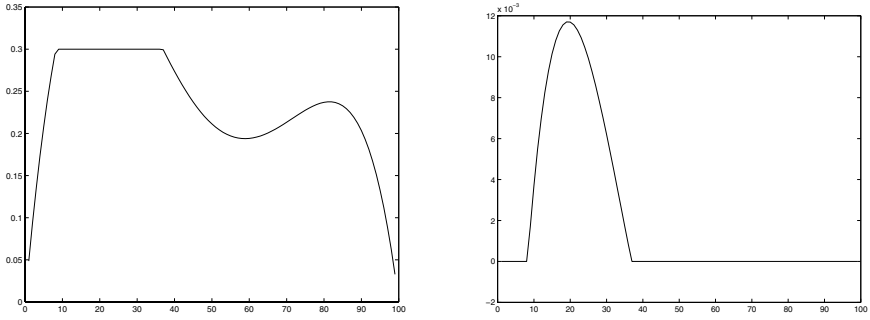

Fig. 1. Example 1: optimal control and its multiplier

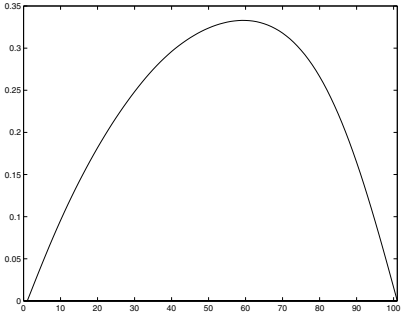

Fig. 2. Example 1: final controlled state

Table 1. Comparison for Example 1

Grid	Init. it	it SQP	it pd	Time	it PD	it sqp	Time	it SS	Time
20	4	2	7	0.0880	2	2	0.0439	6	0.0498
30	4	2	6	0.1074	2	2	0.0614	6	0.0698
40	4	2	6	0.1493	2	2	0.0830	6	0.0913
50	5	1	4	0.1833	2	1	0.1190	6	0.1249
60	4	2	6	0.3013	2	2	0.1541	6	0.1645
70	4	2	6	0.3884	2	2	0.2048	6	0.2262
80	4	2	6	0.5225	2	2	0.2825	6	0.2787
90	4	2	6	0.6700	2	2	0.3401	6	0.3537
100	5	1	4	0.7021	2	1	0.4442	6	0.4513
200	4	2	7	5.1414	2	2	2.4061	6	2.4572

Table 2. Convergence order for Example 1

SQP-PD $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|^2}$	PD-SQP $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$	SSN $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$
0.0305	0.1785	0.1581
0.0116	0.1054	0.4896
0.0598	0.0588	0.1482
0.0232	-	0.0424
3.02×10^{10}	-	7.65×10^{-7}
-	-	1.2×10^{-6}

6.2 Example 2

For this example we make two changes in the data settings from those of the previous example. First we decrease the value of α which, from previous experience (cf. [3]), is known to increase the number of iterations of the primal-dual active set method applied to linear problems. Second, we change the desired state to be nonconstant.

The parameter values used here are: $\mathrm{Re} = 10,\ z_d = \sin(13x),\ b = 0.3$. The parameter α takes two values: either $\alpha = 0.01$ or $\alpha = 0.0001$. Figure 3

