CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 301

Optimality Conditions for Semivectorial Bilevel Convex Optimal Control Problems

Henry Bonnel and Jacqueline Morgan

December 2011

University of Naples Federico II   University of Salerno   Bocconi University, Milan

WORKING PAPER NO. 301

Optimality Conditions for Semivectorial Bilevel Convex Optimal Control Problems

Henry Bonnel[*] and Jacqueline Morgan[**]

Abstract

We present optimality conditions for bilevel optimal control problems where the upper level, to be solved by a leader, is a scalar optimal control problem and the lower level, to be solved by several followers, is a multiobjective convex optimal control problem. Multiobjective optimal control problems arise in many application areas where several conflicting objectives need to be considered. Minimize several objective functionals leads to solutions such that none of the objective functional values can be improved further without deteriorating another. The set of all such solutions is referred to as efficient (also called Pareto optimal, noninferior, or nondominated) set of solutions. The lower level of the semivectorial bilevel optimal control problems can be considered to be associated to a "grande coalition" of a p-player cooperative differential game, every player having its own objective and control function. We consider situations in which these p players react as "followers" to every decision imposed by a "leader" (who acts at the so-called upper level). The best reply correspondence of the followers being in general non uniquely determined, the leader cannot predict the followers choice simply on the basis of his rational behavior. So, the choice of the best strategy from the leader point of view depends of how the followers choose a strategy among his best responses. In this paper, we will consider two (extreme) possibilities: (i) the optimistic situation, when for every decison of the leader, the followers will choose a strategy amongst the efficient controls which minimizes the (scalar) objective of the leader; in this case the leader will choose a strategy which minimizes the best he can obtain amongst all the best responses of the followers: (ii) the pessimistic situation, when the followers can choose amongst the efficient controls one which maximizes the (scalar) objective of the leader; in this case the leader will choose a strategy which minimizes the worst he could obtain amongst all the best responses of the followers. This paper continues the research initiated in [17] where existence results for these problems have been obtained.

[*] Université de la Nouvelle-Calédonie. E-mail: bonnel@univ-nc.nc

[**] Università di Napoli Federico II and CSEF. E-mail: morgan@unina.it

Table of contents

1 Introduction

The aim of this paper is to obtain optimality conditions for the semivectorial bilevel optimal control problems introduced in [17] where existence results have been established.
Semivectorial bilevel optimal control problems are bilevel problems where the upper level corresponds to a scalar optimization problem and the lower level to a multiobjective optimal control problem. Multiobjective optimal control problems arise in many application areas where several conflicting objectives need to be considered. Minimize several objective functionals leads to solutions such that none of the objective functional values can be improved further without deteriorating another. The set of all such solutions is referred to as efficient (also called Pareto optimal, noninferior, or nondominated) set of solutions (see e.g.[38]). The lower level of the semivectorial bilevel optimal control problems can be considered to be associated to a "grande coalition" of a p-player cooperative differential game, every player having its own objective and control function. We consider situations in which these p players react as "followers" to every decision imposed by a "leader" (who acts at the so-called upper level). The best reply correspondence of the followers being in general non uniquely determined, the leader cannot predict the followers choice simply on the basis of his rational behavior. So, the choice of the best strategy from the leader point of view depends of how the followers choose a strategy among his best responses. In this paper, we will consider two (extreme) possibilities:

(i) the optimistic situation, when for every decison of the leader, the followers will choose a strategy amongst the efficient controls which minimizes the (scalar) objective of the leader; in this case the leader will choose a strategy which minimizes the best he can obtain amongst all the best responses of the followers;

(ii) the pessimistic situation, when the followers can choose amongst the efficient controls one which maximizes the (scalar) objective of the leader; in this case the leader will choose a strategy which minimizes the worst he could obtain amongst all the best responses of the followers.

