

Relaxation methods for optimal control problems

Nikolaos S. Papageorgiou

*Department of Mathematics, National Technical University
Zografou Campus, Athens 15780, Greece*

*Institute of Mathematics, Physics and Mechanics
1000 Ljubljana, Slovenia
npapg@math.ntua.gr*

Vicențiu D. Rădulescu*

*Faculty of Applied Mathematics
AGH University of Science and Technology
al. Adama Mickiewicza 30, 30-059 Kraków, Poland
Department of Mathematics, University of Craiova
200585 Craiova, Romania*

*Institute of Mathematics, Physics and Mechanics
1000 Ljubljana, Slovenia
radulescu@inf.ucv.ro*

Dušan D. Repovš

*Faculty of Education and Faculty of Mathematics and Physics
University of Ljubljana, 1000 Ljubljana, Slovenia
Institute of Mathematics, Physics and Mechanics
1000 Ljubljana, Slovenia
dusan.repovs@guest.arnes.si*

Received 7 October 2019

Accepted 20 December 2019

Published 25 February 2020

Communicated by Ahmed Alsaedi

We consider a nonlinear optimal control problem with dynamics described by a differential inclusion involving a maximal monotone map $A : \mathbb{R}^N \rightarrow 2^{\mathbb{R}^N}$. We do not assume that $D(A) = \mathbb{R}^N$, incorporating in this way systems with unilateral constraints in our framework. We present two relaxation methods. The first one is an outgrowth of the reduction

*Corresponding author.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC BY) License which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

method from the existence theory, while the second method uses Young measures. We show that the two relaxation methods are equivalent and admissible.

Keywords: Admissible relaxation; maximal monotone map; Young measure; convex conjugate; weak norm.

Mathematics Subject Classification: 49J20, 34A60

1. Introduction

We consider the following optimal control problem:

$$\left\{ \begin{array}{l} J(x, u) = \int_0^b L(t, x(t), u(t))dt \rightarrow \inf = m, \\ \text{subject to: } -x'(t) \in A(x(t)) + f(t, x(t))u(t) \text{ for almost all } t \in T = [0, b], \\ x(0) = x_0, u(t) \in U(t, x(t)) \text{ for almost all } t \in T. \end{array} \right\} \quad (1)$$

In this problem, the dynamics of the system are described by a differential inclusion involving a maximal monotone map $A : \mathbb{R}^N \rightarrow 2^{\mathbb{R}^N}$. We do not assume that $D(A) = \mathbb{R}^N$ (recall that $D(A) = \{x \in \mathbb{R}^N : A(x) \neq \emptyset\}$ is the domain of $A(\cdot)$). This way, we incorporate in our framework systems with unilateral constraints (differential variational inequalities). In addition, the control constraint set $U(t, x)$ is state-dependent, that is, the system has *a priori* feedback, a setting which is of interest in engineering and economic problems. The existence theory of such problems is based on the so-called “reduction technique”, which was developed in the pioneering works of Cesari [10, 11] and Berkovitz [3] (see also the books of Berkovitz [4] and Cesari [12]). According to this method, the original optimal control problem is reduced to a calculus of variations problem with multivalued dynamics. This problem is obtained by elimination of the control variable $u \in \mathbb{R}^m$. For this approach to work, we need to have enough convex structure in the problem, usually expressed in terms of the “property Q” of Cesari. In the absence of such a convex structure, a minimizing sequence of state-control pairs need not converge to an admissible pair. To rectify this, we need to augment the system and pass to a “convexified” version known as the “relaxed problem”, which captures the asymptotic behavior of the minimizing sequences. The process of relaxation is a delicate one since we have to strike a sensitive balance between competing requirements. We want that the relaxed problem exhibits the following three fundamental properties:

- (a) Every original state is also a relaxed state (that is, the original problem is embedded in the relaxed one).
- (b) Every relaxed state can be approximated by the original ones (that is, we want to make sure that we did not augment the system too much).
- (c) The values of the relaxed and original problems are equal and the relaxed problem has a solution (that is, there exists an optimal state-control pair).

Note that the first two requirements concern the dynamics of the system, while the third one concerns the cost functional. Any relaxation method which meets these three requirements is said to be “admissible”.

In this paper, under general conditions on the data of problem (1), we present two such admissible relaxation methods. The first one is an outgrowth of the reduction method from the existence theory, while the second method uses Young measures.

Relaxation methods for different classes of optimal control problems can be found in the works of Avgerinos and Papageorgiou [1], Buttazzo [8], Buttazzo and Dal Maso [9], Emamizadeh *et al.* [14], Hu and Papageorgiou [21], Liu *et al.* [22], Papageorgiou *et al.* [23–27], Roubiček [30], Tolstonogov [31] and Warga [33]. However, none of the aforementioned works covers the case of systems with unilateral constraints.

2. Mathematical Background and Hypotheses

Let $A : \mathbb{R}^N \rightarrow 2^{\mathbb{R}^N}$. The domain of $A(\cdot)$ is the set

$$D(A) = \{x \in \mathbb{R}^N : A(x) \neq \emptyset\}$$

and the graph of $A(\cdot)$ is the set

$$\text{Gr } A = \{(x, x^*) \in \mathbb{R}^N \times \mathbb{R}^N : x^* \in A(x)\}.$$

We say that $A(\cdot)$ is “monotone”, if

$$(x^* - y^*, x - y)_{\mathbb{R}^N} \geq 0 \quad \text{for all } (x, x^*), (y, y^*) \in \text{Gr } A.$$

If $A(\cdot)$ satisfies

$$(x^* - y^*, x - y)_{\mathbb{R}^N} > 0 \quad \text{for all } (x, x^*), (y, y^*) \in \text{Gr } A \text{ with } x \neq y,$$

then we say that $A(\cdot)$ is “strictly monotone”. We say that $A(\cdot)$ is “maximal monotone”, if $\text{Gr } A$ is maximal with respect to inclusion among the graphs of all monotone maps. This is equivalent to saying that

$$\text{“if } (x^* - y^*, x - y)_{\mathbb{R}^N} \geq 0 \quad \text{for all } (x, x^*) \in \text{Gr } A, \text{ then } (y, y^*) \in \text{Gr } A\text{”}.$$

Suppose that Y, Z are Banach spaces and $V : Y \rightarrow Z$. We say that $V(\cdot)$ is “compact”, if it is continuous and maps bounded sets in Y onto relatively compact subsets of Z . Also, we say that $V(\cdot)$ is “completely continuous”, if $y_n \xrightarrow{w} y$ in Y , implies that $V(x_n) \rightarrow V(x)$ in Z . In general, these two notions are distinct. However, if Y is reflexive, then complete continuity implies compactness. Moreover, if in addition, $V(\cdot)$ is linear, then the two notions coincide.

From the fixed point theory, we will need the so-called “Leray–Schauder alternative theorem”, which we recall here (see [17, 28]).

Theorem 1. *If Y is a Banach space, $V : Y \rightarrow Y$ is a compact map, and*

$$S = \{y \in Y : y = \lambda V(y) \text{ for some } 0 < \lambda < 1\},$$

then one of the following two statements is true:

- (a) S is unbounded;
- (b) V has a fixed point.

Now, let (Ω, Σ, μ) be a finite measure space and X a separable Banach space. We introduce the following families of subsets of X :

$$P_{f(c)}(X) = \{A \subseteq X : A \text{ is nonempty, closed (and convex)}\},$$

$$P_{(w)k(c)}(X) = \{A \subseteq X : A \text{ is nonempty, } (w\text{-}) \text{ compact (and convex)}\}.$$

A multifunction (set-valued function) $F : \Omega \rightarrow P_f(X)$ is said to be “measurable”, if for all $x \in X$ the function

$$\omega \mapsto d(x, F(\omega)) = \inf\{\|x - v\|_X : v \in F(\omega)\}$$

is measurable. If $F(\cdot)$ is measurable, then

$$\text{Gr } F = \{(\omega, v) \in \Omega \times X : v \in F(\omega)\} \in \Sigma \otimes B(X)$$

and the converse is true if Σ is μ -complete. In general, a multifunction $F : \Omega \rightarrow 2^X \setminus \{\emptyset\}$ is said to be “graph measurable”, if $\text{Gr } F \in \Sigma \otimes B(X)$.

Given $1 \leq p \leq \infty$ and a multifunction $F : \Omega \rightarrow 2^X \setminus \{\emptyset\}$, we introduce the set

$$S_F^p = \{f \in L^p(\Omega, X) : f(\omega) \in F(\omega) \mu - \text{a.e. in } \Omega\}.$$

This set can be empty. For a graph measurable multifunction $F(\cdot)$, $S_F^p \neq \emptyset$ if and only if $w \mapsto \inf\{\|v\|_X : v \in F(\omega)\}$ belongs in $L^p(\Omega)$. The set S_F^p is “decomposable”, in the sense that for every $(A, f, g) \in \Sigma \times S_F^p \times S_F^p$ we have

$$\chi_A f + \chi_{A^c} g \in S_F^p.$$

Here, for every $C \in \Sigma$, χ_C is the characteristic function of C and C^c is the complement of C (that is, $C^c = X \setminus C$). Since $\chi_{A^c} = 1 - \chi_A$, the notion of decomposability formally looks like that of convexity. Only now the coefficients are functions. Nevertheless, decomposability is a good substitute of convexity and several results valid for convex sets have their counterparts for decomposable sets (see Fryszkowski [16]).

Let Y be Hausdorff topological space and $G : Y \rightarrow 2^X \setminus \{\emptyset\}$ a multifunction. We say that $G(\cdot)$ is “lower semicontinuous” (lsc for short), respectively, “upper semicontinuous” (usc for short), if for every $U \subseteq X$ open, the set

$$F^-(U) = \{y \in Y : F(y) \cap U \neq \emptyset\},$$

$$\text{respectively, } F^+(U) = \{y \in Y : F(y) \subseteq U\}$$

is open.