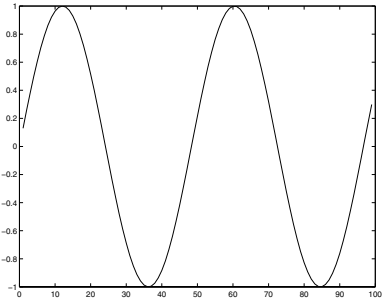

Fig. 3. Example 2: desired state

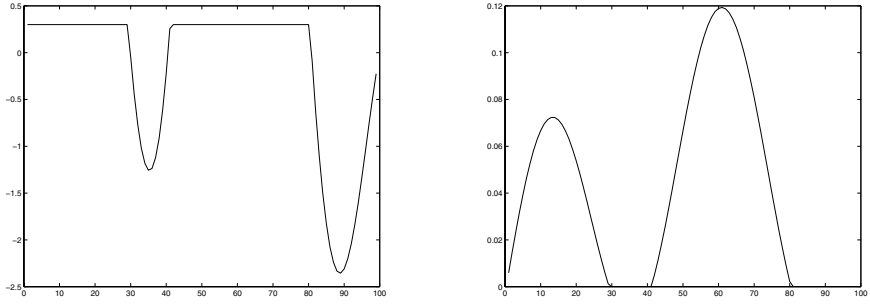

Fig. 4. Example 2: optimal control and its multiplier, $\alpha = 0.01$

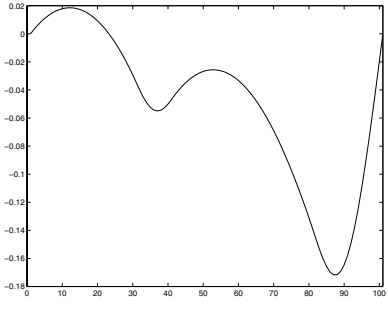

Fig. 5. Example 2: final controlled state, $\alpha = 0.01$

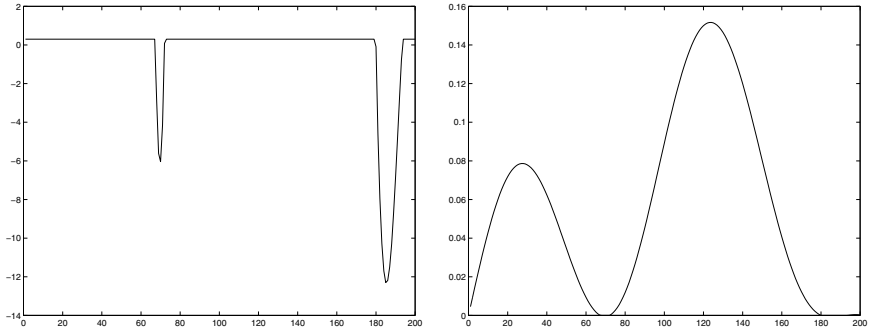

Fig. 6. Example 2: optimal control and its multiplier, $\alpha = 0.0001$

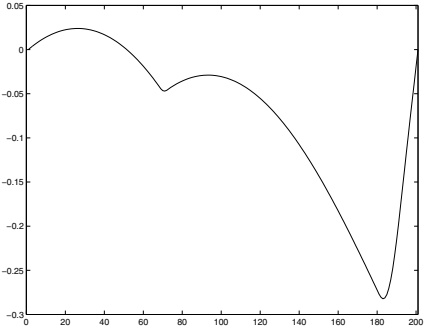

Fig. 7. Example 2: final controlled state, $\alpha = 0.0001$

Table 3. Comparison for Example 2, $\alpha = 0.01$

Grid	Init. it	it SQP	it pd	Time	it PD	it sqp	Time	it SS	Time
20	4	3	11	0.0905	2	2	0.0468	7	0.0566
30	5	2	10	0.1396	2	2	0.0693	7	0.0761
40	5	2	8	0.1767	2	2	0.0936	7	0.1021
50	5	2	10	0.2788	2	2	0.1287	7	0.1371
60	5	2	11	0.4035	2	1	0.1532	7	0.1813
70	5	2	8	0.4535	2	2	0.2262	7	0.2490
80	5	2	11	0.7318	2	2	0.2931	7	0.3032
90	5	2	11	0.9156	2	2	0.3780	7	0.3838
100	5	2	9	1.0376	2	2	0.4680	7	0.4838
200	5	2	10	5.7660	2	2	2.3817	7	2.4446

shows the graph of the desired state and Fig. 4 the graph of the control and its multiplier, when $\alpha = 0.01$. The corresponding optimal state is given in Fig. 5. For $\alpha = 0.0001$ these results are depicted in Figs. 6 and 7. The performance of the methods is presented in Table 3 for $\alpha = 0.01$ and in Table 4 for $\alpha = 0.0001$. The data for the convergence order are recorded in Tables 5 and 6.