The semivectorial bilevel control problems which model these two situations, and which will be described in the next section, include the following problems which have been intensively studied in the last decades so we will give essentially a few earlier references,

- optimizing a scalar valued function over the efficient set associated to a multiobjective optimization (mathematical programming) problem; (introduced in [45] and investigated in [8, 9, 10, 11, 12, 13, 25, 26, 27, 33, 36, 37] and [50] for a survey)

- optimizing a scalar valued function over an efficient control set associated to a multiobjective optimal control problem (introduced and investigated in [15], followed by [18]);

- semivectorial bilevel static problems (introduced and investigated in [16], followed by [14, 3, 22, 31, 51, 30], for the optimistic case);

- Stackelberg problems (introduced in [49] and investigated e.g. in [5, 43, 40]);

- Bilevel optimization problems (e.g. [46, 42, 41, 28, 24] and [29] for an extensive bibliography);

- Stackelberg dynamic problems (introduced in [23, 48] and investigated e.g. in [6, 46, 44, 47] and [5], book with an extensive bibliography).

In this paper, we rewrite the optimistic and pessimistic semivectorial bilevel control problems as bilevel problems where the lower level is a scalar optimization problem which admits a unique solution, using scalarization techniques as in [17]. So we are able to give optimality conditions for the lower level problem in the general case (supposing that the leader's controls are bounded) using Pontryagin maximum principle. This theoretically allows to obtain under suitable conditions the dependence of the optimal control on the leader's variables. However, this approach is very difficult to apply because one needs to solve a bilocal problem. That is why we consider the particular but important case when the followers problem is linear-quadratic. In this case we show that using a resolvent matrix obtained from data, we can explicitly solve the bilocal problem and express the optimal control and the state as functions of leader's variables, and   we show that these dependences are continuously differentiable. Finally we present optimality conditions for the upper levels of the optimistic and pessimistic problems .

2   Preliminaries and problem statement

All the assumptions and notations considered in this section and introduced in [17] will be kept throughout this paper.

For the *leader* we denote by J_l the scalar objective, by u_l the control function, and by $\mathcal{U}_l$ the set of admissible controls. For the followers we denote by $\mathbf{J_f} = (J_1, \ldots, J_p)$ the vector objective (p-scalar objectives), by $\mathbf{u_f} = (u_1, \ldots, u_p)$ the control function whose values belong to the set $\mathbf{U_f} = U_1 \times \cdots \times U_p \subseteq \mathbb{R}^{m_f} = \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_p}$. $\mathbf{U_f}$ is assumed to be nonempty, closed, convex, and $0 \in \mathbf{U_f}$. Real numbers t_0, T are fixed ($t_0 < T$) and represent respectively the initial time and an upper bound of the final time. The set of final time values $\mathcal{T} = [\underline{t}, \overline{t}] \subset]t_0, T[$, where $\underline{t} \leq \overline{t}$. The final time, denoted $t_1 \in \mathcal{T}$, may be variable and it is decided by the leader, hence t_1 is fixed in the followers problem. We assume that

$$\mathcal{U}_l \subset L_2^{m_l}([t_0, T]) \quad \text{is closed, nonempty and convex.} \tag{1}$$

 For each fixed $(t_1, u_l) \in \mathcal{T} \times \mathcal{U}_l$, the followers have to solve the following parametric multi-objective control problem, called lower level problem:

$$(\mathbf{LL})_{(t_1, u_l)} \qquad \begin{cases} \underset{(\mathbf{u_f}, x)}{\mathbf{MIN}} \; \mathbf{J_f}(t_1, u_l, \mathbf{u_f}, x) \\ \text{subject to } (\mathbf{u_f}, x) \text{ verifies (2)-(5)} \end{cases}$$

$$\begin{align}
\mathbf{u_f}(t) &\in \mathbf{U_f} \text{ a.e. on } [t_0, T], \quad \mathbf{u_f}(t) = 0 \text{ a.e. on } [t_1, T] \tag{2} \\
\dot{x}(t) &= A(t)\,x(t) + B_l(t)\,u_l(t) + \mathbf{B_f}(t)\mathbf{u_f}(t) \quad \text{a.e. on } [t_0, t_1] \tag{3} \\
x(t_0) &= x_0 \tag{4} \\
x(t_1) &\in \mathcal{F}, \tag{5}
\end{align}$$

where $A : [t_0, T] \to \mathbb{R}^{n \times n}$, $B_l : [t_0, T] \to \mathbb{R}^{n \times m_l}$, $\mathbf{B_f} : [t_0, T] \to \mathbb{R}^{n \times m_f}$ are continuous matrix-valued functions, and the control function $\mathbf{u_f} = (u_1, \ldots, u_p) \in L_2^{m_f}([t_0, T]) = L_2^{m_1}([t_0, T]) \times \cdots \times L_2^{m_p}([t_0, T])$.