Recall that on $P_f(X)$ we can define a generalized metric, known as the “Hausdorff metric”, by

$$h(A, C) = \max\{\sup\{d(a, C) : a \in A\}, \sup\{d(c, A) : c \in C\}\}$$

for all $A, C \in P_f(X)$. A multifunction $F : Y \rightarrow P_f(X)$ is said to be “ h -continuous”, if it is continuous from Y into the metric space $(P_f(X), h)$.

The next theorem, due to Bressan and Colombo [6] and Fryszkowski [15], is an illustration of how decomposability can serve as a substitute of convexity. It extends the celebrated Michael selection theorem (see, for example, Hu and Papageorgiou [20]).

Theorem 2. *If Y is a separable metric space and $F : Y \rightarrow P_f(L^1(\Omega, X))$ is a lsc multifunction with decomposable values, then there exists a continuous map $g : Y \rightarrow L^1(\Omega, X)$ such that*

$$g(y) \in F(y) \quad \text{for all } y \in Y.$$

Now, let $\Omega = T = [0, b]$. On $L^1(T, X)$ we introduce the “weak norm” defined by

$$\|u\|_w = \sup \left\{ \left\| \int_s^t u(\tau) d\tau \right\|_X : 0 \leq s \leq t \leq b \right\}$$

or, equivalently, by

$$\|u\|_w = \sup \left\{ \left\| \int_0^t u(\tau) d\tau \right\|_X : 0 \leq t \leq b \right\}.$$

From Hu and Papageorgiou [21, p. 24], we have the following result.

Proposition 3. *If X is reflexive, $\{u_n, u\}_{n \geq 1} \subseteq L^p(T, X)$, $1 \leq p < \infty$, $u_n \xrightarrow{\|\cdot\|_w} u$ and*

$$\sup_{n \geq 1} \|u_n\|_p < \infty,$$

then $u_n \xrightarrow{w} u$ in $L^p(T, X)$.

Let (Ω, Σ, μ) be a complete finite measure space and Y a Polish space. Recall that this means that Y is a separable Hausdorff topological space and there is a metric d on Y compatible with the topology of Y such that (Y, d) is complete. By $P(Y)$ we denote the set of all probability measures on Y endowed with the narrow topology τ_n (see Papageorgiou and Winkert [29, p. 375]). Let $B(Y)$ be a Borel σ -field on Y , $ca(\Sigma \otimes B(Y))$ the space of all \mathbb{R} -valued signed measures on $\Sigma \otimes B(Y)$ and $p_\Omega : \Omega \times Y \rightarrow \Omega$ the projection map. Given $\lambda \in ca(\Sigma \otimes B(Y))$ and $\mu = \lambda \circ p_\Omega^{-1}$, the disintegration theorem says that there exists a Σ -measurable map $\hat{\lambda} : \Omega \rightarrow P(Y)$ such that

$$\lambda(A \times C) = \int_A \hat{\lambda}(\omega)(C) d\mu \quad \text{for all } A \in \Sigma, C \in B(Y).$$

By $ca_+(\Sigma \otimes B(Y))$ we denote the \mathbb{R}_+ -valued elements of $ca(\Sigma \times B(Y))$.

A “Young measure” on $\Omega \times Y$ is a $\lambda \in ca_+(\Sigma \otimes B(Y))$ such that

$$\mu = \lambda \circ p_\Omega^{-1}$$

(that is, $\mu(A) = \lambda(A \times Y)$ for all $A \in \Sigma$). The space of Young measures on $\Omega \times Y$ is denoted by $\mathcal{Y}(\Omega \times Y)$. On account of the disintegration theorem mentioned above,

we can identify $\lambda \in \mathcal{Y}(\Omega \times Y)$ with its disintegration $\hat{\lambda}(\cdot)$. So, we say that a Young measure is a measurable map $\hat{\lambda} : \Omega \rightarrow P(Y)$. Such maps are also known as “transition measures”. The space of transition measures is denoted by $R(\Omega, Y)$. We know that the following statements are equivalent:

- (a) $\hat{\lambda} \in \mathbb{R}(\Omega, Y)$.
- (b) For every $C \in B(Y)$, the map $\omega \mapsto \xi_C(\omega) = \hat{\lambda}(\omega)(C)$ is Σ -measurable (see Papageorgiou and Winkert [29, p. 387]). Given a Σ -measurable function $u : \Omega \rightarrow Y$, the “Young measure associated with u ”, is the transition probability defined by

$$\hat{\lambda}^u(\omega) = \delta_{u(\omega)}$$

with $\delta_{u(\omega)}(\cdot)$ being the Dirac measure defined by

$$\delta_{u(\omega)}(C) = \begin{cases} 1 & \text{if } u(\omega) \in C \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } C \in B(Y).$$

Let $\varphi : \Omega \times Y \rightarrow \mathbb{R}$ be a “Carathéodory function”, that is, for all $y \in Y$, the mapping $\omega \mapsto \varphi(\omega, y)$ is Σ -measurable and for μ -a.e. $\omega \in \Omega$, the mapping $y \mapsto \varphi(\omega, y)$ is continuous. We know that such function is $\Sigma \otimes B(Y)$ -measurable (see Hu and Papageorgiou [20, p. 142]). We say that a Carathéodory function is “ L^1 -bounded”, if there exists $k \in L^1(\Omega)$ such that $|\varphi(\omega, y)| \leq k(\omega)$ μ -a.e. on Ω , for all $y \in Y$. By $\text{Car}_b(\Omega \times Y)$ we denote the space of all L^1 -bounded Carathéodory functions.

The “Young narrow topology” on $\mathcal{Y}(\Omega \times Y)$, is the weakest topology on $\mathcal{Y}(\Omega \times Y)$ for which the maps

$$\lambda \mapsto I_\varphi(\lambda) = \int_{\Omega \times Y} \varphi(\omega, y) d\lambda = \int_{\Omega} \left(\int_Y \varphi(\omega, y) \hat{\lambda}(\omega)(dy) \right) d\mu$$

with $\hat{\lambda}(\cdot)$ being the disintegration of λ and $\varphi \in \text{Car}_b(\Omega \times Y)$, are all continuous. This topology on $\mathcal{Y}(\Omega \times Y)$ is denoted by τ_n^y .

Now, we introduce the hypotheses on the data of problem (1).

$H(A)$: $A : D(A) \subseteq \mathbb{R}^N \rightarrow 2^{\mathbb{R}^N}$ is a maximal monotone map such that $0 \in A(0)$.

Remark 1. We do not require that $D(A) = \mathbb{R}^N$. In this way, our framework incorporates systems with unilateral constraints (differential variational inequalities).

$H(f)$: $f : T \times \mathbb{R}^N \rightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^N)$ is a map such that

- (i) for all $x \in \mathbb{R}^N$ and all $u \in \mathbb{R}^m$, $t \mapsto f(t, x)u$ is measurable;
- (ii) for every $r > 0$, there exists $l_r \in L^1(T)$ such that

$$\|f(t, x) - f(t, y)\|_{\mathcal{L}} \leq l_r(t)|x - y| \quad \text{for almost all } t \in T \text{ and all } |x|, |y| \leq r;$$

- (iii) $\|f(t, x)\|_{\mathcal{L}} \leq a(t)(1 + |x|)$ for almost all $t \in T$ and all $x \in \mathbb{R}^N$, with $a \in L^1(T)$.

$H(U): U : T \times \mathbb{R}^N \rightarrow P_k(\mathbb{R}^m)$ is a multifunction such that

- (i) for all $x \in \mathbb{R}^N$, $t \mapsto U(t, x)$ is measurable;
- (ii) there exists $k \in L^\infty(T)$ such that

$$h(U(t, x), U(t, y)) \leq k(t)|x - y| \quad \text{for almost all } t \in T \text{ and all } x, y \in \mathbb{R}^N;$$

- (iii) $|U(t, x)| = \sup\{|u| : u \in U(t, x)\} \leq a_0(t)$ for almost all $t \in T$ and all $x \in \mathbb{R}^N$, with $a_0 \in L^\infty(T)$.

$H(L): L : T \times \mathbb{R}^N \times \mathbb{R}^m \rightarrow \mathbb{R}$ is a function such that

- (i) for all $(x, u) \in \mathbb{R}^N \times \mathbb{R}^m$, $t \mapsto L(t, x, u)$ is measurable;
- (ii) for every $r > 0$, there exists $\vartheta_r \in L^1(T)$ such that

$$|L(t, x, u) - L(t, y, v)| \leq \vartheta_r(t)(|x - y| + |u - v|)$$

for almost all $t \in T$ and all $|x|, |y|, |u|, |v| \leq r$;

- (iii) for every $r > 0$, there exists $a_r \in L^1(T)$ such that

$$|L(t, x, u)| \leq a_r(t) \quad \text{for almost all } t \in T \text{ and all } |x|, |u| \leq r.$$

$H_0: x_0 \in \overline{D(A)}$.

We introduce the “convexified” dynamics of problem (1), namely the following control system:

$$\left\{ \begin{array}{l} -x'(t) \in A(x(t)) + f(t, x(t))u(t) \text{ for almost all } t \in T = [0, b], \\ x(0) = x_0, u(t) \in \text{conv } U(t, x(t)) \text{ for almost all } t \in T. \end{array} \right\}. \quad (2)$$

Then we define the following two sets:

$$P = \{(x, u) \in C(T, \mathbb{R}^N) \times L^1(T, \mathbb{R}^m) : (x, u) \text{ is an admissible state-control pair for (1)}\},$$

$$P_c = \{(x, u) \in C(T, \mathbb{R}^N) \times L^1(T, \mathbb{R}^m) : (x, u) \text{ is an admissible state-control pair for (2)}\}.$$

Also, we set

$$\mathcal{S} = \text{proj}_{C(T, \mathbb{R}^N)} P \text{ and } \mathcal{S}_c = \text{proj}_{C(T, \mathbb{R}^N)} P_c.$$

These are the sets of admissible trajectories for the original system (the set \mathcal{S}) and for the convexified system (the set \mathcal{S}_c).

