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

I. Chryssoverghi<sup>1</sup>, J. Geiser<sup>2</sup>, J. Al-Hawasy<sup>1</sup>

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

#### Abstract

We consider an optimal control problem described by a second order elliptic boundary value problem, jointly nonlinear in the state and control, with control and state constraints, where the state constraints and cost functionals involve also the state gradient. Since this problem may have no classical solutions, it is also formulated in the relaxed form. The classical problem is discretized by using a finite element method for state approximation, while the controls are approximated by elementwise constant, or linear, or multilinear, controls. Various necessary conditions for optimality are given for the classical and the relaxed problem, in the continuous and the discrete case. We then study the behavior in the limit of discrete optimality, and of discrete problem, and also a progressively refining version of this method to the continuous classical problem. We prove that accumulation points of sequences generated by the first method are extremal for the discrete problem, and that strong classical (resp. relaxed) accumulation points of sequences of discrete controls generated by the second method are admissible and weakly extremal classical (resp. relaxed) for the continuous classical (resp. relaxed) problem. Finally, numerical examples are given.

**Keywords.** Optimal control, nonlinear elliptic systems, state constraints, discretization, finite elements, discrete penalized gradient projection method, progressive refining.

### **1** Introduction

We consider an optimal control problem described by a second order elliptic boundary value problem, which is jointly nonlinear in the state and control, with control and state constraints, where the state constraints and cost functionals involve also the gradient of the state. The problem is discretized by using a Galerkin finite element method with continuous elementwise linear basis functions for state approximation, while the controls are approximated by (not necessarily continuous) elementwise constant, or linear, or multilinear, controls. Various necessary conditions for optimality are given for the classical and the relaxed problem, in the continuous and the discrete case. Under appropriate assumptions, we prove that strong accumulation points in  $L^2$  of sequences of optimal (resp. admissible and extremal) discrete controls are optimal (resp. admissible and weakly extremal classical) for the continuous classical problem, and that relaxed accumulation points of sequences of optimal (resp. admissible and extremal) discrete controls are optimal (resp. admissible and weakly extremal relaxed) for the continuous relaxed problem. We then apply a penalized gradient projection method to each discrete problem, and also a corresponding discrete method to the continuous classical 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 first method are extremal for the discrete problem, and that strong classical (resp. relaxed) accumulation points of sequences of discrete controls generated by the second method are admissible and weakly extremal classical (resp. relaxed) for the continuous classical (resp. relaxed) problem. Finally, numerical examples are given. For approximation and optimization methods applied to distributed optimal control problems, see e.g. [2], [5,6], [8-12], [16-18], and the references therein.

### 2. The continuous optimal control problems

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

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

(2.2) y(x) = 0 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  $\Omega$ , where U is a compact subset of  $\mathbb{R}^{\nu}$ , the state constraints are

(2.4) 
$$G_m(w) \coloneqq \int_{\Omega} g_m(x, y(x), \nabla y(x), w(x)) dx = 0, \quad m = 1, ..., p,$$
  
(2.5)  $G_m(w) \coloneqq \int g_m(x, y(x), \nabla y(x), w(x)) dx \le 0, \quad m = p + 1, ..., q,$ 

and the cost functional is

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

The state equation will be interpreted in the following weak form

(2.7) 
$$y \in V$$
, and  $a(y,v) + \int_{\Omega} b(x, y(x), w(x))v(x)dx = \int_{\Omega} f(x, w(x))v(x)dx$ ,  $\forall v \in V$ ,  
where  $a(\cdot, \cdot)$  is the usual bilinear form associated with A and defined on  $V \times V$ 

(2.8) 
$$a(y,v) \coloneqq \sum_{i,j=1}^{d} \int_{\Omega} a_{ij}(x) \frac{\partial y}{\partial x_i} \frac{\partial v}{\partial x_j} dx.$$

Defining the set of classical controls

(2.9)  $W := \{w : x \mapsto w(x) \mid w \text{ measurable from } \Omega \text{ to } U\} \subset L^{\infty}(\Omega) \subset L^{2}(\Omega),$ 

the continuous classical optimal control problem P is to minimize  $G_0$  subject to  $w \in W$  and to the above state constraints.

It is well known that such nonconvex optimal control problems may have no classical solutions, but reformulated in the so-called relaxed form, they 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], [15]) (2.10)  $R := \{r : \Omega \to M_1(U) \mid r \text{ weakly measurable}\} \subset L^{\infty}_w(\Omega, M(U)) \equiv L^1(\Omega, C(U))^*$ , 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 a given  $\phi \in L^1(\Omega; C(U)) = L^1(\overline{\Omega}; C(U))$  (or  $\phi \in B(\overline{\Omega}, U; \mathbb{R})$ , where  $B(\overline{\Omega}, U; \mathbb{R})$  is the set of Caratheodory functions in the sense of Warga [19]) and  $r \in L^{\infty}_w(\Omega, M(U))$  (in particular, for  $r \in R$ ), we shall use the *notation* (2.11)  $\phi(x, r(x)) \coloneqq \int_U \phi(x, u) r(x) (du)$ ,

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

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

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

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

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

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

(2.14) 
$$G_m(r) := \int_{\Omega} g_m(x, y(x), \nabla y(x), r(x)) dx, \quad m = 0, ...q$$

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

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

(2.16) 
$$\sum_{i,j=1}^{d} a_{ij}(x) z_i z_j \ge \alpha_0 \sum_{i=1}^{d} z_i^2, \quad \forall z_i, z_j \in \mathbb{R}, \ x \in \Omega,$$

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

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

(2.18) 
$$|f(x,0,u)| \le \phi_0(x), \quad \forall (x,u) \in \Omega \times U$$

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

(2.19)  $0 \le f_{y}(x, y, u) \le \phi_{1}(x) \eta_{1}(|y|), \quad \forall (x, y, u) \in \Omega \times \mathbb{R} \times U,$ 

where  $\eta_1$  is an increasing function from  $[0, +\infty)$  to  $[0, +\infty)$ ,  $\phi_1 \in L^{\infty}(\Omega)$  if the functionals  $G_m$  depend on  $\nabla y$ , 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

(2.20) 
$$|g_m(x, y, y', u)| \leq \psi_{0m}(x) + \beta_{0m} |y|^2,$$
  
 $\forall (x, y, y', u) \in \Omega \times \mathbb{R}^{d+1} \times U \text{ with } |y| \leq C',$ 

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

Assumptions 2.4 The functions  $g_m, g_{my}, g_{my'}$  are defined on  $\Omega \times \mathbb{R}^{d+1} \times U'$ , where U' is an open set containing the compact set U, measurable on  $\Omega$  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  $g_{my}, g_{my'}$  satisfy

- (2.21)  $|g_{my}(x, y, y', u)| \le \psi_{1m}(x) + \beta_{1m} |y'|^{\frac{2(\rho-1)}{\rho}},$
- (2.22)  $|g_{my'}(x, y, y', u)| \le \psi_{2m}(x) + \beta_{2m}|y'|,$
- (2.23)  $|g_{mu}(x, y, y', u)| \leq \psi_{3m}(x) + \beta_{3m} |y'|,$  $\forall (x, y, y', u) \in \Omega \times \mathbb{R}^{d+1} \times U', \text{ with } |y| \leq C',$

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

The following theorem follows directly form Theorem 3.1 in [3].

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

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

The following proposition is a simple generalization of Proposition 2.1 in [7], and will be useful in the sequel. It can be proved by using the (possible) convergence  $r^k \rightarrow r$ , the fact that a converging sequence in  $L^s$  is dominated (in norm a.e. in  $\Omega$ , and up to a subsequence) by a fixed function in  $L^s$ , Hölder's inequality, Egorov's theorem, and Lebesgue's dominated convergence theorem.

**Proposition 2.1** For i = 1, ..., K,  $K \ge 0$ , let  $s_i \in [1, +\infty]$ ,  $\sigma_i \in [0, s_i]$  if  $s_i < +\infty$ ,  $\sigma_i \coloneqq 0$ if  $s_i = +\infty$ , with  $\frac{1}{s_0} + \sum_{i=1}^{K} \frac{\sigma_i}{s_i} \le 1$ ,  $\frac{1}{s_i} \coloneqq 0$  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

(2.25) 
$$|F(x, y, u)| \le \Phi(x) + \Psi(x) \prod_{i=1}^{K} \xi_i(||y_i||),$$
  
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

(2.26) 
$$\lim_{k \to \infty} \int_{\Omega} F(x, y^{k}(x), r^{k}(x)) dx = \int_{\Omega} F(x, y(x), r(x)) dx.$$

**Theorem 2.2** Under Assumptions 2.1-3, the operator  $r \mapsto y_r$  (resp.  $w \mapsto y_w$ ), from *R* (resp. *W* with the relative topology of  $L^2(\Omega; \mathbb{R}^{\nu})$ , hence of  $L^{\infty}(\Omega; \mathbb{R}^{\nu})$ ) to *V*, and

to  $C_0(\Omega)$ , and the functionals  $r \mapsto G_m(r)$  on R (resp.  $w \mapsto G_m(w)$  on W with the same topologies) are continuous. If the relaxed problem has an admissible control (i.e. satisfying all the constraints), then it has a solution.

