Discretization-Optimization Methods for Nonlinear Elliptic Relaxed Optimal Control Problems with State Constraints

I. Chryssoverghi[1], J. Coletsos[1], J. Geiser[2], B. Kokkinis[1]

([1]) Department of Mathematics, School of Applied Mathematics and Physics
National Technical University of Athens (NTUA)
Zografou Campus, 15780 Athens, Greece
e-mail: ichris@central.ntua.gr
([2]) Weierstrass Institute for Applied Analysis and Stochastics (WIAS), Mohrenstrasse 39, D-10117 Berlin, Germany
e-mail: geiser@wias-berlin.de

Abstract

We consider an optimal control problem described by a second order elliptic boundary value problem, jointly nonlinear in the state and control with high monotone nonlinearity in the state, with control and state constraints, where the state constraints and cost functional involve also the state gradient. Since no convexity assumptions are made, the problem may have no classical solutions, and so it is reformulated in the relaxed form using Young measures. Existence of an optimal control and necessary conditions for optimality are established for the relaxed problem. The relaxed problem is then discretized by using a finite element method, while the controls are approximated by elementwise constant Young measures. We show that relaxed accumulation points of sequences of optimal (resp. admissible and extremal) discrete controls are optimal (resp. admissible and extremal) for the continuous relaxed problem. We then apply a penalized conditional descent method to each discrete problem, and also a progressively refining version of this method to the continuous relaxed problem. We prove that accumulation points of sequences generated by the first method are admissible and extremal for the discrete relaxed problem, and that accumulation points of sequences of discrete controls generated by the second method are admissible and extremal for the continuous relaxed problem. Finally, numerical examples are given.

Keywords. Optimal control, nonlinear elliptic systems, state constraints, relaxed controls, discretization, finite elements, discrete penalized conditional gradient method, progressive refining.

1 Introduction

We consider an optimal control problem described by a second order elliptic boundary value problem, jointly nonlinear in the state and control with high monotone nonlinearity in the state, with control and state constraints, where the state constraints and cost functional depend also on the state gradient. Since no convexity assumptions are made, the problem may have no classical solutions, and so it is reformulated in the relaxed form using Young measures. Existence of an optimal control and necessary conditions for optimality are established for the relaxed problem. The relaxed problem is then discretized by using a Galerkin finite element method with continuous elementwise linear basis functions, while the controls are approximated by elementwise constant Young measures. We first state the necessary conditions for optimality for the continuous and the discrete relaxed problems. Under appropriate assumptions, we show that accumulation points of sequences of optimal (resp. admissible and extremal) discrete controls are optimal (resp. admissible and extremal) for the continuous relaxed problem. We then apply a penalized conditional descent method to each discrete problem, and also a corresponding discrete method to the continuous relaxed problem, which progressively refines the discretization during the iterations, thus reducing computing time and memory. We prove that accumulation points of sequences generated by the fixed discretization method are admissible and extremal for the discrete relaxed problem, and that accumulation points of sequences of discrete controls generated by the progressively refining method are admissible and extremal for the continuous relaxed problem. Finally, numerical examples are given. For the theory and various approximation and optimization methods in optimal control and variational problems, see [2-4], [6-13], [15-21], and the references therein.

2 The continuous optimal control problem

Let Ω be a bounded domain in $\mathbb{R}^d$, with Lipschitz boundary Γ. Consider the nonlinear elliptic state equation

(2.1) $\quad Ay + f(x, y(x), w(x)) = 0 \quad \text{in } \Omega,$

(2.2) $\quad y(x) = 0 \quad \text{on } \Gamma,$

where A is the formal second order elliptic differential operator

(2.3) $$Ay := -\sum_{j=1}^{d}\sum_{i=1}^{d}(\partial/\partial x_i)[a_{ij}(x)\partial y/\partial x_j].$$

The constraints on the control are $w(x) \in U$ in Ω, where U is a compact, *not necessarily convex*, subset of $\mathbb{R}^\nu$, the state constraints are

(2.4) $$G_m(w) := \int_\Omega g_m(x, y(x), \nabla y(x), w(x))dx = 0, \quad m = 1, ..., p,$$

(2.5) $$G_m(w) := \int_\Omega g_m(x, y(x), \nabla y(x), w(x))dx \le 0, \quad m = p+1, ..., q,$$

and the cost functional to be minimized is

(2.6) $$G_0(w) := \int_\Omega g_0(x, y(x), \nabla y(x), w(x))dx.$$

Defining the set of *classical controls*

(2.7) $$W := \{w : x \mapsto w(x) \,|\, w \text{ measurable from } \Omega \text{ to } U\},$$

the continuous classical optimal control problem is to minimize G_0 subject to the above constraints.

It is well known that such nonconvex optimal control problems may have no classical solutions. Existence can be proved under some convexity assumptions, which are often unrealistic for nonlinear systems. Reformulated in the so-called relaxed form, these problems are convexified in some sense and have a solution in an extended space under weak assumptions.

Next, we define the set of *relaxed controls* (or Young measures; for the relevant theory, see [19], [20])

(2.8) $$R := \{r : \Omega \to M_1(U) \,|\, r \text{ weakly measurable}\} \subset L^\infty_w(\Omega, M(U)) \equiv L^1(\Omega, C(U))^*,$$

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, and R is convex, metrizable and compact. If each classical control $w(\cdot)$ is identified with its associated Dirac relaxed control $r(\cdot) := \delta_{w(\cdot)}$, then W *may also* be considered as a subset of R, and W is *thus* dense in R. For $\phi \in L^1(\Omega; C(U)) = L^1(\bar{\Omega}; C(U))$ (or $\phi \in B(\bar{\Omega}, U; \mathbb{R})$, where $B(\bar{\Omega}, U; \mathbb{R})$ is the set of Caratheodory functions in the sense of Warga [20]) and $r \in L^\infty_w(\Omega, M(U))$ (in particular, for $r \in R$), we shall use the *notation*

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

and $\phi(x,r(x))$ is thus *linear* (under convex combinations, for $r\in R$) in r. A sequence (r_k) converges to $r\in R$ in R iff

(2.10) $\lim_{k\to\infty}\int_\Omega \phi(x,r_k(x))dx=\int_\Omega \phi(x,r(x))dx,$

for every $\phi\in L^1(\Omega;C(U))$, or $\phi\in B(\bar\Omega,U;\mathbb{R})$, or $\phi\in C(\bar\Omega\times U)$.

We denote by $|\cdot|$ the Euclidean norm in $\mathbb{R}^n$, by $\|\cdot\|_\infty$ the norm in $L^\infty(\Omega,\mathbb{R}^n)$, by $(\cdot,\cdot)$ and $\|\cdot\|$ the inner product and norm in $L^2(\Omega,\mathbb{R}^n)$, and by $(\cdot,\cdot)_1$ and $\|\cdot\|_1$ the inner product and norm in the Sobolev space $V:=H_0^1(\Omega)$. We can now formulate the relaxed problem as follows. The relaxed state equation (in weak form) is given by

(2.11) $y\in V$ and $a(y,v)+\int_\Omega f(x,y(x),r(x))v(x)dx=0,\quad \forall v\in V,$

where $a(\cdot,\cdot)$ is the usual bilinear form associated with A and defined on $V\times V$

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

the control constraint is $r\in R$, and the relaxed functionals are

(2.13) $G_m(r):=\int_\Omega g_m(x,y(x),\nabla y(x),r(x))dx,\quad m=0,...q.$

The continuous relaxed optimal control Problem P is to minimize $G_0(r)$ subject to the constraints

$r\in R,\quad G_m(r)=0,\ m=1,...,p,\quad G_m(r)\le 0,\ m=p+1,...,q.$

In the sequel, we shall make some of the following assumptions.

Assumptions 2.1 The coefficients a_{ij} satisfy the ellipticity condition

(2.14) $\sum_{i,j=1}^{d}a_{ij}(x)z_iz_j\ge \alpha_0\sum_{i=1}^{d}z_i^2,\quad \forall z_i,z_j\in\mathbb{R},\ x\in\Omega,$

with $\alpha_0>0$, $a_{ij}\in L^\infty(\Omega)$, which implies that

(2.15) $|a(y,v)|\le \alpha_1\|y\|_1\|v\|_1,\quad a(v,v)\ge\alpha_2\|v\|_1^2,\quad \forall y,v\in V,$

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

Assumptions 2.2 The functions f and f_y are defined on $\Omega\times\mathbb{R}\times U$, measurable for fixed y,u, continuous for fixed x, and satisfy

(2.16) $|f(x,0,u)|\le\phi_0(x),\quad \forall(x,u)\in\Omega\times U,$

where $\phi_0\in L^s(\Omega)$, $s\ge 2$, $s\ge n/2$ (e.g. $s=2$, for $n=1,2,3$), and

(2.17) $0\le f_y(x,y,u)\le\phi_1(x)\eta_1(|y|),\quad \forall(x,y,u)\in\Omega\times\mathbb{R}\times U,$

where η_1 is an increasing function from $[0,+\infty)$ to $[0,+\infty)$, $\phi_1\in L^\infty(\Omega)$ if the functionals G_m depend on ∇y, and $\phi_1\in L^s(\Omega)$ otherwise.