3. First Relaxation Method

The first relaxation method is motivated by the “reduction method” of the existence theory and it uses the convexified control system (2).

Proposition 4. *If hypotheses $H(A), H(f), H(U), H_0$ hold, then $P \neq \emptyset$ and there exists $\hat{c} > 0$ such that $\|x\|_{C(T, \mathbb{R}^N)} \leq \hat{c}$ for all $x \in \mathcal{S}_c$.*

Proof. Consider the following orientor field:

$$F(t, x) = f(t, x)U(t, x) = \bigcup_{u \in U(t, x)} f(t, x)u \in P_k(\mathbb{R}^N).$$

Claim 1. *For every $x \in \mathbb{R}^N$, the map $t \mapsto F(t, x)$ is measurable.*

We fix $x \in \mathbb{R}^N$ and consider the multifunction $t \rightarrow U(t, x)$. The measurability of this multifunction (see hypothesis $H(U)(i)$) implies that we can find a sequence $\{u_n^x\}_{n \geq 1} \subseteq S_{U(\cdot, x)} = \{u : T \rightarrow \mathbb{R}^m \text{ measurable and } u(t) \in U(t, x) \text{ for almost all } t \in T\}$ such that

$$U(t, x) = \{\overline{u_n^x(t)}\}_{n \geq 1} \quad \text{for almost all } t \in T$$

(see Hu and Papageorgiou [20, Theorem 24, p. 156]). Then we have

$$\begin{aligned} F(t, x) &= f(t, x)\{\overline{u_n^x(t)}\}_{n \geq 1} = \{\overline{f(t, x)u_n^x(t)}\}_{n \geq 1}, \\ &\Rightarrow t \mapsto F(t, x) \text{ is measurable.} \end{aligned}$$

This proves Claim 1.

Claim 2. *For almost all $t \in T$, the map $x \mapsto F(t, x)$ is locally h -Lipschitz.*

Let $r > 0$ and consider $x, y \in \mathbb{R}^N$ such that $|x|, |y| \leq r$. If $\xi \in F(t, x)$, then

$$\xi = f(t, x)u \quad \text{with } u \in U(t, x).$$

Given $\epsilon > 0$, we choose $v \in U(t, y)$ such that

$$\begin{aligned} |u - v| &\leq d(u, U(t, y)) + \epsilon \\ &\leq h(U(t, x), U(t, y)) + \epsilon \\ &\leq k(t)|x - y| + \epsilon \quad \text{for almost all } t \in T \end{aligned} \tag{3}$$

(see hypothesis $H(U)(ii)$).

Then we have

$$\begin{aligned} d(\xi, F(t, y)) &= d(f(t, x)u, F(t, y)) \\ &\leq |f(t, x)u - f(t, y)v| \\ &\leq \|f(t, x)\|_{\mathcal{L}}|u - v| + |f(t, x)v - f(t, y)v| \\ &\leq a_r(t)|u - v| + l_r(t)|x - y| \quad (\text{with } a_r(\cdot) = a(\cdot)(1 + r)) \end{aligned}$$

$$\begin{aligned} &\leq [a_r(t)k(t) + l_r(t)]|x - y| + a_r(t)\epsilon \quad \text{for almost all } t \in T \text{ (see (3))} \\ &\Rightarrow h(F(t, x), F(t, y)) \leq k_r(t)|x - y| + a_r(t)\epsilon \\ &\text{for almost all } t \in T, \text{ with } k_r \in L^1(T). \end{aligned}$$

Let $\epsilon \rightarrow 0^+$, and conclude that for almost all $t \in T$, the multifunction $x \rightarrow F(t, x)$ is locally h -Lipschitz. This proves Claim 2.

From Claims 1 and 2 it follows that $(t, x) \rightarrow F(t, x)$ is measurable (see Hu and Papageorgiou [20, Proposition 7.9, p. 229]). Thus, we can infer that for every measurable function $x : T \rightarrow \mathbb{R}^N$, the multifunction $t \rightarrow F(t, x(t))$ is measurable (superpositional measurability).

We consider the following multivalued Cauchy problem:

$$-x'(t) \in A(x(t)) + F(t, x(t)) \quad \text{for almost all } t \in T, \quad x(0) = x_0. \quad (4)$$

We will show the existence of solutions for this problem. To this end, let $h \in L^1(T, \mathbb{R}^N)$ and consider the following Cauchy problem:

$$-x'(t) \in A(x(t)) + h(t) \quad \text{for almost all } t \in T, \quad x(0) = x_0. \quad (5)$$

By virtue of the Bénéilan–Brezis theorem (see Brezis [7, Proposition 3.8, p. 82]), we know that problem (5) has a unique solution $x \in W^{1,1}((0, b), \mathbb{R}^N) = AC^1(T, \mathbb{R}^N)$. So, we can define the solution map $K : L^1(T, \mathbb{R}^N) \rightarrow C(T, \mathbb{R}^N)$, which to each $h \in L^1(T, \mathbb{R}^N)$ assigns the unique solution $K(h) \in AC^1(T, \mathbb{R}^N) \subseteq C(T, \mathbb{R}^N)$ of (5).

Claim 3. *The solution map $K : L^1(T, \mathbb{R}^N) \rightarrow C(T, \mathbb{R}^N)$ is completely continuous.*

Suppose that $h_n \xrightarrow{w} h$ in $L^1(T, \mathbb{R}^N)$ and let $x_n = K(h_n)$ for all $n \in \mathbb{N}$ and $x = K(h)$. From Baras [2], we know that

$$\{x_n\}_{n \geq 1} \subseteq C(T, \mathbb{R}^N) \text{ is relatively compact.}$$

So, up to a subsequence, we may assume that

$$x_n \rightarrow x^* \quad \text{in } C(T, \mathbb{R}^N). \quad (6)$$

Exploiting the monotonicity of $A(\cdot)$, we obtain

$$(x'_n(t) - x'(t), x_n(t) - x(t))_{\mathbb{R}^N} + (h_n(t) - h(t), x_n(t) - x(t))_{\mathbb{R}^N} \leq 0$$

for almost all $t \in T$,

$$\Rightarrow \frac{1}{2} \frac{d}{dt} |x_n(t) - x(t)|^2$$

$$\leq (h_n(t) - h(t), x(t) - x_n(t))_{\mathbb{R}^N} \quad \text{for almost all } t \in T,$$

$$\Rightarrow |x_n(t) - x(t)|^2 \leq 2 \int_0^t (h_n(s) - h(s),$$

$$x(s) - x_n(s))_{\mathbb{R}^N} ds \quad \text{for all } t \in T, n \in \mathbb{N},$$

$$\begin{aligned} &\Rightarrow |x^*(t) - x(t)|^2 \leq 0 \quad \text{for all } t \in T \\ &\quad (\text{see (6) and recall that } h_n \xrightarrow{w} h \text{ in } L^1(T, \mathbb{R}^N)), \\ &\Rightarrow x = x^*. \end{aligned}$$

We deduce that for the original sequence we have

$$\begin{aligned} x_n &\rightarrow x \text{ in } C(T, \mathbb{R}^N), \\ &\Rightarrow K : L^1(T, \mathbb{R}^N) \rightarrow C(T, \mathbb{R}^N) \text{ is completely continuous.} \end{aligned}$$

This proves Claim 3.

Let $N_F : C(T, \mathbb{R}^N) \rightarrow 2^{L^1(T, \mathbb{R}^N)}$ be the multivalued Nemytski operator corresponding to the multifunction $F(t, x)$, that is,

$$N_F(x) = S_{F(\cdot, x(\cdot))}^1 \text{ for all } x \in C(T, \mathbb{R}^N).$$

The measurability of $t \rightarrow F(t, x(t))$ and hypotheses $H(f)$ (iii) and $H(U)$ (iii) imply that

$$N_F(x) \in P_{wk}(L^1(T, \mathbb{R}^N)) \quad \text{for all } x \in C(T, \mathbb{R}^N).$$

On account of Hu and Papageorgiou [20, Theorems 7.26 and 7.27, p. 237], the multifunction $x \rightarrow N_F(x)$ is h -continuous. In particular, it is also lsc (see Hu and Papageorgiou [20, Proposition 2.66, p. 61]). So, we can apply Theorem 2 and produce a continuous map $g : C(T, \mathbb{R}^N) \rightarrow L^1(T, \mathbb{R}^N)$ such that

$$g(x) \in N_F(x) \quad \text{for all } x \in C(T, \mathbb{R}^N). \tag{7}$$

Evidently, the map

$$x \rightarrow (K \circ g)(x) \text{ is continuous.} \tag{8}$$

Hypotheses $H(f)$ (iii) and $H(U)$ (iii) and the complete continuity of $K(\cdot)$, imply that $(K \circ g)(\cdot)$ maps bounded sets in $C(T, \mathbb{R}^N)$ to relatively compact sets in $C(T, \mathbb{R}^N)$. Therefore

$$x \rightarrow (K \circ g)(x) \text{ is compact.} \tag{9}$$

Consider the set

$$C = \{x \in C(T, \mathbb{R}^N) : x = \lambda(K \circ g)(x), 0 < \lambda < 1\}.$$

Claim 4. *The set $C \subseteq C(T, \mathbb{R}^N)$ is bounded.*

Let $x \in C$. Then

$$\frac{1}{\lambda}x = K(g(x)),$$

$$\Rightarrow -x'(t) \in \lambda A(\lambda^{-1}x(t)) + \lambda g(x)(t) \quad \text{for almost all } t \in T, \quad x(0) = \lambda x_0.$$

We take inner product with $x(t)$ and use hypothesis $H(A)$. Then

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |x(t)|^2 + \lambda(g(x)(t), x(t))_{\mathbb{R}^N} &\leq 0 \quad \text{for almost all } t \in T, \\ \Rightarrow \frac{1}{2} |x(t)|^2 &\leq \frac{1}{2} |x_0|^2 + \int_0^t |g(x)(s)| |x(s)| ds \quad (\text{since } 0 < \lambda < 1). \end{aligned}$$

Invoking Lemma A.5 of Brezis [7, p. 157], we obtain

$$\begin{aligned} |x(t)| &\leq |x_0| + \int_0^t |g(x)(s)| ds \\ &\leq |x_0| + \int_0^t \hat{a}(s)(1 + |x(s)|) ds \quad \text{for all } t \in T, \text{ with } \hat{a} \in L^1(T) \\ &\quad (\text{see (7) and hypotheses } H(F)(\text{iii}), H(U)(\text{iii})), \\ \Rightarrow |x(t)| &\leq c_1 \quad \text{for all } t \in T, \text{ all } x \in C, \\ &\quad \text{and some } C_1 > 0 \text{ (by Gronwall's inequality)}. \end{aligned}$$

Therefore $C \subseteq C(T, \mathbb{R}^N)$ is bounded. This proves Claim 4.

