Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Progressively refining penalized gradient projection method for semilinear parabolic optimal control problems*

Ion Chryssoverghi [1], Jürgen Geiser[2] and Jamil Al-Hawasy[1]

submitted: 9th November 2005

[1] Department of Mathemarics,
School of Applied Mathematics and Physics,
National Technical University of Athens
Zografou Campus
15780 Athens
Greek
E-Mail: ichris@central.ntua.gr

[2] Weierstrass Institute
for Applied Analysis
and Stochastics
Mohrenstrasse 39
D-10117 Berlin
Germany
E-Mail: geiser@wias-berlin.de

No. 1065
Berlin 2005

2000 *Mathematics Subject Classification.* 93C83 74S05 35K20 47A05.

Key words and phrases. Optimal Control, Semilinear Parabolic Systems, Discretisation, Finite Element Method, Theta-Scheme, Discrete Penalized Gradient Projection Methods.

2003 *Physics Abstract Classifications.* 02.60.Cb 44.05.+e.

*This work has been supported by the DFG Research Center MATHEON – "Mathematics for key technologies" (FZT 86) in Berlin.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax:	+ 49 30 2044975
E-Mail:	preprint@wias-berlin.de
World Wide Web:	http://www.wias-berlin.de/

Abstract

We consider an optimal control problem defined by semilinear parabolic partial differential equations, with control and state constraints, where the state constraints and cost functional involve also the state gradient. The problem is discretized by using a finite element method in space and an implicit θ-scheme in time for state approximation, while the controls are approximated by blockwise constant ones. We propose a discrete penalized gradient projection method, which is applied to the continuous problem and progressively refines the discretization during the iterations, thus reducing computing time and memory. We prove that strong accumulation points in L^2 of sequences generated by this method are admissible and weakly extremal for the continuous problem. Finally, numerical examples are given.

1 Introduction

We consider an optimal distributed control problem for systems described by a semilinear parabolic boundary value problem, with control and state constraints, where the state constraints and cost functional involve also the state gradient. The problem is discretized by using a Galerkin finite element method with continuous elementwise linear basis functions in space and an implicit θ-scheme in time for state approximation, while the controls are approximated by blockwise constant ones. We first state the weak necessary conditions for optimality for the continuous problem. We then propose a discrete penalized gradient projection method, which is applied to the continuous problem and progressively refines the discretization during the iterations, thus reducing computing time and memory. We prove that strong accumulation points in L^2 (if they exist) of sequences of discrete controls generated by this method are admissible and weakly extremal for the continuous problem. Finally, numerical examples are given. For discretization and optimization methods applied to distributed optimal control problems, see e.g. [1-7], [13-15] and the references therein.

2 The continuous optimal control problems

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

(2.1) $\quad y_t + A(t)y = f(x,t,y(x,t),w(x,t)) \quad \text{in } Q = \Omega \times I,$

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

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

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

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

The constraints on the control are $w(x,t) \in U$ in Q, where U is a convex and compact subset of $\mathbb{R}^s$, the state constraints are

(2.5) $\quad G_m(w) := \int_Q g_m(x,t,y,\nabla y,w)dxdt = 0, \quad m = 1,...,p,$

(2.6) $\quad G_m(w) := \int_Q g_m(x,t,y,\nabla y,w)dxdt \le 0, \quad m = p+1,...,q,$

and the cost functional is

(2.7) $\quad G_0(w) := \int_Q g_0(x,t,y,\nabla y,w)dxdt.$

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

We define the set of controls

(2.8) $\quad W := \{w \in L^2(Q,\mathbb{R}^s) \,|\, w : Q \to U\}$,

endowed with the relative strong topology of $L^2(Q,\mathbb{R}^s)$. We denote by $|\cdot|$ the Euclidean norm in $\mathbb{R}^n$, by $(\cdot,\cdot)$ and $\|\cdot\|$ the inner product and norm in $L^2(\Omega,\mathbb{R}^n)$, by $(\cdot,\cdot)_Q$ and $\|\cdot\|_Q$ the inner product and norm in $L^2(Q,\mathbb{R}^n)$, by $(\cdot,\cdot)_1$ and $\|\cdot\|_1$ the inner product and norm in the Sobolev space $V := H_0^1(\Omega)$, and by $<\cdot,\cdot>$ the duality bracket between the dual $V^* = H^{-1}(\Omega)$ and V. We also define the usual bilinear form associated with $A(t)$ and defined on $V \times V$

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

In the following, we shall make various assumptions on the data.

Assumption 2.1 The coefficients a_{ij} satisfy the ellipticity conditions

(2.10) $\quad \sum_{j=1}^{d}\sum_{i=1}^{d}a_{ij}(x,t)z_i z_j \ge \alpha\sum_{i=1}^{d}z_i^2, \quad \forall z_i \in \mathbb{R},$ a.e. in $Q,$

with $\alpha > 0$, $a_{ij} \in L^\infty(Q)$.

Assumption 2.1 implies that

(2.11) $\quad |a(t,y,v)| \le \alpha_1 \|y\|_1 \|v\|_1, \quad a(t,v,v) \ge \alpha_2 \|v\|_1^2, \quad t \in I, \ \ y \in V, \ v \in V,$

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

Assumption 2.2 The function f is defined on $Q \times \mathbb{R} \times U$, measurable for fixed y,u, continuous for fixed x,t (Caratheodory function), and satisfies the following condition

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

with $\psi \in L^2(Q)$, $\beta \ge 0$, and the Lipschitz condition

(2.13) $\quad |f(x,t,y_1,u) - f(x,t,y_2,u)| \le L|y_1 - y_2|, \quad \forall (x,t,y_1,y_2,u) \in Q \times \mathbb{R}^2 \times U.$

The state equation will be interpreted in the following weak form

(2.14) $\quad < y_t, v > + a(t,y,v) = (f(t,y,w),v), \quad \forall v \in V,$ a.e. in $I,$

(2.15) $\quad y(t) \in V$ a.e. in $I, \quad y(0) = y^0.$

The following result is classical.

Proposition 2.1 Under Assumptions 2.1-2, for every control $w \in W$ and $y^0 \in L^2(\Omega)$, the state equation has a unique solution $y = y_w$ such that $y \in L^2(I,V)$ and

$y_t \in L^2(I,V^*)$. Moreover, y is essentially equal to a function in $C(\bar{I}, L^2(\Omega))$, and thus the initial condition is well defined.