**Proof.** Let  $(w_k)$  be a sequence that converges to w in W, with the relative topology of  $L^2(\Omega, \mathbb{R}^v)$ . Since the corresponding sequence of states  $(y_k)$  is bounded in V and in  $C_0^{\alpha}(\overline{\Omega})$ , for some  $\alpha \in (0,1)$ , and since the injection of  $C_0^{\alpha}(\overline{\Omega})$  into  $C_0(\overline{\Omega})$  is compact, there exists a subsequence (same notation) converging to some y in Vweakly and in  $C_0(\overline{\Omega})$  strongly. Let any  $v \in V$  be given. By the state equation  $(2.27) \quad a(y_k, v) + \int_{\Omega} f(x, y_k(x), w_k(x))v dx = 0.$ 

By the mean value theorem and since  $\eta_1$  is increasing, we have, for every y with  $|y| \le C$  (C defined in Theorem 2.1), and for some  $\mu(x) \in [0,1]$ 

$$(2.28) |f(x, y, u)v| \leq |f(x, 0, u)v| + |f(x, y, u)v - f(x, 0, u)v| = |f(x, 0, u)v| + |f_y(x, \mu(x)y, u)yv| \leq |\phi_0(x)| + |\phi_1(x)||v|\eta_1(|\mu(x)y|)|y| \leq |\phi_0(x)| + |\phi_1(x)||v|\eta_1(C)C ,$$

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

(2.29) 
$$\alpha_2 \|y^n - y\|_1^2 \le a(y^n - y, y^n - y)$$
  
=  $-(f(y^n, w^n), y^n) - a(y, y^n) - a(y^n - y, y) \to 0,$ 

since  $y_k \to y$  in *V* weakly and  $(f(y^n, w^n), y^n) \to (f(y, w), y)$  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  $G_m$  follows then from Proposition 2.1. The proofs for  $r \mapsto y_r$  and  $G_m : R \to \mathbb{R}$  are similar. The existence of an optimal relaxed control follows from the compactness of *R* and the continuity of the functionals  $G_m$  (the set of admissible controls is a closed subset of *R*).

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

(2.30) 
$$c_R := \min_{\text{constraints on } r} G_0(r) \le \inf_{\text{constraints on } w} G_0(w) := c_W,$$

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 [19], 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, defined on R (resp. W, with U convex) is *l*-differentiable at r (resp. w) for every integer l, i.e. for every l and any choice of l controls  $r_i \in R$  (resp.  $w_i \in W$ ), i = 1, ..., l, we have

$$(2.31) \quad G(r + \sum_{i=1}^{l} \varepsilon_i(r_i - r)) - G(r) = \sum_{i=1}^{l} DG(r, r_i - r)\varepsilon_i + o(\sum_{i=1}^{l} |\varepsilon_i|),$$

$$(\text{resp. } G(w + \sum_{i=1}^{l} \varepsilon_i(w_i - w)) - G(w) = \sum_{i=1}^{l} DG(w, w_i - w)\varepsilon_i + o(\sum_{i=1}^{l} |\varepsilon_i|)),$$

$$\text{for } \varepsilon_i \ge 0, \ \sum_{i=1}^{l} \varepsilon_i \le 1,$$
with  $DG(r, r_i - r) \coloneqq \int_{\Omega} H(x, y, \nabla y, z, r_i(x) - r(x)) dx,$ 

$$(\text{resp. } DG(w, w_i - w) \coloneqq \int_{\Omega} H_u(x, y, \nabla y, z, w)(w_i(x) - w(x)) dx),$$
where the Hamiltonian is defined by
$$(2.32) \quad H(x, y, y', z, u) \coloneqq -z f(y, x, u) + g(x, y, y', u),$$
and the adjoint state  $z \coloneqq z_r \in V$  satisfies the linear adjoint equation
$$(2.33) \quad a(v, z) + (f_y(y, v)z, v) = (g_y(y, \nabla y, r), v) + (g_{y'}(y, \nabla y, r), \nabla v),$$

$$(\text{resp. } a(v, z) + (f_y(y, w)z, v) = (g_y(y, \nabla y, w), v) + (g_{y'}(y, \nabla y, w), \nabla v)),$$

$$\forall v \in V, \text{ with } y \coloneqq y_r (\text{resp. } y \coloneqq y_w).$$

In particular, the directional derivative of the functional G defined on R (resp. W, with U convex) is given by

$$(2.34) \quad DG(r,\overline{r}-r) = \lim_{\alpha \to 0^+} \frac{G(r+\alpha(\overline{r}-r)) - G(r)}{\alpha}$$
$$= \int_{\Omega} H(x, y(x), \nabla y(x), z(x), r'(x) - r(x)) dx,$$
$$(\text{resp. } DG(r, \overline{w} - w) = \lim_{\alpha \to 0^+} \frac{G(r+\alpha(\overline{w} - w)) - G(w)}{\alpha}$$
$$= \int_{\Omega} H_u(x, y(x), \nabla y(x), z(x), w(x))(\overline{w}(x) - w(x)) dx ).$$

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

**Proof.** We shall prove the *l*-differentiability for classical controls only; we could also prove the Fréchet differentiability in this case, but the proof will be thus similar to the proof for relaxed ones. We first remark that, by our assumptions and since the injection  $V \subset L^{\rho}$  is continuous, the functional

(2.35)  $v \mapsto (g_v(y, \nabla y, w), v) + (g_{v'}(y, \nabla y, w), \nabla v)$ 

belongs to the dual  $V^*$  of V, and  $f_y(y,w) \in L^s(\Omega)$ ,  $2 \le s \le \infty$ ,  $f_y(y,w) \ge 0$ . Hence the linear adjoint equation has a unique solution  $z \in V$ , for every  $w \in W$ , by the Lax-Milgram theorem (if  $s = \infty$ ), or by Lemma 3.2 in [3] (if  $s < \infty$ , no y' in g). Now let

$$w \in W, \ w_i \in W, \ \varepsilon_i \in (0,1), \ i = 1, ..., l, \ \varepsilon := (\varepsilon_1, ..., \varepsilon_l), \ \text{with} \ \left|\varepsilon\right| := \sum_{i=1}^l \left|\varepsilon_i\right| \le 1, \ \text{and set}$$

$$(2.36) \ w_{\varepsilon} := w + \sum_{i=1}^l \varepsilon_i (w_i - w), \ \delta w_i := w_i - w, \ y := y_w, \ y_{\varepsilon} := y_{w_{\varepsilon}}, \ \delta_{\varepsilon} y := y_{\varepsilon} - y.$$

From the state equation, we have

$$(2.37) \quad a(\delta y_{\varepsilon}, v) + (f(y_{\varepsilon}, w_{\varepsilon}) - f(y, w), v) = a(\delta y_{\varepsilon}, v) + (f(y_{\varepsilon}, w_{\varepsilon}) - f(y, w_{\varepsilon}), v) + (f(y, w_{\varepsilon}) - f(y, w), v) = 0$$

Using the mean value theorem, we see that  $\delta y_{\varepsilon}$  satisfies the linear equation

(2.38) 
$$a(\delta y_{\varepsilon}, v) + (f_{y}(y + \mu \delta y_{\varepsilon}) \delta y_{\varepsilon}, v)) = -\sum_{i=1}^{l} (\varepsilon_{i} f_{u}(y, w + \mu \delta w_{\varepsilon}) \delta w_{i}, v), \quad \forall v \in V,$$

where the functions

(2.39) 
$$\overline{a} \coloneqq f_y(y + \mu \delta y_{\varepsilon})$$
 (with  $\overline{a} \ge 0$ ),  $\overline{f} \coloneqq -\sum_{i=1}^l \varepsilon_i f_u(y, w + \mu \delta w_{\varepsilon}) \delta w_i$ ,

belong to  $L^{\infty}(\Omega)$  (or  $L^{s}$ ) and  $L^{s}(\Omega)$ , respectively, by our assumptions. It then follows from Lemma 3.2 in [3] that