Then (9) and Claim 4 enable us to apply Theorem 1 (the Leray–Schauder alternative theorem). So, we can find $\hat{x} \in C(T, \mathbb{R}^N)$ such that

$$\begin{aligned} \hat{x} &= (K \circ g)(\hat{x}), \\ \Rightarrow -\hat{x}'(t) &\in A(\hat{x}(t)) + g(\hat{x})(t) \in A(\hat{x}(t)) + F(t, \hat{x}(t)) \quad \text{for almost all } t \in T, \\ \hat{x}(0) &= x_0. \end{aligned}$$

Consider the multifunction

$$\begin{aligned} E(t) &= \{u \in U(t, \hat{x}(t)) : g(\hat{x})(t) = f(t, \hat{x}(t))u\}, \\ \Rightarrow \text{Gr}E &\in \mathcal{L}_T \otimes B(\mathbb{R}^m), \end{aligned}$$

with \mathcal{L}_T being the Lebesgue σ -field of T and $B(\mathbb{R}^m)$ the Borel σ -field of \mathbb{R}^m . Applying the Yankov–von Neumann–Aumann selection theorem (see Hu and Papageorgiou [20, Theorem 2.14, p. 158]), we obtain a measurable map $\hat{u} : T \rightarrow \mathbb{R}^m$ such that

$$\begin{aligned} \hat{u}(t) &\in E(t) \quad \text{for almost all } t \in T, \\ \Rightarrow g(\hat{x})(t) &= f(t, \hat{x}(t))\hat{u}(t) \quad \text{for almost all } t \in T, \\ \Rightarrow (\hat{x}, \hat{u}) &\in P. \end{aligned}$$

Moreover, as in the proof of Claim 4, we can show that there exists $\hat{c} > 0$ such that

$$\|x\|_{C(T, \mathbb{R}^N)} \leq \hat{c} \quad \text{for all } x \in \mathcal{S}_C.$$

The proof is now complete. □

In what follows, we denote by $L^1(T, \mathbb{R}^m)_w$ the Lebesgue space $L^1(T, \mathbb{R}^N)$ equipped with the weak topology and by $L^1_w(T, \mathbb{R}^N)$ the same space furnished with the weak norm.

Proposition 5. *If hypotheses $H(A), H(f), H(U), H_0$ hold then $P_c \subseteq C(T, \mathbb{R}^N) \times L^1(T, \mathbb{R}^m)_w$ is sequentially compact.*

Proof. Let $\{(x_n, u_n)\}_{n \geq 1} \subseteq P_c$. Hypothesis $H(U)$ (iii) implies that by passing to a subsequence if necessary, we may assume that

$$u_n \xrightarrow{w} u \quad \text{in } L^1(T, \mathbb{R}^m). \tag{10}$$

Recall that $\|x_n\|_{C(T, \mathbb{R}^N)} \leq \hat{c}$ for all $n \in \mathbb{N}$. Hence, by hypothesis $H(f)$ (iii) and the Dunford–Pettis theorem, we see that

$$\{g_n(\cdot) = f(\cdot, x_n(\cdot))u_n(\cdot)\}_{n \geq 1} \subseteq L^1(T, \mathbb{R}^n)$$

is relatively w -compact. So, we may assume that for some $g \in L^1(T, \mathbb{R}^N)$, we have

$$g_n \xrightarrow{w} g \quad \text{in } L^1(T, \mathbb{R}^N). \tag{11}$$

By Claim 4 in the proof of Proposition 4, we have

$$x_n = K(g_n) \rightarrow K(g) = x \quad \text{in } C(T, \mathbb{R}^N). \tag{12}$$

For every $h \in L^\infty(T, \mathbb{R}^N) = L^1(T, \mathbb{R}^N)^*$, we have

$$\int_0^b (g_n, h)_{\mathbb{R}^N} dt = \int_0^b (u_n, f(t, x_n)^* h)_{\mathbb{R}^m} dt. \tag{13}$$

Note that

$$\begin{aligned} \|f(t, x_n(t))^* - f(t, x(t))^*\|_{\mathcal{L}} &= \|f(t, x_n(t)) - f(t, x(t))\|_{\mathcal{L}} \\ &\leq l_{\hat{c}}(t) \|x_n - x\|_{C(T, \mathbb{R}^N)} \quad \text{for almost all } t \in T, \\ &\Rightarrow \|f(\cdot, x_n(\cdot))^* h(\cdot) - f(\cdot, x(\cdot))^* h(\cdot)\|_{\infty} \\ &\rightarrow 0 \text{ as } n \rightarrow \infty \quad (\text{see (12)}). \end{aligned} \tag{14}$$

If in (13), we pass to the limit as $n \rightarrow \infty$ and use (10), (11), (14), then

$$\begin{aligned} \int_0^b (g, h)_{\mathbb{R}^N} dt &= \int_0^b (f(t, x)u, h)_{\mathbb{R}^N} dt \quad \text{for all } h \in L^\infty(T, \mathbb{R}^N), \\ &\Rightarrow g(\cdot) = f(\cdot, x(\cdot))u(\cdot). \end{aligned} \tag{15}$$

From Proposition 3.6 of Brezis [7, p. 70], we have

$$\begin{aligned} \frac{1}{2}|x_n(t) - y|^2 &\leq \frac{1}{2}|x_n(s) - y|^2 + \int_s^t (g_n(\tau) - y^*, x_n(\tau) - y)_{\mathbb{R}^N} dt \\ &\text{for all } 0 \leq s \leq t \leq b, (y, y^*) \in \text{Gr } A, n \in \mathbb{N}, \end{aligned}$$

$$\Rightarrow \frac{1}{2}|x(t) - y|^2 \leq \frac{1}{2}|x(s) - y|^2 + \int_s^t (g(\tau) - y^*, x(\tau) - y)_{\mathbb{R}^N} dt$$

for all $0 \leq s \leq t \leq b, (y, y^*) \in \text{Gr } A$.

Invoking once again Proposition 3.6 of Brezis [7, p. 70], we have

$$\begin{aligned} -x'(t) &\in A(x(t)) + f(t, x(t))u(t) \quad \text{for almost all } t \in T, \\ x(0) = x_0, u(t) &\in \text{conv } U(t, x(t)) \quad \text{for almost all } t \in T, \end{aligned}$$

(see (15) and hypothesis $H(U)$ (ii)). Therefore

$$(x, u) \in P_c$$

and so we can conclude that P_c is sequentially compact in $C(T, \mathbb{R}^N) \times L^1(T, \mathbb{R}^m)_w$. □

Let $V = \mathbb{R}^m \times \mathbb{R}$ and consider the multifunction $\Gamma : T \times \mathbb{R}^N \rightarrow P_f(V)$ defined by

$$\Gamma(t, x) = \{(u, \eta) \in \mathbb{R}^m \times \mathbb{R} : u \in U(t, x), L(t, x, u) \leq \eta \leq a_{\hat{c}}(t)\}.$$

where $\hat{c} > 0$ is from Proposition 4 and $a_{\hat{c}}(\cdot)$ from $H(L)$ (iii). (16)

Proposition 6. *If hypotheses $H(U), H(L)$ hold, then*

- (a) *the multifunction $(t, x) \rightarrow \Gamma(t, x)$ is graph measurable;*
- (b) *for almost all $t \in T$, the map $x \rightarrow \Gamma(t, x)$ is usc.*

Proof. (a) Let $\overline{B}(t) = \{(x, u, \eta) \in \mathbb{R}^N \times \mathbb{R}^m \times \mathbb{R} : |u| \leq a_0(t), |\eta| \leq a_{\hat{c}}(t)\}$. Then

$$\text{Gr } \Gamma = (\text{Gr } U \times \mathbb{R}) \cap \text{epi } L \cap \overline{B}.$$

Recall that $\text{epi } L = \{(t, x, u, \eta) \in T \times \mathbb{R}^N \times \mathbb{R}^m \times \mathbb{R} : L(t, x, u) \leq \eta\}$. Hypotheses $H(L)$ (i), (ii) imply that $(t, x, u) \rightarrow L(t, x, u)$ is measurable. So, it follows that $\text{epi } L \in \mathcal{L}_T \otimes B(\mathbb{R}^N) \otimes B(\mathbb{R}^m) \otimes B(\mathbb{R})$. Similarly, hypotheses $H(U)$ (i), (ii) imply that $(t, x) \rightarrow U(t, x)$ is graph measurable. Therefore

$$\text{Gr } U \times \mathbb{R} \in \mathcal{L}_T \otimes B(\mathbb{R}^N) \otimes B(\mathbb{R}^m) \otimes B(\mathbb{R}).$$