Assumptions 2.3 The functions g_m are defined on $\Omega\times\mathbb{R}^{d+1}\times U$, measurable for fixed y,y',u, continuous for fixed x, and satisfy

(2.18) $|g_m(x,y,y',u)|\le\psi_{0m}(x)+\beta_{0m}|y'|^2,\quad \forall(x,y,y',u)\in\Omega\times\mathbb{R}^{d+1}\times U$ with $|y|\le C'$,

for some $C'>C$, where $\psi_{0m}\in L^1(\Omega)$, $\beta_{0m}\ge 0$.

Assumptions 2.4 The functions $g_{my}, g_{my'}$ are defined on $\Omega \times \mathbb{R}^{d+1} \times U$, measurable on Ω for fixed $(y, y', u) \in \mathbb{R}^{d+1} \times U$, continuous on $\mathbb{R}^{d+1} \times U$ for fixed $x \in \Omega$, and satisfy

(2.19) $\left|g_{my}(x, y, y', u)\right| \le \psi_{1m}(x) + \beta_{1m} \left|y'\right|^{\frac{2(\rho-1)}{\rho}}$,

(2.20) $\left|g_{my'}(x, y, y', u)\right| \le \psi_{2m}(x) + \beta_{2m} \left|y'\right|$,

$\forall (x, y, y', u) \in \Omega \times \mathbb{R}^{d+1} \times U,$' with $|y| \le C'$,

where $C' < C$, $\psi_{im} \in L^2(\Omega)$, $\beta_{im} \ge 0$, $\rho \in [1, \infty)$ if $n = 1$ or 2, $\rho < \sigma := \dfrac{2n}{n-2}$ if $n \ge 3$.

The following theorem follows directly form Theorem 3.1 in [4].

Theorem 2.1 Under Assumptions 2.1-2, for every relaxed control $r \in R$, the state equation has a unique solution $y := y_r \in V \cap C^{\alpha}(\bar{\Omega})$, for some $\alpha \in (0,1)$. Moreover, there exists constants $C, \bar{C}$ such that

(2.21) $\|y_r\|_1 + \|y_r\|_\infty \le C$, $\|y_r\|_{C^\alpha} \le \bar{C}$, for every $r \in R$.

The following proposition generalizes Proposition 2.1 in [9], with a simpler proof, and will be useful in the sequel.

Proposition 2.1 For $i = 1, ..., K$, $K \ge 0$, let $s_i \in [1, +\infty]$, $\sigma_i \in [0, s_i]$ if $s_i < +\infty$, $\sigma_i := 0$ if $s_i = +\infty$, with $\dfrac{1}{s_0} + \sum_{i=1}^{K} \dfrac{\sigma_i}{s_i} \le 1$, $\dfrac{1}{s_i} := 0$ if $s_i = +\infty$. Let F be a function defined on $\Omega \times (\mathbb{R}^N)^K \times U$, measurable for every y, u fixed, continuous for every x fixed, and satisfying

(2.22) $\left|F(x, y, u)\right| \le \Phi(x) + \Psi(x) \prod_{i=1}^{K} \xi_i(\|y_i\|)$,

for every $(x, y, u) \in \Omega \times (\mathbb{R}^N)^K \times U$, with $\|y_i\| \le C_i$ if $s_i = +\infty$,

where $y := (y_1, ..., y_K)$, $\Phi \in L^1(\Omega)$, $\Psi \in L^{s_0}(\Omega)$, $\xi_i(\|y_i\|) := \|y_i\|^{\sigma_i}$ if $s_i < +\infty$, $\xi_i(\|y_i\|) := 1$ if $s_i = +\infty$. If (y_i^k) converges to y_i in $L^{s_i}(\Omega; \mathbb{R}^N)$ strongly, $i = 1, ..., K$, with $\|y_i^k\|_\infty \le C_i$ (for k sufficiently large) if $s_i = +\infty$, and (r^k) converges to r in R, then

(2.23) $\lim_{k \to \infty} \int_\Omega F(x, y^k(x), r^k(x)) dx = \int_\Omega F(x, y(x), r(x)) dx$.

Proof. We have

(2.24) $\int_\Omega F(x, y^k, r^k) dx - \int_\Omega F(x, y, r) dx = A_k + B_k$,

(2.25) $A_k := \int_\Omega F(x, y^k, r^k) dx - \int_\Omega F(x, y, r^k) dx$,

(2.26) $B_k := \int_\Omega F(x, y, r^k) dx - \int_\Omega F(x, y, r) dx.$

Since (r_k) converges to r in R, we have $B_k \to 0$. Since $y_i^k \to y_i$ in $L^{s_i}(\Omega;\mathbb{R}^N)$ strongly, $i=1,...,K$, we have also $\|y_i^k\| \to \|y_i\|$ in $L^{s_i}(\Omega)$ strongly, hence (see [5]) there exist functions $\bar{y}_i \in L^{s_i}(\Omega)$ such that (up to subsequences in k, same notation)

(2.27) $\|y_i^k(x)\| \le \bar{y}_i(x)$, in $\Omega - S_i^k$, $i=1,...,K$, if $s_i < +\infty$,

with $\mathrm{meas}(S_i^k)=0$. If $s_i = +\infty$, we have also, for $k \ge k_0$ (for some k_0)

(2.28) $\|y_i^k(x)\| \le C_i$, in $\Omega - S_i^k$,

with $\mathrm{meas}(S_i^k)=0$. We then have, for every $k \ge k_0$

(2.29) $$\left|F(x,y^k(x),r^k(x))\right| \le \Phi(x) + \Psi(x)\prod_{i=1}^{K}\xi_i(\|\bar{y}_i(x)\|) := \bar{F}(x),$$

in $\Omega - \underset{1\le i\le K,\, k\ge k_0}{\cup} S_i^k$, i.e. a.e. in Ω,

where $\bar{F} \in L^1(\Omega)$, by the (multiple) Hölder inequality. On the other hand, by Egorov's theorem, there exist subsequences (same notation) such that $y_i^k(x) \to y_i(x)$ $i=1,...,K$, a.a. in Ω. By the uniform continuity of F, for x fixed, on the compact set $B(x)\times U$, where $B(x)$ is a closed ball in $\mathbb{R}^{NK}$ with center $y(x)$ and containing $y^k(x)$ for every k (or for $k \ge k'$), we have, since $r^k \in M_1(U)$

(2.30) $$\begin{aligned}&\left|F(x,y^k(x),r^k(x)) - F(x,y(x),r^k(x))\right| \\ &\left|\int_U [F(x,y^k(x),u) - F(x,y(x),u)]r^k(du)\right| \\ &\int_U \left|F(x,y^k(x),u) - F(x,y(x),u)\right| r^k(du) \\ &\le \max_{u\in U}\left|F(x,y^k(x),u) - F(x,y(x),u)\right| \to 0, \quad \text{a.e. in } \Omega.\end{aligned}$$

The result follows then from Lebesgue's dominated convergence theorem and the uniqueness of the limit.

Theorem 2.2 Under Assumptions 2.1-3, the operator $r \mapsto y_r$ from R to V, and to $C_0(\bar{\Omega})$, and the functionals $r \mapsto G_m(r)$ on R, are continuous. If the relaxed problem has an admissible control (i.e. satisfying all the constraints), then it has a solution.