$L_2^m([t_0, T])$ stands for the usual Hilbert space of equivalence classes (two functions are equivalent iff they coincide a.e.) of (Lebesgue) measurable functions u from $[t_0, T]$ to $\mathbb{R}^m$, such that the function $t \mapsto u^T(t)u(t)$ is (Lebesgue) integrable over $[t_0, T]$ endowed with the norm $\|u\|_2 := \left(\int_{t_0}^{T} u^T(t)u(t) \, dt \right)^{1/2}$. The target set $\mathcal{F} \subset \mathbb{R}^n$ is assumed to be closed, convex and nonempty.

The initial state $x_0 \in \mathbb{R}^n$ is specified.

For each $u = (t_1, u_l, \mathbf{u_f}) \in \mathcal{T} \times L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T])$, there exists a unique solution (in the sense of Carathéodory) x_u of the Cauchy problem (3)-(4), and $x_u \in H_1^n([t_0, t_1])$. $H_1^n([t_0, t_1])$ stands for the Hilbert space of absolutely continuous functions from $[t_0, t_1]$ to $\mathbb{R}^n$ with derivative in $L_2^n([t_0, t_1])$ endowed with the norm $x \mapsto \|x\| := (\|\dot{x}\|_2^2 + \|x\|_2^2)^{1/2}$.

The *feasible set* $\mathcal{S}(t_1, u_l)$ for the problem $(\mathbf{LL})_{(t_1, u_l)}$ is defined in the following way

$$S(t_1, u_l) = \{(\mathbf{u_f}, x) \in L_2^{m_f}([t_0, T]) \times H_1^n([t_0, t_1]) | \; (\mathbf{u_f}, x) \;\; \text{verifies relations (2)-(5)}\}. \qquad (6)$$

Thus, problem $(\mathbf{LL})_{(t_1, u_l)}$ can be written as

$$(\mathbf{LL})_{(t_1, u_l)} \qquad\qquad \underset{(\mathbf{u_f}, x) \in S(t_1, u_l)}{\mathbf{MIN}} \quad \mathbf{J_f}(t_1, u_l, \mathbf{u_f}, x).$$

Next we give the following standard definitions.

DEFINITION 1. For problem $(\mathbf{LL})_{(t_1, u_l)}$ the element $(\bar{\mathbf{u}}_\mathbf{f}, \bar{x}) \in \mathcal{S}(t_1, u_l)$ is said to be

- *an efficient (or Pareto) control process* if there is no element $(\mathbf{u_f}, x) \in \mathcal{S}(t_1, u_l)$ satisfying

$$\forall i \in \{1, \ldots, p\} \qquad J_i(t_1, u_l, \mathbf{u_f}, x) \leq J_i(t_1, u_l, \bar{\mathbf{u}}_\mathbf{f}, \bar{x})$$

  and

$$\exists i_0 \in \{1, \ldots, p\} \qquad J_{i_0}(t_1, u_l, \mathbf{u_f}, x) < J_{i_0}(t_1, u_l, \bar{\mathbf{u}}_\mathbf{f}, \bar{x}).$$

- *a weakly efficient (or weakly Pareto) control process* if there is no element $(\mathbf{u_f}, x) \in \mathcal{S}(t_1, u_l)$ satisfying

$$\forall i \in \{1, \ldots, p\} \qquad J_i(t_1, u_l, \mathbf{u_f}, x) < J_i(t_1, u_l, \bar{\mathbf{u}}_\mathbf{f}, \bar{x}).$$

- *a properly efficient* (or *properly Pareto*) *control process* (see [34], or [20], [38] for generalizations) if it is an efficient control process and there exists a real number $M > 0$ so that for every $i \in \{1, \ldots, p\}$ and every $(\mathbf{u_f}, x) \in \mathcal{S}(t_1, u_l)$ with $J_i(t_1, u_l, \mathbf{u_f}, x) < J_i(t_1, u_l, \bar{\mathbf{u}}_\mathbf{f}, \bar{x})$ at least one $k \in \{1, \ldots, p\}$ exists with $J_k(t_1, u_l, \mathbf{u_f}, x) > J_k(t_1, u_l, \bar{\mathbf{u}}_\mathbf{f}, \bar{x})$ and

$$\frac{J_i(t_1, u_l, \bar{\mathbf{u}}_\mathbf{f}, \bar{x}) - J_i(t_1, u_l, \mathbf{u_f}, x)}{J_k(t_1, u_l, \mathbf{u_f}, x) - J_k(t_1, u_l, \bar{\mathbf{u}}_\mathbf{f}, \bar{x})} \leq M.$$