Table 4. Comparison for Example 2, $\alpha = 0.0001$

Grid	Init. it	it SQP	it pd	Time	it PD	it sqp	Time	it SS	Time
20	6	3	20	0.1472	2	3	0.0640	9	0.0692
30	7	3	30	0.2915	2	3	0.0936	10	0.1045
40	7	3	20	0.3159	2	3	0.1238	10	0.1357
50	8	2	15	0.3812	2	1	0.1585	10	0.1810
60	8	3	21	0.6474	2	2	0.2258	11	0.2587
70	8	2	12	0.6012	2	2	0.2935	10	0.3050
80	9	2	14	0.8810	2	1	0.3820	11	0.4271
90	8	2	14	1.0586	2	2	0.4689	10	0.4807
100	8	2	14	1.3004	2	2	0.5690	10	0.5866
200	8	2	12	5.2132	2	2	2.3883	10	2.4222

Table 5. Convergence order for Example 2, $\alpha = 0.01$

SQP-PD $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|^2}$	PD-SQP $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$	SSN $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$
0.0143	1.5984	2.8779
0.0077	0.3548	0.3496
0.0090	0.2266	0.2287
0.0204	0.0449	0.0448
7.5×10^{11}	-	9.2×10^{-6}
-	-	1.5×10^{-8}
-	-	0.8331

Table 6. Convergence order for Example 2, $\alpha = 0.0001$

SQP-PD $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|^2}$	PD-SQP $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$	SSN $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$
0.0161	3.6781	4.6968
0.0027	1.7695	1.7142
0.0012	0.9250	0.9307
0.0084	0.5810	0.5904
7.7×10^{6}	0.5549	0.5516
-	0.4307	0.4361
-	0.1370	0.1361
-	-	3.0×10^{-4}
-	-	3.2×10^{-8}
-	-	9.7×10^{-4}

6.3 Example 3

This test deals with the control problem in the case where the constraint and the desired state are not constant functions. Figure 8 shows the graph of the control and its multiplier and Fig. 9 the final optimal state.

The parameter values are: Re $= 10$, $\alpha = 0.01$, $z_d = \sin(13x)$ and $b = \sin(x^2 + 4x)$. The behavior of the methods is depicted in Table 7 and the convergence rates are shown in Table 8.

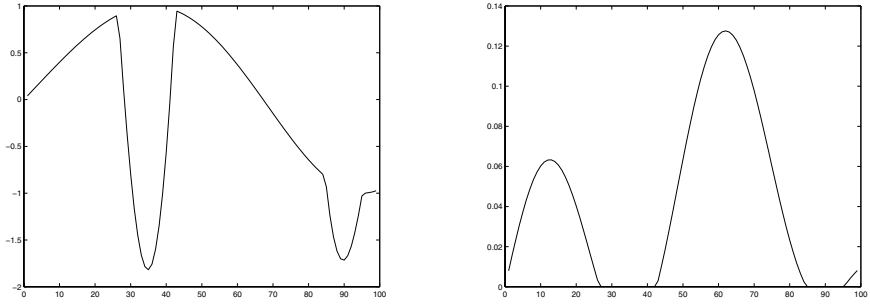

Fig. 8. Example 3: optimal control and its multiplier

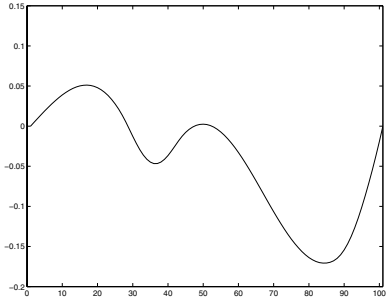

Fig. 9. Example 3: final controlled state

Table 7. Comparison for Example 3

Grid	Init. it	it SQP	it pd	Time	it PD	it sqp	Time	it SS	Time
20	5	2	10	0.0948	2	2	0.0497	7	0.0566
30	5	2	12	0.1544	2	2	0.0712	7	0.0772
40	4	3	15	0.2418	2	3	0.0897	7	0.1016
50	5	2	11	0.2905	2	2	0.1279	7	0.1367
60	5	2	10	0.3754	2	2	0.1718	7	0.1813
70	5	2	10	0.5018	2	2	0.2252	7	0.2362
80	5	2	10	0.6618	2	2	0.2899	7	0.3034
90	6	1	7	0.6979	2	1	0.3780	7	0.3799
100	5	2	10	1.0433	2	2	0.4530	7	0.4686
200	6	1	7	4.4075	2	1	2.2880	7	2.3054

Table 8. Convergence order for Example 3

SQP-PD $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|^2}$	PD-SQP $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$	SSN $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$
0.0175	1.4244	2.6958
0.0091	0.3846	0.3806
0.0130	0.2591	0.2610
0.0179	0.0585	0.0584
3.9×10^{8}	-	1.9×10^{-5}
-	-	2.6×10^{-8}
-	-	0.3379

6.4 Example 4

This last example deals with the control problem under a more difficult constraint function and a higher Reynolds number. The parameter values are: $\mathrm{Re} = 100,\ \alpha = 0.01,\ z_d = 0.3,\ b = x^2$.

