Control and Cybernetics

vol. **28** (1999) No. 2

Discrete relaxed method for semilinear parabolic optimal control problems

by

I. Chryssoverghi, J. Coletsos and B. Kokkinis

Department of Mathematics,
National Technical University, Athens, Greece

We consider an optimal control problem for systems governed by semilinear parabolic partial differential equations with control and state constraints, without any convexity assumptions. A discrete optimization method is proposed to solve this problem in its relaxed form which combines a penalized Armijo type method with a finite element discretization and constructs sequences of discrete Gamkrelidze relaxed controls. Under appropriate assumptions, we prove that accumulation points of these sequences satisfy the relaxed Pontryagin necessary conditions for optimality. Moreover, we show that the Gamkrelidze controls thus generated can be replaced by simulating piecewise constant classical controls.

Keywords: Optimal control, semilinear parabolic systems, discretization, penalty method, relaxed controls

1. Introduction

It is well known that optimal control problems without convexity assumptions have no classical solutions in general, and these problems are usually studied by considering their corresponding relaxed formulations. As a consequence, relaxation theory has been used to develop existence theory, necessary conditions for optimality and also numerical approximation and iterative optimization methods (see Warga, 1972, 1977, 1982; Ekeland, 1972; Teo, 1983; Chryssoverghi, 1986; Chryssoverghi et al., 1993, 1997, 1998; Fattorini, 1994; Casas, 1996; Roubíček, 1991, 1997).

In this paper, we consider an optimal distributed control problem for systems governed by semilinear parabolic partial differential equations with control and global state constraints. In order to solve nonconvex optimal control problems, it seems more natural to apply some appropriate optimization method directly on the relaxed problem, thus exploiting its nonconvex structure at each iteration (see Chryssoverghi et al. 1997). Note that even classical methods often converge to solutions of the relaxed problem (see Williamson and Polak, 1976). On the other hand, to solve numerically these problems one must discretize them in a way or another. We propose here a simultaneous optimization and discretization method to solve this problem in its relaxed form (see Polak, 1997, for related classical methods applied to lumped parameter problems). This method combines in one algorithm a penalized conditional gradient Armijo type method with a finite element discretization, and constructs sequences of discrete Gamkrelidze relaxed controls. The method does not consider separate discrete problems. The discretization of the state equation and functionals involves here an additional numerical integration procedure for implementation reasons (see Chryssoverghi et al., 1993). Under appropriate assumptions, we prove that accumulation points of sequences constructed by this method satisfy the strong relaxed Pontryagin necessary conditions for optimality. Moreover, we show that any converging subsequence thus constructed can be replaced by a sequence of simulating piecewise constant classical controls which converges to the same limit. The main advantages of the discrete relaxed method are the following:

(i) improved implementability of the algorithm due to the Armijo type step search and the numerical integration procedure used,

(ii) reduction of computations due to the progressive refinement of the discretization accordingly to the closer approximation of the continuous extremal control,

(iii) avoidance of the rather strict assumptions on the discrete state constraints assuring the convergence of separate discrete problems,

(iv) the algorithm involves single instead of the double (or triple) infinite procedures involved when some optimization method is applied to each discrete problem separately and the relaxed controls are approximated by classical ones.

2. The continuous optimal control problems

Let Ω be a bounded domain in $\mathbf{R}^d$ with a Lipschitz boundary Γ, and let $I := (0,T)$, $0 < T < \infty$. Consider the following semilinear parabolic state equations

$$\partial y/\partial t + A(t)y = f(x,t,y(x,t),w(x,t)), \quad \text{in } Q := \Omega \times I, \tag{1}$$

$$y(x,t) = 0, \quad \text{in } \Sigma := \Gamma \times I, \tag{2}$$

$$y(x,0) = y^0(x), \quad \text{in } \Omega, \tag{3}$$

where $A(t)$ is the second order differential operator

$$A(t)y := -\sum_{j=1}^{d}\sum_{i=1}^{d}(\partial/\partial x_i)\left[a_{ij}(x,t)(\partial y/\partial x_j)\right].$$

The constraints on the control variable w are

$$w(x,t) \in U, \quad \text{in } Q,$$

where U is a compact, not necessarily convex, subset of $\mathbf{R}^{d'}$. The constraints on the state and the control variables y, w are

$$J_m(w) := \int_Q g_m(x,t,y(x,t),w(x,t))dx\,dt = 0, \quad 1 \le m \le p,$$

$$J_m(w) := \int_Q g_m(x,t,y(x,t),w(x,t))dx\,dt \le 0, \quad p < m \le q,$$

and the cost functional

$$J_0(w) := \int_Q g_0(x,t,y(x,t),w(x,t))dx\,dt,$$

where $y := y_w$ is the solution of (1) for the control w. The optimal control problem is to minimize $J_0(w)$ subject to the above constraints.

Since such problems have no classical solutions in general, without additional convexity assumptions on the data, it is classical (see Warga, 1972, and Roubíček, 1997) to work on the relaxed formulation of the problem, which we define below.

Define the set of classical controls

$$W := \{w : (x,t) \to w(x,t) | w \text{ measurable from } Q \text{ to } U\},$$

and the set of relaxed controls

$$R := \{r : (x,t) \to r(x,t) | r \text{ measurable from } Q \text{ to } M_1(U)\},$$

where the set $M_1(U)$ of probability measures on U is a subset of the space $M(U) \equiv C(U)^*$ of Radon measures on U, and has here the relative weak star topology. We have

$$R \subset L^\infty_w(Q, M(U)) \equiv L^1(Q, C(U))^*,$$

where $L^\infty_w(Q, M(U))$ is the set of (equivalence classes of) functions from Q to $M(U)$ which are measurable w.r.t. a weak norm topology on $M(U)$ (which coincides on $M_1(U)$ with the weak star topology) and essentially bounded w.r.t. the strong dual norm on $M(U)$. The space $L^1(Q, C(U))$ is naturally isomorphic here to the space of Caratheodory functions on $Q \times U$, with an integrability property (see Warga, 1972). The subset R is endowed with the relative weak star topology. The sets $M_1(U)$ and R are convex and, with their respective topology, metrizable and compact. For $\phi \in L^1(Q, C(U))$ and $r \in span\,(R)$, we use the notation

$$\phi(x,t,r(x,t)) := \int_U \phi(x,t,u) r(x,t)(du).$$

Note that this expression is linear w.r.t. r. A sequence $\{r_k\}$ converges to r in R if

$$\lim_{k\to\infty} \int_Q \phi(x,t,r_k(x,t))dx\,dt = \int_Q \phi(x,t,r(x,t))dx\,dt,$$

for every $\phi \in L^1(Q, C(U))$, or equivalently for every function of the form $\phi = \chi\psi$, with $\chi \in C(\overline{Q})$ and $\psi \in C(U)$. In addition, we identify every classical control $w(\cdot,\cdot)$ with its associated Dirac relaxed control $\delta_{w(\cdot,\cdot)}$. Thus, we have $W \subset R$ and it is proved in Warga (1972) that W is dense in R.

