Mixed Discretization-Optimization Methods for Relaxed Optimal Control of Nonlinear Parabolic Systems

ION CHRYSSOVERGHI
Department of Mathematics, School of Applied Mathematics and Physics
National Technical University of Athens
Zografou Campus, 15780 Athens
GREECE

Abstract: - A nonconvex optimal control problem is considered, for systems governed by a parabolic partial differential equation, nonlinear in the state and control variables, with control and state constraints. Since this problem may have no classical solutions, it is reformulated in the relaxed form. The relaxed problem is discretized by using a finite element method in space and an implicit theta-scheme in time, while the controls are approximated by blockwise constant relaxed controls. Results are obtained on the behavior in the limit of discrete optimality, and of discrete admissibility and extremality. We then propose a penalized conditional descent method, applied to the discrete relaxed problem, and a progressively refining version of this method, applied to the continuous relaxed problem, that reduces computing time and memory. The behavior in the limit of sequences constructed by these methods is examined. Finally, numerical examples are given.

Key-Words: - Optimal control, nonlinear parabolic systems, state constraints, relaxed controls, discretization, finite elements, theta-scheme, discrete penalized conditional descent method, progressive refining.

1 Introduction

A nonconvex optimal control problem is considered, for systems governed by a parabolic partial differential equation, nonlinear in the state and control variables, with control and state constraints. Since this problem may have no classical solutions, it is reformulated in the relaxed form. The relaxed problem is discretized by using a Galerkin finite element method with continuous piecewise linear, or multilinear, basis functions in space and an implicit theta-scheme in time for space approximation, while the controls are approximated by blockwise constant relaxed controls. It is shown that, under appropriate assumptions, the relaxed accumulation points of sequences of optimal (resp. admissible and extremal) discrete relaxed controls are optimal (resp. admissible and extremal) for the continuous relaxed problem. We then propose a penalized conditional descent method, applied to the discrete relaxed problem, and a corresponding mixed discretization-optimization method, applied to the continuous relaxed problem, that progressively refines the discretization during the iterations, thus reducing computing time and memory. The result here is that the accumulation points of sequences generated by the first (resp. second) method are admissible and extremal for the discrete (resp. continuous) relaxed problem. The computed Gamkrelidze relaxed controls can then be approximated by piecewise constant classical ones using a simple procedure. Finally, numerical examples are given. For approximation and optimization methods in nonconvex optimal control and variational problems, see [1,3-8,10,11,14], and the references therein.

2 The continuous optimal control problems

Let Ω be a bounded domain in $\mathbb{R}^d$, with boundary Γ, $I := (0,T)$, $T < \infty$, an interval, and consider the semilinear parabolic state equation

(1) $y_t + A(t)y + a_0(x,t)^T \nabla y + b(x,t,y(x,t),w(x,t))$
$= f(x,t,y(x,t),w(x,t))$ in $Q := \Omega \times I$,

(2) $y(x,t) = 0$ on $\Sigma := \Gamma \times I$, $y(x,0) = y^0(x)$ in Ω,

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

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

This equation will be interpreted in the weak form

(4) $< y_t, v > + a(t,y,v) + (a_0(t)^T, \nabla y) + (b(t,y,w),v)$
$= (f(t,y,w),v)$, $\forall v \in V$, $y(t) \in V$ a.e. in I,
$y(0) = y^0$,

where $< \cdot,\cdot >$ denotes the duality bracket between the dual V^* and $V := H_0^1(\Omega)$, $(\cdot,\cdot)$ the usual inner product on $L^2(\Omega)$, and $a(t,\cdot,\cdot)$ the usual bilinear form on $V \times V$ associated with $A(t)$

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

Define the set of *classical controls*

(6) $W := \{w : Q \to U \mid w \text{ measurable}\}$,

where U is a compact, not necessarily convex, subset of $\mathbb{R}^{d'}$, and the functionals

(7) $G_m(w) := \int_Q g_m(x,t,y,w)dxdt,\ m = 0,...,q.$

The continuous classical optimal control problem is to minimize $G_0(w)$ subject to the constraints

(8) $w \in W$, $G_m(w) = 0$, $m = 1,..., p$,

(9) $G_m(w) \le 0$, $m = p+1,...,q$.