Proof. Let (r_k) be a sequence converging to r in R. Since the corresponding sequence of states (y_k) is bounded in V and in $C_0^\alpha(\bar{\Omega})$, for some $\alpha \in (0,1)$, and since the injection of $C_0^\alpha(\bar{\Omega})$ into $C_0(\bar{\Omega})$ is compact, there exists a subsequence (same notation) converging to some y in V weakly and in $C_0(\bar{\Omega})$ strongly. Let any $v \in V$ be given. By the state equation

(2.31) $$a(y_k,v) + \int_\Omega f(x,y_k(x),r_k(x))v\,dx = 0.$$

By the mean value theorem and since η_1 is increasing, we have, for every y with $|y| \le C$ (C defined in Theorem 2.1), and for some $\mu(x) \in [0,1]$

(2.32) $$\begin{aligned}\left|f(x,y,u)v\right| &\le \left|f(x,0,u)v\right| + \left|f(x,y,u)v - f(x,0,u)v\right| \\ &= \left|f(x,0,u)v\right| + \left|f_y(x,\mu(x)y,u)yv\right| \le |\phi_0(x)| + |\phi_1(x)||v|\eta_1(|\mu(x)y|)|y| \\ &\le |\phi_0(x)| + |\phi_1(x)||v|\eta_1(C)C\ ,\end{aligned}$$

Since $\phi_0 \in L^s$, $\phi_1 \in L^s$ or L^∞, $v \in V \subset L^2$, and $y_k \to y$ in L^∞, we can apply Proposition 2.1 to pass to the limit in the state equation for y_k and find that $y = y_r$. Next, we have

(2.33) $\alpha_2 \left\| y^n - y \right\|_1^2 \le a(y^n - y, y^n - y) = -(f(y^n, r^n), y^n) - a(y, y^n) - a(y^n - y, y) \to 0,$

since $y_k \to y$ in V weakly and $(f(y^n, r^n), y^n) \to (f(y, r), y)$ by Proposition 2.1, which shows that $y^n \to y$ in V strongly. The convergence of the original sequence follows from the uniqueness of the limit. The continuity of the functionals G_m on R follows then from Proposition 2.1. The existence of an optimal relaxed control follows then from the compactness of R and the continuity of the functionals G_m (the set of admissible controls a closed subset of R).

Note that the classical problem may have no classical solution, and since $W \subset R$, we generally have

(2.34) $c_R := \min_{\text{constraints on } r} G_0(r) \le \inf_{\text{constraints on } w} G_0(w) := c_W$,

where the equality holds, in particular, if there are no state constraints, as W is dense in R. Since usually approximation methods slightly violate the state constraints, approximating an optimal relaxed control by a relaxed or a classical control, hence the possibly lower relaxed optimal cost c_R, is not a drawback in practice (see [20], p. 259).

Lemma 2.1 Under Assumptions 2.1-4, dropping the index m in g_m, G_m, for $r, r' \in R$, the functional G is l-*differentiable* at r for every integer l, i.e. for every l and any choice of l controls $r_i \in R$, $i = 1, ..., l$, we have

(2.35) $G(r + \sum_{i=1}^{l} \varepsilon_i (r_i - r)) - G(r) = \sum_{i=1}^{l} DG(r, r_i - r)\varepsilon_i + o(\sum_{i=1}^{l} |\varepsilon_i|),$

for $\varepsilon_i \ge 0$, $\sum_{i=1}^{l} \varepsilon_i \le 1$,

with $DG(r, r_i - r) := \int_\Omega H(x, y, \nabla y, z, r_i(x) - r(x))dx,$

where the Hamiltonian is defined by

(2.36) $H(x, y, y', z, u) := -z\, f(y, x, u) + g(x, y, y', u),$

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

(2.37) $a(v, z) + (f_y(y, r)z, v) = (g_y(y, \nabla y, r), v) + (g_{y'}(y, \nabla y, r), \nabla v),$

$\forall v \in V$, with $y := y_r$.

In particular, the directional derivative of the functional G is given by

(2.38) $DG(r, \overline{r} - r) = \lim_{\alpha \to 0^+} \dfrac{G(r + \alpha(\overline{r} - r)) - G(r)}{\alpha}$

$= \int_\Omega H(x, y(x), \nabla y(x), z(x), r'(x) - r(x))dx$

Moreover, the operator $r \mapsto z_r$, from R to V, and the functional $(r, \overline{r}) \mapsto DG(r, \overline{r} - r)$, on $R \times R$, are continuous.

Proof: We first remark that by our assumptions, and since the injection $V \subset L^\rho$ is continuous, the functional

(2.39) $v \mapsto (g_y(y, \nabla y, r), v) + (g_{y'}(y, \nabla y, r), \nabla v)$

belongs to the dual V^* of V, and $f_y(y,r) \in L^s(\Omega)$, $2 \le s \le \infty$, $f_y(y,r) \ge 0$. Hence the linear adjoint equation has a unique solution $z \in V$, for every $r \in R$, by the Lax-Milgram theorem (if $s = \infty$), or by Lemma 3.2 in Ref. 3 (if $s < \infty$, no y' in g). Now let $r \in R$, $r_i \in R$, $\varepsilon_i \in (0,1)$, $i = 1,...,l$, $\varepsilon := (\varepsilon_1,...,\varepsilon_l)$, with $|\varepsilon| := \sum_{i=1}^{l} |\varepsilon_i| \le 1$, and set $r_\varepsilon := r + \sum_{i=1}^{l} \varepsilon_i (r_i - r)$, $y := y_r$, $y_\varepsilon := y_{r_\varepsilon}$, $\delta_\varepsilon y := y_\varepsilon - y$. From the state equation, we have

(2.40) $a(\delta y_\varepsilon, v) + (f(y_\varepsilon, r_\varepsilon) - f(y,r), v)$
$$= a(\delta y_\varepsilon, v) + (f(y_\varepsilon, r_\varepsilon) - f(y, r_\varepsilon), v) + (f(y, r_\varepsilon) - f(y,r), v) = 0.$$

By the mean value theorem, we see that δy_ε satisfies the linear equation

(2.41) $a(\delta y_\varepsilon, v) + (f_y(y + \mu \delta y_\varepsilon) \delta y_\varepsilon, v)) = -\sum_{i=1}^{l} (\varepsilon_i f(y, r_i - r), v)$, $\quad \forall v \in V$,

where the functions $\bar{a} := f_y(y + \mu \delta y_\varepsilon)$ (with $\bar{a} \ge 0$) and $\bar{f} := -\sum_{i=1}^{l} \varepsilon_i f(y, r_i - r)$ belong to $L^\infty(\Omega)$ (or L^s) and $L^s(\Omega)$, respectively, by our assumptions. It then follows from Lemma 3.2 in Ref. 3 that

(2.42) $\|\delta_\varepsilon y\|_1 + \|\delta_\varepsilon y\|_\infty \le c \|\bar{f}\|_{L^s} \le c' |\varepsilon|$.

Now, by our assumptions, for fixed $r \in R$, the functional on the open subset $Y \times L^2(\Omega, \mathbb{R}^d)$ of $L^\infty(\Omega) \times L^2(\Omega, \mathbb{R}^d)$

(2.43) $\Phi(y, y', r) := \int_\Omega g(x, y, y', r) dx,$

where

(2.44) $Y := \left\{ \phi \in L^\infty(\Omega) \,\middle|\, \|\phi\|_\infty < C' \right\}$,

is Fréchet differentiable *uniformly* in r, i.e.

(2.45) $\Phi(y + \delta y, y' + \delta y', r) - \Phi(y, y', r)$
$$= \int_\Omega [g_y(x, y, y', r) \delta y + g_{y'}(x, y, y', r) \delta y'] dx + \theta(\delta y, \delta y')(\|\delta y\|_\infty + \|\delta y'\|),$$

where $\theta(\delta y, \delta y') \to 0$ as $\|\delta y\|_\infty + \|\delta y'\| \to 0$, with θ *independent* of the control $r \in R$. This can be shown under our assumptions by using the mean value theorem in max-form, Hölder's inequality, and Proposition 2.1 for a fixed control. Using then the above estimate on $\delta_\varepsilon y$, we obtain

(2.46) $G(r_\varepsilon) - G(r)$
$$= \int_\Omega g_y(y, \nabla y, r_\varepsilon) \delta_\varepsilon y dx + \int_\Omega g_{y'}(y, \nabla y, r_\varepsilon) \nabla \delta_\varepsilon y dx + \sum_{i=1}^{l} \varepsilon_i \int_\Omega g(y, \nabla y, r_i - r) dx$$
$$+ \theta(\delta_\varepsilon y, \nabla \delta_\varepsilon y)(\|\delta_\varepsilon y\|_\infty + \|\nabla \delta_\varepsilon y\|)$$
$$= \int_\Omega g_y(y, \nabla y, r) \delta_\varepsilon y dx + \int_\Omega g_{y'}(y, \nabla y, r) \nabla \delta_\varepsilon y dx + \sum_{i=1}^{l} \varepsilon_i \int_\Omega g(y, \nabla y, r_i - r) dx$$
$$+ \theta(\delta_\varepsilon y, \nabla \delta_\varepsilon y)(\|\delta_\varepsilon y\|_\infty + \|\nabla \delta_\varepsilon y\|) + o(|\varepsilon|)$$

$$= \int_\Omega g_y \delta_\varepsilon y dx + \int_\Omega g_{y'} \nabla \delta_\varepsilon y dx + \sum_{i=1}^{l} \varepsilon_i \int_\Omega g(y, \nabla y, r_i - r) dx + o(|\varepsilon|).$$

Similarly, the state equation yields by linearization

(2.47) $$a(\delta_\varepsilon y, z) + (f_y(y,r)\delta_\varepsilon y, z) + \sum_{i=1}^{l} \varepsilon_i (f(y, r_i - r)\delta_\varepsilon y, z) + o(|\varepsilon|) = 0.$$

On the other hand, the adjoint equation yields

(2.48) $$a(\delta_\varepsilon y, z) + (f_y(y,r)z, \delta_\varepsilon y) = (g_y(y,r), \delta_\varepsilon y) + (g_{y'}(y,r), \nabla \delta_\varepsilon y).$$

Gathering the above results, we obtain

(2.49) $$G(r_\varepsilon) - G(r) = \sum_{i=1}^{l} \varepsilon_i \int_\Omega [-z\, f(y, r_i - r) + g(y, \nabla y, r_i - r)] dx + o(|\varepsilon|)$$

$$= \sum_{i=1}^{l} \varepsilon_i \int_\Omega H(x, y(x), \nabla y(x), z(x), r_i(x) - r(x)) dx + o(|\varepsilon|).$$

Finally, the continuity of the operator $r \mapsto z_r$ is proved by using the continuity of $r \mapsto y_r$, from R to L^∞, the compact injection $V \subset L^2$, and Proposition 2.1. The continuity of the functional $(r, \overline{r}) \mapsto DG(r, \overline{r} - r)$ follows from the above continuities.