Figure 10 shows the optimal control and the multiplier for the inequality constraint. Figure 11 depicts the final controlled state and Table 9 gives the information about the performance of the methods. The convergence rates are shown in Table 10.

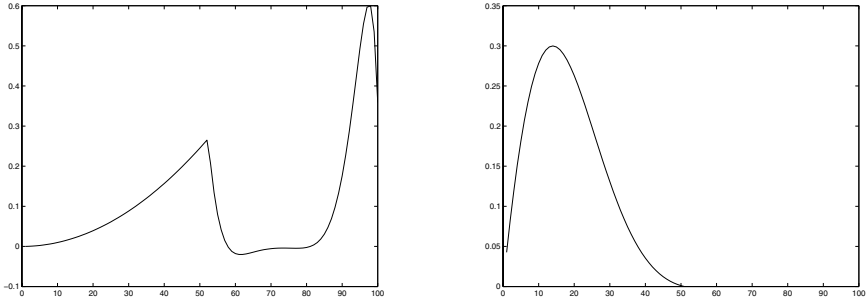

Fig. 10. Example 4: optimal control and its multiplier

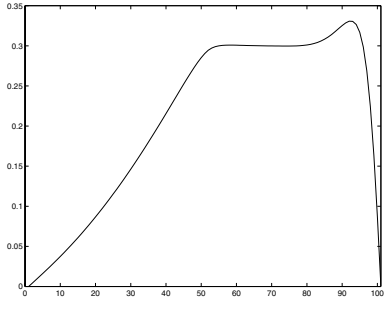

Fig. 11. Example 4: final controlled state

Table 9. Comparison for Example 4

Grid	Init. it	it SQP	it pd	Time	it PD	it sqp	Time	it SS	Time
20	8	3	9	0.1141	2	2	0.0692	11	0.0829
30	9	3	8	0.1667	2	2	0.1070	12	0.1219
40	9	2	4	0.1804	2	2	0.1394	11	0.1463
50	9	2	4	0.2469	2	2	0.1892	11	0.1970
60	8	3	7	0.3855	2	3	0.2460	11	0.2597
70	9	3	8	0.5427	2	3	0.3395	12	0.3559
80	10	2	6	0.6390	2	2	0.4407	12	0.4558
90	10	2	6	0.9201	2	2	0.6324	12	0.6450
100	10	2	6	1.1906	2	2	0.8090	12	0.8247
200	10	2	6	6.2728	2	2	4.1618	12	4.2057

Table 10. Convergence order for Example 4

SQP-PD $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|^2}$	PD-SQP $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$	SSN $\frac{\|u_{k+1}-u^*\|}{\|u_k-u^*\|}$
0.0827	0.5839	0.9499
0.0840	0.7627	0.9827
0.0954	0.7069	0.8023
0.1104	0.5751	0.7531
0.0242	0.4441	0.6995
0.4962	0.2376	0.6126
1.9×10^{14}	-	0.5088
-	-	0.3611
-	-	0.1238
-	-	2.4×10^{-4}
-	-	1.4×10^{-7}
-	-	0.0035

7 Conclusions

The numerical comparison of the methods explained shows that the PD-SQP and the SSN methods are at least as efficient as the SQP-PD method. In all the cases tested, the time taken by the PD-SQP algorithm was shorter than that needed by the other two methods. The difference to the SSN method is not as significant as that to the SQP-PD method. Thus PD-SQP is a strong competitor to SQP-PD and SSN, which may appear, at first sight, to be the more "natural" choices.