It is well known that the classical problem may have no solutions. Next, define the set of *relaxed controls* (for the relevant theory, see [13,12])

(10) $R := \{r : Q \to M_1(U) \mid r \text{ weakly measurable}\}$
$\subset L_w^\infty(Q, M(U)) \equiv L^1(Q, C(U))^*,$

where $M(U)$ (resp. $M_1(U)$) is the set of Radon (resp. probability) measures on U. The set R is endowed with the relative weak star topology of $L^1(Q,C(U))^*$. The set R is convex, metrizable and compact. If every classical control $w(\cdot)$ is identified with its associated Dirac relaxed control $r(\cdot) = \delta_{w(\cdot)}$, then W may be regarded as a subset of R, and W is thus dense in R. For $\phi \in L^1(Q,C(U))$ (or $\phi \in B(\bar{Q},U;\mathbb{R})$, where $B(\bar{Q},U;\mathbb{R})$ is the set of Caratheodory functions in the sense of Warga [13]) and $r \in R$, we shall write for simplicity

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

The continuous relaxed optimal control problem is then defined by replacing w by r (with the above notation) and W by R in the classical problem.

The results of this paper can be proved by using the techniques of [2,3,5-7,13].

We shall make the following assumptions. The boundary Γ is Lipschitz if $b = 0$; else, Γ is C^1 and $n \le 3$. The a_{ij} satisfy the ellipticity condition

(12) $\sum_{j=1}^{d}\sum_{i=1}^{d} a_{ij}(x,t)z_i z_j \ge \alpha_0 \sum_{i=1}^{d} z_i^2,\ \forall z_i, z_j \in \mathbb{R}$, in Q,

with $\alpha_0 > 0$, $a_{ij} \in L^\infty(Q)$. We have $a_0 \in L^\infty(Q)^d$, and the functions b, f, g_m, b_y, f_y, g_m, g_{my} are defined on $Q \times \mathbb{R} \times U$, measurable for fixed (y,u), continuous for fixed (x,t), and satisfy the conditions

(13) $|b(x,t,y,u)| \le \phi(x,t) + \beta y^2$, $b(x,t,y,u)y \ge 0$,

(14) $|f(x,t,y,u)| \le \psi(x,t) + \gamma|y|$,
$\forall (x,t,y,u) \in Q \times \mathbb{R} \times U$,

(15) $|f(x,t,y,u) - f(x,t,y',u)| \le L|y - y'|$,
$\forall (x,t,y,y',u) \in Q \times \mathbb{R}^2 \times U$,

(16) $b(x,t,y,u) \le b(x,t,y',u)$,
$\forall (x,t,y,y',u) \in Q \times \mathbb{R}^2 \times U$, with $y \le y'$,

(17) $|g_m(x,t,y,u)| \le \zeta_m(x,t) + \delta_m y^2$,
$|b_y(x,t,y,u)| \le \xi(x,t) + \eta|y|$, $|f_y(x,t,y,u)| \le L_1$,
$|g_{my}(x,t,y,u)| \le \zeta_{m1}(x,t) + \delta_{m1}|y|$,
$\forall (x,t,y,u) \in Q \times \mathbb{R} \times U$,

where $\phi, \psi, \xi, \zeta_{m1} \in L^2(Q)$, $\zeta_m \in L^1(Q)$, $\beta, \gamma, \eta \ge 0$, $\delta_m, \delta_{m1}, \delta_{m2} \ge 0$.

For every control $r \in R$ and $y^0 \in L^2(\Omega)$, the relaxed state equation has a unique solution $y := y_r$ such that $y \in L^2(I,V)$, $y_t \in L^2(I,V^*)$; moreover, y is essentially equal to a function in $C(\bar{I}, L^2(\Omega))$.

Theorem 1 (Continuity - Existence) The operator $r \mapsto y_r$, from R to $L^2(I,V)$, and to $L^2(I, L^4(\Omega))$ if $b \ne 0$, and the functionals $r \mapsto G_m(r)$, $m = 0,...,q$, from R to $\mathbb{R}$, are continuous. If the relaxed problem is feasible, then it has a solution.

Lemma 1 (Functional Directional Derivative) Dropping the index m in the functionals, the directional derivative of G is given, for $r, r' \in R$, by

(18) $DG(r, r'-r) := \lim_{\varepsilon \to 0^+} \frac{G(r + \varepsilon(r'-r)) - G(r)}{\varepsilon}$

(19) $= \int_Q H(x,t,y,z,r'(x,t) - r(x,t))dxdt,$

where the Hamiltonian H is defined by

(20) $H(x,t,y,z,u)$
$:= z[f(x,t,y,u) - b(x,t,y,u)] + g(x,t,y,u),$

and the adjoint state $z := z_r$ satisfies the linear adjoint equation

(21) $-< z_t, v > + a(t,v,z) + (zb_y(y,r), v)$
$= (zf_y(y,r) + g_y(y,r), v),\ \forall v \in V,\ z(t) \in V$,
a.e. in I, $z(T) = 0$,

with $y := y_r$. The mappings $r \mapsto z_r$, from R to $L^2(Q)$, and $(r, r') \mapsto DG(r, r'-r)$, from $R \times R$ to $\mathbb{R}$, are continuous.