Assumption 2.3 The functions g_m, $m=0,...,q$, are defined on $Q\times\mathbb{R}^{d+1}\times U$, measurable for fixed $y,\bar{y},u$, continuous for fixed x,t, and satisfy

(2.16) $\quad |g_m(x,t,y,y',u)| \le \zeta_m(x,t) + \eta_m y^2 + \eta'_m |y'|^2, \quad \forall (x,t,y,y',u)\in Q\times\mathbb{R}^{d+1}\times U,$

with $\zeta_m \in L^1(Q)$, $\eta_m \ge 0$, $\eta'_m \ge 0$.

The following results are proved by using the techniques of [4], [6], [16] (see also [9]).

Lemma 2.1 Under Assumptions 2.1-2, the operator $w \mapsto y_w$, from W to $L^2(I,V)$, is continuous, and under Assumptions 2.1-3, the functionals $w\mapsto G_m(w)$, $m=0,...,q$, defined on W, are continuous.

Note that the above continuous optimal control problem may have no classical solutions. The existence of an optimal control can be proved under some convexity assumptions on the data (Cesari property). For nonconvex problems, where these (not realistic, if f in nonlinear w.r.t. u) assumptions are avoided, and the relevant relaxation theory, see [3-7], [9].

Assumption 2.4 The functions f, f_y, f_u (resp. g_{my}, g_{mu}) are defined on $Q\times\mathbb{R}\times\tilde{U}$ (resp. $Q\times\mathbb{R}^{d+1}\times\tilde{U}$), where $\tilde{U}$ is an open set containing the compact set U, measurable on Q for fixed $(y,u)\in\mathbb{R}\times U$ (resp. $(y,y',u)\in\mathbb{R}^{d+1}\times U$) and continuous on $\mathbb{R}\times U$ (resp. $\mathbb{R}^{d+1}\times U$) for fixed $(x,t)\in Q$, and satisfy

(2.17) $\quad |g_{my}(x,t,y,y',u)| \le \zeta_{m1}(x,t) + \eta_{m1}|y| + \eta'_{m1}|y'|, \quad \forall (x,t,y,y',u)\in Q\times\mathbb{R}^{d+1}\times U,$

(2.18) $\quad |g_{my'}(x,t,y,y',u)| \le \zeta_{m2}(x,t) + \eta_{m2}|y| + \eta'_{m2}|y'|, \quad \forall (x,t,y,y',u)\in Q\times\mathbb{R}^{d+1}\times U,$

(2.19) $\quad |g_{mu}(x,t,y,y',u)| \le \zeta_{m3}(x,t) + \eta_{m3}|y| + \eta'_{m3}|y'|, \quad \forall (x,t,y,y',u)\in Q\times\mathbb{R}^{d+1}\times U,$

with $\zeta_{m1},\zeta_{m2},\zeta_{m3} \in L^2(Q)$, $\eta_{m1},\eta_{m2},\eta'_{m1},\eta'_{m2},\eta'_{m3} \ge 0$, and

(2.20) $\quad |f_y(x,t,y,u)| \le L_1, \quad \forall (x,t,y,u)\in Q\times\mathbb{R}\times U,$

(2.21) $\quad |f_u(x,t,y,u)| \le \zeta(x,t) + \eta|y|, \quad \forall (x,t,y,u)\in Q\times\mathbb{R}\times U,$

with $\zeta\in L^2(Q)$, $\eta \ge 0$.

Lemma 2.2 We drop the index m in G_m, g_m. Under Assumptions 2.1-4, for $w,w'\in W$, the directional derivative of G is given by

(2.22) $$DG(w,\bar{w}-w) := \lim_{\varepsilon\to 0^+} \frac{G(w+\varepsilon(\bar{w}-w)) - G(w)}{\varepsilon}$$
$$= \int_Q H_u(x,t,y,\nabla y, z, w)(\bar{w}-w)\,dxdt,$$

where the Hamiltonian is defined by

(2.23) $\quad H(x,t,y,y',z,u) := zf(x,t,y,u) + g(x,t,y,y',u),$

and the adjoint $z := z_w$ satisfies the adjoint equation

(2.24) $\quad -< z_t, v > +a(t,v,z) = (zf_y(t,y,w) + g_y(t,y,w), v) + (g_{y'}(t,y,\nabla y,w), \nabla v),$

$\forall v \in V$, a.e. in I, with $y := y_w$,

(2.25) $\quad z(t) \in V$ a.e. in I, $z(T) = 0.$

The mappings $w \mapsto z_w$, from W to $L^2(Q)$, and $(w,\overline{w}) \mapsto DG(w,\overline{w}-w)$, from $W \times W$ to $\mathbb{R}$, are continuous.

Theorem 2.1 Under Assumptions 2.1-4, if $w \in W$ is optimal for the control problem, then w is *weakly extremal,* i.e. there exist multipliers $\lambda_m \in \mathbb{R}$, $m = 0,...,q$, with

(2.26) $\quad \lambda_0 \ge 0$, $\lambda_m \ge 0$, $m = p+1,...,q$, $\sum_{m=0}^{q} |\lambda_m| = 1$,

such that

(2.27) $\quad \sum_{m=0}^{q} \lambda_m DG_m(w,\overline{w}-w) \ge 0, \quad \forall \overline{w} \in W$,

and

(2.28) $\quad \lambda_m G_m(w) = 0, \quad m = p+1,...,q$ (*transversality conditions*).