We denote by $(\cdot,\cdot)$ and $\| \cdot \|$ the inner product and the norm in $L^2(\Omega)$, by $(\cdot,\cdot)_1$ and $\| \cdot \|_1$ the usual inner product and the norm in the Sobolev space $V := H_0^1(\Omega)$ and by $< \cdot,\cdot >$ the duality bracket between V and its dual V^*. Define the family of bilinear forms on V

$$a(t,y,v) := \sum_{j=1}^{d} \sum_{i=1}^{d} \int_\Omega a_{ij}(x,t)(\partial y/\partial x_j)(\partial v/\partial x_i)dx.$$

Now we can define the weak and relaxed form of the state equation

$$\langle y', v\rangle + a(t,y,v) = \int_\Omega f(x,t,y(x,t),r(x,t))v(x)dx, \text{ for every } v \in V, \quad (4)$$

$$y(x,0) = y^0(x), \quad \text{in } \Omega, \quad (5)$$

where the derivative is taken in the sense of distributions on I with values in V. Defining the functionals

$$J_m(r) := \int_Q g_m(x,t,y(x,t),r(x,t))dx\,dt, \quad 0 \le m \le q,$$

the continuous relaxed optimal control problem (CRP) is to minimize $J_0(r)$ subject to the constraints $r \in R$, $J_m(r) = 0$, $1 \le m \le p$, and $J_m(r) \le 0$, $p < m \le q$, where $y = y_r$ is the (unique) solution of (4)-(5). The continuous classical problem (CCP) is the problem (CRP) with additional constraint $r \in W$.

We suppose that the operators $A(t)$ satisfy the following conditions

$$\sum_{j=1}^{d} \sum_{i=1}^{d} a_{ij}(x,t) z_i z_j \ge \alpha \sum_{i=1}^{d} z_i^2, \quad (x,t) \in \overline{Q}, \quad z_i \in \mathbf{R}, \quad 1 \le i \le d,$$

with $\alpha > 0$, $a_{ij} \in L^\infty(Q)$, $i,j = 1,\ldots,d$, which imply that

$$|a(t,y,v)| \le \alpha_1 \| y \|_1 \| v \|_1, \quad t \in I, \quad y,v \in V,$$

$$a(t,v,v) \ge \alpha_2 \| v \|_1^2, \quad t \in I, \quad v \in V,$$

for some $\alpha_1 \ge \alpha_2 > 0$.

We suppose also that the function f is defined on $Q \times \mathbf{R} \times U$, measurable for fixed y, u, continuous for fixed x, t, and satisfies

$$|f(x,t,y,u)| \le F(x,t) + \beta|y|, \quad \text{for every } x,t,y,u,$$

with $F \in L^2(Q)$, $\beta \ge 0$, and

$$|f(x,t,y_1,u) - f(x,t,y_2,u)| \le L|y_1 - y_2|, \quad \text{for every } x,t,y_1,y_2,u.$$

Then, for every $r \in R$, $y^0 \in L^2(\Omega)$, it can be proved that equations (4)-(5) have a unique solution $y = y_r$ such that $y \in L^2(I, V)$ and $y' \in L^2(I, V^*)$ (see e.g. Zeidler, 1985-1990). It follows that y is essentially equal to a function in $C(\overline{I}, L^2(\Omega))$, and thus the initial condition (5) makes sense.

For completeness, we state here some theoretical results concerning the continuous problems, whose proofs may be found in Chryssoverghi et al. (1993).

For the existence of an optimal relaxed control, we suppose in addition that the functions g_m, $0 \le m \le q$, are measurable for fixed y, u, continuous for fixed x, t, and satisfy

$$|g_m(x,t,y,u)| \le G_m(x,t) + \gamma_m |y|^2, \quad \text{for every } x,t,y,u,$$

with $G_m \in L^1(Q)$, $\gamma_m \ge 0$, $0 \le m \le q$.

LEMMA 2.1 *The mapping $r \mapsto y_r$, from R to $L^2(Q)$, and the functionals J_m, $0 \le m \le q$, are continuous.*

THEOREM 2.1 *If there exists an admissible control, i.e. a control satisfying the constraints, then there exists an optimal control for the CRP.*

For the relaxed necessary condition for optimality, we suppose in addition that f_y' and g_{my}', $0 \le m \le q$, exist, are measurable for fixed y, u, continuous for fixed x, t, and satisfy

$$|g_{my}'(x,t,y,u)| \le G_{1m}(x,t) + \gamma_{1m}|y|, \quad \text{for every } x,t,y,u,$$

with $G_{1m} \in L^2(Q)$, $\gamma_{1m} \ge 0$, $0 \le m \le q$.
Since f is Lipschitzian, we have also

$$|f_y'(x,t,y,u)| \le L, \quad \text{for every } x,t,y,u.$$

LEMMA 2.2 *Dropping the index m, for $r, r' \in R$, the directional derivative of J is given by*

$$\begin{aligned} DJ(r, r' - r) &:= \lim_{\varepsilon \to 0^+} \left[J(r + \varepsilon(r' - r)) - J(r) \right] / \varepsilon \\ &= \int_Q H(x,t,y(x,t),z(x,t),r'(x,t) - r(x,t)) dx\, dt, \end{aligned}$$

where, for each function g, the general Hamiltonian H is defined by

$$H(x,t,y,z,u) := z f(x,t,y,u) + g(x,t,y,u), \tag{6}$$

and the general adjoint state $z := z_r$ satisfies

$$-z' + A^*(t)z = f_y'(y,r)z + g_y'(y,r), \quad \text{in } Q, \tag{7}$$

$$z(x,t) = 0, \quad \text{in } \Sigma, \tag{8}$$

$$z(x,T) = 0, \quad \text{in } \Omega, \tag{9}$$

where $y := y_r$. Moreover, the mappings $r \mapsto z_r$, from R to $L^2(Q)$, and $(r, r') \mapsto DJ(r, r' - r)$, from $R \times R$ to $\mathbf{R}$, are continuous.

THEOREM 2.2 *If $r \in R$ is optimal for either the CRP or the CCP, then r is extremal, i.e. there exist multipliers $\lambda_m \in \mathbf{R}$, $0 \le m < q$, with $\lambda_0 \ge 0$, $\lambda_m \ge 0$, $p < m \le q$, and $\sum_0^q |\lambda_m| \ne 0$, such that*

$$H(x,t,y(x,t),z(x,t),r(x,t)) \\ = \min_{u \in U} H(x,t,y(x,t),z(x,t),u), \quad \text{a.e. in } Q, \ (\textit{Minimum Principle}),$$

and

$$\lambda_m J_m(r) = 0, \quad p < m \le q, \quad (\textit{Transversality Conditions}),$$

where H and z are defined by (6-9), with g replaced by $\sum_0^q \lambda_m g_m$.

Remark. The above minimum principle is in fact equivalent to the global condition

$$\int_Q H(x,t,y(x,t),z(x,t),r^{'}(x,t) - r(x,t))dx\,dt \ge 0, \quad \text{for every } r^{'} \in R,$$

i.e.

$$\sum_0^q \lambda_m DJ_m(r, r^{'} - r) \ge 0, \quad \text{for every } r^{'} \in R,$$

(see Warga, 1972, p. 361), and Theorem 2.2 follows then from the general (nonqualified) multiplier Theorem V.2.3 from Warga (1972) (see also Ioffe and Tikhomirov, 1979).