(2.40) 
$$\|\delta_{\varepsilon} y\|_{1} + \|\delta_{\varepsilon} y\|_{\infty} \le c \|\overline{f}\|_{L^{s}} \le c' |\varepsilon|.$$

Now, by our assumptions, the functional on the open subset  $Y \times L^2(\Omega, \mathbb{R}^d) \times W'$  of  $L^{\infty}(\Omega) \times L^2(\Omega, \mathbb{R}^d) \times L^{\infty}(\Omega, \mathbb{R}^v)$ 

(2.41) 
$$\Phi(y, y', w) := \int_{\Omega} g(x, y, y', w) dx,$$

where

(2.42) 
$$Y := \left\{ \phi \in L^{\infty}(\Omega) \, \middle| \, \|\phi\|_{\infty} < C' \right\}, \quad W' := \left\{ \psi \in L^{\infty}(\Omega) \, \middle| \, \psi : \Omega \to U' \right\},$$

has the Fréchet derivative defined by (2.43)  $\Phi'(x, y, y', w)(\delta y, \delta y', \delta w)$ 

$$= \int_{\Omega} [g_y(x, y, y', w) \delta y + g_{y'}(x, y, y', w) \delta y' + g_u(x, y, y', w) \delta w] dx$$

This can be shown under our assumptions by using the mean value theorem in maxform, the Cauchy-Schwartz inequality, and Proposition 2.1. Using then the above estimate on  $\delta_{\varepsilon} y$ , we have

(2.44) 
$$o(\|\delta_{\varepsilon} y\|_{\infty} + \|\nabla \delta_{\varepsilon} y\| + \|\sum_{i=1}^{l} \varepsilon_{i} \delta w_{i}\|_{\infty}) = o(|\varepsilon|),$$

hence

$$(2.45) \quad G(w_{\varepsilon}) - G(w) = \int_{\Omega} g_{y}(y, \nabla y, w) \delta_{\varepsilon} y dx + \int_{\Omega} g_{y'}(y, \nabla y, w) \nabla \delta_{\varepsilon} y dx + \sum_{i=1}^{l} \varepsilon_{i} \int_{\Omega} g_{u}(y, \nabla y, w) \delta w_{i} dx + o(|\varepsilon|).$$

Similarly, the state equation, for v := z, yields by linearization

(2.46) 
$$a(\delta_{\varepsilon}y,z) + (f_{y}(y,w)\delta_{\varepsilon}y,z) + \sum_{i=1}^{r} \varepsilon_{i}(f_{u}(y,w)\delta w_{i},z) + o(|\varepsilon|) = 0.$$

On the other hand, the adjoint equation, for  $v := \delta y_{\varepsilon}$ , yields (2.47)  $a(\delta_{\varepsilon} y, z) + (f_{y}(y, w)z, \delta_{\varepsilon} y) = (g_{y}(y, w), \delta_{\varepsilon} y) + (g_{y'}(y, w), \nabla \delta_{\varepsilon} y).$ Gathering the above results, we obtain

$$(2.48) \quad G(w_{\varepsilon}) - G(w) = \sum_{i=1}^{l} \varepsilon_{i} \int_{\Omega} [-z f_{u}(y, w) + g_{u}(y, \nabla y, w)] \delta w_{i} dx + o(|\varepsilon|)$$
$$= \sum_{i=1}^{l} \varepsilon_{i} \int_{\Omega} H_{u}(x, y(x), \nabla y(x), z(x), w(x)) \delta w_{i} dx + o(|\varepsilon|).$$

Finally, the continuity of the operator  $w \mapsto z_w$  is proved by using the continuity of  $w \mapsto y_w$ , from W to  $L^{\infty}$ , the compact injection  $V \subset L^2$ , and Proposition 2.1. The continuity of the functional  $(w, \overline{w}) \mapsto DG(r, \overline{w} - w)$  follows from the above continuities. The continuity proofs for relaxed controls are similar.

The following theorem states various continuous necessary conditions for optimality.

**Theorem 2.3** Under Assumptions 2.1-4, if  $r \in R$  (resp.  $w \in W$ , with U convex) is optimal for Problem  $\overline{P}$  or P (resp. Problem P), then r (resp. w) is strongly extremal relaxed (resp. weakly extremal classical), i.e. there exist multipliers  $\lambda_m \in \mathbb{R}$ , m = 0, ..., q, with

(2.49) 
$$\lambda_0 \ge 0$$
,  $\lambda_m \ge 0$ ,  $m = p + 1, ..., q$ ,  $\sum_{m=0}^{q} |\lambda_m| = 1$ ,

such that

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

(2.51)  $\lambda_m G_m(r) = 0$ , m = p + 1, ..., q (relaxed transversality conditions). (resp.

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

(2.53)  $\lambda_m G_m(w) = 0$ , m = p + 1, ..., q (classical transversality conditions)).

The global condition (2.50) is equivalent to the *strong relaxed pointwise minimum* principle

(2.54) 
$$H(x, y(x), \nabla y(x), z(x), r(x)) = \min_{u \in U} H(x, y(x), \nabla y(x), z(x), u)$$
, a.e. in  $\Omega$ ,

where the complete Hamiltonian and adjoint H, z are defined with g replaced by

 $\sum_{m=0}^{q} \lambda_m g_m$ . If U is convex, then this principle implies the weak relaxed pointwise

*minimum principle* 

(2.55) 
$$H_u(x, y, z, r(x))r(x) = \min_{i} H_u(x, y, z, r(x))\phi(x, r(x)),$$
 a.e. in  $\Omega$ 

where the minimum is taken over the set  $B(\overline{\Omega}, U; U)$  of Caratheodory functions (see [18]), which in turn implies the *global weak relaxed condition* 

(2.56) 
$$\int_{\Omega} H_u(x, y, z, r(x)) [\phi(x, r(x)) - r(x)] dx \ge 0, \quad \forall \phi \in B(\Omega, U; U).$$

A control r satisfying this condition and the above transversality conditions is called *weakly extremal relaxed*. The global condition (2.52) is equivalent to the *weak classical pointwise minimum principle* 

(2.57) 
$$H_u(x, y(x), \nabla y(x), z(x), w(x))w(x) = \min_{u \in U} H_u(x, y(x), \nabla y(x), z(x), w(x))u,$$

a.e. in  $\Omega$ .

**Proof.** The functionals  $G_m$ , m = 0, ..., q, defined on R (resp. W) are continuous (Theorem 2.1) and, by Lemma 2.2, (p+1)-differentiable (cost and p equality state constraints) at r (resp. w). The global condition (i) (resp. (iii)) and the transversality conditions (ii) (resp. (iv)) follow then from the general multiplier theorem V.2.3 (resp. V.3.2) in [19] ( $G_m$  depends here on the control *only*, since  $y_r$  or  $y_w$  is unique for every r or w). The equivalence of the global and pointwise conditions is standard, in both cases (see e.g. [19]) since U is closed (it has a dense denumerable subset). Now, the strong relaxed pointwise minimum principle can be written, for a.a.  $x \in \Omega$ , x *fixed* 

(2.58)  $\int_U H(u)r(du) \leq H(u), \ \forall u \in U.$ 

Let  $\phi \in B(\overline{\Omega}, U; U)$  be any Caratheodory function. Since U is convex here, we have (2.59)  $\int_{U} H(u)r(du) \leq H(u + \varepsilon(\phi(u) - u)), \quad \forall u \in U, \forall \varepsilon \in [0,1],$ hence (2.60)  $\int_{U} H(u)r(du) \leq \int_{U} H(u + \varepsilon(\phi(u) - u))r(du).$ By the Mean Value Theorem and the uniform continuity of H in u(2.61)  $0 \leq \int_{U} \frac{H(u + \varepsilon(\phi(u) - u)) - H(u)}{\varepsilon} r(du)$   $= \int_{U} H_u(u + \varepsilon\mu(u)(\phi(u) - u))(\phi(u) - u)r(du) \quad (0 \leq \mu(u) \leq 1)$   $= \int_{U} H_u(u)(\phi(u) - u)r(du) + \alpha(\varepsilon),$ where  $\alpha(\varepsilon) \to 0$  as  $\varepsilon \to 0$ , hence (2.62)  $\int_{U} H_u(u)(\phi(u) - u)r(du) = H'_u(r)(\phi(r) - r) \geq 0,$ for every  $\phi \in B(\Omega, U; U)$ , a.e. in  $\Omega$ , which is the weak relaxed minimum principle. By integration, we get the global weak relaxed condition (2.63)  $\int_{\Omega} H_u(r)(\phi(r) - r)dxdt \geq 0, \quad \forall \phi \in B(\Omega, U; U).$ 