We also point out that PD-SQP and SQP-PD have the feature of the primal-dual strategy that the active set is determined in finitely many steps. The stopping criterion is utilized only for the accuracy of computing the nonlinearity in the state equation. Mesh independence of SQP and SSN was verified numerically and we also have strong evidence to conjecture the mesh independence of the nonlinear primal-dual active set method.

In all the cases, as α decreases, the number of iterations increases for all three strategies. Also the number of iterations increases for all the methods with increase in the difficulty of the structure of z_d. For all four examples we observed that strict complementarity holds, i.e., either $\lambda = 0$ or the constraint is active, but these two features do not occur simultaneously.

References

1. Alt, W.: The Lagrange-Newton method for infinite dimensional optimization problems. Numer. Funct. Anal. Optim. **11**, 201–224 (1990)
2. Bergounioux, M., Ito, K., Kunisch, K.: Primal-dual strategy for constrained optimal control problems. SIAM J. Control Optim. **37**, 1176–1194 (1999)
3. Bergounioux, M., Hadou, M., Hintermüller, M., Kunisch, K.: A comparison of a Moreau-Yosida and based active set strategy interior point methods for constrained optimal control problems. SIAM J. Optim. **11**, 495–521 (2000)
4. Burns, J.A., Kang, S.: A control problem for Burgers equation with bounded input-output. NASA CR-182062 Langley, VA: NASA 1990
5. Burns, J.A., Kang, S.: A stabilization problem for Burgers equation with unbounded control and observation. In: Desch, W. et al. (eds.): Estimation and control of distributed parameter systems. (International Series of Numerical Mathematics **100**). Basel: Birkhäuser 1991, pp. 51–72
6. Chen, G., Wang, H.K., Weerakoon, S.: An initial value control problem for the Burgers equation. In: Kappel, F. et al. (eds.): Distributed parameter systems. (Lecture Notes in Control and Information Sciences **75**). Berlin: Springer 1985, pp. 52–76
7. Dontchev, A.L.: Local analysis of a Newton-type method based on partial linearization. In: Renegar, J. et al (eds.): The mathematics of numerical analysis. (Lectures in Applied Mathematics **32**) Providence, RI: AMS 1996, pp. 295–306
8. Fletcher, C.A.J.: Computational techniques for fluid dynamics. 1. Fundamental and general techniques. 2nd ed. Berlin: Springer 1991
9. Griebel, M., Dornseifer, Neunhoeffer, T.: Numerische Simulation in der Strömungslehre. Wiesbaden: Vieweg 1995
10. Hager, W.W., Hearn, D.W.: Application of the dual active set algorithm to quadratic network problems. Comput. Optim. Appl. **1**, 349–373 (1993)
11. Hintermüller, M., Ito, K., Kunisch, K.: The primal dual active set strategy as a semismooth Newton method. SIAM J. Optim. **13**, 865–888 (2003)
12. Hopf, E.: The partial differential equation $u_t + uu_x = \mu u_{xx}$. Comm. Pure Appl. Math. **3**, 201–230 (1950)
13. Ito, K., Kang, S.: A dissipative feedback control synthesis for systems arising in fluid dynamics. SIAM J. Control Optim. **32**, 831–854 (1994)
14. Ito, K., Kunisch, K.: The primal-dual active set method for nonlinear optimal control problems with bilateral constraints. SIAM. J. Control Optim. **43**, 357–376 (2004)
15. Kunisch, K., Rösch, A.: Primal-dual active set strategy for a general class of constrained optimal control problems. SIAM J. Optim. **13**, 321–334 (2002)
16. Maurer, H., Zowe, J.: First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems. Math. Programming **16**, 98–110 (1979)
17. de los Reyes, J.C.: Constrained optimal control of stationary viscous incompressible fluids by primal-dual active set methods. Dissertation. Graz: Karl-Franzens-Universität Graz 2003

18. Temam, R.: Navier-Stokes equations: Theory and numerical analysis. Rev. ed. Amsterdam: North-Holland 1979
19. Tröltzsch, F., Volkwein, S.: The SQP method for control constrained optimal control of the Burgers equation. ESAIM Control. Optim. Calc. Var. **6**, 649–674 (2001)
20. Volkwein, S.: Mesh independence of an augmented Lagrangian-SQP method in Hilbert spaces and control problems for the Burgers equation. Dissertation. Berlin: Technische Universität Berlin 1997