Theorem 2 (Optimality Conditions) If $r \in R$ is optimal for *either* the relaxed *or* the classical optimal control problem, then r is *extremal*, i.e. there exist multipliers $\lambda_m \in \mathbb{R}$, $m = 0,...,q$, with $\lambda_0 \geq 0$, $\lambda_m \geq 0$, $m = p+1,...,q$, $\sum_{m=0}^{q} |\lambda_m| = 1$, such that

(22) $\sum_{m=0}^{q} \lambda_m DG_m(r, r'-r) \geq 0, \ \forall r' \in R,$

(23) $\lambda_m G_m(r) = 0, \ m = p+1,...,q$,

(transversality conditions).

The global condition (22) is equivalent to the strong relaxed pointwise minimum principle

(24) $H(x,t,y(x,t),z(x,t),r(x,t))$

$= \min_{u \in U} H(x,t,y(x,t),z(x,t),u)$, a.e. in Q,

where the complete Hamiltonian H and adjoint z are defined with $g := \sum_{m=0}^{q} \lambda_m g_m$.

3 Discretization

We suppose now that Γ is appropriately piecewise C^1 if $b = 0$, Γ is C^1 and $n \leq 3$ if $b \neq 0$, a, a_0 are independent of t (for simplicity), a is symmetric if $\theta \neq 1$ in the θ-scheme below, b, b_y, b_u, f, f_y, f_u, g_m, g_{my}, g_{mu} are continuous (possibly finitely t-piecewise) on the closure of their domains of definition, and $y^0 \in V$. For each $n \geq 0$, let Ω^n be a subdomain of Ω with polyhedral boundary Γ^n such that $\text{dist}(\Gamma^n, \Gamma) = o(h^n)$, $\{E_i^n\}_{i=1}^{M^n}$ an admissible regular quasi-uniform partition of $\bar{\Omega}^n$ into closed elements (e.g. d-simplices), with $h^n = \max_i [\text{diam}(E_i^n)] \to 0$ as $n \to \infty$, and $\{I_j^n\}_{j=1}^{N^n}$, a subdivision of the interval $\bar{I}$ into closed intervals $I_j^n = [t_{j-1}^n, t_j^n]$, of equal length Δt^n, with $\Delta t^n \to 0$ as $n \to \infty$. Define the blocks $Q_{ij}^n = E_i^n \times I_j^n$, and the subspace $V^n \subset V$ of functions that are continuous on $\bar{\Omega}$, are multilinear (linear, for d-simplices) on each E_i^n, and vanish on $\Omega - \Omega^n$. Let u_0 be any given fixed point in U. The set of *discrete classical controls* $W^n \subset W$ is the subset of classical controls that are constant on the interior of each block Q_{ij}^n and equal to u_0 on $Q - (\bar{I} \times \bar{\Omega}^n)$. The set of *discrete relaxed controls* $R^n \subset R$ is the subset of relaxed controls that are equal to a constant measure in $M_1(U)$ on the interior of each block Q_{ij}^n and equal to δ_{u_0} on $Q - (\bar{I} \times \bar{\Omega}^n)$. Clearly, we have $W^n \subset R^n$.

For a given discrete control $r^n \in R^n$, and $\theta \in [1/2, 1]$ if $b = 0$, $\theta = 1$ if $b \neq 0$, the corresponding discrete state $y^n := (y_0^n, ..., y_N^n)$ is given by the discrete state equation (implicit θ-scheme)

(25) $(1/\Delta t^n)(y_j^n - y_{j-1}^n, v) + a(y_{j\theta}^n, v)$

$+(b(t_{j\theta}^n, y_{j\theta}^n, r_j^n), v) = (f(t_{j\theta}^n, y_{j\theta}^n, r_j^n), v),$

$\forall v \in V^n, \ j = 1,...,N,$

(26) $(y_0^n - y^0, v)_1 = 0, \ \forall v \in V^n,$

(27) $y_j^n \in V^n, \ j = 1,...,N,$

(28) $y_{j\theta}^n := (1-\theta) y_{j-1}^n + \theta y_j^n, \ \ t_{j\theta}^n := (1-\theta) t_{j-1}^n + \theta t_j^n$.

If $\Delta t^n \leq c'$ (resp. $\Delta t^n \leq c'(h^n)^2$), for some c' sufficiently small, independent of n and r^n, if $b = 0$ (resp. $b \neq 0$), then, for every n and every control r^n, the discrete state equation has a unique solution y^n such that $\|y_j^n\| \leq c$, $j = 0,...,N$, with c independent of n and r^n. The solution y_j^n can be computed by the predictor-corrector method, using the linearized semi-implicit predictor scheme.