**Remark.** In the absence of equality state constraints, it can be shown that if the optimal control *w* is *regular*, i.e. there exists  $w' \in W$  such that (2.64)  $G_m(w) + DG_m(w, w'-w) < 0$ , 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 Sections 3 and 4 that  $\Omega$  is a polyhedron (for simplicity). For each integer  $n \ge 0$ , let  $\{E_i^n\}_{i=1}^{N^n}$  be an admissible regular partition of  $\overline{\Omega}$  into elements (e.g. d-simplices), with  $h^n = \max_i [\operatorname{diam}(E_i^n)] \to 0$  as  $n \to \infty$ . Let  $V^n \subset V$  be the subspace of functions that are continuous on  $\overline{\Omega}$  and linear (or multilinear) on each element  $E_i^n$ . The set of discrete controls  $W^n \subset W$  is defined as the subset of (not necessarily continuous) controls  $w^n$  that are (optionally) constant, or linear, or multilinear, on each element  $E_i^n$ , and (optionally) such that  $\|\nabla w^n\|_{\infty} \le L$ , with L independent of n (this reduces to a finite number of linear constraints on the coefficients defining  $w^n$ ). We endow  $W^n$  with the Euclidean topology.

**Remark.** If  $\Omega$  has an appropriately piecewise  $C^1$  boundary  $\Gamma$ , one can approximate  $\Gamma$  by a polyhedral one  $\Gamma^n$ , with domain  $\Omega^n$ , 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  $\Omega^n$  and zero values on  $\Gamma^n$ ).

The following assumptions are stronger than Assumptions 2.2-4.

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

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

- (3.2)  $0 \le f_y(x, y, u) \le c_2(1 + |y|^{\rho-2}),$
- (3.3)  $|f_u(x, y, u)| \le c_3(1+|y|^{\rho-1}),$  $\forall (x, y, u) \in \Omega \times \mathbb{R} \times U',$
- (3.4)  $|g_m(x, y, y', u)| \le c_4(1 + |y|^{\rho} + |y'|^2),$

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

- (3.6)  $|g_{my'}(x, y, y', u)| \le c_6(1+|y|^{\frac{\rho}{2}}+|y'|),$
- (3.7)  $|g_{mu}(x, y, y', u)| \le c_7 (1 + |y|^{\frac{\rho}{2}} + |y'|),$  $\forall (x, y, y', u) \in \Omega \times \mathbb{R}^{d+1} \times U',$

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

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

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

The following theorem can be proved by using the techniques in [13] (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  $w^n \in W^n$ , 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, defined on  $W^n$ , are given by

(3.9) 
$$G_m^n(r^n) = \int_{\Omega} g_m(y^n, \nabla y^n, w^n) dx, \quad m = 0, ..., q.$$

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

(3.10) Case (a)  $|G_m^n(w^n)| \le \varepsilon_m^n, m = 1,..., p,$ 

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

and

(3.12)  $G_m^n(w^n) \le \varepsilon_m^n, \ \varepsilon_m^n \ge 0, \ m = p+1,...,q,$ 

where the *feasibility perturbations*  $\varepsilon_m^n$  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  $G_m^n(w^n)$  subject to  $w^n \in W^n$  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  $w^n \mapsto y^n$ , from  $W^n$  to  $V^n$ , are continuous. Under assumptions 2.1 and 3.1 (on  $f, f_y, g_m$ ), the functionals  $w^n \mapsto G_m^n(w^n)$ , on  $W^n$ , 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  $w^n, \overline{w}^n \in W^n$  by

(3.13)  $DG^n(w^n, \overline{w}^n - w^n) = \int_{\Omega} H_u(x, y^n, \nabla y^n, z^n, w^n)(\overline{w}^n - w^n)dx,$ 

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

(3.14) 
$$a(z^n, v^n) + (z^n f_y(y^n, w^n), v^n) = (g_y(y^n, \nabla y^n, w^n), v^n) + (g_{y'}(y^n, \nabla y^n, w^n), \nabla v^n),$$
  
 $\forall v^n \in V^n, \text{ where } y^n := y^n_{w^n}.$ 

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

**Theorem 3.3** Under Assumptions 2.1 and 3.1, if  $w^n \in W^n$  is optimal for Problem  $P_b^n$ , then  $w^n$  is weakly discrete extremal classical (or discrete extremal), i.e. there exist multipliers  $\lambda_m^n \in \mathbb{R}$ , m = 0, ..., q, with  $\lambda_m^n \ge 0$ ,  $\lambda_m^n \ge 0$ , m = p + 1, ..., q,  $\sum_{m=0}^q |\lambda_m^n| = 1$ , such that

$$(3.15) \quad \sum_{m=0}^{q} \lambda_m^n DG_m^n(w^n, \overline{w}^n - w^n) = \int_{\Omega} H^n(y^n, \nabla y^n, z^n, \overline{w}^n - w^n) dx \ge 0, \quad \forall \overline{w}^n \in W^n,$$
  
(3.16) 
$$\lambda_m^n(G_m(w^n) - \varepsilon_m^n) = 0, \quad m = p+1, ..., q,$$

where  $H^n$  and  $z^n$  are defined with g replaced by  $\sum_{m=0}^q \lambda_m^n g_m$ . The global condition (3.17) is equivalent to the strong discrete classical elementwise minimum principle (3.18)  $\int_{E_i^n} H_u^n(y^n, \nabla y^n, z^n, w^n) w^n dx = \min_{u \in U} \int_{E_i^n} H_u^n(y^n, \nabla y^n, z^n, w^n) u dx, \quad i = 1, ..., N^n$ .

Let  $\overline{W}^n$  be the set of elementwise constant discrete controls. Clearly,  $\overline{W}^n \subset W^n$  in all cases. The following control approximation result (i) (resp. (ii)) is proved similarly to the corresponding result in [8] (resp. [13]).

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

(ii) For every  $w \in W$ , there exists a sequence  $(w^n \in \overline{W}^n)$  of discrete classical controls, considered as relaxed ones, that converges to w in  $L^2$  strongly.

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  $(w^n \in W^n)$  converges to  $r \in R$  in R (resp. to  $w \in W$  in  $L^2$  strongly), then  $y^n \to y_r$  (resp.  $y^n \to y_w$ ) in V strongly,  $G^n(w^n) \to G(r)$  (resp.  $G^n(w^n) \to G(w)$ ), and  $z^n \to z_r$  (resp.  $z^n \to z_w$ ) in  $L^{\rho}(\Omega)$  strongly (and in V strongly, if the functionals do not depend on  $\nabla y$ ).

(ii) If the sequences  $(w^n \in W^n)$  and  $(\overline{w}^n \in W^n)$  converge to w and  $\overline{w}$ , respectively, in W, then

(3.19)  $DG^n(w^n, \overline{w}^n - w^n) \rightarrow DG(w, \overline{w} - w).$ 

**Proof.** (i) Suppose that  $w^n \to r$  in R. From the discrete state equation, we have (3.20)  $a(y^n, y^n) + (f(y^n, w^n) - (f(0, w^n), y^n - 0) = -(f(0, w^n), y^n),$ and since f is increasing in y

(3.21) 
$$\alpha_2 \|y^n\|_1^2 \le a(y^n, y^n) \le |(f(0, w^n), y^n)| \le \|f(0, w^n)\|_s \|y^n\| \le c \|y^n\|_1,$$

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 Ref. 20), we can suppose that  $y^n \to y$  in  $L^{\rho}(\Omega)$  strongly. For any given  $v \in C_0^1(\overline{\Omega})$ , let  $(v^n \in V^n)$  be the sequence of interpolates of v at the vertices of the partition of  $\Omega$ . This sequence converges to v in  $C_0^1(\overline{\Omega})$  (hence in V) strongly. We have

(3.22)  $a(y^n, v^n) + (f(y^n, w^n), v^n) = 0.$ 

Since  $w^n \to r$  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 (3.23) a(y,v)+(f(y,r),v)=0,

which holds also for every  $v \in V \subset L^s$ , as  $C_0^1(\overline{\Omega})$  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

(3.24) 
$$\alpha_2 \left\| y^n - y \right\|_1^2 \le a(y^n - y, y^n - y) = (-f(y^n, w^n), y^n) - a(y, y^n) - a(y^n - y, y).$$