3. Discretizations

We shall now define some discrete approximations of controls, states and functionals. We suppose here for simplicity that the domain Ω is a polyhedron and that $a(t,v,w) = a(v,w)$. In addition, we suppose that the functions $f, f_y^{'}$ and $g_m, g_{my}^{'}$, $0 \le m \le q$, are continuous on $\overline{Q} \times \mathbf{R} \times U$ and that the functions F and G_m, G_{1m}, $0 \le m \le q$, are constant.

For every integer $n \ge 1$, let $\{S_i^n\}_{i=1}^{M(n)}$ be an admissible regular quasi-uniform triangulation (see Ciarlet, 1978, or Temam, 1977) of $\overline{\Omega}$ into closed d-simplices, with $h_n := \max_i[diam\,(S_i^n)] \to 0$ as $n \to \infty$, and $\{I_j^n\}_{j=0}^{N(n)-1}$ a subdivision of the interval $\overline{I}$ into $N(n)$ intervals $I_j^n := [t_j^n, t_{j+1}^n]$, of length Δt_j^n, with $\Delta t^n := \max_j \Delta t_j^n \to 0$ as $n \to \infty$. We suppose also for simplicity that each $S_{i'}^{n+1}$ is a subset of some S_i^n and that each $I_{j'}^{n+1}$ is a subset of some I_j^n. Set $Q_{ij}^n := S_i^n \times I_j^n$.

Let $V^n \subset V$ be the subspace of functions which are continuous on $\overline{\Omega}$ and affine on each S_i^n. Let $R^n \subset R$ be the set of discrete relaxed controls

$$R^n := \left\{ r^n \in R \middle| r^n(x,t) := r_{ij}^n \in M_1(U), \text{ on } \overset{o}{Q}{}_{ij}^n, 1 \le i \le M, 0 \le j \le N-1 \right\},$$

and $W^n := R^n \cap W$ the set of discrete classical controls. Note that we have $R^{n+1} \subset R^n$. R^n is endowed with the product weak star topology of $[M_1(U)]^{MN}$.

For numerical reasons and also in order that the algorithm described in Section 5 be implementable, we shall carry here the discretizations of states and functionals a step further than in Chryssoverghi et al. (1993). We shall use the following notation, for a function ϕ defined on $\overline{\Omega}$

$$[\phi]_\Omega^n := \sum_{i=1}^{M} \sum_{k=1}^{L} C^{nk} meas\,(S_i^n)\phi(x_i^{nk}),$$

where the x_i^{nk} and C^{nk} are nodes and coefficients of some integration rule in the simplex S_i^n for each n, with $\sum_{k=1}^{L} C^{nk} = 1$, $C^{nk} \ge 0$, $1 \le k \le L(n)$, and $L(n) \le L'$ for every n. For a given discrete relaxed control

$$r^n := (r_j^n, 0 \le j \le N-1), \quad \text{with } r_j^n := (r_{ij}^n, 1 \le i \le M),$$

the corresponding discrete state $y^n := (y_j^n, 0 \le j \le N)$ is given by

$$(1/\Delta t_j)(y_{j+1}^n - y_j^n, v) + a(y_{j+1}, v) = [f(t_j^n, y_j^n, r_j^n)v]_\Omega^n,$$
$$\text{for every } v \in V^n,\ 0 \le j \le N\text{–}1, \tag{10}$$
$$y_0^n \text{ given}, \quad y_j^n \in V^n, \quad 0 \le j \le N. \tag{11}$$

It is clearly understood that, for each j, $r^n = r_{ij}^n$ (constant measure) on the closed simplex S_i^n in the expression $[f(t_j^n, y_j^n, r_j^n)v]_\Omega^n$. Choosing a basis in V^n, equation (10) clearly reduces to a regular linear system for each j. The discrete functionals are defined by

$$J_m^n(r^n) := \sum_{j=0}^{N-1} \Delta t_j^n [g_m(t_j^n, y_j^n, r_j^n)]_\Omega^n, \quad 0 \le m \le q. \tag{12}$$

LEMMA 3.1 *Let $v \in V^n$ and define for each $1 \le k \le L$ the piecewise constant function a.e. in $\overline{\Omega}$*

$$\overline{v}^k(x) = v\left(x_i^{nk}\right) \quad in \ \overset{o}{S}{}_i^n, \quad 1 \le i \le M.$$

Then

$$\| v - \overline{v}^k \| \le h_n \| v \|_1, \tag{13}$$

and

$$\| \overline{v}^k \| = \left\{ \sum_{i=1}^{M} meas(S_i^n) \left[v(x_i^{nk}) \right]^2 \right\}^{1/2} \leq c \| v \|, \tag{14}$$

where the constant c is independent of n.

Proof: By Taylor's formula and since v is linear on each S_i^n, we have

$$\overline{v}^k(x) = v(x_i^{nk}) = v(x) + (x_i^{nk} - x)^T \bigtriangledown v(x), \quad \text{in } S_i^n,$$

hence

$$|\overline{v}^k(x) - v(x)| \leq \| x_i^{nk} - x \|_2 \| \bigtriangledown v(x) \|_2 \leq h_n \| \bigtriangledown v(x) \|_2, \quad \text{in } S_i^n,$$

where $\| \cdot \|_2$ denotes the euclidean norm in $\mathbf{R}^d$. Therefore

$$\| \overline{v}^k - v \| \leq h_n \| \bigtriangledown v \| \leq h_n \| v \|_1 .$$

It follows that

$$\| \overline{v}^k \| \leq \| v \| + \| \overline{v}^k - v \| \leq \| v \| + h_n \| v \|_1,$$

and using the inverse inequality (see Ciarlet, 1978)

$$\| \overline{v}^k \| \leq c \| v \| .$$

■

LEMMA 3.2 *The functionals* $r^n \mapsto y_j^n$, $1 \leq j \leq N$, *and* $r^n \mapsto J_m^n(r^n)$, $0 \leq m \leq q$, *on* $R^n \equiv [M_1(U)]^{MN}$ *are continuous.*

Proof: We first remark that if a function ϕ defined on $\mathbf{R} \times U$ is continuous, then the natural extension of ϕ to $\mathbf{R} \times M_1(U)$

$$\phi(y, \sigma) := \int_U \phi(y, u) \sigma(du)$$

is continuous. The continuity of the mappings $r^n \mapsto y_j^n$, $1 \leq j \leq N$, is then easily proved by induction on j by choosing a basis in V^n. The continuity of the functionals $r^n \mapsto J_m^n(r^n)$ follows.