The discrete control constraint is $r^n \in R^n$ and the discrete functionals are

(29) $G_m^n(r^n) := \Delta t^n \sum_{j=0}^{N-1} \int_\Omega g_m(t_{j\theta}^n, y_{j\theta}^n, r_j^n) dx,$

$m = 0,...,q.$

The discrete state constraints are *either* of the *two* following ones

(30) Case (a) $\left| G_m^n(r^n) \right| \leq \varepsilon_m^n, \ m = 1,...,p,$

(31) Case (b) $G_m^n(r^n) = \varepsilon_m^n, \ m = 1,...,p,$

and

(32) $G_m^n(r^n) \leq \varepsilon_m^n, \ \varepsilon_m^n \geq 0, \ m = p+1,...,q,$

where the feasibility perturbations ε_m^n are given numbers converging to zero, to be defined later. The discrete cost functional to be minimized is $G_0^n(r^n)$.

Theorem 3 (Discrete Continuity - Existence) The mappings $r^n \mapsto y_j^n$ and $r^n \mapsto G_m^n(r^n)$, defined on R^n, are continuous. If any of the discrete problems is feasible, then it has a solution.

Lemma 2 (Discrete Functional Directional Derivative) Dropping m, the directional derivative of the functional G^n is given, for $r^n, r'^n \in R^n$, by

(33) $DG^n(r^n, r'^n - r^n)$

$= \Delta t^n \sum_{j=0}^{N-1} \int_\Omega H(t_{j\theta}^n, y_{j\theta}^n, z_{j,1-\theta}^n, r'^n_j - r_j^n) dx$, where the discrete adjoint z^n is given by the linear adjoint scheme

(34) $-(1/\Delta t^n)(z_j^n - z_{j-1}^n, v)$
$+a(v, z_{j,1-\theta}^n) + (z_{j,1-\theta}^n b_y(t_{j\theta}^n, y_{j\theta}^n, r_j^n), v)$
$= (z_{j,1-\theta}^n f_y(t_{j\theta}^n, y_{j\theta}^n, r_j^n) + g_y(t_{j\theta}^n, y_{j\theta}^n, r_j^n), v),$
$\forall v \in V^n,\ j = N,...,1,\ z_N^n = 0,\ z_j^n \in V^n,$

which has a unique solution z_{j-1}^n for each j, for Δt^n sufficiently small. Moreover, the mappings $r^n \mapsto z^n$ and $(r^n, r'^n) \mapsto DG^n(r^n, r'^n - r^n)$ are continuous.

Theorem 4 (Discrete Optimality Conditions) If $r^n \in R^n$ is optimal for the discrete problem (state constraints Case (b)), then r^n is *extremal*, i.e. there exist multipliers $\lambda_m^n \in \mathbb{R}$, $m = 0,...,q$, with $\lambda_m^n \ge 0$, $\lambda_m^n \ge 0$, $m = p+1,...,q$, $\sum_{m=0}^{q} |\lambda_m^n| = 1$, such that

(35) $\sum_{m=0}^{q} \lambda_m^n DG_m^n(r^n, r'^n - r^n)$
$= \Delta t^n \sum_{j=1}^{N} \int_\Omega H^n(t_{j\theta}^n, y_{j\theta}^n, \nabla y_{j\theta}^n, z_{j,1-\theta}^n, r'^n_j - r_j^n) dx \ge 0,$
$\forall r'^n \in R^n,$

(36) $\lambda_m^n [G_m(r^n) - \varepsilon_m^n] = 0,\ m = p+1,...,q,$

where H^n and z^n are defined with $g := \sum_{m=0}^{q} \lambda_m^n g_m$.

The condition (35) is equivalent to the strong discrete blockwise minimum principle

(37) $\int_\Omega H^n(t_j^n, y_{j\theta}^n, \nabla y_{j\theta}^n, z_{j,1-\theta}^n, r_{ij}^n) dx$
$= \min_{u \in U} \int_\Omega H^n(t_{j\theta}^n, y_{j\theta}^n, \nabla y_{j\theta}^n, z_{j,1-\theta}^n, u) dx,$
$i = 1,...,M,\ j = 1,...,N.$

4 Behavior in the limit

Proposition 1 (Control Approximation) For every $r \in R$, there exists a sequence $(w^n \in W^n \subset R)$ that converges to r in R.

Lemma 3 (Stability) If Δt is sufficiently small, for every $r^n \in R^n$, the following inequalities hold, where the constants c are independent of n and r^n

(38) $\|y_k^n\| \le c,\ k = 0,...,N,\ \sum_{j=1}^{N} \|y_j^n - y_{j-1}^n\|^2 \le c,$

(39) $\Delta t^n \sum_{j=1}^{N} \|y_{j\theta}^n\|_1^2 \le c,$

(40) $\Delta t^n \sum_{j=0}^{N} \|y_j^n\|_1^2 \le c$ (under the condition $\Delta t^n \le C(h^n)^2$, for some constant C independent of n, if $\theta = 1/2$),

(41) $\Delta t^n \sum_{j=1}^{N} \|y_j - y_{j-1}\|_1^2 \le c$ (under the condition $\Delta t^n \le C(h^n)^2$).