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

$$(3.25) \quad \alpha_{2} \left\| z^{n} \right\|_{1}^{2} \leq a(z^{n}, z^{n}) + (f_{y}(y^{n}, r^{n})z^{n}, z^{n}) \\ \leq \left| (g_{y}(y^{n}, \nabla y^{n}, z^{n}, r^{n}), z^{n}) \right| + \left| (g_{y'}(y^{n}, \nabla y^{n}, z^{n}, r^{n}), \nabla z^{n}) \right| \\ \leq (c_{4}(1 + \left| y^{n} \right|^{p-1} + \left| \nabla y^{n} \right|^{\frac{2(p-1)}{p}}), \left| z^{n} \right|) + (c_{5}(1 + \left| y^{n} \right|^{\frac{p}{2}} + \left| \nabla y^{n} \right|), \left| z^{n} \right|)$$

$$\leq c'_{4}(1 + \left\|y^{n}\right\|_{L^{\rho}}^{\rho-1} + \left\|\nabla y^{n}\right\|^{\frac{2(\rho-1)}{\rho}})\left\|z^{n}\right\|_{L^{\rho}} + c'_{5}(1 + \left\|y^{n}\right\|_{L^{\rho}}^{\frac{\rho}{2}} + \left\|\nabla y^{n}\right\|)\left\|z^{n}\right\|_{1} \leq c\left\|z^{n}\right\|_{1},$$

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  $\nabla y$ , 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. The above proofs are similar if  $w^n \to w$  in  $L^2$  strongly.

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

**Remark.** Suppose that  $w^n \to w$  in  $L^2$  strongly and that Assumptions 2.1, 2.2 (instead of 3.1) are satisfied. Under the strong assumptions of [4], it can be shown that  $y_{w^m}^n \to y_{w^m}$  uniformly as  $n \to \infty$ , for *m* fixed. On the other hand,  $y_{w^m} \to y_w$  also uniformly by Theorem 2.2. Therefore, under all these assumptions,  $y_{w^n}^n \to y_w$  uniformly as  $n \to \infty$ . Similar remarks hold for the convergence of the functionals, adjoints and functional derivatives.

We suppose in the sequel that the considered continuous Problem P or  $\overline{P}$  is feasible. The following (theoretical, in the presence of state constraints) theorem addresses the behavior in the limit of optimal discrete controls.

**Theorem 3.4** In the presence of state constraints, we suppose that the sequences  $(\mathcal{E}_m^n)$  in the discrete state constraints (Case (a)) converge to zero as  $n \to \infty$  and satisfy (3.26)  $|G_m^n(\tilde{w}^n)| \le \mathcal{E}_m^n, \quad m = 1, ..., p, \quad G_m^n(\tilde{w}^n) \le \mathcal{E}_m^n, \quad \mathcal{E}_m^n \ge 0, \quad m = p + 1, ..., q,$ 

for every *n*, where  $(\tilde{w}^n \in W^n \subset R)$  is a sequence converging in *R* (resp in  $L^2$  strongly) to an optimal control  $\tilde{r} \in R$  (resp.  $\tilde{w} \in W$ ) of Problem  $\overline{P}$  (resp. *P*), which always exists (resp. if it exists). For each *n*, let  $w^n$  be optimal for Problem  $P_a^n$ . Then every strong relaxed (resp. classical) accumulation point of  $(w^n)$ , which always exists (resp. if it exists), is optimal for Problem  $\overline{P}$  (resp. *P*).

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

(3.27)  $G_0^n(w^n) \le G_0^n(\tilde{w}^n), \quad |G_m^n(w^n)| \le \varepsilon_m^n, \ m = 1, ..., p, \quad G_m^n(w^n) \le \varepsilon_m^n, \ m = p+1, ..., q.$ Passing to the limit and using Lemma 3.2, we obtain that r (resp. w) is optimal for Problem  $\overline{P}$  (resp. P). If there are no state constraints, by taking a sequence converging to some optimal control, we also obtain that the limit control is optimal.

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  $(\varepsilon_m^n)$  that converge to zero and such that the discrete problem is feasible for every n. Let  $\overline{w}^n$  be any solution of the following *auxiliary minimization problem without state constraints* 

(3.28) 
$$c^n := \min_{w^n \in W^n} \{ \sum_{m=1}^p [G_m^n(w^n)]^2 + \sum_{m=p+1}^q [\max(0, G_m^n(w^n))]^2 \},\$$

and set

(3.29)  $\varepsilon_m^n \coloneqq G_m^n(\overline{w}^n), \quad m = 1, \dots, p, \quad \varepsilon_m^n \coloneqq \max(0, G_m^n(\overline{w}^n)), \quad m = p + 1, \dots, q.$ 

Let  $\tilde{v} \in R$  (resp.  $\tilde{v} \in W$ ) be an admissible control for Problem  $\overline{P}$  (resp. *P*), and  $(\tilde{w}^n \in W^n)$  a sequence converging to  $\tilde{v}$  in *R* (resp. in  $L^2$  strongly), by Proposition 3.1. We then have

(3.30) 
$$\lim_{n\to\infty} [G_m^n(\tilde{w}^n)]^2 = [G_m(\tilde{v})]^2 = 0, \quad m = 1,...,p,$$

(3.31) 
$$\lim_{n \to \infty} [\max(0, G_m^n(\tilde{w}^n))]^2 = [\max(0, G_m(\tilde{v}))]^2 = 0, \quad m = p + 1, ..., q_n$$

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  $\varepsilon_m^n$ . We suppose in the sequel that the perturbations  $\varepsilon_m^n$  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  $\varepsilon_m^n$  vanish, i.e. the discrete problem with zero perturbations is feasible. Also, see [11] for a study on how the perturbations  $\varepsilon_m^n$  can be practically chosen to be zero, if there are only inequality state constraints.

**Theorem 3.5** For each n, let  $w^n$  be admissible and extremal for Problem  $P_b^n$ . Then (i) Every strong accumulation point of the sequence  $(w^n)$  in  $L^2(\Omega)$  (if it exists) is admissible and weakly extremal classical for Problem P.

(ii) Suppose that  $\|\nabla v^n\|_{\infty} \leq L$  for every  $v^n \in W^n$ , with *L* independent of *n*, in the definition of  $W^n$ . Then every accumulation point of  $(w^n)$  in *R* is admissible and weakly extremal relaxed for Problem  $\overline{P}$ .

**Proof.** (i) Suppose that some subsequence  $(w^n)$  (same notation) converges to some  $w \in W$  in  $L^2(\Omega)$  strongly. For each n, let  $\lambda_m^n$ , m = 0, ..., q be multipliers as in Theorem 3.2. Since  $\sum_{m=0}^{q} |\lambda_m^n| = 1$ , the sequences  $(\lambda_m^n)$  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  $\overline{w} \in W$  and  $\overline{w}^n \to \overline{w}$  (Proposition 3.1) (3.32)  $\sum_{m=0}^{q} \lambda_m DG(w, \overline{w} - w) = \lim_{m=0}^{q} \lambda_m^n DG_m^n(w^n, \overline{w}^n - w^n) \ge 0$ ,

(3.32) 
$$\sum_{m=0}^{n} \lambda_m DO_m(w, w - w) - \lim_{n \to \infty} \sum_{m=0}^{n} \lambda_m DO_m(w, w - w) \ge 0$$
  
(3.33) 
$$\lambda_m G_m(w) = \lim \lambda_m^n [G_m^n(w^n) - \varepsilon_m^n] = 0, \quad m = p+1, ..., q,$$

(3.34) 
$$G_m(w) = \lim_{n \to \infty} [G_m^n(w^n) - \varepsilon_m^n] = 0, \quad m = 1, ..., p,$$

(3.35) 
$$G_m(w) = \lim_{n \to \infty} [G_m^n(w^n) - \varepsilon_m^n] \le 0, \quad m = p + 1, ..., q,$$

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

(ii) Since *R* is compact and  $\sum_{m=0}^{q} |\lambda_m^n| = 1$ , let  $(w^n)$ ,  $(\lambda_m^n)$ , m = 0, ..., q, be subsequences such that  $w^n \to r$  in *R* and  $\lambda_m^n \to \lambda_m$ , m = 0, ..., q, and consider the discrete optimality condition in global form, which can be written (3.36)  $\int_{\Omega} H_u^n(x, y^n, z^n, w^n)(\overline{w}^n - w^n) dx dt \ge 0$ , for every  $\overline{w}^n \in W^n$ ,

where  $H^n, z^n$  are defined with  $g := \sum_{m=0}^q \lambda_m^n g_m$ . Define the elementwise constant vector functions

(3.37)  $\tilde{x}^n(x) :=$  barycenter of  $E_i^n$ , for  $x \in E_i^{o^n}$ , i = 1, ..., M,

(3.38)  $\tilde{w}^n(x) := w^n(\tilde{x}^n(x))$ , for  $x \in E_i^{o^n}$ , i = 1, ..., M.