The following theorem states the continuous relaxed necessary conditions for optimality.

Theorem 2.3 Under Assumptions 2.1-4, if $r \in R$ is optimal for Problem P, then r is *extremal*, i.e. there exist multipliers $\lambda_m \in \mathbb{R}$, $m = 0,...,q$, with

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

such that

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

(2.52) $\lambda_m G_m(r) = 0, \quad m = p+1,...,q$ (transversality conditions).

The global condition (2.51) is equivalent to the *strong relaxed pointwise minimum principle*

(2.53) $H(x, y(x), \nabla y(x), z(x), r(x)) = \min_{u \in U} H(x, y(x), \nabla y(x), z(x), u)$, a.e. in Ω,

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

Proof: The functionals G_m, $m = 0,...,q$, are continuous on R (Theorem 2.1) and, by Lemma 2.2, $(p+1)$-differentiable (cost and p equality constraints) at r. The global condition (i) and the transversality conditions (ii) follow then from the general multiplier theorem V.2.3 in [20] in unqualified form (G_m depends here on the control *only*, since y_r is unique for every r). The equivalence of the global and pointwise conditions is standard (see [20]) since U is closed (it has a dense denumerable subset).

Remark. In the absence of equality state constraints, it can be shown that, if the optimal control r is *regular*, i.e. there exists $r' \in R$ such that

(2.54) $G_m(r) + DG_m(r, r' - r) < 0$, $\quad m = p+1, ..., q$,

(*Slater condition*), then $\lambda_0 \neq 0$ for any multipliers as in Theorem 2.3.

3 Discretizations and behavior in the limit

We suppose in the sequel (Sections 3 and 4) that Ω is a polyhedron for simplicity. For each integer $n \geq 0$, let $\{E_i^n\}_{i=1}^{N^n}$ be an admissible regular partition of $\bar{\Omega}$ into d-simplices (elements), with $h^n = \max_i[\text{diam}(E_i^n)] \to 0$ as $n \to \infty$. Let $V^n \subset V$ be the subspace of functions that are continuous on $\bar{\Omega}$ and linear on each element E_i^n. The set of *discrete relaxed* (resp. *classical*) *controls* $R^n \subset R$ (resp. $W^n \subset W$) is defined as the subset of relaxed (resp. classical) controls that are equal, on each element $\overset{o}{E_i^n}$, to a constant probability measure on U (resp. a constant value in U). Clearly $W^n \subset R^n$. We endow R^n with the weak star topology of $M_1(U)^N$.

Remark. If Ω has an appropriately piecewise C^1 boundary Γ, one can approximate Γ by a polyhedral one Γ^n, with domain Ω^n, up to $o(h^n)$; the results of this section still hold in this case, with slight modifications in the definitions of V^n, W^n and in the proof of Lemma 3.2 (interpolation inside Ω^n and zero values on Γ^n).

The following assumptions are stronger than Assumptions 2.2-4.

Assumptions 3.1 The functions f, f_y (resp $g_m, g_{my}, g_{my'}$) are defined on $\Omega \times \mathbb{R} \times U$ (resp. on $\Omega \times \mathbb{R}^{d+1} \times U$), measurable for fixed y, u (resp. y, y', u), continuous for fixed x, and satisfy

(3.1) $\quad |f(x, y, u)| \leq c_1(1 + |y|^{\rho-1})$,

(3.2) $\quad 0 \leq f_y(x, y, u) \leq c_2(1 + |y|^{\rho-2})$,

(3.3) $\quad |g_m(x, y, y', u)| \leq c_3(1 + |y|^{\rho} + |y'|^2)$,

(3.4) $\quad |g_{my}(x, y, y', u)| \leq c_4(1 + |y|^{\rho-1} + |y'|^{\frac{2(\rho-1)}{\rho}})$,

(3.5) $\quad |g_{my'}(x, y, y', u)| \leq c_5(1 + |y|^{\frac{\rho}{2}} + |y'|)$,

$\quad \forall (x, y, y', u) \in \Omega \times \mathbb{R}^{d+1} \times U$,

where $c_i \geq 0$, $\rho \in [1, \infty)$ if $n = 1$ or 2, $\rho < \sigma := \dfrac{2n}{n-2}$ if $n \geq 3$. Note that each of the above inequalities is also satisfied if it holds for some $\bar{c}_i \geq 0$ and $\bar{\rho} \in [1, \rho)$.

For a given discrete control $r^n \in R^n$, the discrete state $y^n := y_{r^n}^n \in V^n$ is the solution of the discrete state equation

(3.6) $\quad a(y^n, v^n) + (f(y^n, r^n), v^n) = 0, \quad \forall v^n \in V^n$.

The following theorem can be proved by using the techniques in [14] (via Brouwer's fixed point theorem), under our coercivity, monotonicity and continuity assumptions.

Theorem 3.1 Under Assumptions 2.1 and 3.1 (on f, f_y), for every control $r^n \in R^n$, the discrete state equation has a unique solution $y^n \in V^n$.

The discrete state equation, which is a nonlinear system, can be solved by iterative methods. The discrete functionals are defined by

(3.7) $\quad G_m^n(r^n) = \int_\Omega g_m(y^n, \nabla y^n, r^n)dx, \quad m = 0,...,q.$

The discrete control constraint is $r^n \in R^n$ and the discrete state constraints are *either* of the two following ones

(3.8) $\quad$ Case (a) $\quad \left|G_m^n(r^n)\right| \le \varepsilon_m^n, \;\; m = 1,..., p,$

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

and

(3.10) $\quad G_m^n(r^n) \le \varepsilon_m^n, \;\; \varepsilon_m^n \ge 0, \;\; m = p+1,...,q,$

where the *feasibility perturbations* ε_m^n are chosen numbers converging to zero, to be defined later. The *discrete relaxed optimal control Problem* P_a^n (resp. P_b^n) is to minimize $G_m^n(r^n)$ subject to $r^n \in R^n$ and the above state constraints, Case (a) (resp. Case (b)).

The proof of the following theorem parallels that of Theorem 2.1, noting that all norms are equivalent in the finite dimensional space V^n.

Theorem 3.2 Under Assumptions 2.1 and 3.1 (on f, f_y), the operator $r^n \mapsto y^n$, from R^n to V^n, are continuous. Under Assumptions 2.1 and 3.1 (on f, f_y, g_m), the functionals $r^n \mapsto G_m^n(r^n)$, on R^n, are continuous, and for every n, if Problem P_a^n, or P_b^n, is feasible, then it has a solution.

The proofs of the following lemma and theorem also parallel the continuous case.

Lemma 3.1 Under Assumptions 2.1 and 3.1, dropping m in the functionals, G^n is l-differentiable for every l, and its directional derivative is given for $r^n, \overline{r}^n \in R^n$ by

(3.11) $\quad DG^n(r^n, \overline{r}^n - r^n) = \int_\Omega H(x, y^n, \nabla y^n, z^n, \overline{r}^n - r^n)dx,$

where the discrete adjoint state $z^n := z^n{}_{r^n} \in V^n$ satisfies the linear discrete adjoint equation

(3.12) $\quad a(z^n, v^n) + (z^n f_y(y^n, r^n), v^n) = (g_y(y^n, \nabla y^n, r^n), v^n) + (g_{y'}(y^n, \nabla y^n, r^n), \nabla v^n),$

$\qquad \forall v^n \in V^n, \quad$ where $y^n := y^n{}_{r^n}$.