Since $\overline{B}(\cdot)$ is clearly graph measurable, we conclude that

$$\text{Gr } \Gamma \in \mathcal{L}_T \otimes B(\mathbb{R}^N) \otimes B(\mathbb{R}^m) \otimes B(\mathbb{R}).$$

(b) Suppose that

$$\begin{aligned} x_n \rightarrow x \text{ in } \mathbb{R}^N, \quad u_n \rightarrow u \text{ in } \mathbb{R}^m, \quad \eta_n \rightarrow \eta \text{ in } \mathbb{R}, \\ (u_n, \eta_n) \in \Gamma(t, x_n) \quad \text{for all } n \in \mathbb{N}. \end{aligned}$$

We have

$$\begin{aligned}
 u_n &\in U(t, x_n), L(t, x_n, u_n) \leq \eta_n \leq a_{\hat{c}}(t) \\
 &\Rightarrow u \in U(t, x), L(t, x, u) \leq \eta \leq a_{\hat{c}}(t) \\
 &\quad (\text{see hypotheses } H(U)(\text{ii}), H(L)(\text{ii})), \\
 &\Rightarrow (u, \eta) \in \text{Gr}\Gamma(t, \cdot), \\
 &\Rightarrow x \rightarrow \Gamma(t, x) \text{ is usc (see Hu and Papageorgiou [20, Proposition 2.23, p. 43])}.
 \end{aligned}$$

The proof is now complete. □

We introduce the integrand $\hat{L} : T \times \mathbb{R}^N \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ defined by

$$\hat{L}(t, x, u) = \begin{cases} L(t, x, u) & \text{if } u \in U(t, x), \\ +\infty & \text{otherwise.} \end{cases} \tag{17}$$

By $\hat{L}^{**}(t, x, u)$ we denote the second convex conjugate of the function $u \rightarrow \hat{L}(t, x, u)$ (see, for example, Gasinski and Papageorgiou [17, p. 512]).

Then the relaxed optimal control problem for (1) is

$$\left\{ \begin{aligned} J_r(x, u) &= \int_0^b \hat{L}^{**}(t, x(t), u(t)) dt \rightarrow \inf = m_r, \\ &\text{subject to: } (x, u) \in P_c. \end{aligned} \right\} \tag{18}$$

In what follows, given a function $\varphi : \mathbb{R}^m \rightarrow \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$, we set

$$\text{dom } \varphi = \{u \in \mathbb{R}^m : \varphi(u) < +\infty\}$$

(the effective domain of φ).

Proposition 7. *If hypotheses $H(U), H(L)$ hold, then*

- (a) $\text{dom } \hat{L}^{**}(t, x, \cdot) = \text{conv } U(t, x)$ for all $(t, x) \in T \times \mathbb{R}^N$;
- (b) $\hat{L}^{**}(t, x, u) = \min\{\eta \in \mathbb{R} : (u, \eta) \in \overline{\text{conv}} \Gamma(t, x)\}$ for almost all $t \in T$ and all $(x, u) \in \mathbb{R}^N \times \mathbb{R}^m$.

Proof. (a) From Proposition 3.2 of Ekeland and Temam [13, p. 16], we know that

$$\text{epi } \hat{L}^{**}(t, x, \cdot) = \overline{\text{conv}} \text{epi } \hat{L}(t, x, \cdot) \quad \text{for almost all } t \in T \text{ and all } x \in \mathbb{R}^N. \tag{19}$$

Note that $\text{dom } \hat{L}(t, x, \cdot) = U(t, x)$ (see (17)). So, from (19) we infer that

$$\text{dom } \hat{L}^{**}(t, x, \cdot) = \text{conv } U(t, x) \quad \text{for almost all } t \in T \text{ and all } x \in \mathbb{R}^N.$$

(b) Similarly, from (19) and (16), we see that

$$\hat{L}^{**}(t, x, u) = \inf\{\eta \in \mathbb{R} : (u, \eta) \in \text{conv } \Gamma(t, x)\}$$

$$\text{for almost all } t \in T \text{ and all } (x, u) \in \mathbb{R}^N \times \mathbb{R}^m.$$

The proof is now complete. □

We will also need the following result about the integrable selectors of a graph measurable multifunction, which is itself of independent interest.

Proposition 8. *If $G : T \rightarrow P_k(\mathbb{R}^d)$ ($d \geq 1$) is graph measurable and $S_G^1 \neq \emptyset$ and is uniformly integrable, then $\overline{S_G^1}^{\|\cdot\|_w} = S_{\text{conv } G}^1$.*

Proof. From Proposition 3.30 of Hu and Papageorgiou [20, p. 185], we have

$$S_{\text{conv } G}^1 = \overline{S_G^1}^w \quad \text{in } L^1(\Omega, \mathbb{R}^N). \tag{20}$$

On the other hand, Proposition 7.16 of Hu and Papageorgiou [20, p. 232], asserts that given any $\epsilon > 0$, we can find $T_\epsilon \subseteq T$ closed such that

$$|T \setminus T_\epsilon|_1 < \epsilon \quad \text{and} \quad G|_{T_\epsilon} \text{ is } h\text{-continuous}$$

(here by $|\cdot|$, we denote the Lebesgue measurable on \mathbb{R}). Then we have $G(T_\epsilon) \in P_k(\mathbb{R}^d)$. Moreover, since S_G^1 is uniformly integrable, so is $\overline{\text{conv } S_G^1} = S_{\text{conv } G}^1$. Hence $S_{\text{conv } G}^1$ has property U of Bourgain [5] and so applying the theorem of Gutman [19] (see also Hu and Papageorgiou [20, Proposition 4.14, p. 195]), we infer that on $S_{\text{conv } G}^1$ the w -topology and the $\|\cdot\|_w$ -topology coincide. So, from (20), we have

$$S_{\text{conv } G}^1 = \overline{S_G^1}^{\|\cdot\|_w}.$$

The proof is now complete. □

Using this proposition, we can get the following approximation result which is critical in establishing the admissibility of the relaxation (18).

Proposition 9. *If hypotheses $H(A), H(f), H(U), H(L), H_0$ hold and $(\hat{x}, \hat{u}) \in P_c$, then we can find $\hat{u}_n \in S_{U(\cdot, \hat{x}(\cdot))}^1$, $n \in \mathbb{N}$, such that*

$$\begin{aligned} \hat{u}_n &\xrightarrow{\|\cdot\|_w} \hat{u} \quad \text{in } L^1(T, \mathbb{R}^m) \\ L(\cdot, \hat{x}_n(\cdot), \hat{u}_n(\cdot)) &\xrightarrow{\|\cdot\|_w} \hat{L}^{**}(\cdot, \hat{x}(\cdot), \hat{u}(\cdot)) \quad \text{in } L^1(T). \end{aligned}$$

Proof. Since $(\hat{x}, \hat{u}) \in P_c$, from Proposition 7, we have

$$\begin{aligned} (\hat{u}(t), \hat{L}^{**}(t, \hat{x}(t), \hat{u}(t))) &\in \overline{\text{conv } \Gamma(t, \hat{x}(t))} \quad \text{for almost all } t \in T, \\ \Rightarrow \hat{w}(\cdot) = (\hat{u}(\cdot), \hat{L}^{**}(\cdot, \hat{x}(\cdot), \hat{u}(\cdot))) &\in S_{\text{conv } \Gamma(\cdot, \hat{x}(\cdot))}^1. \end{aligned}$$

On account of Proposition 8, we can find $w_n \in S_{\Gamma(\cdot, \hat{x}(\cdot))}^1$ such that

$$w_n \xrightarrow{\|\cdot\|_w} w \quad \text{in } L^1(T, \mathbb{R}^m).$$

But $w_n(\cdot) = (\hat{u}_n(\cdot), \eta_n(\cdot))$ with $\hat{u}_n \in S_{U(\cdot, \hat{x}(\cdot))}^1$, $L(t, \hat{x}(t), \hat{u}_n(t)) \leq \eta_n(t) \leq a_{\hat{z}}(t)$ for almost all $t \in T$, and all $n \in \mathbb{N}$. Therefore

$$\hat{u}_n \xrightarrow{\|\cdot\|_w} \hat{u} \quad \text{in } L^1(T, \mathbb{R}^m) \quad \text{and} \quad \sup_{t \in T} \int_0^t (\eta_n(s) - \hat{L}^{**}(s, \hat{x}(s), \hat{u}_n(s))) ds \rightarrow 0,$$

$$\begin{aligned} &\Rightarrow \sup_{t \in T} \int_0^t (L(s, \hat{x}(s), \hat{u}_n(s)) - \hat{L}^{**}(s, \hat{x}(s), \hat{u}_n(s))) ds \rightarrow 0, \\ &\Rightarrow L(\cdot, \hat{x}(\cdot), \hat{u}_n(\cdot)) \xrightarrow{\|\cdot\|_w} \hat{L}^{**}(\cdot, \hat{x}(\cdot), \hat{u}(\cdot)) \quad \text{in } L^1(\Omega). \end{aligned}$$

The proof is now complete. □

Now we can establish the density of P in P_c in the space $C(T, \mathbb{R}^N) \times L^1(T, \mathbb{R}^m)_w$.