Clearly,  $\tilde{x}^n - x \to 0$  uniformly on  $\Omega$ , and by our assumption on  $W^n$  and the mean value theorem,  $\|\tilde{w}^n - w^n\|_{\infty} \leq Lh^n \to 0$ . For every function  $\phi \in C(\overline{\Omega} \times U; U)$ , we then have

(3.39) 
$$\int_{\Omega} H_{u}^{n}(x, y^{n}, z^{n}, w^{n}) [\phi(\tilde{x}^{n}, \tilde{w}^{n}) - w^{n}] dx$$
$$= \int_{\Omega} H_{u}^{n}(x, y^{n}, z^{n}, w^{n}) [\phi(x^{n}, w^{n}) - w^{n}] dx$$
$$+ \int_{\Omega} H_{u}^{n}(x, y^{n}, z^{n}, w^{n}) [\phi(\tilde{x}^{n}, \tilde{w}^{n}) - \phi(x^{n}, w^{n})] dx \ge 0.$$

Using the above convergences, Lemma 3.2, the uniform continuity of  $\phi$ , and Proposition 2.1 ( $|\phi(..)| \le c$  since  $\phi(..) \in U$ ), we can pass to the limit and obtain

$$(3.40) \quad \int_{\Omega} H_u(x, y, z, r(x)) [\phi(x, r(x)) - r(x)] dx$$
  
$$\coloneqq \int_{\Omega} \int_U H_u(x, y, z, r(x)) [\phi(x, u) - u] r(du) dx \ge 0, \quad \forall \phi \in C(\overline{\Omega} \times U; U),$$

where H, z are defined with  $g := \sum_{m=0}^{q} \lambda_m g_m$ , and with multipliers  $\lambda_m$  as in the continuous optimality conditions (see (i)). Now let  $\phi \in B(\overline{\Omega}, U; U)$  be any Caratheodory function, or equivalently,  $\phi \in L^1(\Omega, C(U; U))$  (see [19]), and let  $(\phi_k)$  be a sequence in  $C(\overline{\Omega}; C(U; U)) \equiv C(\overline{\Omega} \times U; U)$  converging to  $\phi$  in  $L^1(\Omega, C(U; U))$  strongly. By Egorof's theorem, we can suppose that  $\phi_k \to \phi$  a.e. in  $\Omega$ , with values in C(U; U), hence a.e. in  $\Omega \times U$ , with values in U. Replacing  $\phi$  by  $\phi_k$  in the above inequality and using Lebesgue's dominated convergence theorem, we can pass to the limit as  $n \to \infty$  in the transversality conditions and the state constraints similarly to (i). Therefore, r is weakly extremal relaxed and admissible for Problem  $\overline{P}$ .

## 4 Discrete penalized gradient projection 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) 
$$G^{nl}(w^n) := G^n_0(w^n) + \{\sum_{m=1}^p M^l_m [G^n_m(w^n)]^2 + \sum_{m=p+1}^q M^l_m [\max(0, G^n_m(w^n))]^2\} / 2$$

Let  $\gamma \ge 0$ ,  $b', c' \in (0,1)$ , and let  $(\beta^l)$ ,  $(\zeta_k)$  be positive sequences, with  $(\beta^l)$  decreasing and converging to zero, and  $\zeta_k \le 1$ . The algorithm described below contains two options. In the case of the progressively refining version, we suppose that each element  $E_{i'}^{n+1}$  is a subset of some element  $E_i^n$ . In this case, 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 cost derivatives for this control, but with the possibly finer discretization n+1, makes sense. (In this version, and if  $\Omega$  is not polyhedral, one has to modify slightly near the boundary the control  $w_k^{nl}$ , at the end of Step 3, before going to Step 2, and if the discretization has been refined). The *discrete relaxed penalized gradient projection methods* are described by the following Algorithm.

#### Algorithm

Step 1. Set k := 0, l := 1, choose a value of n and an initial control  $w_0^{n1} \in W^n$ . Step 2. Find  $v_k^{nl} \in W^n$  such that

(4.2) 
$$e_{k} := DG^{nl}(w_{k}^{nl}, v_{k}^{nl} - w_{k}^{nl}) + \frac{\gamma}{2} \|v_{k}^{nl} - w_{k}^{nl}\|^{2}$$
$$= \min_{\overline{v}^{n} \in W^{n}} [DG^{nl}(w_{k}^{nl}, \overline{v}^{n} - w_{k}^{nl}) + \frac{\gamma}{2} \|\overline{v}^{n} - w_{k}^{nl}\|^{2}]$$

and set  $d_k := DG^{nl}(w_k^{nl}, v_k^{nl} - w_k^{nl})$ .

Step 3. If  $|d_k| \le \beta^l$ , set  $w^{nl} := w_k^{nl}$ ,  $v^{nl} := v_k^{nl}$ ,  $d^l := d_k$ ,  $e^l := e_k$ , 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  $\alpha(s) = c^{is} \zeta_k \in (0,1]$  and  $\alpha(s)$  satisfies the inequality

(4.3) 
$$G^{nl}(w_k^{nl} + \alpha(s)(v_k^{nl} - w_k^{nl})) - G^{nl}(w_k^{nl}) \le \alpha(s)b'd_k$$
,  
and then set  $\alpha_k := \alpha(\overline{s})$ .

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

In the above 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  $\varepsilon_m^n$  (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.

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, the projection  $v_{ki}^{nl}$  of the function in  $L^2(E_i^n)$ 

(4.4) 
$$u_{ki}^{nl} := w_{ki}^{nl} - (1/\gamma) H_u^n (y_{ki}^{nl}, \nabla y_{ki}^{nl}, z_{ki}^{nl}, w_{ki}^{nl}),$$

onto the convex subset of  $L^2(E_i^n)$  of constant, or linear, or multilinear, functions  $v_i^{ml}$ , with values in U, which reduces to the minimization of a quadratic function of the coefficients of the control  $v_i^{ml}$ . The parameter  $\gamma$  is chosen here experimentally to yield a good rate of convergence. If  $\gamma = 0$ , the above Algorithm is a *penalized conditional gradient method*, and Step 2 reduces similarly to the minimization of a linear function, for each *i*. On the other hand, by the definition of the directional derivative and since  $b' \in (0,1)$ , clearly the Armijo step  $\alpha_k$  in Step 4 can be found for every *k*.

A (continuous classical or relaxed, or discrete) extremal (or weakly extremal) control is called *abnormal* if there exist multipliers as in the corresponding optimality conditions, with  $\lambda_0 = 0$  (or  $\lambda_0^n = 0$ ). A control is admissible *and* abnormal extremal in very exceptional, degenerate, situations (see [19]).

(4.5) With  $w^{nl}$  defined in Step 3, define the sequences of multipliers  $\lambda_m^{nl} := M_m^l G_m^n(w^{nl}), \ m = 1, ..., p, \quad \lambda_m^{nl} := M_m^l \max(0, G_m^n(w^{nl})), \ m = p+1, ..., q.$ 

**Theorem 4.1** (i) In Version B, if  $(w^{nl})$  is a subsequence of the sequence generated by the Algorithm in Step 3 that converges to some  $w \in W$  in  $L^2$  strongly, as  $l \to \infty$  (hence  $n \to \infty$ ). If the sequences  $(\lambda_m^{nl})$  are bounded, then w is admissible and weakly extremal classical for Problem P.

(ii) Suppose that  $\|\nabla v^n\|_{\infty} \leq L$  for every  $v^n \in W^n$ , with *L* independent of *n*, in the definition of  $W^n$ . In Version B, let  $(w^{nl})$  be a subsequence, considered as a sequence in *R*, of the sequence generated by the Algorithm in Step 3 that converges to some *r* in the compact set *R*, as  $l \to \infty$  (hence  $n \to \infty$ ). If the sequences  $(\lambda_m^{nl})$  are bounded, then *r* is admissible and weakly extremal relaxed for Problem  $\overline{P}$ .

(iii) In Version A, let  $(w^{nl})$ , *n* fixed, be a subsequence of the sequence generated by the Algorithm in Step 3 that converges to some  $w^n \in W^n$  as  $l \to \infty$ . If the sequences  $(\lambda_m^{nl})$  are bounded, then  $w^n$  is admissible and extremal for Problem  $P^n$ .