Moreover, the operator $r^n \mapsto z^n{}_{r^n}$, from R^n to V^n, and the functional $(r^n, \overline{r}^n) \mapsto DG^n(r^n, \overline{r}^n - r^n)$, on $R^n \times R^n$, are continuous.

Theorem 3.3 Under Assumptions 2.1 and 3.1, if $r^n \in R^n$ is optimal for Problem P_b^n, then r^n is *discrete extremal*, i.e. there exist multipliers $\lambda_m^n \in \mathbb{R}$, $m = 0,...,q$, with $\lambda_m^n \geq 0$, $\lambda_m^n \geq 0$, $m = p+1,...,q$, $\sum_{m=0}^{q} \left|\lambda_m^n\right| = 1$, such that

(3.14) $\sum_{m=0}^{q} \lambda_m^n DG_m^n(r^n, \bar{r}^n - r^n) = \int_\Omega H^n(y^n, \nabla y^n, z^n, \bar{r}^n - r^n)dx \geq 0, \quad \forall \bar{r}^n \in R^n,$

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

where H^n and z^n are defined with $g := \sum_{m=0}^{q} \lambda_m^n g_m$. The global condition (3.14) is equivalent to the *strong discrete relaxed elementwise minimum principle*

(3.16) $\int_{E_i^n} H^n(y^n, \nabla y^n, z^n, r^n)dx = \min_{u \in U} \int_{E_i^n} H^n(y^n, \nabla y^n, z^n, u)dx, \quad i = 1,..., N^n.$

The following control approximation result is proved similarly to [10] (see also [18]).

Proposition 3.2 For every $r \in R$, there exists a sequence $(w^n \in W^n \subset R^n)$ of discrete classical controls, considered as relaxed ones, that converges to r in R.

The following key lemma gives consistency results.

Lemma 3.2 We suppose that Assumptions 2.1 and 3.1 are satisfied and drop m in the functionals.

(i) If the sequence $(r^n \in R^n)$ converges to $r \in R$ in R, then $y^n \to y_r$ in V strongly, $G^n(r^n) \to G(r)$, and $z^n \to z_r$ in $L^\rho(\Omega)$ strongly (and in V strongly, if the functionals do not depend on ∇y).

(ii) If the sequences $(r^n \in R^n)$ and $(\bar{r}^n \in R^n)$ converge to r and $\bar{r}$, respectively, in R, then

(3.17) $DG^n(r^n, \bar{r}^n - r^n) \to DG(r, \bar{r} - r).$

Proof: (i) From the discrete state equation, we have

(3.18) $a(y^n, y^n) + (f(y^n, r^n) - (f(0, r^n), y^n - 0) = -(f(0, r^n), y^n),$

and since f is increasing in y

(3.19) $\alpha_2 \left\|y^n\right\|_1^2 \leq a(y^n, y^n) \leq \left|(f(0, r^n), y^n)\right| \leq \left\|f(0, r^n)\right\|_s \left\|y^n\right\| \leq c \left\|y^n\right\|_1,$

which shows that the sequence (y^n) is bounded in V. By Alaoglu's theorem, there exists a subsequence (same notation) that converges weakly in V to some $y \in V$, and since the injection of V in $L^\rho(\Omega)$ is compact (see [1]), we can suppose that $y^n \to y$ in $L^\rho(\Omega)$ strongly. For any given $v \in C_0^1(\bar{\Omega})$, let $(v^n \in V^n)$ be the sequence of interpolates of v at the vertices of the partition of Ω. This sequence converges to v in $C_0^1(\bar{\Omega})$ (hence in V) strongly. We have

(3.20) $a(y^n, v^n) + (f(y^n, r^n), v^n) = 0.$

Since $r^n \to r$ in R and $y^n \to y$ in V strongly, hence in $L^\rho(\Omega)$ strongly, by Proposition 2.1 and our assumptions, we can pass to the limit in this equation and find

(3.21) $a(y,v)+(f(y,r),v)=0,$

which holds also for every $v\in V\subset L^s$, as $C_0^1(\bar{\Omega})$ is dense in V. Therefore $y=y_r$. The convergence in $L^\rho(\Omega)$ strongly of the initial sequence follows then from the uniqueness of the limit. Next, we have

(3.22) $\alpha_2\left\|y^n-y\right\|_1^2\le a(y^n-y,y^n-y)=(-f(y^n,r^n),y^n)-a(y,y^n)-a(y^n-y,y).$

By Proposition 2.1 and the above convergences of (y^n), the last expression converges to zero; hence $y^n\to y$ in V strongly. The convergence $G^n(r^n)\to G(r)$ follows from the above convergences and the same proposition. From the adjoint equation, we have

(3.23)
$$\begin{aligned}\alpha_2\left\|z^n\right\|_1^2 &\le a(z^n,z^n)+(f_y(y^n,r^n)z^n,z^n)\\ &\le\left|(g_y(y^n,\nabla y^n,z^n,r^n),z^n)\right|+\left|(g_{y'}(y^n,\nabla y^n,z^n,r^n),\nabla z^n)\right|\\ &\le(c_4(1+\left|y^n\right|^{p-1}+\left|\nabla y^n\right|^{\frac{2(p-1)}{p}}),\left|z^n\right|)+(c_5(1+\left|y^n\right|^{\frac{p}{2}}+\left|\nabla y^n\right|),\left|z^n\right|)\\ &\le c'_4(1+\left\|y^n\right\|_{L^\rho}^{\rho-1}+\left\|\nabla y^n\right\|^{\frac{2(\rho-1)}{\rho}})\left\|z^n\right\|_{L^\rho}+c'_5(1+\left\|y^n\right\|_{L^\rho}^{\frac{\rho}{2}}+\left\|\nabla y^n\right\|)\left\|z^n\right\|_1\le c\left\|z^n\right\|_1,\end{aligned}$$

which shows that (z^n) is bounded in V. The continuity of $r\mapsto z_r$ from R to $L^2(\Omega)$ is then shown similarly to that of $r\mapsto y_r$, using also that continuity and Proposition 2.1. If the functionals do not depend on $y':=\nabla y$, then the continuity $r\mapsto z_r$ from R to V is proved similarly to the continuity of $r\mapsto y_r$ from R to V.

(ii) The convergence here follows from the results of (i) and Proposition 2.1.

Remark. Suppose that $r^n\to r$ in R and that Assumptions 2.1, 2.2 (instead of 3.1) are satisfied. Under the strong assumptions of [8], it can be shown that $y^n_{r^m}\to y_{r^m}$ uniformly as $n\to\infty$, for m fixed. On the other hand, $y_{r^m}\to y_r$ also uniformly by Theorem 2.2. Therefore, under all the above assumptions, $y^n_{r^n}\to y_r$ uniformly as $n\to\infty$. Similar remarks hold for the convergence of the functionals, adjoints and functional derivatives.

We suppose in the following that Problem P is feasible. The following (theoretical, in the presence of state constraints) theorem examines the behavior in the limit of optimal discrete controls.

Theorem 3.4 If there are state constraints, we suppose that the sequences (ε_m^n) in the discrete state constraints (Case (a)) converge to zero as $n\to\infty$ and satisfy

(3.24) $\left|G_m^n(\tilde{r}^n)\right|\le\varepsilon_m^n,\quad m=1,...,p,\quad G_m^n(\tilde{r}^n)\le\varepsilon_m^n,\quad \varepsilon_m^n\ge 0,\ m=p+1,...,q,$

for every n, where $(\tilde{r}^n\in W^n)$ is a sequence converging in R to an optimal control $\tilde{r}\in R$ of Problem P_a^n. For each n, let r^n be optimal for Problem P_a^n. Then every accumulation point of (r^n) is optimal for Problem P.

Proof: Note that our assumption implies that the discrete problems are feasible for every n. Let (r^n) be a subsequence (same notation) that converges to some $r\in R$ (Proposition 3.2). Since r^n is optimal, hence admissible, and $\tilde{r}^n$ is admissible, for Problem P_a^n, we have

(3.25) $G_0^n(r^n) \le G_0^n(\tilde{r}^n), \quad \left|G_m^n(r^n)\right| \le \varepsilon_m^n, \; m = 1,..., p, \quad G_m^n(r^n) \le \varepsilon_m^n, \; m = p+1,..., q.$

Passing to the limit and using Lemma 3.2, we see that r is optimal for Problem P. In the absence of state constraints, by taking a sequence converging to some optimal control of Problem P, we readily obtain in the limit that r is optimal for Problem P.