The global condition (2.27) is equivalent to the *weak pointwise minimum principle*

(2.29) $\quad H_u(x,t,y,\nabla y,z,w(x,t))w(x,t) = \min_{u \in U} H_u(x,t,y,\nabla y,z,w(x,t))u,$ a.e. in Q,

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

3 The discrete optimal control problem

In the sequel, we suppose that the domain Ω is a polyhedron (for simplicity), $a(t,u,v)$ is independent of t and symmetric, the functions f, f_y, f_u, g_m, g_{my}, $g_{my'}$, g_{mu} are continuous in all their arguments (possibly finitely piecewise in t), the functions $\psi, \zeta_m, \eta_m, \eta'_m$, $\zeta_{m1}, \zeta_{m2}, \zeta_{m3}$, $\eta_{m1}, \eta_{m2}, \eta_{m3}$, $\eta'_{m1}, \eta'_{m2}, \eta'_{m3}$ are constant, and $y^0 \in V := H_0^1(\Omega)$. For each integer $n \ge 0$, let $\{E_i^n\}_{i=1}^{M^n}$ be an admissible regular quasi-uniform triangulation of $\overline{\Omega}$ into closed elements (e.g. d-simplices), with $h^n = \max_i[\mathrm{diam}(E_i^n)] \to 0$ as $n \to \infty$, and $\{I_j^n\}_{j=1}^{N^n}$, a subdivision of the interval $\overline{I}$ 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 $n \to \infty$. We define the *blocks* $Q_{ij}^n := E_i^n \times I_j^n$. Let V^n be the subspace of functions that are continuous on $\overline{\Omega}$ and linear (or multilinear) on each E_i^n. We define the set of (blockwise constant) *discrete controls*

(3.1) $\quad W^n := \{w^n \in W \mid w^n(x,t) = w_{ij}^n, \text{ on } \overset{o}{Q}{}_{ij}^n\}$,

endowed with the relative (Euclidean here) topology of W.

Remark. For implementation reasons, we could alternatively use a coarser partition for the discrete controls, that is, use discrete controls that are constant on hyperblocks

$Q'^{n}_{i'j'} := E'^{n}_{i'} \times I'^{n}_{j'}$, where the $E'^{n}_{i'}$ are appropriate unions of some elements E^n_i and $I'^{n}_{j'}$ are appropriate unions of some intervals I^n_j.

For a given discrete control $w^n := (w^n_0, ..., w^n_{N-1}) \in W^n$, with $w^n_j := (w^n_{0j}, ..., w^n_{Mj})$, and $\theta \in [1/2, 1]$, the corresponding discrete state $y^n := (y^n_0, ..., y^n_N)$ is given by the following discrete state equation (implicit θ-scheme)

(3.2) $\quad (1/\Delta t^n)(y^n_j - y^n_{j-1}, v) + a(y^n_{j\theta}, v) = (f(t^n_{j\theta}, y^n_{j\theta}, w^n_j), v),$

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

(3.3) $\quad (y^n_0 - y^0, v)_1 = 0, \quad$ for every $v \in V^n$, $\quad y^n_j \in V^n$, $j = 1, ..., N$,

where we set $y^n_{j\theta} := (1-\theta) y^n_{j-1} + \theta y^n_j$, $t^n_{j\theta} := (1-\theta) t^n_{j-1} + \theta t^n_j$. The discrete control constraints are $w^n \in W^n$ and the discrete functionals

(3.4) $$G^n_m(w^n) := \Delta t^n \sum_{j=0}^{N-1} \int_\Omega g_m(x, t^n_{j\theta}, y^n_{j\theta}, \nabla y^n_{j\theta}, w^n_j) dx.$$

Under Assumptions 2.1-2, 3.1, for Δt^n sufficiently small, depending on the Lipschitz constant L of f, and for each j, the above θ-scheme has a unique solution y^n_j, which can be computed by the standard predictor-corrector method, where regular linear systems are involved, and where the corrector scheme is contractive.

Lemma 3.1 Under Assumptions 2.1-3, the mappings $w^n \mapsto y^n_j$ and $w^n \mapsto G^n_m(w^n)$, $m = 0, ..., q$, defined on W^n, are continuous.

Proof: The continuity of the operators $w^n \mapsto y^n_j$ is easily proved either by induction on j using the Lipschitz continuity of f, or by using the discrete Bellman inequality (see [12]). The continuity of $w^n \mapsto G^n_m(w^n)$ follows from the continuity of g_m.

Lemma 3.2 We drop the index m. Under Assumptions 2.1-4, the directional derivative of the functional G^n is given by

(3.5) $$DG^n(w^n, \overline{w}^n - w^n) = \Delta t^n \sum_{j=0}^{N-1} (H_u(t^n_{j\theta}, y^n_{j\theta}, \nabla y^n_{j\theta}, z^n_{j,1-\theta}, w^n_j), \overline{w}^n_j - w^n_j),$$

where the discrete adjoint system z^n is given by

(3.6) $$-(1/\Delta t^n)(z^n_j - z^n_{j-1}, v) + a(v, z^n_{j,1-\theta})$$
$$= (z^n_{j,1-\theta} f_y(t^n_{j\theta}, y^n_{j\theta}, w^n_j) + g_y(t^n_{j\theta}, y^n_{j\theta}, \nabla y^n_{j\theta}, w^n_j), v) + (g_{y'}(t^n_{j\theta}, y^n_{j\theta}, \nabla y^n_{j\theta}, w^n_j), \nabla v),$$

for every $v \in V^n$, $j = N, ..., 1$, $\quad z^n_N = 0$, $\quad z^n_j \in V^n$,

The mappings $w^n \mapsto z^n$ and $(w^n, \overline{w}^n) \mapsto DG^n(w^n, \overline{w}^n - w^n)$ are continuous. For Δt^n sufficiently small, and for each j, the linear discrete adjoint scheme has a unique solution z^n_{j-1}.

The following classical control approximation result is proved similarly to the lumped parameter case (see [10]).

Proposition 3.1 Under Assumption 3.1, for every $w \in W$, there exists a sequence $(w^n \in W^n)$ that converges to w in L^2 strongly.

The next stability lemma gives useful a priori estimates.

Lemma 3.3 (Stability) Under Assumptions 2.1-2, if Δt is sufficiently small, then for every $w^n \in W^n$, we have the following inequalities, where the constants c are independent of n

(3.7) $\quad \|y_k^n\| \le c, \quad k = 0, ..., N,$

(3.8) $\quad \sum_{j=1}^{N} \|y_j^n - y_{j-1}^n\|^2 \le c,$

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

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

Proof. Dropping the index n for simplicity of notation, setting $v = 2\theta\Delta t y_j$ in the discrete equation, and using our assumptions on a and f, we have