■

We state here a useful straightforword generalization of the discrete Gronwall inequality (see Thomee, 1997). Let $\{\phi_j\}_0^N$, $\{\psi_j\}_0^N$, $\{\delta_j\}_0^{N-1}$ be sequences of nonnegative numbers, with $\{\psi_j\}$ nondecreasing. If

$$\phi_k \leq \psi_k + \sum_0^{k-1} \delta_j \phi_j, \quad 0 \leq k \leq N,$$

then

$$\phi_k \leq \psi_k e^{\sum_0^{k-1} \delta_j}, \qquad 0 \leq k \leq N.$$

LEMMA 3.3 *Dropping the index n, for $r, r' \in R^n$, $\varepsilon \in (0,1]$, set*

$$r_\varepsilon := r + \varepsilon(r' - r), \quad y := y_r, \quad y_\varepsilon := y_{r_\varepsilon}, \quad \Delta y := y_\varepsilon - y.$$

We have

$$\| \Delta y_j \| \leq c\varepsilon, \quad 0 \leq j \leq N.$$

Proof: By the discrete state equation, we have

$$\begin{aligned}
&(\Delta y_{j+1} - \Delta y_j, \Delta y_{j+1}) + \Delta t_j a(\Delta y_{j+1}, \Delta y_{j+1}) \\
&= \Delta t_j \left[\{ f(y_{\varepsilon j}, r_{\varepsilon j}) - f(y_j, r_j) \} \Delta y_{j+1} \right]_\Omega^n .
\end{aligned}$$

By our assumptions on f, c denoting various constants and for $0 < \theta \leq 1$

$$\begin{aligned}
&\| \Delta y_{j+1} - \Delta y_j \|^2 + \| \Delta y_{j+1} \|^2 - \| \Delta y_j \|^2 + 2\alpha_2 \Delta t_j \| \Delta y_{j+1} \|_1^2 \\
&\leq 2\Delta t_j \left[\left\{ |f(y_{\varepsilon j}, r_{\varepsilon j}) - f(y_j, r_{\varepsilon j})| + |\varepsilon f(y_j, r_j' - r_j)| \right\} |\Delta y_{j+1}| \right]_\Omega^n \\
&\leq c\Delta t_j \left[\{ |\Delta y_j| + \varepsilon(1 + |y_j|) \} |\Delta y_{j+1}| \right]_\Omega^n \\
&\leq c\Delta t_j \left[(|\Delta y_j| + \varepsilon)(|\Delta y_j| + |\Delta y_{j+1} - \Delta y_j|) \right]_\Omega^n \\
&\leq c\Delta t_j \left[\varepsilon^2/\theta + |\Delta y_j|^2/\theta + \theta |\Delta y_{j+1} - \Delta y_j|^2 \right]_\Omega^n \\
&\leq c\Delta t_j \sum_{k=1}^{L} C^{nk} \left(\varepsilon^2/\theta + \| \overline{\Delta y}_j^k \|^2 /\theta + \theta \| \overline{\Delta y}_{j+1}^k - \overline{\Delta y}_j^k \|^2 \right) \\
&\leq c\Delta t_j \left(\varepsilon^2/\theta + \| \Delta y_j \|^2 /\theta + \theta \| \Delta y_{j+1} - \Delta y_j \|^2 \right) \quad \text{(by (14))}
\end{aligned}$$

Choosing θ sufficiently small, it follows that

$$\begin{aligned}
&\| \Delta y_{j+1} - \Delta y_j \|^2 /2 + \| \Delta y_{j+1} \|^2 - \| \Delta y_j \|^2 + 2\alpha_2 \Delta t_j \| \Delta y_{j+1} \|_1^2 \\
&\leq c\Delta t_j \left(\varepsilon^2 + \| \Delta y_j \|^2 \right).
\end{aligned}$$

By summation and the discrete Gronwall inequality, we obtain (in particular), since $\Delta y_0 = 0$ and $\sum_j \Delta t_j = T$

$$\| \Delta y_j \| \leq c\varepsilon, \quad 0 \leq j \leq N.$$

■

THEOREM 3.1 *Dropping the index m, define the general discrete relaxed adjoint state $z^n = (z_j^n, 0 \leq j \leq N)$ by*

$$\begin{aligned}
&(1/\Delta t_j^n)(z_j^n - z_{j+1}^n, v) + a(z_j^n, v) \\
&= \left[f_y'(t_j^n, y_j^n, r_j^n) z_{j+1}^n v \right]_\Omega^n + \left[g_{yj}'(t_j^n, y_j^n, r_j^n) v \right]_\Omega^n, \quad \textit{for every } v \in V^n, \\
&j = N-1, \ldots, 0, \quad z_N^n = 0, \quad z_j^n \in V^n, \quad 0 \leq j \leq N. \qquad (15)
\end{aligned}$$

The directional derivative of the functional J^n on R^n is given by

$$DJ^n(r^n, r'^n - r^n) = \sum_{j=0}^{N-1} \Delta t_j^n \left[H(t_j^n, y_j^n, z_{j+1}^n, r_j'^n - r_j^n) \right]_\Omega^n,$$

where H is defined by (6). Moreover, the functional on $R^n \times R^n$

$$(r^n, r'^n) \mapsto DJ^n(r^n, r'^n - r^n)$$

is continuous.

Proof: By Lemma 3.1, (14), we have for $0 \le j \le N-1$, $1 \le k \le L$

$$\| \overline{\Delta y}_j^k \| := \left\{ \sum_{i=1}^{M} meas\,(S_i^n) |\Delta y_j(x_i^{nk})|^2 \right\}^{1/2} \le c \, \| \Delta y_j \| .$$

Using the Mean Value Theorem and the assumptions on g, it can then be easily proved that the discrete functional on V^n

$$G(y) := [g(y, r_j)]_\Omega^n$$

has Fréchet derivative

$$G'(y)\Delta y = \left[g_y'(y, r_j)\Delta y \right]_\Omega^n .$$

By Lemma 3.3, dropping the index n, we have

$$J(r_\varepsilon) - J(r) = \sum_{j=0}^{N-1} \Delta t_j \,[g(y_{\varepsilon j}, r_{\varepsilon j})]_\Omega^n - \sum_{j=0}^{N-1} \Delta t_j [g(y_j, r_j)]_\Omega^n$$

$$= \sum_{j=0}^{N-1} \Delta t_j \,[g(y_{\varepsilon j}, r_{\varepsilon j})]_\Omega^n - \sum_{j=0}^{N-1} \Delta t_j [g(y_j, r_{\varepsilon j})]_\Omega^n + \sum_{j=0}^{N-1} \Delta t_j \left[\varepsilon g(y_j, r_j' - r_j) \right]_\Omega^n$$

$$= \sum_{j=0}^{N-1} \Delta t_j \left[g_y'(y_j, r_{\varepsilon j})\Delta y_j \right]_\Omega^n + o(|\Delta y_j|) + \varepsilon \sum_{j=0}^{N-1} \Delta t_j \left[g(y_j, r_j' - r_j) \right]_\Omega^n$$

$$= \sum_{j=0}^{N-1} \Delta t_j \left[g_y'(y_j, r_j)\Delta y_j \right]_\Omega^n + \varepsilon \sum_{j=0}^{N-1} \Delta t_j \left[g_y'(y_j, r_j' - r_j)\Delta y_j \right]_\Omega^n + o(\varepsilon)$$

$$+ \varepsilon \sum_{j=0}^{N-1} \Delta t_j \left[g(y_j, r_j' - r_j) \right]_\Omega^n$$

$$= \sum_{j=0}^{N-1} \Delta t_j \left[g_y'(y_j, r_j)\Delta y_j \right]_\Omega^n + \varepsilon \sum_{j=0}^{N-1} \Delta t_j \left[g(y_j, r_j' - r_j) \right]_\Omega^n + o(\varepsilon).$$