In the sequel the symbol $\sigma \in \{e, we, pe\}$ stands for "efficient" when $\sigma = e$, "weakly efficient" when $\sigma = we$ and "properly efficient" when $\sigma = pe$.

The set of all σ-control processes associated to problem $(\mathbf{LL})_{(t_1, u_l)}$ will be denoted by $\mathcal{P}_\sigma(t_1, u_l)$.

3

Finally we consider the following *semivectorial bilevel optimal control* problems:

$$(\text{OSVBC})_\sigma \qquad \min_{(t_1,u_l)\in\mathcal{T}\times\mathcal{U}_l} \min_{(\mathbf{u_f},x)\in\mathcal{P}_\sigma(t_1,u_l)} J_l(t_1,u_l,\mathbf{u_f},x)$$

called *"optimistic semivectorial bilevel control problem"*, and

$$(\text{PSVBC})_\sigma \qquad \min_{(t_1,u_l)\in\mathcal{T}\times\mathcal{U}_l} \sup_{(\mathbf{u_f},x)\in\mathcal{P}_\sigma(t_1,u_l)} J_l(t_1,u_l,\mathbf{u_f},x)$$

called *"pessimistic semivectorial bilevel control problem"*.

REMARK 1. Note that the terminal time t_1 is fixed for the lower level problem, but it is a decision variable for the leader. Of course, a particular case can be obtained when the terminal time t_1 is fixed for the leader too, i.e. when $\mathcal{T} = \{t_1\}$.

REMARK 2. $(\mathbf{LL})_{(t_1,u_l)}$ may be also considered as the problem to solve by the *"grande coalition"* of a p-player cooperative differential game, (see [35] and its extensive references list) where the functional J_i and the control u_i represent the payoff and the control of the player number i, $i \in \{1, \ldots, p\}$. Then, our optimistic semivectorial bilevel problem corresponds to a strong Stackelberg problem in which, for any choice of (t_1, u_l), the leader can force the followers to choose amongst the σ-control processes one which minimizes the leader payoff. On the other hand, the pessimistic semivectorial bilevel problem corresponds to a weak Stackelberg problem in which, for any choice of the leader variables (t_1, u_l), the followers could choose amongst the σ-control processes one which is the worst for the leader.

We assume that for all $t_1 \in [t_0, T]$ and all $(u_l, \mathbf{u_f}, x) \in L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T]) \times H_1^n([t_0, t_1])$, we have

$$J_l(t_1, u_l, \mathbf{u_f}, x) = \int_{t_0}^{t_1} f_l(t, u_l(t), \mathbf{u_f}(t), x(t))dt,$$

and also, for all $i \in \{1, \ldots, p\}$

$$J_i(t_1, u_l, \mathbf{u_f}, x) = \psi_i(x(t_1)) + \int_{t_0}^{t_1} f_i(t, u_l(t), \mathbf{u_f}(t), x(t))dt,$$

where, for all $i \in \{1, \ldots, p\}$, the functions $\psi_i, \psi_l : \mathbb{R}^n \to \mathbb{R}$, $f_i, f_l : [t_0, T] \times \mathbb{R}^{m_l} \times \mathbb{R}^{m_f} \times \mathbb{R}^n \to \mathbb{R}$ verify the following *preliminary assumptions*