(iv) In any of the three convergence cases (i), (ii) (with the additional assumption), or (iii), suppose that the (discrete  $P_b^n$  or continuous) limit problem has no admissible, abnormal extremal, controls. If the limit control is admissible, then the sequences of multipliers are bounded, and this control is extremal for the limit problem 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  $W^n$  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)  $e := \lim_{k \to \infty, k \in K} e_k = DG(\tilde{w}, \tilde{v} - \tilde{w}) + (\gamma/2) \left\| \tilde{v} - \tilde{w} \right\|^2 \le 0,$  $d := \lim_{k \to \infty, k \in K} d_k = DG(\tilde{w}, \tilde{v} - \tilde{w}) \le \lim_{k \to \infty, k \in K} e_k = e \le 0.$ (4.7)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,  $\Phi$  is differentiable on (0,1) and has derivative  $\Phi'(\alpha) = DG(w + \alpha(v - w), v - w).$ (4.8)Using the mean value theorem, we have, for each  $\alpha \in (0,1]$ (4.9)  $G(w_k + \alpha(v_k - w_k)) - G(w_k) = \alpha DG(w_k + \alpha'(v_k - w_k), v_k - w_k),$ for some  $\alpha' \in (0, \alpha)$ . Therefore, for  $\alpha \in [0, 1]$ , by Lemma 3 (4.10)  $G(w_k + \alpha(v_k - w_k)) - G(w_k) = \alpha(d + \varepsilon_{k\alpha}),$ 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)  $d + \varepsilon_{k\alpha} \leq b(d + \eta_k) = b'd_k$ , for  $\alpha \in [0, \overline{\alpha}]$ , for some  $\overline{\alpha} > 0$ , and  $k \ge \overline{k}$ ,  $k \in K$ . Hence  $(4.12) \quad G(w_k + \alpha(v_k - w_k)) - G(w_k) \leq \alpha b' d_k,$ for  $\alpha \in [0,\overline{\alpha}]$ , for some  $\overline{\alpha} > 0$ , and  $k \ge \overline{k}$ ,  $k \in K$ . It follows from the choice of the Armijo step  $\alpha_k$  in Step 4 that  $\alpha_k \ge c\overline{\alpha}$ , for  $k \ge \overline{k}$ ,  $k \in K$ . Hence (4.13)  $G(w_{k+1}) - G(w_k) = G(w_k + \alpha_k(v_k - w_k)) - G(w_k)$  $\leq \alpha_k b' d_k \leq c \overline{\alpha} b' d_k \leq c \alpha b' d / 2,$ 

for  $k \ge \overline{k}$ ,  $k \in K$ . 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 Lemma 3.1. Therefore, we must have d = e = 0, and  $d_k \to 0$ ,  $e_k \to 0$ , for the whole sequences, 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  $(w^{nl})$  be a subsequence (same notation) of the sequence generated by the Algorithm in Step 3 that converges to some  $w \in W$  in  $L^2$  strongly as  $l, n \to \infty$ . Suppose that the sequences  $(\lambda_m^{nl})$  are bounded and (up to subsequences) that  $\lambda_m^{nl} \to \lambda_m$ . By Lemma 3.2, we have

(4.14) 
$$0 = \lim_{l \to \infty} \frac{\lambda_m^{nl}}{M_m^l} = \lim_{l \to \infty} G_m^n(w^{nl}) = G_m(w), \quad m = 1, ..., p,$$
  
(4.15) 
$$0 = \lim_{l \to \infty} \frac{\lambda_m^{nl}}{M_m^l} = \lim_{l \to \infty} [mean(0, C_m^n(w^{nl}))] = mean(0, C_m(w)), \quad m = n + 1.$$

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

which show that w is admissible. Now let any  $\tilde{v} \in W$  and, by Proposition 1, let  $(\tilde{v}^n \in W^n)$  be a sequence of elementwise constant discrete controls that converges to  $\tilde{v}$  in  $L^2$  strongly. Set  $\lambda_0^{nl} = 1$ , for every n, l, and let  $(\lambda_m^{nl})$  be subsequences such that  $\lambda_m^{nl} \rightarrow \lambda_m$ . By Step 2, we have

(4.16) 
$$\sum_{m=0}^{q} \lambda_m^{nl} DG_m^n(w^{nl}, \tilde{v}^n - w^{nl}) + (\gamma/2) \int_{\Omega} \left| \tilde{v}^n - w^{nl} \right|^2 dx \ge e^l,$$

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

(4.17) 
$$\sum_{m=0}^{q} \lambda_m DG_m(w, \tilde{v} - w) + (\gamma/2) \int_{\Omega} \left| \tilde{v} - w \right|^2 dx \ge 0, \quad \forall \tilde{v} \in W,$$

Replacing  $\tilde{v}$  by  $w + \mu(\tilde{v} - w)$ , with  $\mu \in (0,1]$ , dividing by  $\mu$ , and then taking the limit as  $\mu \to 0^+$ , we get

(4.18) 
$$\sum_{m=0}^{q} \lambda_m DG_m(w, \tilde{v} - w) \ge 0, \quad \forall \tilde{v} \in W.$$

By construction of the  $\lambda_m^{nl}$ , 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  $G_m(w) < 0$ , for some index  $m \in [p+1,q]$ , then for sufficiently large l we have  $G_m^{nl}(w^{nl}) < 0$  and  $\lambda_m^l = 0$ , hence  $\lambda_m = 0$ , i.e. the transversality conditions hold. Therefore, w is also weakly extremal classical.

(ii) Let  $(w^{nl})$  be a subsequence (same notation), considered as a sequence in R, of the sequence generated in Step 3, that converges to some  $r \in R$  as  $l, n \to \infty$ . The admissibility of r is shown similarly to (i). Suppose that the sequences  $(\lambda_m^{nl})$  are bounded and (up to subsequences) that  $\lambda_m^{nl} \to \lambda_m$ . Now, by Steps 2 and 3 we have, for every  $\tilde{v}^n \in W^n$ 

$$(4.19) \quad DG^{nl}(w^{nl}, \tilde{v}^n - w^{nl}) + (\gamma/2) \left\| \tilde{v}^n - w^{nl} \right\|^2 \\ = DG_0^n(w^{nl}, \tilde{v}^n - w^{nl}) + \sum_{m=1}^p \lambda_m^{nl} DG_m^n(w^{nl}, \tilde{v}^n - w^{nl}) + \sum_{m=p+1}^q \lambda_m^{nl} DG_m^n(w^{nl}, \tilde{v}^n - w^{nl}) \\ + (\gamma/2) \left\| \tilde{v}^n - \overline{w}^{nl} \right\|^2 \\ = \int_{\Omega} H_u^{nl}(x, y^{nl}, z^{nl}, w^{nl}) (\tilde{v}^n - w^{nl}) dx dt + (\gamma/2) \int_{\Omega} \left| \tilde{v}^n - w^{nl} \right|^2 dx dt \ge e^l,$$

where  $H^{nl}, z^{nl}$  are defined with  $g := \sum_{m=0}^{q} \lambda_m^{nl} g_m$ . By an argument similar to the proof of Theorem 3.5 (ii), and since  $e^l \to 0$  by Step 3, we obtain here

(4.20)  $\int_{\Omega} H_u(x, y, z, r(x)) [\phi(x, r(x)) - r(x)] dx$ 

$$+(\gamma/2)\int_{\Omega} [\phi(x,r(x)) - r(x)]^2 dx \ge 0, \quad \text{for every } \phi \in B(\overline{\Omega},U;U),$$

where H, z are defined with  $g := \sum_{m=0}^{q} \lambda_m g_m$ , and with the  $\lambda_m$  as in the continuous optimality conditions (see (i)). Replacing  $\phi$  by  $u + \mu(\phi - u)$ , with  $\mu \in (0,1]$ , dividing by  $\mu$ , and then taking the limit as  $\mu \to 0^+$ , we obtain the global weak relaxed condition

(4.21) 
$$\int_{\Omega} H_u(x, y, z, r(x)) [\phi(x, r(x)) - r(x)] dx \ge 0, \text{ for every } \phi \in B(\overline{\Omega}, U; U).$$

Finally, the transversality conditions are shown as in (i). Therefore, r is also weakly extremal relaxed.

(iii) The admissibility of the limit control  $w^n$  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 3.2, we obtain, similarly to (i)

(4.22) 
$$\sum_{m=0}^{q} \lambda_m D\tilde{G}_m^n(w^n, v^{n} - w^n) = \sum_{m=0}^{q} \lambda_m DG_m^n(w^n, v^{n} - w^n) \ge 0, \quad \forall v^{n} \in W^n,$$

and the discrete transversality conditions

(4.23) 
$$\lambda_m^n G_m^n(w^n) = \lambda_m^n [G_m^n(w^n) - \varepsilon_m^n] = 0, \quad m = p+1,...,q,$$

with multipliers  $\lambda_m^n$  as in the discrete optimality conditions.