Proposition 10. *If hypotheses $H(A), H(f), H(U), H(L), H_0$ hold and $(\hat{x}, \hat{u}) \in P_c$, then we can find $\{(\hat{x}_n, \hat{v}_n)\}_{n \geq 1} \subseteq P$ such that*

$$\begin{aligned} \hat{x}_n &\rightarrow \hat{x} \quad \text{in } C(T, \mathbb{R}^N), \\ v_n &\xrightarrow{\|\cdot\|_w} \hat{u} \quad \text{in } L^1(T, \mathbb{R}^m), \\ L(\cdot, \hat{x}_n(\cdot), \hat{v}_n(\cdot)) &\xrightarrow{\|\cdot\|_w} \hat{L}^{**}(\cdot, \hat{x}(\cdot), \hat{u}(\cdot)). \end{aligned}$$

Proof. According to Proposition 9, we can find $\{\hat{u}_n\}_{n \geq 1} \subseteq S_{U(\cdot, \hat{x}(\cdot))}^1$ such that

$$\begin{aligned} \hat{u}_n &\xrightarrow{\|\cdot\|_w} \hat{u} \quad \text{in } L^1(T, \mathbb{R}^m) \quad \text{and} \\ L(\cdot, \hat{x}(\cdot), \hat{u}_n(\cdot)) &\xrightarrow{\|\cdot\|_w} \hat{L}^{**}(\cdot, \hat{x}(\cdot), \hat{u}(\cdot)) \quad \text{in } L^1(T). \end{aligned} \tag{21}$$

Since $\{\hat{u}\}_{n \geq 1} \subseteq L^1(T, \mathbb{R}^m)$ is bounded (see hypothesis $H(U)$ (iii)), by Proposition 3, we have

$$\hat{u}_n \xrightarrow{w} \hat{u} \quad \text{in } L^1(T, \mathbb{R}^m). \tag{22}$$

Consider the multifunction $V_n : T \times \mathbb{R}^N \rightarrow 2^{\mathbb{R}^m} \setminus \{\emptyset\}$ defined by

$$V_n(t, x) = \left\{ \xi \in \mathbb{R}^m : |\hat{u}_n(t) - \xi| \leq k(t)|\hat{x}(t) - x| + \frac{1}{n}, \xi \in U(t, x) \right\} \tag{23}$$

(see hypothesis $H(U)$ (ii)).

We consider the following control system:

$$\begin{cases} -x'(t) \in A(x(t)) + f(t, x(t))v(t) \text{ for almost all } t \in T = [0, b], \\ x(0) = x_0, v \in S_{V_n(\cdot, x(\cdot))}^1. \end{cases} \tag{24n}$$

Reasoning as in the proof of Proposition 4, we can show that for every $n \in \mathbb{N}$ problem (24n) has admissible state-control pairs. So, let (\hat{x}_n, \hat{v}_n) be such a pair for (24n). Then

$$(\hat{x}_n, \hat{v}_n) \in P \quad \text{for all } n \in \mathbb{N}.$$

As in the proof of Proposition 4, using the theorem of Baras [2], we see that

$$\{\hat{x}_n\}_{n \geq 1} \subseteq C(T, \mathbb{R}^N) \text{ is relatively compact.}$$

So, we may assume that

$$\hat{x}_n \rightarrow \hat{x}^* \quad \text{in } C(T, \mathbb{R}^N). \tag{25}$$

We have

$$\begin{aligned}
 -\hat{x}'_n(t) &\in A(\hat{x}_n(t)) + f(t, \hat{x}_n(t))v_n(t) \quad \text{for almost all } t \in T, \hat{x}_n(0) = x_0, n \in \mathbb{N}, \\
 -\hat{x}'(t) &\in A(\hat{x}(t)) + f(t, \hat{x}(t))\hat{u}(t) \quad \text{for almost all } t \in T, \hat{x}(0) = x_0.
 \end{aligned}$$

Exploiting the monotonicity of $A(\cdot)$, we obtain

$$\begin{aligned}
 &(\hat{x}'_n(t) - \hat{x}'(t), \hat{x}_n(t) - \hat{x}(t))_{\mathbb{R}^N} + (f(t, \hat{x}_n(t))\hat{v}_n(t) - f(t, \hat{x}(t))\hat{u}(t), \hat{x}_n(t) \\
 &\quad - \hat{x}(t))_{\mathbb{R}^N} \leq 0 \quad \text{for almost all } t \in T, \\
 &\Rightarrow \frac{1}{2} \frac{d}{dt} |\hat{x}_n(t) - \hat{x}(t)|^2 + (f(t, \hat{x}_n(t))\hat{v}_n(t) \\
 &\quad - f(t, \hat{x}(t))\hat{u}(t), \hat{x}_n(t) - \hat{x}(t))_{\mathbb{R}^N} \leq 0 \\
 &\quad \text{for almost all } t \in T, \\
 &\Rightarrow \frac{1}{2} |\hat{x}_n(t) - \hat{x}(t)|^2 + \int_0^1 (f(s, \hat{x}_n(s))\hat{v}_n(s) - f(s, \hat{x}(s))\hat{u}(s), \hat{x}_n(s) \\
 &\quad - \hat{x}(s))_{\mathbb{R}^N} ds \leq 0 \quad \text{for all } n \in \mathbb{N}. \tag{26}
 \end{aligned}$$

We estimate the integral in the left-hand side of (26). Then

$$\begin{aligned}
 &\left| \int_0^t (f(s, \hat{x}_n(s))\hat{v}_n(s) - f(s, \hat{x}(s))\hat{u}(s), \hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^N} ds \right| \\
 &\leq \left| \int_0^t (f(s, \hat{x}_n(s))\hat{v}_n(s) - f(s, \hat{x}_n(s))\hat{u}_n(s), \hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^N} ds \right| \\
 &\quad + \left| \int_0^t (f(s, \hat{x}_n(s))\hat{u}_n(s) - f(s, \hat{x}_n(s))\hat{u}(s), \hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^N} ds \right| \\
 &\quad + \left| \int_0^t (f(s, \hat{x}_n(s))\hat{u}(s) - f(s, \hat{x}(s))\hat{u}(s), \hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^N} ds \right|. \tag{27}
 \end{aligned}$$

We examine each summand in the right-hand side of (27). We have

$$\begin{aligned}
 &\left| \int_0^t (f(s, \hat{x}_n(s))\hat{v}_n(s) - f(s, \hat{x}_n(s))\hat{u}_n(s), \hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^N} ds \right| \\
 &= \left| \int_0^t (\hat{v}_n(s) - \hat{u}_n(s), f(s, \hat{x}_n(s))^* (\hat{x}_n(s) - \hat{x}(s)))_{\mathbb{R}^N} ds \right| \\
 &\leq \left| \int_0^t |\hat{v}_n(s) - \hat{u}_n(s)| \|f(s, \hat{x}_n(s))^*\|_{\mathcal{L}} |\hat{x}_n(s) - \hat{x}(s)| ds \right| \\
 &= \left| \int_0^t |\hat{v}_n(s) - \hat{u}_n(s)| \|f(s, \hat{x}_n(s))\|_{\mathcal{L}} |\hat{x}_n(s) - \hat{x}(s)| ds \right| \\
 &\leq \int_0^t [k(s)|\hat{x}_n(s) - \hat{x}(s)| + \frac{1}{n}] a_1(s) |\hat{x}_n(s) - \hat{x}(s)| ds
 \end{aligned}$$

for some $a_1 \in L^1(T)$, all $n \in \mathbb{N}$ (see (23), (25) and hypothesis $H(f)$ (iii))

$$\leq \int_0^t a_2(s) |\hat{x}_n(s) - \hat{x}(s)|^2 ds + \frac{c_2}{n} \|\hat{x}_n - \hat{x}\|_{C(T, \mathbb{R}^N)}$$

for some $a_2 \in L^1(T)$, $c_2 > 0$ and all $n \in \mathbb{N}$. (28)

For the second summand we have

$$\begin{aligned} & \left| \int_0^t (f(s, \hat{x}_n(s))\hat{u}_n(s) - f(s, \hat{x}_n(s))\hat{u}(s), \hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^N} ds \right| \\ &= \left| \int_0^t (\hat{u}_n(s) - \hat{u}(s), f(s, \hat{x}_n(s))^*(\hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^m} ds \right|. \end{aligned} \tag{29}$$

We know that bounded sets in $L^\infty(T, \mathbb{R}^m)$ furnished with the w^* -topology, are metrizable (see [29, p. 230]). Hence hypothesis $H(U)$ (iii) and (25) imply that

$$\hat{u}_n \xrightarrow{w^*} \hat{u} \quad \text{in } L^\infty(T, \mathbb{R}^m). \tag{30}$$

Also note that

$$f(s, \hat{x}_n(s))^*(\hat{x}_n(s) - \hat{x}(s)) \rightarrow f(s, \hat{x}^*(s))^*(\hat{x}^*(s) - \hat{x}(s)) \quad \text{for almost all } s \in T.$$

So, by the Lebesgue dominated convergence theorem, we have

$$f(\cdot, \hat{x}_n(\cdot))^*(\hat{x}_n - \hat{x})(\cdot) \rightarrow f(\cdot, \hat{x}^*(\cdot))^*(\hat{x}^* - \hat{x})(\cdot) \quad \text{in } L^1(T, \mathbb{R}^m). \tag{31}$$

From (29)–(31) it follows that

$$\left| \int_0^t (f(s, \hat{x}_n(s))\hat{u}_n(s) - f(s, \hat{x}_n(s))\hat{u}(s), \hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^N} ds \right| \rightarrow 0. \tag{32}$$

Finally, for the third summand, we have

$$\begin{aligned} & \left| \int_0^t (f(s, \hat{x}_n(s))\hat{u}(s) - f(s, \hat{x}_n(s))\hat{u}(s), \hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^N} ds \right| \\ & \leq c_3 \int_0^t l_r(s) |\hat{x}_n(s) - \hat{x}(s)|^2 ds \quad \text{with } r = \sup_{n \geq 1} \|\hat{x}_n\|_{C(T, \mathbb{R}^N)} < \infty, \\ & \text{for some } c_3 > 0 \text{ (see hypotheses } H(f)\text{(ii), } H(U)\text{(iii))}. \end{aligned} \tag{33}$$

We return to (27), pass to the limit as $n \rightarrow \infty$ and use (28), (32), (33) and (25).