(3.11)
$$\theta(\|y_j - y_{j-1}\|^2 + \|y_j\|^2 - \|y_{j-1}\|^2)$$
$$+\Delta t[a(y_{j\theta}, y_{j\theta}) + \theta^2 a(y_j, y_j) - (1-\theta)^2 a(y_{j-1}, y_{j-1})]$$
$$\le 2\theta\Delta t |(f(t_{j\theta}, y_{j\theta}, w_j), y_j)| \le c\Delta t(1 + \|y_j\| + \|y_{j-1}\|)\|y_j\|$$
$$\le c\Delta t(1 + \|y_{j-1}\|^2 + \|y_j\|^2) \le c\Delta t(1 + \|y_{j-1}\|^2 + \|y_j - y_{j-1}\|^2),$$

hence, for $\Delta t \le \dfrac{\theta}{2c}$

(3.12)
$$\theta(\frac{1}{2}\|y_j - y_{j-1}\|^2 + \|y_j\|^2 - \|y_{j-1}\|^2)$$
$$+\Delta t[a(y_{j\theta}, y_{j\theta}) + \theta^2 a(y_j, y_j) - (1-\theta)^2 a(y_{j-1}, y_{j-1})] \le c\Delta t(1 + \|y_{j-1}\|^2).$$

By summation over j, $j = 1, ..., k$, we obtain, for $\theta > 1/2$

(3.13)
$$\theta(\sum_{j=1}^{k} \frac{1}{2}\|y_j - y_{j-1}\|^2 + \|y_k\|^2) + \alpha_2\Delta t \sum_{j=1}^{k} \|y_{j\theta}\|_1^2 + \alpha_2\Delta t c' \sum_{j=1}^{k} \|y_j\|_1^2$$
$$\le \theta\|y_0\|^2 + \alpha_1\Delta t(1-\theta)^2 \|y_0\|_1^2 + c\Delta t \sum_{j=1}^{k} (1 + \|y_{j-1}\|^2), \quad \text{with } c' > 0,$$

and for $\theta = 1/2$

(3.14)
$$\frac{1}{2}(\sum_{j=1}^{k} \frac{1}{2}\|y_j - y_{j-1}\|^2 + \|y_k\|^2) + \alpha_2\Delta t \sum_{j=1}^{k} \|y_{j\theta}\|_1^2$$
$$\le \frac{1}{2}\|y_0\|^2 + \alpha_1 \frac{\Delta t}{4}\|y_0\|_1^2 + c\Delta t \sum_{j=1}^{k} (1 + \|y_{j-1}\|^2).$$

Since $\|y_0\|$ and $\|y_0\|_1$ remain bounded, using the discrete Bellman-Gronwall inequality (see [12]), we obtain inequality (3.14). The inequalities (3.8), (3.9), and (3.10) if $\theta > 1/2$, follow. If $\theta = 1/2$, by the inverse inequality (see [8]), the condition $\Delta t^n \le C(h^n)^2$, and inequality (3.8), we get

(3.15) $\Delta t\sum_{j=1}^{N}\|y_j-y_{j-1}\|_1^2\le\frac{\Delta t}{h^2}\sum_{j=1}^{N}\|y_j-y_{j-1}\|^2\le C\sum_{j=1}^{N}\|y_j-y_{j-1}\|^2\le c.$

Inequality (3.10) follows from this inequality and inequality (3.9), in this case.

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

(3.16) $v_-(t):=v_{j-1},\quad v_+(t):=v_j,\quad v_\theta(t):=(1-\theta)v_{j-1}+\theta v_j,\quad t\in \overset{o}{I}{}_j^n,\ j=1,...,N,$

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

In the sequel, we suppose that $\Delta t^n\le C(h^n)^2$, for some constant C independent of n, if $\theta=1/2$.

Theorem 3.1 (Consistency of states and functionals) Under Assumptions 2.1-3, if $w^n\to w\in W$ in L^2 strongly, then the corresponding discrete states $y_-^n,y_+^n,y_\theta^n,y_\wedge^n$ converge to y_w in $L^2(Q)$ strongly, $y_\theta^n\to y_w$ in $L^2(I,V)$ strongly, and

(3.18) $\lim_{n\to\infty}G_m^n(w^n)=G_m(w)$, $\quad m=0,...,q.$

Proof. By Lemma 3.3 (estimate (3.8) multiplied by Δt), $y_+^n-y_-^n\to 0$ in $L^2(Q)$ strongly. Since, by equation (3.9) in Lemma 3.3, y_-^n and y_+^n are bounded in $L^2(I,V)$, it follows that $y_\wedge^n$ and y_θ^n are also bounded in $L^2(I,V)$. By extracting subsequences, we can suppose that $y_\wedge^n\to y$ and $y_\theta^n\to y$ in $L^2(I,V)$ weakly (hence in $L^2(Q)$ weakly), for the same y. The discrete state equation can be written in the form

(3.19) $\frac{d}{dt}(y_\wedge^n(t),v^n)=(\psi^n(t),v^n)_1,\quad \forall v^n\in V^n$, a.e. in $(0,T)$,

in the scalar distribution sense, where the piecewise constant function ψ^n is defined, using Riesz's representation theorem, by

(3.20) $(\psi_j^n(t),v^n)_1:=-a(y_{j\theta}^n,v^n)+(f(t_{j\theta}^n,y_{j\theta}^n,w_j^n),v^n),\quad$ in $\overset{o}{I}{}_j^n$, $j=1,...,N.$

By our assumptions, we have, for $j=1,...,N$

(3.21) $\left|(\psi_j^n,v^n)_1\right|\le c[\|y_{j\theta}^n\|_1\|v^n\|_1+(1+\|y_{j\theta}^n\|)\|v^n\|]\le c(1+\|y_{j\theta}^n\|_1)\|v^n\|_1$,

hence

(3.22) $\|\psi_j^n\|_1\le c(1+\|y_{j\theta}^n\|_1)$ and $\|\psi_j^n\|_1^2\le c(1+\|y_{j\theta}^n\|_1^2).$

Therefore, using equation (3.9) in Lemma 3.3

(3.23) $\int_0^T\|\psi^n(t)\|_1^2\,dt\le c(1+\int_0^T\|y_\theta^n\|_1^2\,dt)\le c,$

which shows that ψ^n belongs to $L^2(I,V)$, hence to $L^1(I,V)$. Following the proof of Lemma 5.6 in [11], it can then be shown that