Next, we study the behavior in the limit of extremal discrete controls. Consider the discrete problems P_b^n. We shall construct sequences of perturbations (ε_m^n) that converge to zero and such that the discrete problem is feasible for every n. Let $r'^n \in R^n$ be any solution of the following *auxiliary minimization problem without state constraints*

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

and set

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

Let $\tilde{r}$ be an admissible control for Problem P, and $(\tilde{r}^n \in R^n)$ a sequence converging to $\tilde{r}$ (Proposition 3.2). We have

(3.28) $\lim_{n\to\infty} [G_m^n(\tilde{r}^n)]^2 = [G_m(\tilde{r})]^2 = 0, \quad m = 1,..., p,$

(3.29) $\lim_{n\to\infty} [\max(0, G_m^n(\tilde{r}^n))]^2 = [\max(0, G_m(\tilde{r}))]^2 = 0, \quad m = p+1,..., q,$

which imply a fortiori that $c^n \to 0$, hence $\varepsilon_m^n \to 0$, $m = 1,..., q$. Then clearly Problem P_b^n is feasible for every n, for these ε_m^n.

We suppose in the following that the perturbations ε_m^n are chosen as in the above *minimum feasibility* procedure. Note that in practice we usually have $c^n = 0$, for sufficiently large n, due to sufficient discrete controllability, in which case the perturbations ε_m^n vanish, i.e. the discrete problem with zero perturbations is feasible. Also, see [7] and [13] for a study on how the perturbations ε_m^n can be practically chosen to be zero, if there are only inequality state constraints.

Theorem 3.5 For each n, let r^n be admissible and extremal for Problem P_b^n. Then every accumulation point of the sequence (r^n) is admissible and extremal for Problem P.

Proof: Suppose that a subsequence (r^n) (same notation) converges to some $r \in R$. For each n, let λ_m^n, $m = 0,..., q$ be multipliers as in Theorem 3.2. Since $\sum_{m=0}^{q} \left|\lambda_m^n\right| = 1$, the sequences (λ_m^n) are bounded, and by extracting a subsequence, we can suppose that $\lambda_m^n \to \lambda_m$, $m = 0,..., q$. By Lemma 3.2 and Proposition 2.1, we then obtain, for any given $\overline{r} \in R$ and $\overline{r}^n \to \overline{r}$

(3.30) $\sum_{m=0}^{q} \lambda_m DG_m(r, \overline{r} - r) = \lim_{n\to\infty} \sum_{m=0}^{q} \lambda_m^n DG_m^n(r^n, \overline{r}^n - r^n) \ge 0,$

(3.31) $\lambda_m G_m(r) = \lim_{n\to\infty} \lambda_m^n [G_m^n(r^n) - \varepsilon_m^n] = 0, \quad m = p+1,..., q,$

(3.32) $G_m(r)=\lim_{n\to\infty}[G_m^n(r^n)-\varepsilon_m^n]=0, \quad m=1,...,p,$

(3.33) $G_m(r)=\lim_{n\to\infty}[G_m^n(r^n)-\varepsilon_m^n]\le 0, \quad m=p+1,...,q,$

and $\lambda_0\ge 0$, $\lambda_m\ge 0$, $m=p+1,...,q$, $\sum_{m=0}^{q}|\lambda_m|=1$, which show that r is admissible and extremal for Problem P.

4 Discrete relaxed penalized conditional 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*

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

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 case of the progressively refining version, we suppose that each element $E_{i'}^{n+1}$ is a subset of some element E_i^n (if Ω is polyhedral). In this case, we have $R^n\subset R^{n+1}$, and thus a control $r^n\in R^n$ may be considered also as belonging to R^{n+1} (see Step 3), hence the computation of states, adjoints and cost derivatives for this control, but with the possibly finer discretization $n+1$, makes sense. (In this version, and if Ω is not polyhedral, one has to modify slightly near the boundary the control r_k^{nl}, at the end of Step 3, before going to Step 2, and if the discretization has been refined). The *discrete relaxed penalized conditional descent methods* are described by the following Algorithm.

Algorithm

Step 1. Set $k:=0$, $l:=1$, choose a value of n and an initial control $r_0^{n1}\in R^n$.

Step 2. Find $\overline{r}_k^{nl}\in R^n$ such that

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

Step 3. If $|d_k|\le\beta^l$, set $r^{nl}:=r_k^{nl}$, $\overline{r}^{nl}:=\overline{r}_k^{nl}$, $d^l:=d_k$, $l:=l+1$, $[n:=n+1]$, and go to Step 2.

Step 4. (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 the inequality

(4.3) $G^{nl}(r_k^{nl}+\alpha(s)(\overline{r}_k^{nl}-r_k^{nl}))-G^{nl}(r_k^{nl})\le\alpha(s)b'd_k$,

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

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

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

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

In this Algorithm, we consider two versions:

Version A. $[n = n+1]$ is *skipped* in Step 3: n is a constant integer chosen in Step 1, i.e. we choose a *fixed discretization*, and replace the discrete functionals G_m^n by the perturbed ones $\tilde{G}_m^n := G_m^n - \varepsilon_m^n$, in which case the method is applied to Problem P_b^n.
Version B. $[n = n+1]$ is *not skipped* in Step 3: we have a *progressively refining* discrete method, i.e. $n \to \infty$ (see proof of Theorem 4.1 below), in which case we can take $n = 1$ in Step 1, hence $n = l$ in the Algorithm.

The progressively refining version has the advantage of reducing computing time and memory, and also of avoiding the computation of minimum feasibility perturbations ε_m^n (see Section 3). 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.

One can easily see that a *classical* control $\overline{r}_k^{nl}$ in Step 2 can be found for every k by minimizing on U the integral (practically using some numerical integration rule) on the element E_i^n

(4.5) $\quad \int_{E_i^n} H(x, y^n, \nabla y^n, u)dx$

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

A (continuous or discrete) extremal control is called *abnormal* if there exist multipliers as in the corresponding optimality conditions, but with $\lambda_0 = 0$ (or $\lambda_0^n = 0$). A control is admissible *and* abnormal extremal in exceptional, degenerate, situations (see [20]).

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

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

Theorem 4.1 We suppose that Assumptions 2.1, 3.1-4 are satisfied.
(i) In Version B, if (r^{nl}) is a subsequence of the sequence generated by the Algorithm in Step 3 that converges to some $r \in R$, as $l \to \infty$ (hence $n \to \infty$). If the sequences of multipliers (λ_m^{nl}) are bounded, then r is admissible and weakly extremal classical for Problem P.
(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 Problem P_b^n.
(iii) In the above convergence case (i) (resp. (ii)), suppose that Problem P (resp. P_b^n) has no admissible, abnormal extremal, controls. If the limit control is admissible, then the sequences of multipliers are bounded, and this limit control is extremal as above.
Proof. We shall first show that $l \to \infty$ in the Algorithm. Suppose, on the contrary, that l, hence n (in both Versions A, B), remains constant after a finite number of iterations in k, and so we drop here the indices l, n. Let us show that then $d_k \to 0$. Since R^n is compact, let $(r_k)_{k \in K}$, $(\overline{r}_k)_{k \in K}$ be subsequences of the sequences generated in Steps 2 and 5 such that $r_k \to \tilde{r}$, $\overline{r}_k \to \tilde{\overline{r}}$, in R^n, as $k \to \infty$, $k \in K$. Clearly, by Step 2, $d_k \le 0$ for every k, hence

(4.7) $d := \lim_{k\to\infty,\, k\in K} d_k = DG(\tilde{r}, \tilde{\bar{r}} - \tilde{r}) \le 0.$

Suppose that $d < 0$. The function $\Phi(\alpha) := G(r + \alpha(r' - r))$ is continuous on $[0,1]$. Since the directional derivative $DG(r, r' - r)$ is linear w.r.t. $r' - r$, Φ is differentiable on $(0,1)$ and has derivative

(4.8) $\Phi'(\alpha) = DG(r + \alpha(r' - r), r' - r).$

Using the mean value theorem, we have, for each $\alpha \in (0,1]$

(4.9) $G(r_k + \alpha(\bar{r}_k - r_k)) - G(r_k) = \alpha DG(r_k + \alpha'(\bar{r}_k - r_k), \bar{r}_k - r_k),$

for some $\alpha' \in (0, \alpha)$. Therefore, for $\alpha \in [0,1]$, by Lemma 3.1

(4.10) $G(r_k + \alpha(\bar{r}_k - r_k)) - G(r_k) = \alpha(d + \varepsilon_{k\alpha}),$

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) $d + \varepsilon_{k\alpha} \le b(d + \eta_k) = b' d_k,$

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

(4.12) $G(r_k + \alpha(\bar{r}_k - r_k)) - G(r_k) \le \alpha b' d_k,$

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

(4.13) $G(r_{k+1}) - G(r_k) = G(r_k + \alpha_k(\bar{r}_k - r_k)) - G(r_k) \le \alpha_k b' d_k \le c\bar{\alpha} b' d_k \le c\alpha b' d/2,$

for $k \ge \bar{k}$, $k \in K$. It follows that $G(r_k) \to -\infty$ as $k \to \infty$, $k \in K$. This contradicts the fact that $G(r_k) \to G(\tilde{r})$ as $k \to \infty$, $k \in K$, by Lemma 3.1. Therefore, we must have $d = 0$ and $d_k \to 0$ for the whole sequence, since the limit 0 is unique. But Step 3 then implies that $l \to \infty$, which is a contradiction. Therefore, $l \to \infty$. This shows also that $n \to \infty$ in Version B.