Then

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \left| \int_0^t (f(s, \hat{x}_n(s))\hat{u}_n(s) - f(s, \hat{x}(s))\hat{u}(s), \hat{x}_n(s) - \hat{x}(s))_{\mathbb{R}^N} ds \right| \\ & \leq \int_0^t a_3(s) |\hat{x}^*(s) - \hat{x}(s)|^2 ds \quad \text{for some } a_3 \in L^1(T) \text{ (see (25)).} \end{aligned} \tag{34}$$

So, if in (26) we pass to the limit as $n \rightarrow \infty$ and use (25) and (34), then

$$\begin{aligned} |\hat{x}^*(t) - \hat{x}(t)|^2 & \leq \int_0^t 2a_3(s) |\hat{x}^*(s) - \hat{x}(s)|^2 ds \quad \text{for all } t \in T, \\ & \Rightarrow \hat{x}^* = \hat{x} \quad \text{(by Gronwall's inequality).} \end{aligned}$$

Therefore for the original sequence we have

$$\hat{x}_n \rightarrow \hat{x} \quad \text{in } C(T, \mathbb{R}^N). \tag{35}$$

From (23) and (35), we see that

$$\begin{aligned} \|\hat{u}_n - \hat{v}_n\|_1 &\rightarrow 0, \\ \Rightarrow \hat{v}_n &\xrightarrow{\|\cdot\|_w} \hat{u} \quad \text{in } L^1(T, \mathbb{R}^N) \quad (\text{see (21) and (22)}). \end{aligned}$$

Finally, recall (see (21)) that

$$L(\cdot, \hat{x}(\cdot), \hat{u}_n(\cdot)) \xrightarrow{\|\cdot\|_w} \hat{L}^{**}(\cdot, \hat{x}(\cdot), \hat{u}(\cdot)) \quad \text{in } L^1(T). \tag{36}$$

On account of hypotheses $H(L)$ (ii) and (23), (35), we have

$$L(\cdot, \hat{x}_n(\cdot), \hat{v}_n(\cdot)) - L(\cdot, \hat{x}(\cdot), \hat{u}(\cdot)) \xrightarrow{\|\cdot\|_w} 0. \tag{37}$$

From (36) and (37) we obtain

$$L(\cdot, \hat{x}_n(\cdot), \hat{v}_n(\cdot)) \xrightarrow{\|\cdot\|_w} \hat{L}^{**}(\cdot, \hat{x}(\cdot), \hat{v}(\cdot)) \quad \text{in } L^1(T, \mathbb{R}),$$

which concludes the proof. □

Now we are ready to show that our first relaxation method which produces problem (18), is admissible.

Theorem 11. *If hypotheses $H(A), H(f), H(U), H(L), H_0$ hold, then the following properties hold:*

(a) *there exists $(\hat{x}, \hat{u}) \in P_c$ such that*

$$J_r(\hat{x}, \hat{u}) = m_r;$$

(b) $m_r = m$;

(c) *there exists a sequence $\{(\hat{x}_n, \hat{u}_n)\}_{n \geq 1} \subseteq P$ such that*

$$J(\hat{x}_n, \hat{u}_n) \downarrow m_r \quad \text{and} \quad \hat{x}_n \rightarrow \hat{x} \quad \text{in } C(T, \mathbb{R}^N) \quad \text{and} \quad \hat{u}_n \xrightarrow{\|\cdot\|_w} \hat{u} \quad \text{in } L^1(T, \mathbb{R}^m).$$

Proof. (a) Let $\{(\hat{y}_n, \hat{v}_n)\}_{n \geq 1} \subseteq P_c$ be a minimizing sequence for the relaxed problem (18). So, we have

$$J_r(\hat{y}_n, \hat{v}_n) \downarrow m_r.$$

We know that

$$\{(\hat{y}_n, \hat{v}_n)\}_{n \geq 1} \subseteq C(T, \mathbb{R}^N) \times L^1(T, \mathbb{R}^m)_w \quad \text{is relatively compact}$$

(see Proposition 5). So, we may assume that

$$\hat{y}_n \rightarrow \hat{x} \quad \text{in } C(T, \mathbb{R}^N) \quad \text{and} \quad \hat{v}_n \xrightarrow{w} \hat{u} \quad \text{in } L^1(T, \mathbb{R}^m). \tag{38}$$

Then from (38) and Proposition 3.151 of Gasinski and Papageorgiou [18, p. 441], we have

$$J_r(\hat{x}, \hat{u}) \leq \liminf_{n \rightarrow \infty} J_r(\hat{y}_n, \hat{v}_n) = m_r. \tag{39}$$

However, by (38) and Proposition 5, we have

$$(\hat{x}, \hat{u}) \in P_c \Rightarrow J_r(\hat{x}, \hat{u}) = m_r \quad (\text{see (39)}).$$

(b) and (c) From Proposition 10, we know that there exists a sequence $\{(\hat{x}_n, \hat{u}_n)\}_{n \geq 1} \subseteq P$ such that

$$\begin{aligned} \hat{x}_n &\rightarrow \hat{x} \quad \text{in } C(T, \mathbb{R}^N), \quad \hat{u}_n \xrightarrow{\|\cdot\|_w} \hat{u} \quad \text{in } L^1(T, \mathbb{R}^m), \\ L(\cdot, \hat{x}(\cdot), \hat{u}_n(\cdot)) &\xrightarrow{\|\cdot\|_w} \hat{L}^{**}(\cdot, \hat{x}(\cdot), \hat{u}(\cdot)) \quad \text{in } L^1(T). \end{aligned} \tag{40}$$

We have

$$\begin{aligned} &|J_r(\hat{x}_n, \hat{u}_n) - J_r(\hat{x}, \hat{u})| \\ &\leq \|L(\cdot, \hat{x}_n(\cdot), \hat{u}_n(\cdot)) - \hat{L}^{**}(\cdot, \hat{x}(\cdot), \hat{u}(\cdot))\|_w \rightarrow 0 \quad (\text{see (40)}), \\ &\Rightarrow J(\hat{x}, \hat{u}_n) \rightarrow J_r(\hat{x}, \hat{u}) = m_r. \end{aligned}$$

We know that

$$\begin{aligned} m &\leq J(\hat{x}_n, \hat{u}_n) \quad \text{for all } n \in \mathbb{N}, \\ &\Rightarrow m \leq m_r, \\ &\Rightarrow m = m_r \quad (\text{since we always have } m_r \leq m). \end{aligned}$$

The proof is now complete. □

4. Second Relaxation Method

In this section, we present an alternative relaxation method based on Young measures.

So, now the control constraint set for the relaxed system is

$$\Sigma(t, x) = \{\mu \in R(T, \bar{B}_M) : \mu(U(t, x)) = 1\},$$

where $M = \|a_0\|_\infty$ (see hypothesis $H(U)$ (iii)) and $\bar{B}_M = \{v \in \mathbb{R}^m : |v| \leq M\}$. Then, the new relaxed optimal control problem, is the following:

$$\left\{ \begin{aligned} &\hat{J}_r(x, \lambda) = \int_0^b \int_{\bar{B}} L(t, x(t), u) \lambda(t)(du) dt \rightarrow \inf = \hat{m}_r, \\ &\text{subject to: } -x'(t) \in A(x(t)) + \int_{\bar{B}_M} f(t, x(t)) u \lambda(t)(du) \text{ for a.a. } t \in T = [0, b], \\ &x(0) = x_0, \quad \lambda \in S_{\Sigma(\cdot, x(\cdot))}, \end{aligned} \right. \tag{41}$$

where

$$S_{\Sigma(\cdot, x(\cdot))} = \{ \lambda : T \rightarrow P(\bar{B}_M) : \lambda(\cdot) \text{ is measurable, } \\ \lambda(t) \in \Sigma(t, x(t)) \text{ for almost all } t \in T \}.$$

Recall that on $P(\bar{B}_M)$ we consider the narrow topology τ_n . The orientor field for this new convexified control system is

$$\hat{F}(t, x) = \left\{ \int_{\bar{B}_M} f(t, x) u \lambda(t)(du) : \lambda \in S_{\Sigma(\cdot, x)} \right\}.$$

So, if we denote by \hat{P}_c the set of admissible state-control pairs for (41), then \hat{P}_c is compact in $C(T, \mathbb{R}^N) \times \mathcal{Y}(T \times \bar{B}_M)_{\tau_n^y}$.

We show that this new relaxed optimal control problem is equivalent to (18) and hence also admissible.