(3.24) $\int_{-\infty}^{+\infty}|\tau|^{2\rho}\|\hat{y}_\wedge^n(\tau)\|^2\,d\tau\le c,\quad$ for $\rho<1/4,$

where $\hat{y}_\wedge^n$ denotes the Fourier transform of $y_\wedge^n$, with $y_\wedge^n$ extended by 0 outside $[0,T]$. By the 2nd compactness theorem in [11], p. 274, there exists a subsequence (same notation) such that $y_\wedge^n\to\tilde{y}$ in $L^2(Q)$ strongly, for some $\tilde{y}$, and we must have $\tilde{y}=y$, since $\hat{y}^n_\wedge \to y$ also in $L^2(Q)$ weakly. Since, by Lemma 3.3 ((3.8) multiplied by Δt), $y^n_+ - y^n_- \to 0$ in $L^2(Q)$, we get $y^n_\theta \to y$ in $L^2(Q)$ strongly. Similarly to the proof of Lemma 4.3 in [4], we can then pass to the limit in the weak discrete equation, integrated in t, using Proposition 2.1 in [3] for the nonlinear term, and show that $y = y_w$. Next, to prove the strong convergence $y^n_\theta \to y$ in $L^2(I,V)$, we first remark that, by the discrete and continuous state equations, the boundedness of (y^n_N) in $L^2(\Omega)$ by Lemma 3.3 (3.7), the above convergences, Proposition 2.1 in [3], and taking the sequence $(v^n \in V^n)$ of functions interpolating an arbitrary $v \in C^1_0(\Omega)$ (which clearly converges to v in $H^1(\Omega)$ strongly), we have

(3.25) $$(y^n_N, v) = (y^n_N, v - v^n) + (y^n_N, v^n)$$
$$= (y^n_N, v - v^n) + (y^n_0, v^n) + \int_0^T (f(y^n_\theta, w^n), v^n)dt - \int_0^T a(y^n_\theta, v^n)dt$$
$$\to (y^0, v) + \int_0^T (f(y,w), v)dt - \int_0^T a(y,v)dt = (y(T), v),$$

for every $v \in C^1_0(\Omega)$, hence $(y^n_N, v) \to (y(T), v)$ for every $v \in L^2(\Omega)$, since $C^1_0(\Omega)$ is dense in $L^2(\Omega)$, i.e. $y^n_N \to y(T)$ in $L^2(\Omega)$ weakly. We then write

(3.26) $$\alpha_2 \left\| y^n_\theta - y \right\|^2_{L^2(I,V)} \le \int_0^T a(y^n_\theta - y, y^n_\theta - y)dt + \frac{1}{2}\left\| y^n_N - y(T) \right\|^2$$
$$= \frac{1}{2}\left\| y^n_0 \right\|^2 - \frac{1}{2}(y^n_N, y(T)) - \frac{1}{2}(y(T), y^n_N - y(T))$$
$$+ \int_0^T (f(y^n_\theta, w^n), y^n_\theta)dt - \int_0^T a(y^n_\theta, y)dt - \int_0^T a(y, y^n_\theta - y)dt,$$

where the last expression converges to zero. The last convergences follow using also Proposition 2.1 in [3].

Note that the condition $\Delta t^n \le C(h^n)^2$ (in fact, the inverse inequality used to derive inequality (3.10), if $\theta = 1/2$) is a worst case one. In practice, the corresponding sequences of gradients (∇y^n) constructed by the algorithms are often bounded in $L^2(Q)$, or even in $L^\infty(Q)$, and the above condition is not needed for $\theta = 1/2$.

Theorem 3.2 (Consistency of adjoints and functional derivatives) Under Assumptions 2.1-4, if $w^n \to w \in W$ in L^2 strongly, then the corresponding discrete adjoints $z^n_-, z^n_+, z^n_\theta, z^n_\wedge$ converge to z_w in $L^2(Q)$ strongly. If $w^n \to w \in W$ and $\bar{w}^n \to \bar{w} \in W$ in L^2 strongly, then

(3.27) $$\lim_{n\to\infty} DG^n_m(w^n, \bar{w}^n - w^n) = DG_m(w, \bar{w} - w), \quad m = 0,...,q.$$

Proof. The proof is similar to that of Theorem 3.1, using also the consistency of the states.

4 Discrete penalized gradient projection method

Let (M^n_m), $m = 1,...,q$, be nonnegative increasing sequences such that $M^n_m \to \infty$ as $n \to \infty$, and define the *penalized discrete functionals*

(4.1) $G^n(w^n) := G_0^n(w^n) + \{\sum_{m=1}^{p} M_m^n [G_m^n(w^n)]^2 + \sum_{m=p+1}^{q} M_m^n [\max(0, G_m^n(w^n))]^2\}/2.$

Let $\gamma \ge 0$, $b, c \in (0,1)$, and let (β^n), (ζ_k) be positive sequences, with (β^n) decreasing and converging to zero, and $\zeta_k \le 1$.

Assumption 4.1 Each element $E_{i'}^{n+1}$ is a subset of some element E_i^n, and either $N^{n+1} = N^n$ or $N^{n+1} = \kappa N^n$, for some integer $\kappa \ge 2$ (usually $\kappa = 2$).

If Assumption 4.1 holds, then we have $W^n \subset W^{n+1}$, and thus a control $w^n \in W^n$ may be considered also as belonging to W^{n+1}, hence the computation of states, adjoints and functional derivatives for this control, but with the possibly finer discretization $n+1$, makes sense. The *discrete penalized gradient projection method* is described by the following algorithm.

Algorithm

Step 1. Set $k := 0$, $n := 1$, and choose an initial control $w_0^n \in W^n$.

Step 2. Find $v_k^n \in W^n$ such that

(4.2) $$e_k := DG^n(w_k^n, v_k^n - w_k^n) + \frac{\gamma}{2}\|v_k^n - w_k^n\|_Q^2$$

$$= \min_{\bar{v}^n \in W^n}[DG^n(w_k^n, \bar{v}^n - w_k^n) + \frac{\gamma}{2}\|\bar{v}^n - w_k^n\|_Q^2],$$

and set $d_k := DG^n(w_k^n, v_k^n - w_k^n)$.

Step 3. If $|d_k| \le \beta^n$, set $w^n := w_k^n$, $v^n := v_k^n$, $d^n := d_k$, $e^n := e_k$, $n := n+1$, and go to Step 2.

Step 4. (*Armijo step search*) Find the lowest integer value $s \in \mathbb{Z}$, say $\bar{s}$, such that $\alpha(s) := c^s \zeta_k \in (0,1]$ and $\alpha(s)$ satisfy the inequality

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

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

Step 5. Set $w_{k+1}^n := w_k^n + \alpha_k(v_k^n - w_k^n)$, $k := k+1$, and go to Step 2.