Now, the state equation, with $v = z_{j+1}$, yields by similar arguments

$$(\Delta y_{j+1} - \Delta y_j, z_{j+1}) + \Delta t_j a(\Delta y_{j+1}, z_{j+1})$$
$$= \Delta t_j \left[f_y'(y_j, r_j) \Delta y_j z_{j+1} \right]_\Omega^n + \varepsilon \Delta t_j \left[f(y_j, r_j' - r_j) z_{j+1} \right]_\Omega^n + o(\varepsilon).$$

and the adjoint equation, with $v = \Delta y_j$

$$(z_j - z_{j+1}, \Delta y_j) + \Delta t_j a(z_j, \Delta y_j)$$
$$= \Delta t_j \left[f_y'(y_j, r_j) z_{j+1} \Delta y_j \right]_\Omega^n + \Delta t_j \left[g_y'(y_j, r_j) \Delta y_j \right]_\Omega^n.$$

Since $\Delta y_0 = z_N = 0$, the result follows by summation over j. ∎

4. Approximation, stability and consistency

The following approximation lemma is proved in Chryssoverghi et al. (1993).

Lemma 4.1 *(Approximation of R) For every $r \in R$, there exists a sequence of discrete classical controls $\{w^n \in W^n\}$ such that $w^n \to r$ in R.*

Lemma 4.2 *(Stability of states) For every $r^n \in R^n$, if $\{y_0^n\}$ is bounded in V^n, then*

$$\| y_j^n \| \leq c, \quad 0 \leq j \leq N,$$
$$\sum_{j=0}^{N} \Delta t_j^n \| y_j^n \|_1^2 \leq c,$$
$$\sum_{j=0}^{N-1} \| y_{j+1}^n - y_j^n \|^2 \leq c,$$

where the constant c is independent of n.

Proof: The proof is similar to the proof of Lemma 3.3 and is omitted. ∎

For given values v_j^n, $j = 0, \ldots, N$, in a vector space, define the following functions, a.e. on $\overline{I}$

$$v_-^n(t) := v_j^n, \quad t \in \overset{o}{I_j^n}, \quad j = 0, \ldots, N-1,$$
$$v_+^n(t) := v_{j+1}^n, \quad t \in \overset{o}{I_j^n}, \quad j = 0, \ldots, N-1,$$
$v_\wedge^n(t)$:= the function which is affine on each I_j^n and such that

$$v_\wedge^n(t_j) = v_j^n, \quad j = 0, \ldots, N.$$

From now on, we suppose that $y^0 \in V$ and we choose y_0^n to be the projection of y^0 onto V^n in V, which implies that $y_0^n \to y^0$ in V strongly, as $n \to \infty$.

THEOREM 4.1 *(Consistency of states and functionals) If $r^n \to r$ in R, then the corresponding discrete states $y_-^n, y_+^n, y_\wedge^n$ converge to y_r in $L^2(Q)$ strongly, as $n \to \infty$ and*

$$\lim_{n\to\infty} J_m^n(r^n) = J_m(r), \quad 0 \le m \le q.$$

Proof: We sketch the proof of the first assertion, which parallels that of Lemma 4.3 in Chryssoverghi et al. (1993), giving the relevant modifications. Using Lemma 4.2, we first show that the sequences $\{y_-^n\}, \{y_+^n\}$ and $\{y_\wedge^n\}$ are bounded and (up to subsequences) converge to some y in $L^2(I,V)$ weakly. Next, by Riesz's theorem, for every $0 \le j \le N-1$, there exists $\psi_j^n \in V^n$ such that

$$(y_\wedge^{n'}, v^n) = (\psi_j^n, v^n)_1 = (\psi_-^n, v^n)_1, \quad \text{for } t \in I_j^n, \quad 0 \le j \le N-1.$$

From the state equation and Lemma 3.1, (14), we have, for $0 \le j \le N-1$

$$\begin{aligned}
\left|(\psi_j^n, v^n)_1\right| &\le c\left[\| y_{j+1}^n \|_1 \| v^n \|_1 + \left(1 + \sum_{k=1}^{L} C^{nk} \| \overline{y}_j^{nk} \|\right) \| \overline{v}^n \|\right] \\
&\le c\left[\| y_{j+1}^n \|_1 \| v^n \|_1 + (1 + \| y_j^n \|) \| v^n \|\right] \\
&\le c\left(1 + \| y_{j+1}^n \|_1 + \| y_j^n \|\right) \| v^n \|_1,
\end{aligned}$$

hence

$$\| \psi_j^n \|_1 \le c\left(1 + \| y_{j+1}^n \|_1 + \| y_j^n \|\right), \quad 0 \le j \le N-1.$$

Therefore, by Lemma 4.2

$$\begin{aligned}
\int_I \| \psi_-^n(t) \|_1 \, dt &\le c \sum_{j=0}^{N-1} \Delta t_j^n \left(1 + \| y_{j+1}^n \|_1 + \| y_j^n \|\right) \\
&\le c\left\{\left[\sum_{j=0}^{N-1} \Delta t_j^n \| y_{j+1}^n \|_1^2\right]^{1/2} + \sum_{j=0}^{N-1} \Delta t_j^n \left(1 + \| y_j^n \|\right)\right\} \le c.
\end{aligned}$$

It follows (see Lemma 5.6 in Temam, 1976) that

$$\int_{-\infty}^{\infty} |\tau|^{2s} \| \mathcal{F}(y_\wedge^n)(\tau) \|^2 \, d\tau \le c,$$

for $0 < s < 1/4$, where $\mathcal{F}(y_\wedge^n)$ is the Fourier transform of $y_\wedge^n$, and hence, by the Compactness Theorem 2.2 in Temam (1976), that (up to a subsequence) $y_\wedge^n \to y$ in $L^2(Q)$ strongly. The passage to the limit in the first two terms of the state equation in integrated form remains unchanged. For the third term containing f, we proceed as follows. Since $y_-^n \to y$ in $L^2(Q)$ strongly, by Lemma 3.1, (13), we have

$$\| \overline{y}_-^{nk} - y_-^n \|_{L^2(Q)} \le h_n \| y_-^n \|_{L^2(I,V)} \le c h_n \to 0 \text{ as } n \to \infty, \text{ for } 1 \le k \le L,$$

hence $\overline{y}^{nk}_{-} \to y$ in $L^2(Q)$ strongly as $n \to \infty$, for $1 \le k \le L$. Similarly (see notation in Chryssoverghi et al., 1993), $w^n_+ \to w$ implies that $\overline{w}^{nk}_+ \to w$ as $n \to \infty$ in $L^2(Q)$ strongly, for $1 \le k \le L$. Finally, since $\sum_{k=1}^{L(n)} C^{nk} = 1$ for every n, we obtain

$$\lim_{n\to\infty} \sum_{j=0}^{N-1} \Delta t^n_j \left[f(t^n_j, y^n_j, r^n_j) w^n_{j+1}\right]^n_\Omega$$
$$= \lim_{n\to\infty} \sum_{k=1}^{L(n)} C^{nk} \int_Q f\left(\overline{x}^{nk}, t^n_-, \overline{y}^{nk}_-, r^n\right) \overline{w}^{nk}_+ dx\, dt$$
$$= \int_Q f(x,t,y,r) w dx\, dt + \lim_{n\to\infty} \sum_{k=1}^{L(n)} C^{nk} \varepsilon^k_n \quad (\text{with } \varepsilon^k_n \to 0)$$
$$= \int_Q f(x,t,y,r) w dx\, dt.$$

It then follows that $y = y_r$.