$$(\mathcal{PA}) \quad \begin{cases} \bullet \quad \psi_i, f_i, f_l \text{ are continuously differentiable;} \\ \bullet \quad \text{there exist integrable functions } a_i, a_l : [t_0, T] \to \mathbb{R} \text{ and real numbers } b_i, b_l, c_i, \\ \quad c_l, d_i, d_l, \text{ such that, for all } (t, u_l, \mathbf{u_f}, x) \in [t_0, T] \times \mathbb{R}^{m_l} \times \mathbb{R}^{m_f} \times \mathbb{R}^n, \\ \quad f_i(t, u_l, \mathbf{u_f}, x) \geqslant a_i(t) + b_i x^T x + c_i u_l^T u_l + d_i \mathbf{u_f}^T \mathbf{u_f}, \\ \quad f_l(t, u_l, \mathbf{u_f}, x) \geqslant a_l(t) + b_l x^T x + c_l u_l^T u_l + d_l \mathbf{u_f}^T \mathbf{u_f}; \\ \bullet \quad \psi_i \text{ is a convex function;} \\ \bullet \quad \text{for each fixed } t \in [t_0, T], \text{ the function } f_i(t, \cdot, \cdot, \cdot) \text{ is convex on } \mathbb{R}^{m_l} \times \mathbb{R}^{m_f} \times \mathbb{R}^n. \end{cases}$$

3 The lower level problem

Let $t_1 \in \mathcal{T}$ be fixed, and let $\Phi : [t_0, t_1] \times [t_0, t_1] \to \mathbb{R}^{n \times n}$ be the matrix valued function satisfying for each $s \in [t_0, t_1]$,

$$\forall t \in [t_0, t_1] \qquad \frac{\partial \Phi}{\partial t}(t, s) = A(t)\Phi(t, s) \tag{7}$$

$$\Phi(s, s) = I_n \tag{8}$$

where I_n is the identity matrix.

Since, for each $(u_l, \mathbf{u_f}) \in L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T])$, the unique solution $x_{(t_1, u_l, \mathbf{u_f})} \in H_1^n([t_0, t_1])$ of the Cauchy problem (3-4) is given by

$$\forall t \in [t_0, t_1] \qquad x_{(t_1, u_l, \mathbf{u_f})}(t) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, s)(B_l(s)u_l(s) + \mathbf{B_f}(s)\mathbf{u_f}(s))ds,$$

it is clear that the map $(u_l, \mathbf{u_f}) \mapsto x_{(t_1, u_l, \mathbf{u_f})}$ is affine from $L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T])$ to $H_1^n([t_0, t_1])$. Moreover, using Cauchy-Schwarz inequality, we obtain easily that the map $(u_l, \mathbf{u_f}) \mapsto x_{(t_1, u_l, \mathbf{u_f})}$ is also continuous from $L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T])$ to $H_1^n([t_0, t_1])$.

For each $i = 1, \ldots, p$, consider the functional

$$(u_l, \mathbf{u_f}) \mapsto \tilde{J}_i(t_1, u_l, \mathbf{u_f}) := J_i(t_1, u_l, \mathbf{u_f}, x_{(t_1, u_l, \mathbf{u_f})}). \tag{9}$$

Define also

$$(u_l, \mathbf{u_f}) \mapsto \tilde{J}_l(t_1, u_l, \mathbf{u_f}) := J_l(t_1, u_l, \mathbf{u_f}, x_{(t_1, u_l, \mathbf{u_f})}). \tag{10}$$

From [17, Lemma 1 and Lemma 2] and the fact that $x_{(t_1, \cdot, \cdot)}$ is continuous and affine from $L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T])$ to $H_1^n([t_0, t_1])$, we obtain the following.

Lemma 1. *For each $i = 1, \ldots, p$, the functional $\tilde{J}_i(t_1, \cdot, \cdot) : L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T]) \to \mathbb{R} \cup \{+\infty\}$ is well defined, lower semicontinuous and convex.*

Also $\tilde{J}_l(t_1, \cdot, \cdot) : L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T]) \to \mathbb{R} \cup \{+\infty\}$ is well defined and lower semicontinuous.

For each $(t_1, u_l) \in \mathcal{T} \times \mathcal{U}_l$ (see (1)), denote

$$\mathcal{U}_f(t_1, u_l) = \{\mathbf{u_f} \in L_2^{m_f}([t_0, T]) | \ \mathbf{u_f}(t) \in \mathbf{U_f} \text{ a.e. on } [t_0, T], \tag{11}$$
$$\mathbf{u_f}(t) = 0 \text{ a.e. on } [t_1, T], \ x_{(t_1, u_l, \mathbf{u_f})}(t_1) \in \mathcal{F}\}.$$

For each $(t_1, u_l) \in \mathbb{R} \times L_2^{m_l}([t_0, T]) \setminus \mathcal{T} \times \mathcal{U}_l$ we put $\mathcal{U}_f(t_1, u_l) = \emptyset$. Thus $\mathcal{U}_f$ is a set valued function $\mathcal{U}_f : \mathbb{R} \times L_2^{m_l}([t_0, T]) \rightrightarrows L_2^{m_f}([t_0, T])$.