The progressively refining method, as compared to the corresponding fixed finest discretization method, usually yields results of similar accuracy, but has the advantage of reducing computing time and memory. It is justified by the fact that finer discretizations become progressively more efficient as the iterate gets closer to an extremal control, while coarser ones in the early iterations have not much influence on the final results.

If $\gamma > 0$, we have a *penalized strict gradient projection method*, in which case one can easily see by “completing the square” that Step 2 amounts to finding, for each $i = 1,...,M$, $j = 1,...,N$, the projection v_{ij}^n of

(4.4) $$u_{ij}^n := w_{ij}^n - \frac{1}{\gamma\,\mu(E_i^n)} \int_{E_i^n} H_u^n(t_{j\theta}^n, y_{i,j\theta}^n, z_{i,j\theta}^n, w_{ij}^n)dx$$

onto the convex set U, where $\mu(E_i^n)$ is the measure of E_i^n. The parameter γ is chosen experimentally to yield a good convergence rate. If $\gamma = 0$, the above

Algorithm is a *penalized conditional gradient method*, and Step 2 reduces to the minimization of a linear function on U. On the other hand, by the definition of the directional derivative and since $d_k \le e_k \le 0$, $b \in (0,1)$, clearly the Armijo step α_k in Step 4 can be found for every k, if $d_k \ne 0$.

An extremal control is called *abnormal* if there exist multipliers as in the optimality conditions, but with $\lambda_0 = 0$. A control is admissible *and* abnormal extremal in exceptional, degenerate, situations (see [16]).

With w^n as defined in Step 3 of the Algorithm, we define the *sequences of multipliers*

(4.5) $\quad \lambda_m^n = M_m^n G_m^n(w^n),\ \ m = 1,...,p, \quad \lambda_m^n = M_m^n \max(0, G_m^n(w^n)),\ \ m = p+1,...,q,$

Theorem 4.1 We suppose that Assumptions 2.1-4, 4.1 are satisfied.
(i) Let (w^n) be a subsequence (if it exists) of the sequence generated by the Algorithm in Step 3 that converges to some $w \in W$ in L^2 strongly, as $n \to \infty$. If the sequences of multipliers (λ_m^n) are bounded, then w is admissible and weakly extremal for the continuous problem.
(ii) Suppose that the continuous problem has no admissible, abnormal extremal, controls. If the limit control w in (i) is admissible, then the sequences of multipliers (λ_m^n) are bounded, w is extremal as above.

Proof. We shall first show that $n \to \infty$ in the Algorithm. Suppose, on the contrary, that n remains constant after a finite number of iterations in k, and we could also drop the index n in the subsequences. Let us show that $d_k \to 0$. Since W^n is compact, let $(w_k)_{k\in K}$, $(v_k)_{k\in K}$ be subsequences of the sequences generated in Steps 2 and 5 such that $w_k \to \tilde{w},\ \ v_k \to \tilde{v},$ in W^n, as $k \to \infty,\ \ k \in K.$ Clearly, by Step 2, $d_k \le e_k \le 0$ for every k, hence

(4.6) $\quad e := \lim\limits_{k\to\infty,\, k\in K} e_k = DG(\tilde{w}, \tilde{v} - \tilde{w}) + (\gamma/2)\left\| \tilde{v} - \tilde{w} \right\|_Q^2 \le 0,$

(4.7) $\quad d := \lim\limits_{k\to\infty,\, k\in K} d_k = DG(\tilde{w}, \tilde{v} - \tilde{w}) \le \lim\limits_{k\to\infty,\, k\in K} e_k = e \le 0.$

Suppose that $d < 0$. The function $\Phi(\alpha) := G(w + \alpha(v - w))$ is continuous on $[0,1]$. Since the directional derivative $DG(w, v - w)$ is linear w.r.t. $v - w$, Φ is differentiable on $(0,1)$ and has derivative

(4.8) $\quad \Phi'(\alpha) = DG(w + \alpha(v - w), v - w).$

Using the Mean Value Theorem, we have, for each $\alpha \in (0,1]$

(4.9) $\quad G(w_k + \alpha(v_k - w_k)) - G(w_k) = \alpha DG(w_k + \alpha'(v_k - w_k), v_k - w_k),$

for some $\alpha' \in (0, \alpha)$. Therefore, for $\alpha \in [0,1]$, by the continuity of DG (Lemma 3.2)

(4.10) $\quad G(w_k + \alpha(v_k - w_k)) - G(w_k) = \alpha(d + \varepsilon_{k\alpha}),$

where $\varepsilon_{k\alpha} \to 0$ as $k \to \infty$, $k \in K$, and $\alpha \to 0^+$. Now, we have $d_k = d + \eta_k$, where $\eta_k \to 0$ as $k \to \infty$, $k \in K$, and since $b \in (0,1)$

(4.11) $\quad d + \varepsilon_{k\alpha} \le b(d + \eta_k) = bd_k,$

for $\alpha \in [0, \bar{\alpha}]$, for some $\bar{\alpha} > 0$, and $k \ge \bar{k}$, $k \in K$. Hence

(4.12) $\quad G(w_k + \alpha(v_k - w_k)) - G(w_k) \le \alpha b d_k,$

for $\alpha \in [0,\bar{\alpha}]$, for some $\bar{\alpha} > 0$, and $k \geq \bar{k}$, $k \in K$. It follows from the choice of the Armijo step α_k in Step 4 that $\alpha_k \geq c\bar{\alpha}$, for $k \geq \bar{k}$, $k \in K$. Hence

(4.13) $G(w_{k+1}) - G(w_k) = G(w_k + \alpha_k(v_k - w_k)) - G(w_k) \leq \alpha_k b d_k \leq c\bar{\alpha} b d_k \leq c\bar{\alpha} b d / 2,$

for $k \geq \bar{k}$, $k \in K$. It follows that $G(w_k) \to -\infty$ as $k \to \infty$, $k \in K$. This contradicts the fact that $G(w_k) \to G(\tilde{w})$ as $k \to \infty$, $k \in K$, by the continuity of the discrete functional (Lemma 3.1). Therefore, we must have $d = 0$, $e = 0$, and $d_k \to d = 0$, $e_k \to e = 0$, for the whole sequences, since the limit 0 is unique. But Step 3 then implies that $n \to \infty$, which is a contradiction. Therefore, we must have $n \to \infty$.