The convergences $J^n_m(r^n) \to J_m(r)$ are proved by using Proposition 2.1 in Chryssoverghi (1986). ■

THEOREM 4.2 *(Consistency of adjoints and derivatives) If $r^n \to r$ in R, then the corresponding discrete adjoint states $z^n_-, z^n_+, z^n_\wedge$ converge to z_r in $L^2(Q)$ strongly, as $n \to \infty$. Moreover, if $r^n \to r$ and $r'^n \to r'$ in R, then*

$$\lim_{n\to\infty} DJ^n_m(r^n, r'^n - r^n) = DJ_m(r, r' - r), \quad 0 \le m \le q.$$

Proof: The proof is similar to that of Theorem 4.1, using this theorem and Proposition 2.1 in Chryssoverghi (1986). ■

5. Discrete relaxed method

Let $\{\varepsilon_{m\kappa}\}$, $1 \le m \le q$, be positive decreasing sequences converging to zero and define the penalized discrete functionals on $R^n \subset R$

$$J^{n,\kappa}(r) := J^n_0(r) + \sum_{m=1}^{p} [J^n_m(r)]^2/(2\varepsilon_{m\kappa}) + \sum_{m=p+1}^{q} \left[\max(0, J^n_m(r))\right]^2/(2\varepsilon_{m\kappa}).$$

Let $\{\gamma_\kappa\}, \{\delta_\kappa\}$ be positive (decreasing) sequences converging to 0, with $\delta_\kappa \le \gamma_\kappa$ for every κ, let $b, c \in (0,1)$ and consider the following algorithm.

Algorithm

Step 1. Choose $r_1 \in R^1$ and set $\overline{r}_0 := r_1$, $\nu = \kappa = n = 1$.
Step 2. Find $\overline{r}_\nu \in R$ such that

$$D_\nu := DJ^{n,\kappa}(r_\nu, \overline{r}_\nu - r_\nu) = \min_{r' \in R^n} DJ^{n,\kappa}(r_\nu, r' - r_\nu). \tag{16}$$

Step 3. If $|D_\nu| \geq \gamma_\kappa$, go to Step 6.
Step 4. If $|D_\nu| \leq \delta_\kappa$, set $r^n := r_\nu$, $D^n := D_\nu$, $\varepsilon_m^n := \varepsilon_{m\kappa}$, $1 \leq m \leq q$, $J^n := J^{n,\kappa}$, $n := n+1$.
Step 5. Set $\kappa := \kappa + 1$ and go to Step 2.
Step 6. Find the smallest positive integer s such that $\alpha_\nu := c^s$ satisfies the inequality

$$J^{n,\kappa}(r_\nu + \alpha_\nu(\overline{r}_\nu - r_\nu)) - J^{n,\kappa}(r_\nu) \leq \alpha_\nu b D_\nu, \tag{17}$$

and set

$$r'_{\nu+1} := r_\nu + \alpha_\nu(\overline{r}_\nu - r_\nu). \tag{18}$$

Step 7. Choose any $r_{\nu+1} \in R$ such that

$$J^{n,\kappa}(r_{\nu+1}) \leq J^{n,\kappa}(r'_{\nu+1}), \tag{19}$$

set $\nu := \nu + 1$ and go to Step 2.

Using the sequences constructed by the algorithm, define the sequences of multipliers

$$\begin{aligned} \lambda_m^n &:= J_m^n(r^n)/\varepsilon_m^n, \quad 1 \leq m \leq p, \\ \lambda_m^n &:= \max(0, J_m^n(r^n))/\varepsilon_m^n, \quad p < m \leq q. \end{aligned}$$

THEOREM 5.1 *Let $\{r^n\}$ be a subsequence (same notation) of the sequence constructed by the algorithm which converges to some control $r \in R$.*

i) If the sequences $\{\lambda_m^n\}$, $1 \leq m \leq q$, remain bounded, then r is admissible and extremal.

ii) Suppose that the problem CRP has no admissible abnormal (see Warga, 1972) extremal controls. If the limit r is admissible, then the sequences $\{\lambda_m^n\}$, $1 \leq m \leq q$, are bounded and r is extremal.

Proof: We shall show first that if κ and n remain constant, then $D_\nu \to 0$ as $\nu \to \infty$. Note that $D_\nu \leq 0$ for every ν by definition. Suppose that $r_\nu \to \widetilde{r}$ and $\overline{r}_\nu \to \overline{r}$ as $\nu \to \infty$ (up to subsequences) and that

$$D := \lim_{\nu\to\infty} D_\nu = \lim_{\nu\to\infty} DJ^{n,\kappa}(r_\nu, \overline{r}_\nu - r_\nu) = DJ^{n,\kappa}(\widetilde{r}, \overline{r} - \widetilde{r}) < 0.$$

The function $F(\alpha) := J^{n,\kappa}(r+\alpha(r^{'}-r))$ is clearly continuous on [0,1], and since the directional derivative $DJ^{n,\kappa}(r, r^{'}-r)$ is linear w.r.t. $r^{'}-r$, F is differentiable on (0,1) and has the derivative

$$F^{'}(\alpha) = DJ^{n,\kappa}(r + a(r^{'} - r), r^{'} - r).$$

By the Mean Value Theorem, we have for $\alpha \in (0,1)$

$$\begin{aligned}&J^{n,\kappa}(r_\nu + \alpha(\overline{r}_\nu - r_\nu)) - J^{n,\kappa}(r_\nu) = \\ &\alpha DJ^{n,\kappa}(r_\nu + \xi_\alpha(\overline{r}_\nu - r_\nu), \overline{r}_\nu - r_\nu) = \alpha(D + \theta_{\nu\alpha}),\end{aligned}$$

where $\theta_{\nu\alpha} \to 0$ as $\nu \to \infty$ and $\alpha \to 0$. On the other hand

$$D_\nu = DJ^{n,\kappa}(r_\nu, \overline{r}_\nu - r_\nu) = D + \theta_\nu,$$

where $\theta_\nu \to 0$ as $\nu \to \infty$. Since $b \in (0,1)$, it follows that there exist ν_0 and $\alpha_0 \in (0,1)$ such that

$$D + \theta_{\nu\alpha} \le b(D + \theta_\nu),$$

for $\nu \ge \nu_0$ and $\alpha \le \alpha_0$. Hence

$$J^{n,\kappa}(r_\nu + \alpha(\overline{r}_\nu - r_\nu)) - J^{n,\kappa}(r_\nu) \le abD_\nu,$$

for $\nu \ge \nu_0$ and $\alpha \le \alpha_0$. It follows from the choice of α_ν in Step 6, (17), that we must necessarily have $\alpha_\nu \ge c\alpha_0$ for $\nu \ge \nu_0$. Hence, by (18) and (19)