(i) Let (r^{nl}) be a subsequence (same notation) of the sequence generated by the Algorithm in Step 3 that converges to some $r \in R$ as $l, n \to \infty$. Suppose that the sequences (λ_m^{nl}) are bounded and (up to subsequences) that $\lambda_m^{nl} \to \lambda_m$. By Lemma 3.2, we have

(4.14) $0 = \lim_{l\to\infty} \dfrac{\lambda_m^{nl}}{M_m^l} = \lim_{l\to\infty} G_m^n(r^{nl}) = G_m(r), \quad m = 1, \dots, p,$

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

which show that r is admissible. Now let any $\tilde{r} \in R$ and, by Proposition 3.2, let $(\tilde{r}^n \in R^n)$ be a sequence of discrete controls that converges to $\tilde{r}$. Let (λ_m^{nl}) be subsequences (same notation) such that $\lambda_m^{nl} \to \lambda_m$. By Step 2, we have

(4.16) $\sum_{m=0}^{q} \lambda_m^{nl} DG_m^n(r^{nl}, \tilde{r}^n - r^{nl}) \ge d^l,$

where $\lambda_0^{nl} := 1$. Since $|d^l| \le \beta^l$ by Step 3, we have $d^l \to 0$. By Lemma 3.2, we can pass to the limit as $l, n \to \infty$ in the above inequality and obtain the first optimality condition

(4.17) $\sum_{m=0}^{q} \lambda_m DG_m(r, \tilde{r} - r) \ge 0, \quad \forall \tilde{r} \in R.$

By construction of the λ_m^{nl}, we clearly have $\lambda_0 = 1$, $\lambda_m \ge 0$, $m = p+1,...,q$, $\sum_{m=0}^{q} |\lambda_m| := c \ge 1$, and we can suppose that $\sum_{m=0}^{q} |\lambda_m| = 1$, by dividing the above inequality by c. On the other hand, if $G_m(r) < 0$, for some index $m \in [p+1,q]$, then for sufficiently large l we have $G_m^{nl}(r^{nl}) < 0$ and $\lambda_m^l = 0$, hence $\lambda_m = 0$, i.e. the transversality conditions hold. Therefore, r is also extremal.

(ii) The admissibility of the limit control r^n is proved as in (i). Passing here to the limit in the inequality resulting from Step 2, as $l \to \infty$, for n fixed, and using Theorem 3.1 and Lemma 3.1, we obtain, similarly to (i)

(4.18) $\sum_{m=0}^{q} \lambda_m D\tilde{G}_m^n(r^n, \tilde{r}^n - r^n) = \sum_{m=0}^{q} \lambda_m DG_m^n(r^n, \tilde{r}^n - r^n) \ge 0, \quad \forall r'^n \in R^n,$

and the discrete transversality conditions

(4.19) $\lambda_m^n \tilde{G}_m^n(r^n) = \lambda_m^n [G_m^n(r^n) - \varepsilon_m^n] = 0, \quad m = p+1,...,q,$

with multipliers λ_m^n as in the discrete optimality conditions.

(iii) In either of the two above convergence cases (i), (ii), suppose that the limit control is admissible and that the limit problem has no admissible, abnormal extremal, controls. Suppose that the multipliers are not all bounded. Then, dividing the corresponding 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) or (ii), this limit control is extremal as above.

By choosing moderately growing sequences (M_m^l) and a sequence (β^l) relatively fast converging to zero, the resulting sequences of multipliers (λ_m^{nl}) 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 $\zeta_k := \alpha_{k-1}$, for $k \ge 1$.

The Algorithm can be practically implemented as follows. Suppose that the integrals involving f and g_m, $m = 0,...,q$, are calculated with sufficient accuracy by some numerical integration rule, involving (usually a small number) μ of nodes x_{ji}^n, $j = 1,...,\mu$, on each element E_i^n, of the form

(4.20) $\int_{E_i^n} \phi(x)dx \approx \text{meas}(E_i^n) \sum_{j=1}^{\mu} C_j \phi(x_{ji}^n)$.

We first choose the initial discrete control in Step 1 to be of Gamkrelidze type, i.e. equal on each E_i^n to a convex combination of $(\mu+q+1)+1$ Dirac measures on U concentrated at $(\mu+q+1)+1$ points of U. Suppose, by induction, that the control r_k^{nl} computed in the Algorithm is of Gamkrelidze type. Since the control $\overline{r}_k^{nl}$ in Step 2 is chosen to be classical (see above), i.e. elementwise Dirac, the resulting control $\tilde{r}_k^{nl} := (1-\alpha_k) r_k^{nl} + \alpha_k \overline{r}_k^{nl}$ in Step 5 is elementwise equal to a convex combination of $(\mu+q+1)+2$ Dirac measures. Using now a known property of convex hulls of finite vector sets, we can construct a Gamkrelidze control r_{k+1}^{nl} *equivalent* to $\tilde{r}_k^{nl}$, i.e. such that the following $\mu+q+1$ equalities (i.e. equality in $\mathbb{R}^{\mu+q+1}$) hold

(4.21) $f(x_{ji}^n, \tilde{y}^{nl}(x_{ji}^n), r_{k+1,i}^{nl}) = f(x_{ji}^n, \tilde{y}^{nl}(x_{ji}^n), \tilde{r}_{ki}^{nl})$, $\quad j = 1,...,\mu,$

(4.22) $\operatorname{meas}(E_i^n)\sum_{j=1}^{\mu} C_j g_m(x_{ji}^n, \tilde{y}^{nl}(x_{ji}^n), \nabla\tilde{y}^{nl}(x_{ji}^n), r_{k+1,i}^{nl})$

$$= \operatorname{meas}(E_i^n)\sum_{j=1}^{\mu} C_j g_m(x_{ji}^n, \tilde{y}^{nl}(x_{ji}^n), \nabla\tilde{y}^{nl}(x_{ji}^n), \tilde{r}_{ki}^{nl}), \quad m = 0,...,q,$$

for each $i = 1,...,M$, where $\tilde{y}^{nl}$ corresponds to $\tilde{r}_k^{nl}$, by selecting only $(\mu + q + 1) + 1$ appropriate points in U among the $(\mu + q + 1) + 2$ ones defining $\tilde{r}_k^{nl}$, for each i. Then the control r_{k+1}^{nl} clearly yields the *same* discrete state and functionals as $\tilde{r}_k^{nl}$. Therefore, the constructed control r_k^{nl} is of Gamkrelidze type for every k. Finally, discrete Gamkrelidze controls computed as above can then be approximated by piecewise constant classical controls using a standard procedure (see [12]), by subdividing here the elements in appropriate subelements whose measures are proportional to the Gamkrelidze coefficients.

5 Numerical examples

Let $\Omega = (0,1)\times(0,1)$ and consider the following examples.

Example 1. Define the reference classical control and state

(5.1) $\quad \bar{w}(x) := \min(1, -1 + 1.5(x_1 + x_2)), \quad \bar{y}(x) := 8x_1x_2(1 - x_1)(1 - x_2),$

and consider the optimal control problem with state equation

(5.2) $\quad -\Delta y + y^3/3 + (2 + w - \bar{w})y - \bar{y}^3/3 - 2\bar{y} - 16[x_1(1 - x_1) + x_2(1 - x_2)] = 0, \quad \text{in } \Omega,$

(5.3) $\quad y(x) = 0 \quad \text{on } \Gamma,$

nonconvex control constraint set $U := \{-1\} \cup [0,1]$, and nonconvex cost functional to be minimized

(5.4) $\quad G_0(w) := \int_\Omega \{0.5[(y - \bar{y})^2 + \|\nabla y - \nabla\bar{y}\|^2] - w^2 + 1\}dx.$

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

(5.5) $\quad r(x)\{1\} = [\bar{w}(x) - (-1)]/2 = \begin{cases} 1, & \text{if } -1 + 1.5(x_1 + x_2) \ge 1 \\ < 1, & \text{if } -1 + 1.5(x_1 + x_2) < 1 \end{cases}$

(5.6) $\quad r(x)\{-1\} = 1 - r(x)\{1\},$

for $x \in \Omega,$ with optimal state $\bar{y}$ and cost 0. We see that r is concentrated at the two points 1 and -1; r is classical ($\equiv 1$) if $-1 + 1.5(x_1 + x_2) \ge 1$, and non-classical otherwise. Note also that the optimal cost value 0 can be approximated as closely as possible by using a classical control, as W is dense in R, but cannot be attained for such a control because the control values $u \in (-1,0)$ (of $\bar{w}$) do not belong to U.