(i) Let (w^n) be a subsequence (same notation) of the sequence generated in Step 3 that converges to some $w \in W$ in L^2 strongly as $n \to \infty$. Suppose that the sequences (λ_m^n) are bounded and that (up to subsequences) $\lambda_m^n \to \lambda_m$. By Theorem 3.1, we have

(4.14) $$0 = \lim_{n\to\infty} \frac{\lambda_m^n}{M_m^n} = \lim_{n\to\infty} G_m^n(w^n) = G_m(w), \quad m = 1,\dots,p,$$

(4.15) $$0 = \lim_{n\to\infty} \frac{\lambda_m^n}{M_m^n} = \lim_{n\to\infty} [\max(0, G_m^n(w^n))] = \max(0, G_m(w)), \quad m = p+1,\dots,q,$$

which show that w is admissible. Now, let any $\tilde{v} \in W$ and, by Proposition 3.1, $(\tilde{v}^n \in W^n)$ a sequence converging to $\tilde{v}$. By Step 2, we have

(4.16) $$\int_Q H_u^n(x, t_\theta^n, y_\theta^n, z_\theta^n, w^n)(\tilde{v}^n - w^n)dxdt + (\gamma/2)\int_Q \left|\tilde{v}^n - w^n\right|^2 dxdt \geq d^n,$$

where H^n and z^n are defined with $g := \sum_{m=0}^{q} \lambda_m^n g_m$. Using Proposition 2.1 in [3] and Theorems 3.1, 3.2, we can pass to the limit as $n \to \infty$ and obtain

(4.17) $$\int_Q H_u(x,t,y,z,w)(\tilde{v} - w)dxdt + (\gamma/2)\int_Q \left|\tilde{v} - w\right|^2 dxdt \geq 0, \quad \forall \tilde{v} \in W,$$

where H and z are defined with $g := \sum_{m=0}^{q} \lambda_m g_m$. Replacing now $\tilde{v}$ by $w + \alpha(\tilde{v} - w)$, dividing by α, and taking the limit as $\alpha \to 0$, we get

(4.18) $$\int_Q H_u(x,t,y,z,w)(\tilde{v} - w)dxdt \geq 0, \quad \forall \tilde{v} \in W,$$

If $G_m(w) < 0$, for some index $m \in [p+1, q]$, then for sufficiently large n we have $G_m^n(w^n) < 0$ and $\lambda_m^n = 0$, hence $\lambda_m = 0$, i.e. the transversality conditions hold. Therefore, w is weakly extremal.

(ii) Suppose that the limit control w is admissible and that the continuous problem has no admissible, abnormal extremal, controls. Suppose that the multipliers are not all bounded. Then, dividing the inequality resulting from Step 2 by the greatest multiplier norm and passing to the limit for a subsequence, we see that we obtain an optimality inequality where the first multiplier is zero, and that the limit control is abnormal extremal, a contradiction. Therefore, the sequences of multipliers are bounded, and by (i), w is extremal as above.

One can easily see that Theorem 4.1 remains valid if we replace d_k by e_k in Step 4 of the Algorithm. In practice, by choosing moderately growing sequences (M_m^n) and a sequence (β^n) relatively fast converging to zero, the resulting sequences of multipliers (λ_m^n) are often kept bounded. One can choose a fixed $\zeta_k := \zeta \in (0,1]$ in Step 4; a usually faster and adaptive procedure is to set $\zeta_0 := 1$, and then $\zeta_k := \alpha_{k-1}$, for $k \geq 1$.

5 Numerical examples

Let $\Omega := I := (0,1)$.

Example 1. Define the reference control and state

(5.1) $$\bar{w}(x,t) := \begin{cases} -1, & 0 \leq t < 0.5, \\ -1 + 16(t-0.5)x(1-x), & 0.5 \leq t \leq 1, \end{cases} \qquad \bar{y}(x,t) := x(1-x)e^t,$$

and consider the following optimal control problem, with state equation

(5.2) $$y_t - y_{xx} = [x(1-x)+2]e^t + \sin y - \sin \bar{y} + w - \bar{w} \quad \text{in } Q,$$

(5.3) $$y = 0 \quad \text{in } \Sigma, \quad y(x,0) = \bar{y}(x,0) \quad \text{in } \Omega,$$

control constraint set $U := [-1,1]$, and cost functional

(5.4) $$G_0(w) := 0.5\int_Q [(y-\bar{y})^2 + \left|\nabla y - \nabla \bar{y}\right|^2 + (w-\bar{w})^2]dxdt.$$

Clearly, the optimal control and state are $\bar{w}$ and $\bar{y}$, and the optimal cost is zero. The discrete gradient projection method, without penalties, was applied to this problem, with successive step sizes $h = \Delta t = 1/20, 1/40, 1/80$ in three equal iteration periods, θ-scheme parameter $\theta = 0.5$, gradient projection parameter $\gamma = 0.5$, Armijo parameters $b = c = 0.5$, and zero initial control. After 18 iterations, we obtained the following results:

(5.5) $$G_0^n(w_k) = 2.731 \cdot 10^{-9}, \quad d_k = -1.014 \cdot 10^{-11}, \quad \varepsilon_k = 1.998 \cdot 10^{-5}, \quad \eta_k = 1.699 \cdot 10^{-5},$$

where d_k is defined in Step 2 of the Algorithm, ε_k is the discrete state max-error at the vertices of the blocks, and η_k the discrete control max-error at the centers of the blocks. Figure 1 shows the last computed control $w_k \approx \bar{w}$.

Example 2. With the same state equation, cost and parameters as in Example 1, but with $U := [-1, 0.5]$, the control constraints being now strictly active, and zero initial control, we obtained after 18 iterations the control shown in Figure 2 and the results:

(5.6) $$G_0^n(w_k) = 1.234481472177051 \cdot 10^{-3}, \quad d_k = -8.847 \cdot 10^{-14}.$$

Example 3. With the state equation (and the boundary conditions of Example 1)

(5.7) $$y_t - y_{xx} = 3w \quad \text{in } Q,$$

the contraint set $U := [-1, 0.8]$, the additional state constraint

(5.8) $$G_1(w) := \int_Q y(x,t)dxdt = 0,$$

and with the cost and parameters as in Example 1, we obtained, after 90 iterations in k of the penalized gradient projection method, the control and state shown in Figures 3 and 4 and the results:

(5.9) $$G_0^n(w_k) = 0.730014380250449, \quad G_1^n(w_k) = -5.132 \cdot 10^{-5}, \quad d_k = -4.936 \cdot 10^{-5}.$$

Since the state equation and the equality state constraint are linear in (y, w) and the cost is convex in (y, w), the optimality conditions (with $\lambda_0 = 1 > 0$) obtained here are also sufficient, and therefore the method actually approximates the optimal control.