For given values $v_0,...,v_N$ in a vector space, define the piecewise constant and continuous piecewise linear functions

(42) $v_-(t) := v_{j-1},\ v_+(t) := v_j,$

(43) $v_\theta(t) := (1-\theta)v_{j-1} + \theta v_j,\ t \in \overset{o}{I_j^n},$

(44) $v_\wedge(t) := v_{j-1} + \frac{t - t_{j-1}^n}{\Delta t^n}(v_j - v_{j-1}),\ t \in \overset{o}{I_j^n}.$

If $b = 0$ (resp. $b \ne 0$), we suppose in what follows that $\Delta t^n \le C$ (resp. $\Delta t^n \le C(h^n)^2$), with C sufficiently small.

Lemma 4 (Consistency) (i) If $r^n \to r$ in R, then the corresponding discrete states $y_\wedge^n, y_+^n, y_+^n, y_+^n$ converge to y_r in $L^2(I, L^4(\Omega))$ (resp $L^2(Q)$) strongly if $b \ne 0$ (resp. $b = 0$), $y_\theta^n \to y_r$ in $L^2(I, V)$ strongly, and $\lim_{n\to\infty} G_m^n(r^n) = G_m(r)$, $m = 0,...,q.$

(ii) If $r^n \to r$ in R, then the corresponding discrete adjoint states $z_-^n, z_+^n, z_{1-\theta}^n, z_\wedge^n$ converge to z_r in $L^2(I, L^4(\Omega))$ (resp. $L^2(Q)$) strongly if $b \ne 0$ (resp. $b = 0$), and $z_{1-\theta}^n \to z_r$ in $L^2(I, V)$ strongly. If $r^n \to r$ and $r'^n \to r'$, then, for $m = 0,...,q$

(45) $\lim_{n\to\infty} DG_m^n(r^n, r'^n - r^n) = DG_m(r, r' - r).$

In what follows, we suppose that the continuous relaxed problem is feasible. The following theorem addresses the behavior in the limit of optimal discrete controls.

Theorem 5 In the presence of state constraints, suppose that the sequences (ε_m^n) in the discrete state constraints, Case (a), converge to zero and satisfy

(46) $|G_m^n(\tilde{r}^n)| \le \varepsilon_m^n,\ m = 1,..., p,$

(47) $G_m^n(\tilde{r}^n) \le \varepsilon_m^n,\ \varepsilon_m^n \ge 0,\ m = p+1,...,q,$

for every n, where $(\tilde{r}^n \in R^n)$ is a sequence converging in R to an optimal control $\tilde{r} \in R$ of the relaxed problem. For each n, let r^n be optimal for the discrete problem, Case (a). Then every relaxed accumulation point of (r^n) is optimal for the continuous relaxed problem.

Next, we examine the behavior in the limit of extremal discrete controls. Consider the discrete problem with state constraints, Case (b). Sequences of perturbations (ε_m^n), converging to zero and such that the discrete problem is feasible for every n, can be constructed as follows. Let $r'^n \in R^n$ be any solution of the problem without state constraints

(48) $c^n := \min_{r^n \in R^n} \{ \sum_1^p [G_m^n(r^n)]^2 + \sum_{p+1}^q [\max(0, G_m^n(r^n))]^2 \}$,

and set

(49) $\varepsilon_m^n := G_m^n(r'^n)$, $m = 1,...,p$,

(50) $\varepsilon_m^n := \max(0, G_m^n(r'^n))$, $m = p+1,...,q$.

It can be easily shown that $c^n \to 0$, hence $\varepsilon_m^n \to 0$, $m = 1,...,q$. Then clearly the discrete problem, Case (b), for these ε_m^n, is feasible for every n.

In what follows, we suppose that the ε_m^n are chosen as in the above minimum feasibility procedure.

Theorem 6 For each n, let r^n be admissible and extremal for the discrete problem, Case (b). Then every accumulation point of (r^n) in R is admissible and extremal for the continuous relaxed problem.

5 Discrete penalized descent methods

Let (M_m^l), $m = 1,...,q$, be positive increasing sequences such that $M_m^l \to \infty$ as $l \to \infty$, and define the penalized discrete functionals

(51) $G^{nl}(r^n) := G_0^n(r^n) + \{ \sum_{m=1}^p M_m^l [G_m^n(r^n)]^2$

$+ \sum_{m=p+1}^q M_m^l [\max(0, G_m^n(r^n))]^2 \} / 2$.