The Algorithm, without penalties, was first applied to this problem using triangular elements, which are half-squares of fixed edge size $h = 1/80$, the second order 3 edge-midpoints rule for numerical integration, with Armijo parameters $b' = c' = 0.5$, and constant initial control $r_0^n(x) := 0.5(\delta_{-1} + \delta_1)$, $x \in \Omega$, where δ_{-1}, δ_1 are the Dirac measures at -1 and 1. After 90 iterations in k, we obtained the following results:

(5.7) $\quad G_0^n(r_k^n) = 2.966250711804619\cdot 10^{-4}, \quad d_k = -6.733\cdot 10^{-7}, \quad \varepsilon_k = 3.331\cdot 10^{-4},$

where ε_k is the discrete max state error at the vertices of the triangles. Figure 1 shows the last control probability function $p_1(x_1,x_2):=r_k^n(x_1,x_2)\{1\}$, for $x_1=x_2$ (cross-section); we also obtained $p_{-1}(x_1,x_2):=r_k^n(x_1,x_2)\{-1\}=1-p_1(x_1,x_2)$.

Example 2. Introducing the state constraint

(5.8) $\quad G_1(u):=\int_\Omega (y-0.22)dx=0,$

in Example 1 and applying here the penalized Algorithm, we obtained, after 168 iterations in k, the control probability function $p_1(x_1,x_2):=r_k^{nl}(x_1,x_2)\{1\}$, for $x_1=x_2$, shown in Figure 2 and the results:

(5.9) $\quad G_0^n(r_k^{nl})=3.748940978325145\cdot 10^{-4},\quad G_1^n(r_k^{nl})=7.335\cdot 10^{-7},$

$\qquad d_k=-1.127\cdot 10^{-5}.$

Note that the continuous relaxed problem here is feasible, as $G_1(r_{-1})\approx 0.017>0$, $G_1(r_1)\approx -0.0028<0$, where $r_{-1}(x):=\delta_{-1}$, $r_1(x):=\delta_1$, $x\in\Omega$, and the function $\phi(\lambda):=G_1(\lambda r_{-1}+(1-\lambda)r_1)$ is continuous on $[0,1]$.

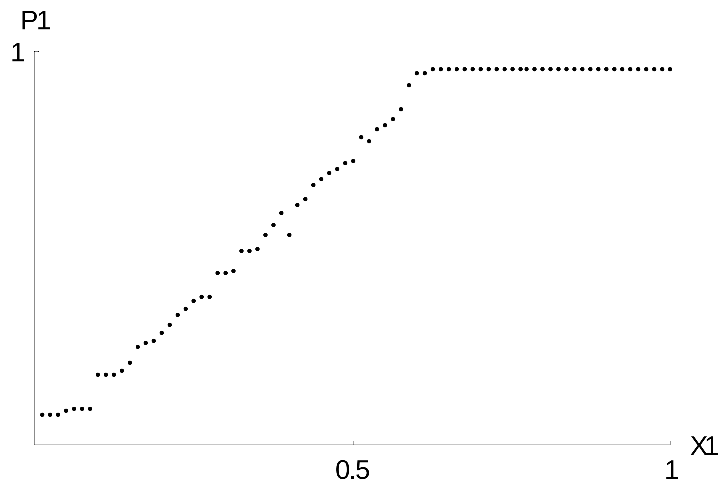

Figure 1. Example 1: Last relaxed control probability p_1, for $x_1=x_2$

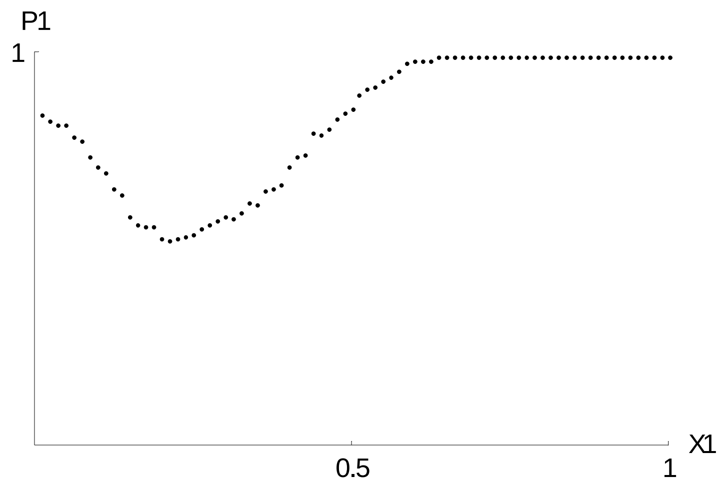

Figure 2. Example 2: Last relaxed control probability p_1, for $x_1=x_2$

Finally, the progressively refining version of the algorithm was also applied to the above problems, with successive step sizes $h = 1/20, 1/40, 1/80$, in three equal iteration periods, and yielded results of similar accuracy, but required here less than half the computing time.

References

[1] Adams, R. A., *Sobolev Spaces*, Academic Press, New York, 1975.
[2] Arnäutu, V., and Neittaanmäki, P., Discretization Estimates for an Elliptic Control Problem, Numer. Funct. Anal. Optim., Vol. 19, pp. 431-464, 1998.
[3] Bartels, S., Adaptive approximation of Young measure solution in scalar non-convex variational problems, SIAM J. Numer. Anal., Vol. 42, pp. 505-629, 2004.
[4] Bonnans, J., and Casas, E., Un principe de Pontryagine pour le contrôle des systèmes semilineaires elliptiques, J. Differ. Equat., Vol. 90, pp. 288-303, 1991.
[5] Brezis, H., *Analyse Fonctionnelle, Théorie et Applications*, Masson, Paris, 1983.
[6] Cartensen, C., and Roubíček, T., Numerical approximation of Young measure in nonconvex variational problems, Numer. Math., Vol. 84, pp. 395-415, 2000.
[7] Casas, E., The relaxation theory applied to Optimal control problems of semilinear elliptic equations, Proc. of the 17^{th} Conf. on Systems, Modeling and Optimization, Prague (1995), Chapman and Hall, 1996.
[8] Casas, E., and Mateos, M., Uniform convergence of the FEM. Applications to state constrained control problems. Comp. Appl. Math. **2002** *21* (1), 67-100.
[9] Chryssoverghi, I., Nonconvex Optimal Control of Nonlinear Monotone Parabolic Systems, Systems Control Lett., Vol. 8, pp. 55-62, 1986.
[10] Chryssoverghi, I., and Bacopoulos, A., Approximation of relaxed nonlinear parabolic optimal control problems, J. Optim. Theory Appl., Vol. 77, No. 1, pp. 31-50, 1993.
[11] Chryssoverghi, I., and Kokkinis, B., Discretization of Nonlinear Elliptic Optimal Control Problems, Systems Control Lett., Vol. 22, pp. 227-234, 1994.
[12] Chryssoverghi, I., Bacopoulos, A., Kokkinis, B., and Coletsos, J., Mixed Frank-Wolfe penalty method with applications to nonconvex optimal control problems, J. Optim. Theory Appl., Vol. 94, No. 2, pp. 311-334, 1997.
[13] Chryssoverghi, I., Bacopoulos, A., Coletsos, J. and Kokkinis, B., Discrete Approximation of Nonconvex Hyperbolic Optimal Control Problems with State Constraints, Control Cybernet., Vol. 27, No. 1, pp. 29-50, 1998.
[14] Lions, J.-L., *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*, Dunod, Paris, 1969 (English translation by Le Van, Holt, Rinehart and Winston, New York, 1972).
[15] Mach, J., Numerical solution of a class of nonconvex variational problems by SQP, Numer. Funct. Anal. Optim., Vol. 23, pp. 573-587, 2002.
[16] Mataché, A.-M., Roubíček, T., and Schwab, C., Higher-order convex approximations of Young measures in optimal control, Adv. Comput. Math., Vol. 19, pp. 73-79, 2003.
[17] Nicolaides, R.A. and Walkington, N.J., Strong convergence of numerical solutions to degenerate variational problems, Math. Comp., Vol. 64, pp. 117-127, 1995.
[18] Roubíček, T., A convergent computational method for constrained optimal relaxed control problems, J. Optim. Theory Appl., Vol. 69, pp. 589-603, 1991.

[19] Roubíček, T., *Relaxation in Optimization Theory and Variational Calculus*, Walter de Gruyter, Berlin, 1997.
[20] Warga, J., *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.
[21] Warga, J., Steepest descent with relaxed controls, SIAM J. Control Optim., Vol. 15, pp. 674-682, 1977.