Recall that

$$\text{dom}\,(\mathcal{U}_f) := \{(t_1, u_l) \in \mathbb{R} \times L_2^{m_l}([t_0, T]) | \ \mathcal{U}_f(t_1, u_l) \neq \emptyset\},$$

and

$$\text{Gr}\,(\mathcal{U}_f) = \{(t_1, u_l, \mathbf{u_f}) \in \mathbb{R} \times L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T]) | \ \mathbf{u_f} \in \mathcal{U}_f(t_1, u_l)\}.$$

We will assume in the sequel that

$$(\mathcal{H}) \qquad\qquad\qquad \text{dom}\,(\mathcal{U}_f) = \mathcal{T} \times \mathcal{U}_l.$$

Proposition 1. *Each of the following is a sufficient condition for* $(\mathcal{H})$.

(a) $\mathcal{F} = \mathbb{R}^n$.

(b) *For each* $t_1 \in \mathcal{T}$, *the linear system*

$$\dot{x}(t) = A(t)x(t) + \mathbf{B_f}(t)\mathbf{u_f}(t), x(t_0) = 0, \quad \mathbf{u_f}(t) \in \mathbf{U_f} \ a.e. \ on \ [t_0, t_1]$$

is controllable, i.e. for any $x_1 \in \mathbb{R}^n$, *there exists* $\mathbf{u_f} \in L_2^{m_f}([t_0, t_1])$ *such that* $\mathbf{u_f}(t) \in \mathbf{U_f}$ *a.e. on* $[t_0, t_1]$, *and the corresponding solution verifies* $x(t_1) = x_1$.

Proof. It is easy to adapt the proof given in [17, Proposition 1], where the initial condition is $x(t_0) = x_0$ (instead of $x(t_0) = 0$ as above). $\square$

It can be easily proved that $\mathcal{U}_f(t_1, u_l)$ is a convex subset of $L_2^{m_f}([t_0, T])$. Thus the problem $(\mathbf{LL})_{(t_1, u_l)}$ can be rewritten as a p-objective convex optimization problem:

$$(\mathbf{M})_{(t_1, u_l)} \qquad \begin{cases} \underset{\mathbf{u_f}}{\mathbf{MIN}}(\tilde{J}_1(t_1, u_l, \mathbf{u_f}), \dots, \tilde{J}_p(t_1, u_l, \mathbf{u_f})) \\ \text{subject to} \quad \mathbf{u_f} \in \mathcal{U}_f(t_1, u_l). \end{cases}$$

DEFINITION 2. Let $\sigma \in \{e, we, pe\}$. An element $\mathbf{u_f} \in L_2^{m_f}([t_0, T])$ will be called σ-*control* of problem $(\mathbf{M})_{(t_1, u_l)}$ iff $(\mathbf{u_f}, x_{(t_1, u_l, \mathbf{u_f})})$ is a σ-control process of problem $(\mathbf{LL})_{(t_1, u_l)}$. We will denote $\mathcal{E}_\sigma(t_1, u_l)$ the set of all σ-controls of the p-objective optimization problem $(\mathbf{M})_{(t_1, u_l)}$.

Thus, using Lemma 1 and the well known scalarization results from vector optimization (see e.g.[38]) we obtain the following.