Let $b', c' \in (0,1)$, and let (β^l), (ζ_k) be positive sequences, with (β^l) decreasing and converging to zero, and $\zeta_k \le 1$. The algorithm described below contains two versions. In the progressively refining version, we suppose that the (possibly) finer discretization for $n+1$ is defined by subdividing the elements E_i^n into subelements and by slightly (up to $o(h^n)$, as Γ is C^1 or piecewise C^1) transforming the resulting boundary elements, if necessary, so as to fit Γ^{n+1}, and then by setting $N^{n+1} = \kappa N^n$, for some integer $\kappa \ge 2$. The discrete relaxed penalized conditional descent methods are described by the following algorithm.

Algorithm

Step 1. Set $k := 0$, $l := 1$, choose an n and $r_0^{nl} \in R^n$.

Step 2. Find $r_k^{nl} \in R^n$ such that

(52) $d_k := DG^{nl}(r_k^{nl}, \overline{r}_k^{nl} - r_k^{nl})$

$= \min_{r'^n \in R^n} DG^{nl}(r_k^{nl}, r'^n - r_k^{nl})$.

Step 3. If $|d_k| > \beta^l$, go to Step 4.

Else, set $r^{nl} := r_k^{nl}$, $\overline{r}^{nl} := \overline{r}_k^{nl}$, $d^l := d_k$, and:

Version A: Set $r_k^{n,l+1} := r_k^{nl}$.

Version B: If the discretization for $n+1$ is finer, set $\tilde{r}_k := r_k^{nl}$ and define $r_k^{n+1,l+1}$ as the modified control resulting from $\tilde{r}_k$ after the slight transformation in the construction of the new boundary elements E_i^{n+1}; else, set $r_k^{n+1,l+1} := r_k^{nl}$. Set $n := n+1$.

In both versions, set $l := l+1$ and go to Step 2.

Step 4. (Modified Armijo Step Search) Find the lowest integer value $s \in \mathbb{Z}$, say $\overline{s}$, such that $\alpha(s) := c'^s \zeta_k \in (0,1]$ and $\alpha(s)$ satisfies

(53) $G^n(w_k^n + \alpha(s)(v_k^n - w_k^n)) - G^n(w_k^n) \le \alpha(s) b' d_k$,

and then set $\alpha_k := \alpha(\overline{s})$.

Step 5. Choose any $r_{k+1}^{nl} \in R^n$ such that

(54) $G^{nl}(r_{k+1}^{nl}) \le G^{nl}(r_k^{nl} + \alpha(s)(\overline{r}_k^{nl} - r_k^{nl}))$,

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

This Algorithm contains two versions:

Version A: n is a constant integer chosen in Step 1, i.e. a *fixed discretization* is chosen and the G_m^n, $m = 1,...,q$, are replaced by $\tilde{G}_m^n := G_m^n - \varepsilon_m^n$.

Version B: This is a *progressively refining* discrete method, i.e. $n \to \infty$, in which case we can set $n = 1$ in Step 1, hence $n = l$ in the Algorithm.

Version B has the advantage of reducing computing time and memory, and also of avoiding the computation of minimum feasibility perturbations ε_m^n. It is justified by the fact that finer discretizations become progressively more essential as the iterate gets closer to an extremal control.

One can see that a *classical* control $\overline{r}_k^{nl}$ can be found in Step 2 by minimizing in $u \in U$ the integral (practically using a numerical integration rule)

(55) $\int_{E_i^n} H(t_{j\theta}^n, x, y_{j\theta}^{nl}, z_{j,1-\theta}^{nl}, u) dxdt$

independently for each $i = 1,...,M$, $j = 1,...,N$. On the other hand, since clearly $d_k \le 0$ and $b' \in (0,1)$, by the definition of the directional derivative the Armijo step α_k in Step 4 can be found for every k.

With r^{nl} as defined in Step 3, define the sequences of multipliers

(56) $\lambda_m^{nl} := M_m^l G_m^n(r^{nl}), \ \ m = 1,..., p,$

(57) $\lambda_m^{nl} := M_m^l \max(0, G_m^n(r^{nl})), \ \ m = p+1,..., q,$

Theorem 7 (i) In Version B, let (r^{nl}) be a subsequence, regarded as a sequence in R, of the sequence generated by the Algorithm in Step 3 that converges to some r in R, as $n,l \to \infty$. If the sequences (λ_m^{nl}) are bounded, then r is admissible and extremal for the continuous relaxed problem.