$$\begin{aligned}&J^{n,\kappa}(r_{\nu+1}) - J^{n,\kappa}(r_\nu) \le J^{n,\kappa}(r_\nu + \alpha_\nu(\overline{r}_\nu - r_\nu)) - J^{n,\kappa}(r_\nu) \le c\alpha_0 bD_\nu \\ &\le c\alpha bD/2,\end{aligned}$$

for $\nu \ge \nu_0$. Since $D < 0$, this shows that $J^{n,\kappa}(r_\nu) \to -\infty$ as $\nu \to \infty$, which contradicts the fact that $J^{n,\kappa}(r_\nu) \to J^{n,\kappa}(\tilde{r}) > -\infty$. Therefore, $D_\nu \to 0$ as $\nu \to \infty$, and this holds for the initial sequence since the limit is unique. This shows also that $\kappa, n \to \infty$ in the algorithm.
Now let $r^{'} \in R$ and $r^{'n} \to r^{'}$, with $r^{'n} \in R^n$ for every n. By Step 2 and 4, we have

$$\begin{aligned}DJ^n(r^n, r^{'n} - r^n) &= DJ_0^n(r^n, r^{'n} - r^n) \\ &+ \sum_{m=1}^{q} \lambda_m^n DJ_m^n(r^n, r^{'n} - r^n) \ge D^n, \\ &\text{for every } n.\end{aligned} \tag{20}$$

Suppose that each sequence $\{\lambda_m^n\}$, $1 \le m \le q$, is bounded, and we can suppose also that $\lambda_m^n \to \lambda_m$ (for a subsequence). We have

$$\begin{aligned}0 &= \lim_{n\to\infty} \varepsilon_m^n \lambda_m^n = \lim_{n\to\infty} J_m^n(r^n) = J_m(r), \quad 1 \le m \le p, \\ 0 &= \lim_{n\to\infty} \varepsilon_m^n \lambda_m^n = \lim_{n\to\infty} \max(0, J_m^n(r^n)) = \max(0, J_m(r)), \quad p < m \le q,\end{aligned}$$

which show that r is admissible. Passing to the limit in inequality (20), we obtain

$$DJ_0(r, r' - r) + \sum_{m=1}^{q} \lambda_m DJ_m(r, r' - r) \geq 0,$$

and this holds for every $r' \in R$. Now if $J_m(r) < 0$ for some $p < m \leq q$, then clearly $\lambda_m^n = 0$ for n sufficiently large and hence $\lambda_m = 0$. Therefore r is extremal.

If r is admissible, suppose that $\lambda_m^n \to \infty$ (for a subsequence), for some m. Dividing inequality (20) by $\max_m |\lambda_m^n|$, setting $\mu_m^n := \lambda_m^n / \max_m |\lambda_m^n|$ and passing to the limit (for a subsequence) in inequality (20), we find

$$\sum_{m=1}^{q} \mu_m DJ_m(r, r' - r) \geq 0, \quad \text{for every } r' \in R,$$

with $\max |\mu_m| = 1$, $\mu_m J_m(r) = 0$ for $p < m \leq q$ (as above), which shows that the admissible control r is abnormal extremal, i.e. contradiction. ■

Implementation of the algorithm

Dropping the index ν, the control $\overline{r} \in R^n$ in Step 2 of the algorithm can be chosen for each ν to be classical (Dirac), since $\overline{r}$ satisfies the relations

$$DJ^{n,\kappa}(r, \overline{r} - r) \leq DJ^{n,\kappa}(r, r' - r), \quad \text{for every } r' \in R^n,$$

i.e.

$$\sum_{j=0}^{N-1} \Delta t_j^n \sum_{i=1}^{M} meas\,(S_i^n) \sum_{k=1}^{L} C^{nk} H_\kappa \left(x_i^{nk}, t_j^n, y_{ij}^{nk}, z_{i,j+1}^{nk}, \overline{r}_{ij} - r'_{ij} \right) \leq 0,$$

for every $r' \in R^n$, if and only if $\overline{r}$ is classical (Dirac), i.e. $\overline{r} :\equiv \overline{w} \in W^n$, and satisfies

$$\sum_{k=1}^{L} C^{nk} H_\kappa \left(x_i^{nk}, t_j^n, y_{ij}^{nk}, z_{i,j+1}^{nk}, \overline{w}_{ij} \right)$$
$$= \min_{u \in U} \sum_{k=1}^{L} C^{nk} H_\kappa \left(x_i^{nk}, t_j^n, y_{ij}^{nk}, z_{i,j+1}^{nk}, u \right), 1 \leq i \leq M, 0 \leq j \leq N-1.$$

Clearly, by the definition of $D_\nu := DJ^{n,\kappa}(r_\nu, \overline{r}_\nu - r_\nu)$ and since $b \in (0,1)$, there exists $\alpha_\nu = c^s$ which satisfies (17).

Next, we show that the control r_ν in Step 7 can be chosen for each ν to be a Gamkrelidze control, i.e. a control of the form

$$r_{\nu ij} = \sum_{\mu=0}^{\rho} \beta_{\nu ij}^{\mu} \overline{r}_{\nu ij}^{\mu}, \quad \text{with } \beta_{\nu ij}^{\mu} \geq 0, \quad 0 \leq \mu \leq \rho, \quad \sum_{\mu=0}^{\rho} \beta_{\nu ij}^{\mu} = 1, \qquad (21)$$

where ρ is constant integer and the $\overline{r}_\nu^\mu$ are classical (Dirac) controls, provided the initial control $\overline{r}_0 := r_1$ is of the form (21) (practically a classical control). Assuming recursively that r_ν is of the form (21), we have by (18)

$$r'_{\nu+1,ij} := r_{\nu ij} + \alpha_\nu(\overline{r}_{\nu ij} - r_{\nu ij}) = \alpha_\nu \overline{r}_{\nu ij} + (1-\alpha_\nu)\sum_{\mu=0}^{\rho} \beta_{\nu ij}^{\mu} \overline{r}_{\nu ij}^{\mu}.$$

Now since the control appears only in a finite and bounded number $\rho^n \leq \rho$ of terms which depends on q and $L(n) \leq L'$ (integration rule), in the state equation and functionals, by Caratheodory's theorem we can replace the control $r'_{\nu+1}$ by an equivalent Gamkrelidze control $r_{\nu+1}$ of the form (21) which yields the same state and values of functionals as $r'_{\nu+1}$. Thus, the control $r_{\nu+1}$ satisfies Step 7.