(iv) In either of the two above convergence cases, 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), (ii) or (iii), this limit control 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  $(M_m^l)$  and a sequence  $(\beta^l)$  relatively fast converging to zero, the resulting sequences of multipliers  $(\lambda_m^{nl})$  are often kept bounded. We 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  $k \ge 1$ .

When directly applied to nonconvex optimal control problems whose solutions are non-classical relaxed controls, the above methods generating classical controls often yield poor convergence (highly oscillating optimal controls). If U is convex, one can then reformulate the problem in Gamkrelidze relaxed form (which is equivalent to the Young measure formulation), using convex combinations of Dirac controls, involving a finite (usually small) number of classical controls. The above methods can then be applied to this extended classical control problem, resulting in better results, since the problem is thus partially convexified in some sense (for details on this approach, see [12]). If U is not convex, one can use methods generating relaxed controls for solving such highly nonconvex problems (see [10]).

### **5** Numerical examples

Let  $\Omega := (0,1)^2$  and consider the following examples.

Example 1. Define the reference controls and state

(5.1)  $\overline{u}(x) \coloneqq x_1 x_2$ ,  $\overline{v}(x) \coloneqq 1 - x_1 x_2$ ,  $\overline{y}(x) \coloneqq 8x_1 x_2 (1 - x_1)(1 - x_2)$ . Consider the following optimal control problem, with state equation (5.2)  $-\Delta y + y^3/3 + (1 + u - \overline{u}))y$   $-\overline{y}^3/3 - \overline{y} - 16[x_1(1 - x_1) + x_2(1 - x_2)] - (v - \overline{v}) = 0$ , in  $\Omega$ , (5.3) y(x) = 0 on  $\Gamma$ , control constraints  $(u(x), v(x)) \in U \coloneqq U_1 \coloneqq [0, 1]^2$ ,  $x \in \Omega$ , and cost functional

(5.4) 
$$G_0(u,v) := 0.5 \int_{\Omega} [(y-\overline{y})^2 + \|\nabla y - \nabla \overline{y}\|^2 + (u-\overline{u})^2 + (v-\overline{v})^2] dx.$$

Clearly, the optimal controls are  $\overline{u}$  and  $\overline{v}$ , the optimal state is  $\overline{y}$ , and the optimal cost is zero. The gradient projection method (without penalties) was first applied to this problem using triangular elements, which are half squares of edge size h = 1/80, and triangle-wise linear discrete controls, with  $\gamma = 0.5$ , Armijo parameters b' = c' = 0.5, and constant initial controls  $u_0^n = v_0^n = 0.5$ . After 15 iterations, we obtained the results

(5.5) 
$$G_0^n(u_k^n, v_k^n) = 2.963 \cdot 10^{-4}, \quad d_k = -7.786 \cdot 10^{-12},$$
  
 $\varepsilon_k = 5.064 \cdot 10^{-5}, \quad \eta_k = 3.052 \cdot 10^{-5},$ 

where  $\varepsilon_k$  (resp.  $\eta_k$ ) is the discrete max error for the state (resp. controls) at the vertices of the triangles (resp. midpoints of the triangle edges). Since  $\varepsilon_k \approx O(h^2)$  and the optimal controls are smooth, the corresponding control  $L^{\infty}$ -error is also  $\approx O(h^2)$ . The computed controls are practically identical to the exact optimal ones and are not shown. Using then trianglewise constant discrete controls, we obtained results of similar accuracy for the cost, state, and control errors (taken here at the barycenters of the triangles), but the control  $L^{\infty}$ -error is now  $\approx O(h)$ .

*Example 2.* With the same state equation, cost and parameters as in Example 1, triangle-wise linear discrete controls, but with the constraint set replaced by the strictly active one  $U := U_2 := [0, 0.6] \times [0.4, 1]$ , we obtained after 15 iterations the results

(5.6)  $G_0^n(u_k^n, v_k^n) = 2.631766265263325 \cdot 10^{-3}, \quad d_k = -3.526 \cdot 10^{-10},$ and the controls shown in Figures 1 and 2.

*Example 3.* With the same state equation and cost and parameters as in Example 2, trianglewise linear discrete controls, but with  $U := U_3 := [0, 0.7] \times [0.3, 1]$  and the additional state constraint

(5.7) 
$$G_1(u,v) := \int_{\Omega} (y - 0.22) dx = 0,$$

and applying the penalized gradient projection method, we obtained after 63 iterations in k the results

(5.8) 
$$G_0^n(u_k^{nl}, v_k^{nl}) = 2.441208146866664 \cdot 10^{-3}, \quad G_1^n(u_k^{nl}, v_k^{nl}) = 4.451 \cdot 10^{-6},$$
  
 $d_k = -2.262 \cdot 10^{-7}.$ 

The controls obtained are shown in Figures 3 and 4, and the state in Figure 5. Note that the continuous relaxed problem here is feasible, as  $G_1(u_0^n, v_0^n) \approx -0.014 < 0$ ,  $G_1(u_k^n, v_k^n) \approx 0.002 > 0$ , where  $(u_0^n, v_0^n)$  (resp.  $(u_k^n, v_k^n)$ ) is the initial (resp. last) control pair in Example 2 (note that  $U_2 \subset U_3$ ), and the function

(5.9) 
$$\phi(\lambda) := G_1(\lambda(u_0^n, v_0^n) + (1 - \lambda)(u_k^n, v_k^n))$$
  
is continuous on [0,1].

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

periods, and yielded results of similar accuracy, but required here less than half the computing time.



Figure 1. Example 2: Last control u



Figure 2. Example 2: Last control v



Figure 3. Example 3: Last control u



Figure 4. Example 3: Last control v



Figure 5. Example (c): Last state y

## 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., **1998**, *19*, 431-464.

[3] Bonnans, J., and Casas, E., Un principe de Pontryagine pour le Contrôle des Systèmes Semilineaires Elliptiques, J. Diff. Equat., **1991** *90*, 288-303.

[4] Casas, E., and Mateos, M., Uniform Convergence of the FEM. Applications to State Constrained Control Problems, Comp. Appl. Math., **2002** *21* (1), 67-100.

[5] Casas, E., and Troeltzsch, F., Error Estimates for the Finite-element Approximation of a Semilinear Elliptic Control Problem, Control Cybern., **2002**, *90*, 695-712.

[6] Casas, E., Mateos, M., and Troeltzsch, F., Error Estimates for the Numerical Approximation of Boundary Semilinear Elliptic Control Problems, Comp. Optim. Appl. **2005**, *31*, 193-220.

[7] Chryssoverghi, I., Nonconvex Optimal Control of Nonlinear Monotone Parabolic Systems, Syst. Control Lett. **1986**, *8*, 55-62.

[8] Chryssoverghi, I., and Bacopoulos, A., Approximation of Relaxed Nonlinear Parabolic Optimal Control Problems, J. Optim. Theory Appl., **1993**, 77 (1), 31-50.

[9] Chryssoverghi, I., and Kokkinis, B, Discretization of Nonlinear Elliptic Optimal Control Problems, Syst. Control Lett., **1994**, *22*, 227-234.

[10] 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., **1997**, *94* (2), 311-334.

[11] Chryssoverghi, I., Bacopoulos, A., Coletsos, J., and Kokkinis, B., Discrete Approximation of Nonconvex Hyperbolic Optimal Control Problems with State Constraints, Control Cybernet., **1998**, *27* (1), 29-50.

[12] Chryssoverghi, I., Coletsos, J., and Kokkinis, B., Discretization Methods for Optimal Control Problems with State Constraints, J. Comput. Appl. Math., **2005**,...

[13] 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).

[14] Polak, E., *Optimization: Algorithms and Consistent Approximations*, Springer, Berlin, 1997.

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

[16] Tiba, D., and Troeltzsch, F., Error Estimates for the Discretization of State Constrained Convex Control Problems, Num. Funct. Anal. Optim., **1996**, *17*, 1005-1028.

[17] Troeltzsch, F., Semidiscrete Ritz-Galerkin Approximation of Nonlinear Parabolic Boundary Control Problem - Strong Convergence of Optimal Controls, Appl. Math. Optim., **1994**, *29*, 309-329.

[18] Troeltzsch, F., An SQP Method for Optimal Control of a Nonlinear Heat Equation, Control Cybern., **1994**, *23*, 267-288.

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