(ii) In Version A, let (r^{nl}), n fixed, be a subsequence of the sequence generated by the Algorithm in Step 3 that converges to some $r^n \in R^n$ as $l \to \infty$. If the sequences (λ_m^{nl}) are bounded, then r^n is admissible and extremal for the discrete problem.

Implementation of the Algorithm

The Algorithm can be practically implemented as follows. Suppose that the integrals on Ω involved in the discrete equations and the functionals are numerically calculated using an s-node integration rule. Choosing the initial discrete control r_0^{n1} in Step 1 to be of Gamkrelidze type, i.e. equal on each block Q_{ij}^n to a convex combination of $(s+q+1)+1$ Dirac measures on U concentrated at $(s+q+1)+1$ points of U, and since the control $\overline{r}_k^{nl}$ in Step 2 is chosen to be classical (see above), it can then be shown by induction that the control r_k^{nl} computed in the Algorithm is also of Gamkrelidze type, for every k. A discrete Gamkrelidze control $r^{nl} := r_k^{nl}$ (Step 3) thus computed can then be approximated, and practically simulated, by a classical one w^{nl} by subdividing each interval I_j^n into subintervals $I_{j\mu}^n$ of lengths proportional to the Gamkrelidze coefficients of r^{nl}, and then defining w^{nl} to be equal, successively on the sub-blocks $E_i^n \times I_{j\mu}^n$, for each fixed i, j, to the classical controls defining r^{nl} (for more details, see [6]).

6 Numerical examples

Example 1. Let $\Omega := I := (0,1)$. Define the functions

(58) $\overline{w}(x) := \begin{cases} 1, \text{ if } \ 0 \le t \le 0.5 \\ 1 - 2(t-0.5)(0.2x+0.4), \text{ if } \ 0.5 < t \le 1 \end{cases}$

(59) $\overline{y}(x) := x(1-x)e^{-t}$,

and consider the problem with state equation

(60) $y_t - y_{xx} + 0.5y|y| + (1 + w - \overline{w})y$
$= 0.5\,\overline{y}|\overline{y}| + \overline{y} + [-x(1-x)+2]e^{-t}$
$+ \sin y - \sin \overline{y} + 3(w - \overline{w})$ in Q,

(61) $y(x,t) = 0$ on Σ, $y(0,x) = x(1-x)$ in Ω,

nonconvex control constraint set

(62) $U := [0, 0.25] \cup [0.75, 1]$

(or $U := \{0,1\}$, two values, on/off type control),

and nonconvex cost functional to be minimized

(63) $G_0(u) := \int_Q [0.5(y-\overline{y})^2 - (w-0.5)^2 + 0.25] dxdt.$

One can easily verify that the unique optimal relaxed control r is given by

(64) $r(x,t)\{1\} := \overline{w}(x,t)$, $r(x,t)\{0\} := 1 - r(x,t)\{1\}$,

with optimal state $\overline{y}$ and cost 0, and we see that r is concentrated at the two points 1 and 0; r is classical for $t \in (0, 0.5)$, and non-classical otherwise. Note also that the optimal cost 0 can be approximated as closely as desired by using a classical control, as W is dense in R, but clearly cannot be attained for such a control. The Algorithm, without penalties (no index l), was applied to this problem using the midpoint integration rule on each interval E_i^n, with step sizes $h = \Delta t = 1/96$, $\theta = 1$, and $b' = c' = 0.5$. After 90 iterations in k, we obtained the results

(65) $G_0^n(r_k^n) = 3.204 \cdot 10^{-6}$, $d_k = -1.020 \cdot 10^{-5}$,
$\eta_k = 5.752 \cdot 10^{-3}$,

where d_k was defined in the Algorithm and η_k is the discrete state maximum error at the points $(ih, j\Delta t)$. Fig.1 shows the last relaxed control probability function $p_1(x,t) := r_k^n(x,t)\{1\}$.

Example 2. Introducing the state constraint

(66) $G_1(w) := \int_Q y\, dxdt = 0,$

in Example 1, and applying here the penalized Algorithm with the same parameters, we obtained after 180 iterations in k the results

(67) $G_0^n(r_k^{nl}) = 6.788323790 \cdot 10^{-3}$,
$G_1^n(r_k^{nl}) = 7.607 \cdot 10^{-7}$, $d_k = -1.388 \cdot 10^{-3}.$

Fig.2 shows the last relaxed control probability $p_1(x,t) := r_k^{nl}(x,t)\{1\}$ and Fig.3 shows the last state.