6. Simulating classical controls

For a given sequence (or subsequence) of Gamkrelidze controls generated by the algorithm and converging to an extremal or optimal relaxed control in R, we shall now construct a sequence of simulating piecewise constant classical controls which converges to the same control in R. Let $\{r^n\}$, with $r^n := (r_{ij}^n,\ 1 \leq i \leq M,\ 0 \leq j \leq N-1) \in R^n$, $r^n \to r$ in R, be such a sequence. We can write

$$r_{ij}^n = \sum_{\mu=0}^{\rho} \beta_{ij}^{n\mu} \overline{r}_{ij}^{n\mu},$$

where the $\overline{r}_{ij}^{n\mu} :\equiv \overline{w}_{ij}^{n\mu}$ are Dirac measures. Now we can partition w.r.t. t each block $Q_{ij}^n := S_i^n \times I_j^n$ into $\rho+1$ slices $Q_{ij}^{n\mu}$ of measures $\beta_{ij}^{n\mu} meas\,(Q_{ij}^n)$, $0 \leq \mu \leq \rho$, and define the sequence of piecewise constant classical controls (considered also as elements of R)

$$w_n(x,t) := \overline{w}_{ij}^{n\mu}, \text{for } (x,t) \in \overset{o}{Q}{}_{ij}^{n\mu}, 1 \leq i \leq M, 0 \leq j \leq N-1, 0 \leq \mu \leq \rho.$$

Note that the simultaneous discretization and optimization allows us here to construct a single sequence of simulating classical controls. For similar approximations, see Chryssoverghi et al. (1993), where a double sequence is constructed, Roubíček (1991), where only discretization is considered, Roubíček (1997), and Warga (1972).

THEOREM 6.1 *The sequence $\{w_n\}$ converges to r in R.*

Proof: Since the linear combinations of functions of the form $\phi = \chi\psi$, with $\chi \in C(\overline{Q})$ and $\psi \in C(U)$, are dense in $L^1(Q, C(U))$, it is sufficient to show that

$$\lim_{n\to\infty} \int_Q \chi(x,t)(\psi(w_n) - \psi(r))dx\,dt = 0, \quad \text{for every } \chi \in C(\overline{Q}),\ \psi \in C(U).$$

Let $\{\chi_n\}$ be a sequence of functions which are constant, $\chi_n(x,t) = \chi_{nij}$, on each $\overset{o}{Q}{}^n_{ij}$, such that $\chi_n \to \chi$ in $L^\infty(Q)$ strongly. We write

$$\int_Q \chi(\psi(w_n) - \psi(r))dx\,dt := a_n + b_n + c_n,$$

where

$$\begin{aligned} a_n &:= \int_Q \chi(\psi(r^n) - \psi(r))dx\,dt, \\ b_n &:= \int_Q (\chi - \chi_n)(\psi(w_n) - \psi(r^n))dx\,dt, \\ c_n &:= \int_Q \chi_n(\psi(w_n) - \psi(r^n))dx\,dt. \end{aligned}$$

We have $a_n \to 0$ since $r^n \to r$ in R, $b_n \to 0$ since $\chi_n \to \chi$ in $L^\infty(Q)$ and ψ is bounded on U, and by construction of w_n and the $Q^{n\mu}_{ij}$

$$\begin{aligned} c_n &= \sum_{i,j} \left[\int_{Q^n_{ij}} \chi_n \psi(w_n)dx\,dt - \int_{Q^n_{ij}} \chi_n \psi(r^n)dx\,dt \right] \\ &= \sum_{i,j} \left[\sum_{\mu} meas\,(Q^{n\mu}_{ij})\chi_{nij}\psi(\overline{w}^{n\mu}_{ij}) - meas\,(Q^n_{ij})\chi_{nij} \sum_{\mu} \beta^{n\mu}_{ij} \psi(\overline{w}^{n\mu}_{ij}) \right] = 0. \end{aligned}$$

■

7. Final comments

A mixed optimization and discretization method has been described for solving an optimal control problem for systems governed by semilinear parabolic differential equations that has many advantages over methods which apply some optimization algorithm to separate discrete problems. This method can be easily extended to more general triangulations, in which case the algorithm has to be slightly modified by including some approximation of the starting control for each n. Of course, other (nonlinear) discrete schemes such as the Crank-Nicholson or the θ-method can also be considered, as well as other relaxed optimization algorithms.

The discrete relaxed method can be applied to several optimal relaxed control problems involving elliptic or hyperbolic partial differential equations, or ordinary differential equations, and also to various problems in the calculus of variations.

Finally, we wish to thank the referees for their useful suggestions.

References

CASAS, E. (1996) *The relaxation theory applied to optimal control problems of semilinear elliptic equations*, Proceedings of the 17th Conference on System Modelling and Optimization, Chapman and Hall, Prague, July 1995.

CHRYSSOVERGHI, I. (1986) Nonconvex optimal control of nonlinear monotone parabolic systems, *Systems and Control Letters*, **8**, 55-62.

CHRYSSOVERGHI, I. and BACOPOULOS, A. (1993) Approximation of relaxed nonlinear parabolic optimal control problems, *Journal of Optimization Theory and Applications*, **77**, 31-50.

CHRYSSOVERGHI, I., BACOPOULOS, A., KOKKINIS, B. and COLETSOS, I. (1997) A mixed Frank-Wolfe penalty method with applications to nonconvex optimal control problems, *Journal of Optimization Theory and Applications*, **94**, 2, 311-334.

CHRYSSOVERGHI, I., BACOPOULOS, A., COLETSOS, I. and KOKKINIS, B. (1998) Discrete approximation of nonconvex hyperbolic optimal control problems with state constraints, *Control and Cybernetics*, **27**, 1, 29-50.

CIARLET, P.G. (1978) *The Finite Element Method for Elliptic Problems*, North-Holland.

EKELAND, I. (1972) Sur le Contrôle Optimal de Systèmes Gouvernés par des Equations Elliptiques, *Journal of Functional Analysis*, **9**, 1-62.

FATTORINI, H.O. (1994) Existence theory and the maximum principle for relaxed infinite-dimensional optimal control problems, *SIAM Journal on Control and Optimization*, **32**, 311-331.

IOFFE, A.D. and TIKHOMIROV, V.M. (1979) *Theory of Extremum Problems*, North-Holland, Amsterdam.

POLAK, E. (1997) *Optimization: Algorithms and Consistent Approximations*, Springer, Berlin.

ROUBÍČEK, T. (1991) A convergent computational method for constrained optimal relaxed control problems, *Journal of Optimization Theory and Applications*, **69**, 589-603.

ROUBÍČEK, T. (1997) *Relaxation in Optimization Theory and Variational Calculus*, Walter de Gruyter, Berlin.

TEMAM, R. (1977) *Navier-Stokes Equations. Theory and Numerical Analysis*, North-Holland, New York.

TEO, K.L. (1983) Convergence of a Conditional Gradient Algorithm for Relaxed Controls Involving First Boundary-Value Problems of Parabolic Type, *Numerical Functional Analysis and Optimization*, **6**, 457-491.

THOMEE, V. (1997) *Galerkin Finite Element Methods for Parabolic Problems*, Springer, Berlin.

WARGA, J. (1972) *Optimal Control of Differential and Functional Equations*, Academic Press, New York.

WARGA, J. (1977) Steepest descent with relaxed controls, *SIAM Journal on*

Control and Optimization, **15**, 674-682.

WARGA, J. (1982) Iterative Procedures for Constrained and Unilateral Optimization Problems, *SIAM Journal on Control and Optimization*, **20**, 3, 360-376.

WILLIAMSON, L.G. and POLAK, E. (1976) Relaxed controls and the convergence of optimal control algorithms, *SIAM Journal on Control*, 14, 4, 737-757.

ZEIDLER, E. (1985-1990) *Functional Analysis and its Applications*, Springer, New York.