Finally, the above results with progressive refining were found to be of practically similar accuracy to those obtained with constant last step sizes $h = \Delta t = 1/80$, but required here less than half the computing time.

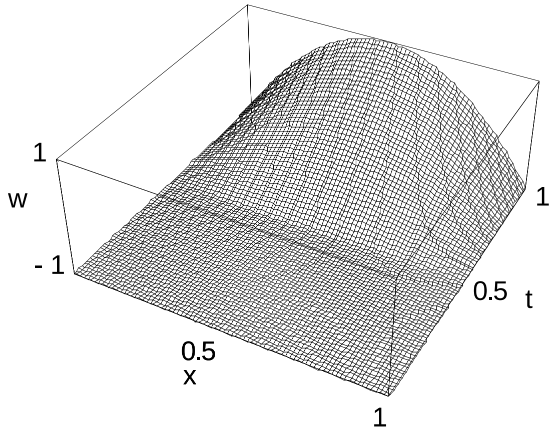

Figure 1. Example 1: Last control

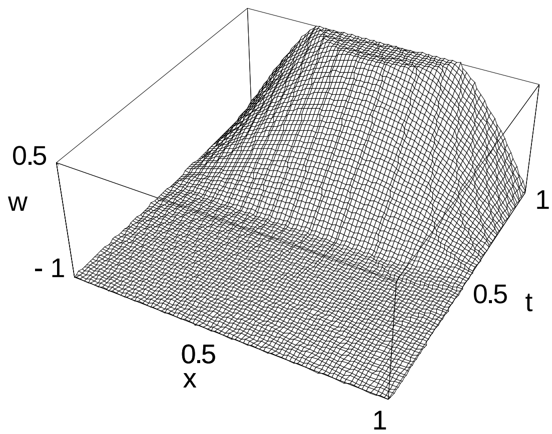

Figure 2. Example 2: Last control

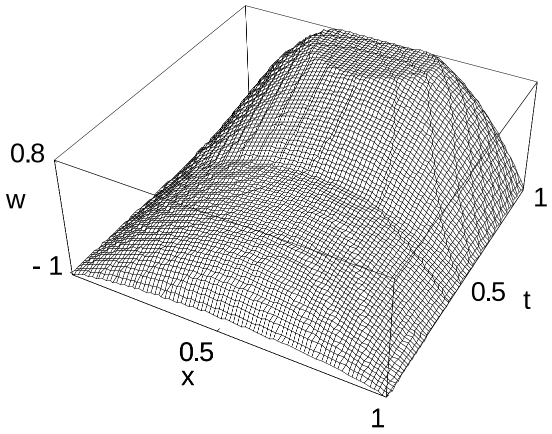

Figure 3. Example 3: Last control

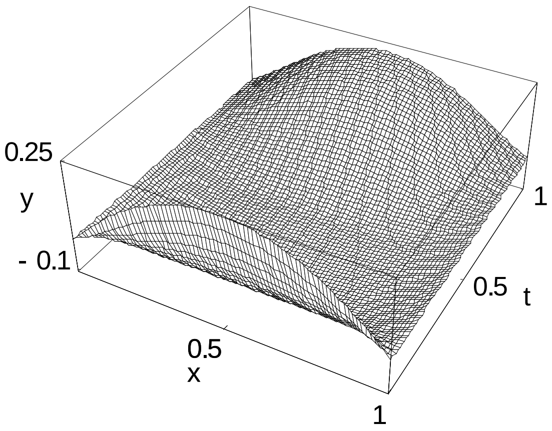

Figure 4. Example 3: Last state

References

[1] E. Casas and F. Troeltzsch, Error estimates for the finite-element approximation of a semilinear elliptic control problem, *Control Cybern.*, **31** (2002), 695-712.
[2] E. Casas, M. Mateos, and F. Troeltzsch, Error estimates for the numerical approximation of boundary semilinear elliptic control problems, *Comp. Optim. Appl.* **31** (2005), 193-220.
[3] I. Chryssoverghi, Nonconvex optimal control of nonlinear monotone parabolic systems, *Systems Control Lett.*, **8** (1986) 55-62.
[4] I. Chryssoverghi and A. Bacopoulos, Approximation of relaxed nonlinear parabolic optimal control problems, *J. Optim. Theory Appl*, **77**, 1 (1993) 31-50.
[5] I. Chryssoverghi, A. Bacopoulos, B. Kokkinis and J. Coletsos, Mixed Frank-Wolfe penalty method with applications to nonconvex optimal control problems, *J. Optim. Theory Appl.*, **94**, 2 (1997) 311-334.
[6] I. Chryssoverghi, A. Bacopoulos, J. Coletsos and B. Kokkinis, Discrete approximation of nonconvex hyperbolic optimal control problems with state constraints, *Control Cybernet.*, **27**, 1 (1998) 29-50.

[7] I. Chryssoverghi, J. Coletsos and B. Kokkinis, Discrete relaxed method for semilinear parabolic optimal control problems, *Control Cybernet.*, **28**, 2 (1999) 157-176.
[8] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, New York, 1978.
[9] H.O. Fattorini, *Infinite Dimensional Optimization Theory and Optimal control*, Cambridge Univ. Press, Cambridge, 1999.
[10] E. Polak, *Optimization: Algorithms and Consistent Approximations*, Springer, Berlin, 1997.
[11] R. Temam, *Navier-Stokes Equations*, North-Holland, New York, 1977.
[12] V. Thomee, *Galerkin Finite Element Methods for Parabolic Problems*, Springer, Berlin, 1997.
[13] D. Tiba and F. Troeltzsch, Error estimates for the discretization of state constrained convex control problems, *Num. Funct. Anal. Optim.*, **17** (1996), 1005-1028.
[14] F. Troeltzsch, Semidiscrete Ritz-Galerkin approximation of nonlinear parabolic boundary control problems - strong convergence of optimal controls, *Appl. Math. Optim.*, **29** (1994), 309-329.
[15] F. Troeltzsch, An SQP method for optimal control of a nonlinear heat equation, *Control Cybern.*, **23** (1994), 267-288.
[16] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.