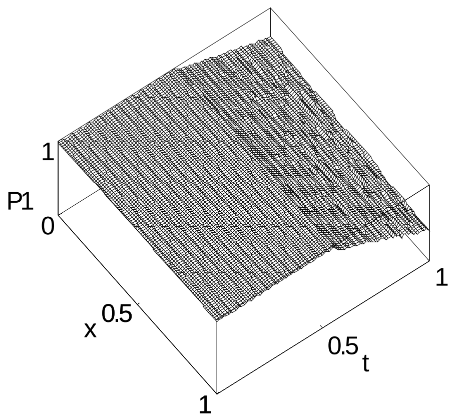

Fig.1

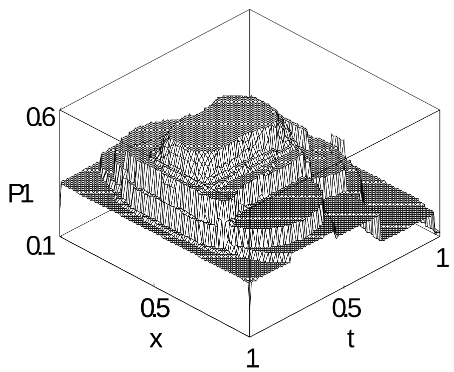

Fig.2

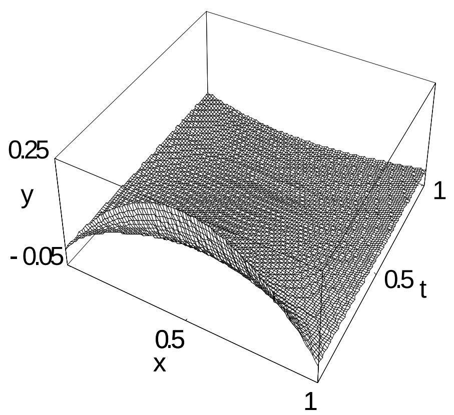

Fig.3

Finally, the progressively refining version of the algorithm was also applied to the above problems, with successive step sizes $h = \Delta t = 1/24, 1/48, 1/96$, in 3 equal iteration periods, and yielded results of similar accuracy, but required here less than half the computing time. Similar results were also obtained using the simulation of the last computed relaxed controls by classical ones (see end of Section 5).

References:

[1] C. Cartensen and T. Roubíček, Numerical Approximation of Young Measure in Nonconvex Variational Problems, *Numerische Mathematik*, Vol. 84, 2000, pp. 395-415.

[2] I. Chryssoverghi, Nonconvex Optimal Control of Nonlinear Monotone Parabolic Systems, *Systems and Control Letters*, Vol. 8, pp. 55-62, 1986.

[3] I. Chryssoverghi and A. Bacopoulos, Approximation of Relaxed Nonlinear Parabolic Optimal Control Problems, *Journal of Optimization Theory and Applications*, Vol. 77, 1993, pp. 31-50.

[4] I. Chryssoverghi, A. Bacopoulos, B. Kokkinis and J. Coletsos, Mixed Frank-Wolfe Penalty Method with Applications to Nonconvex Optimal Control Problems, *J. Optimization Theory and Applications*, Vol. 94, 1997, pp. 311-334.

[5] I. Chryssoverghi, A. Bacopoulos, J. Coletsos and B. Kokkinis, Discrete Approximation of Nonconvex Hyperbolic Optimal Control Problems with State Constraints, *Control and Cybernetics*, Vol. 27, 1998, pp. 29-50.

[6] I. Chryssoverghi, J. Coletsos and B. Kokkinis, Discrete Relaxed Method for Semilinear Parabolic Optimal Control Problems, *Control and Cybernetics*, Vol. 28, 1999, pp. 157-176.

[7] I. Chryssoverghi, J., Coletsos, and B., Kokkinis, Discretization Methods for Nonconvex Optimal Control Problems with State Constraints, *Numerical Functional Analysis and Optimization*, Vol. 26, No. 3, 2005.

[8] I. Chryssoverghi, Discretization Methods for Semilinear Parabolic Optimal Control Problems, *International Journal of Numerical Analysis and Modeling*, Vol. 3, 2006, pp. 437-458.

[9] H.O. Fattorini, *Infinite Dimensional Optimiza-tion Theory and Optimal Control*, Cambridge University Press, 1999.

[10] J. Mach, Numerical Solution of a Class of Nonconvex Variational Problems by SQP, *Numerical Functional Analysis and Optimization*, Vol. 23, 2002, pp. 573-587.

[11] A.-M. Mataché, T. Roubíček and C. Schwab, Higher-order Convex Approximations of Young Measures in Optimal Control, *Advances in Comput. Mathematics*, Vol. 19, 2003, pp. 73-79.

[12] T. Roubíček, *Relaxation in Optimization Theory and Variational Calculus*, Walter de Gruyter, 1997.

[13] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, 1972.

[14] J. Warga, Steepest Descent with Relaxed Controls, *SIAM J. Control and Optimization*, Vol. 15, 1977, pp. 674-682.