Theorem 12. *If hypotheses $H(A), H(f), H(U), G(L), H_0$ hold, then there exists $(\hat{x}, \hat{\lambda}) \in \hat{P}_c$ such that*

$$\hat{J}_r(\hat{x}, \hat{\lambda}) = \hat{n}_r = m = m_r$$

and we can find a sequence $\{(\hat{x}_n, \hat{u}_n)\}_{n \geq 1} \subseteq P$ such that

$$\hat{x}_n \rightarrow \hat{x} \text{ in } C(T, \mathbb{R}^N) \quad \text{and} \quad \delta_{\hat{u}_n} \xrightarrow{\tau_n^y} \hat{\lambda} \text{ in } R(T, \bar{B}_M).$$

Proof. Let $\{(\hat{x}_n, \hat{u}_n)\}_{n \geq 1} \subseteq \hat{P}_c$ be a minimizing sequence for problem (41). Evidently, $\{\hat{\lambda}_n\}_{n \geq 1} \subseteq R(T, \bar{B}_M)$ is tight, so by Prokhorov's theorem (see Papageorgiou and Winkert [29, Theorem 4.7.22, p. 393]), we have at least for a subsequence that

$$\hat{\lambda} \xrightarrow{\tau_n^y} \hat{\lambda} \quad \text{in } R(T, \bar{B}_M). \tag{42}$$

Let $\hat{x}_n \in W^{1,1}((0, b), \mathbb{R}^N) = AC^1(T, \mathbb{R}^N)$ be the relaxed state generated by the relaxed control $\hat{\lambda}_n \in S_{\Sigma(\cdot, x_n(\cdot))}$. We know that

$$\{\hat{x}_n\}_{n \geq 1} \subseteq C(T, \mathbb{R}^N) \text{ is relatively compact (see the proof of Proposition 4).}$$

So, we may assume that

$$\hat{x}_n \rightarrow \hat{x} \quad \text{in } C(T, \mathbb{R}^N). \tag{43}$$

We have

$$(\hat{x}, \hat{u}) \in \hat{P}_c \quad \text{(see (42) and (43)).}$$

Note that

$$\hat{J}_r(\hat{x}_n, \hat{\lambda}_n) = \int_0^b \int_{\bar{B}_M} L(t, \hat{x}_n(t), u) \hat{\lambda}_n(t)(du) dt$$

$$= \int_0^b \int_{\mathbb{R}^N \times \bar{B}_M} L(t, x, u)(\delta_{\hat{x}_n(t)} \otimes \hat{\lambda}_n(t))(dx, du)dt.$$

By the Fiber Product Lemma (see Valadier [32, Lemma 11]), we have

$$\delta_{\hat{x}_n(\cdot)} \otimes \hat{\lambda}_n \xrightarrow{\tau_n^y} \delta_{\hat{x}(\cdot)} \otimes \hat{\lambda} \quad (\text{see (42) and (43)}).$$

Then on account of hypotheses $H(L)$ and the definition of the narrow topology τ_n^y , we have

$$\begin{aligned} \hat{m}_r &= \lim_{n \rightarrow \infty} \hat{J}_r(\hat{x}_n, \hat{\lambda}_n) \\ &= \lim_{n \rightarrow \infty} \int_0^b \int_{\bar{B}_M} L(t, x_n(t), u) \hat{\lambda}_n(t)(du)dt \\ &= \int_0^b \int_{\bar{B}_M} L(t, \hat{x}(t), u) \hat{\lambda}(t)(du)dt \\ &= \hat{J}_r(\hat{x}, \hat{\lambda}) \geq \hat{m}_r \quad (\text{since } (\hat{x}, \hat{\lambda}) \in \hat{P}_c), \\ &\Rightarrow \hat{J}_r(\hat{x}, \hat{\lambda}) = \hat{m}_r. \end{aligned}$$

Due to the density of the Dirac Young measures in $\mathcal{Y}(T \times \bar{B}_M)$ with the Young narrow topology τ_n^y (see Valadier [32, Proposition 8]), we can find $\{\hat{u}_n\}_{n \geq 1} \subseteq S^1_{U(\cdot, \hat{x}(\cdot))}$ such that

$$\delta_{\hat{u}_n(\cdot)} \xrightarrow{\tau_n^y} \hat{\lambda}. \tag{44}$$

As in the proof of Proposition 10, we consider the multifunction $(t, x) \rightarrow V_n(t, x)$ defined by

$$V_n(t, x) = \left\{ \xi \in \mathbb{R}^m : |\hat{u}_n(t) - \xi| \leq k(t)|\hat{x}(t) - x| + \frac{1}{n}, \xi \in U(t, x) \right\}.$$

We use this as control constraint multifunction and consider the control system (24n). We can find admissible state-control pairs $(\hat{x}_n, \hat{v}_n), n \in \mathbb{N}$ for this system such that

$$\hat{x}_n \rightarrow \hat{x} \text{ in } C(T, \mathbb{R}^N), \quad \hat{u}_n - \hat{v}_n \xrightarrow{\|\cdot\|_1 \text{ a.e.}} 0. \tag{45}$$

Using (44) and (45), we have

$$\begin{aligned} J(\hat{x}_n, \hat{v}_n) &\rightarrow \hat{J}_r(\hat{x}, \hat{\lambda}) = \hat{m}_r, \\ &\Rightarrow m \leq \hat{m}_r, \\ &\Rightarrow m = \hat{m}_r = m_r. \end{aligned}$$

The proof is now complete. □

Acknowledgments

This research was supported by the Slovenian Research Agency Grants P1-0292, J1-8131, N1-0114, N1-0064 and N1-0083.

References

- [1] E. Avgerinos and N. S. Papageorgiou, On the sensitivity and relaxability of optimal control problems governed by nonlinear evolution equations with state constraints, *Monatsh. Math.* **109**(1) (1990) 1–23.
- [2] P. Baras, Compacité de l'opérateur $f \mapsto u$ solution d'une équation non linéaire $(du/dt) + Au \ni f$, *C. R. Acad. Sci. Paris Sér. A-B* **286**(23) (1978) A1113–A1116.
- [3] L. D. Berkovitz, Existence theorems in problems of optimal control, *Studia Math.* **44** (1972) 275–285.
- [4] L. D. Berkovitz, *Optimal Control Theory* (Springer-Verlag, New York, 1974).
- [5] J. Bourgain, An averaging result for l^1 -sequences and application to weakly conditionally compact sets in L^1_X , *Israel J. Math.* **32** (1979) 289–298.
- [6] A. Bressan and G. Colombo, Extensions and selections of maps with decomposable values, *Studia Math.* **90** (1988) 69–86.
- [7] H. Brezis, *Operateurs Maximaux Monotones et Semi-Groupes de Contractions Dans Les Espaces de Hilbert* (North-Holland Publishing Co., Amsterdam–London, 1973).
- [8] G. Buttazzo, Some relaxation problems in optimal control theory, *J. Math. Anal. Appl.* **125** (1987) 272–287.
- [9] G. Buttazzo and G. Dal Maso, Γ -convergence and optimal control problems, *J. Optim. Theory Appl.* **38** (1982) 385–407.
- [10] L. Cesari, Existence theorems for weak and usual optimal solutions in Lagrange problems with unilateral constraints I, *Trans. Amer. Math. Soc.* **124** (1966) 369–412.
- [11] L. Cesari, Existence theorems for weak and usual optimal solutions in Lagrange problems with unilateral constraints II, *Trans. Amer. Math. Soc.* **124** (1966) 413–429.
- [12] L. Cesari, *Optimization Theory and Applications* (Springer-Verlag, New York, 1983).
- [13] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems* (North-Holland, Amsterdam, 1976).
- [14] B. Emamizadeh, A. Farjudian and H. Mikayelyan, An elliptic optimal control problem and its two relaxations, *J. Optim. Theory Appl.* **172**(2) (2017) 455–465.
- [15] A. Fryszkowski, Continuous selections for a class non-convex multivalued maps, *Studia Math.* **75** (1983) 163–174.
- [16] A. Fryszkowski, *Fixed Point Theory for Decomposable Sets* (Kluwer Academic Publishers, New York, 2004).
- [17] L. Gasinski and N. S. Papageorgiou, *Nonlinear Analysis* (Chapman & Hall/CRC, Boca Raton, FL, 2006).
- [18] L. Gasinski and N. S. Papageorgiou, *Exercises in Analysis. Part 2: Nonlinear Analysis* (Springer Nature, Cham, 2016).
- [19] S. Gutman, Topological equivalence in the space of integrable vector-valued functions, *Proc. Amer. Math. Soc.* **93** (1985) 40–42.
- [20] S. Hu and N. S. Papageorgiou, *Handbook of Multivalued Analysis. Volume I: Theory* (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997).
- [21] S. Hu and N. S. Papageorgiou, *Handbook of Multivalued Analysis. Volume II: Applications* (Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000).

- [22] X. Liu, Z. Liu and X. Fu, Relaxation in nonconvex optimal control problems described by fractional differential equations, *J. Math. Anal. Appl.* **409** (2014) 446–458.
- [23] N. S. Papageorgiou, V. D. Rădulescu and D. D. Repovš, Sensitivity analysis for optimal control problems governed by nonlinear evolution inclusions, *Adv. Nonlinear Anal.* **6**(3) (2017) 199–235.
- [24] N. S. Papageorgiou, V. D. Rădulescu and D. D. Repovš, Periodic solutions for time-dependent subdifferential evolution inclusions, *Evol. Equ. Control Theory* **6**(3) (2017) 277–297.
- [25] N. S. Papageorgiou, V. D. Rădulescu and D. D. Repovš, Nonlinear elliptic inclusions with unilateral constraint and dependence on the gradient, *Appl. Math. Optim.* **78**(1) (2018) 1–23.
- [26] N. S. Papageorgiou, V. D. Rădulescu and D. D. Repovš, Nonlinear second order evolution inclusions with noncoercive viscosity term, *J. Differential Equations* **264**(7) (2018) 4749–4763.
- [27] N. S. Papageorgiou, V. D. Rădulescu, D. D. Repovš, Periodic solutions for implicit evolution inclusions, *Evol. Equ. Control Theory* **8**(3) (2019) 621–631.
- [28] N. S. Papageorgiou, V. D. Rădulescu and D. D. Repovš, *Modern Nonlinear Analysis — Theory and Methods* (Springer Monographs in Mathematics, Springer Nature, Cham, 2019).
- [29] N. S. Papageorgiou and P. Winkert, *An Introduction to Applied Nonlinear Functional Analysis* (De Gruyter, Berlin, 2018).
- [30] T. Roubíček, *Relaxation in Optimization Theory and Variational Calculus* (De Gruyter, Berlin, 1997).
- [31] A. Tolstonogov, Relaxation in non-convex control problems described by first-order evolution equations, *Sb. Math.* **109** (1999) 1689–1714.
- [32] M. Valadier, A course on Young measures, *Workshop on Measure Theory and Real Analysis (Grado, 1993)*, *Rend. Istit. Mat. Univ. Trieste* **26** (1994) suppl., 349–394 (1995).
- [33] J. Warga, *Optimal Control of Differential and Functional Equations* (Academic Press, New York, 1972).