Theorem 1. *(see [17])* *Let* $(t_1, u_l) \in \mathcal{T} \times \mathcal{U}_l$ *and* $\hat{\mathbf{u}}_f \in \mathcal{U}_f(t_1, u_l)$, *where* $\mathcal{U}_l$ *and* $\mathcal{U}_f$ *are given in (1) and (11) respectively. The control process* $(\hat{\mathbf{u}}_f, x_{(t_1, u_l, \hat{\mathbf{u}}_f)})$ *is weakly (resp. properly) efficient for problem* $(\mathbf{LL})_{(t_1, u_l)}$ *if and only if there exist nonnegative real numbers (resp. positive real numbers)* $\theta_1, \dots, \theta_p$ *with* $\sum_{i=1}^p \theta_i = 1$ *such that* $\hat{\mathbf{u}}_f$ *is an optimal control for the classical scalar optimal control problem :*

$$(\mathbf{S})_{(\theta_1, \dots, \theta_p, t_1, u_l)} \qquad \begin{cases} \underset{\mathbf{u_f}}{\min} \sum_{i=1}^p \theta_i \tilde{J}_i(t_1, u_l, \mathbf{u_f}) \\ \text{subject to} \quad \mathbf{u_f} \in \mathcal{U}_f(t_1, u_l). \end{cases}$$

$\square$

In the sequel we need the following sets

$$\Theta_\sigma = \begin{cases} \{(\theta_1, \dots, \theta_p) \in]0, 1[^p | \sum_{i=1}^p \theta_i = 1\} & \text{if} \ \sigma = pe \\ \\ \{(\theta_1, \dots, \theta_p) \in [0, 1]^p | \sum_{i=1}^p \theta_i = 1\} & \text{if} \ \sigma = we, \end{cases} \qquad (12)$$

and the following hypotheses:

$$H_\sigma(t_1) : \begin{cases} (\exists i \in \{1,\ldots,p\}) \left(\forall(t,v,x) \in [t_0,t_1] \times \mathbb{R}^{m_l} \times \mathbb{R}^n\right) & \\ \mathbf{u_f} \mapsto f_i(t,v,\mathbf{u_f},x) \text{ is strictly convex on } \mathbb{R}^m & \text{if } \sigma = pe \\ & \\ (\forall i \in \{1,\ldots,p\}) \left(\forall(t,v,x) \in [t_0,t_1] \times \mathbb{R}^{m_l} \times \mathbb{R}^n\right) & \\ \mathbf{u_f} \mapsto f_i(t,v,\mathbf{u_f},x) \text{ is strictly convex on } \mathbb{R}^m & \text{if } \sigma = we, \end{cases}$$

and

$$(Hc)_\sigma : \begin{cases} \forall i \in \{1,\ldots,p\} : \ \psi_i \geqslant 0, \ b_i = c_i = 0, \ d_i \geqslant 0, \ \sum_{j=1}^p d_j > 0 & \text{if } \sigma = pe \\ & \\ \forall i \in \{1,\ldots,p\} : \ \psi_i \geqslant 0, \ b_i = c_i = 0, \ d_i > 0 & \text{if } \sigma = we, \end{cases}$$

where b_i, c_i, d_i have been introduced in the preliminary assumptions $(\mathcal{PA})$.

Theorem 2. *(see [17]) Let $\sigma \in \{we, pe\}$ and $(t_1, u_l) \in \mathcal{T} \times \mathcal{U}_l$. Assume that $H_\sigma(t_1)$ holds. Moreover, suppose that at least one of the following hypotheses holds:*

(i) $\mathbf{U_f}$ *is bounded;*

(ii) $(Hc)_\sigma$.

Then, for each $\theta = (\theta_1, \ldots, \theta_p) \in \Theta_\sigma$, there exists a unique optimal control $\mathbf{u_f}(\theta, t_1, u_l, \cdot) \in \mathcal{U}_f(t_1, u_l)$ of the scalar problem $(\mathrm{S})_{(\theta, t_1, u_l)}$.

It is obvious that according to Theorem 1, $\mathbf{u_f}(\theta, t_1, u_l, \cdot)$ is a σ-control for multiobjective problem $(\mathbf{M})_{(t_1, u_l)}$. Moreover, Theorem 1 implies also that for each σ-control $\mathbf{u_f} \in \mathcal{U}_f(t_1, u_l)$ of the multiobjective problem $(\mathbf{M})_{(t_1, u_l)}$, there exists $\theta \in \Theta_\sigma$ such that $\mathbf{u_f}$ is the unique optimal control of the scalar problem $(S)_{(\theta, t_1, u_l)}$.

Thus we can state the following.

Corollary 1. *Let $(t_1, u_l) \in \mathcal{T} \times \mathcal{U}_l$. Under the hypotheses of Theorem 2 we have that the correspondence $\theta \mapsto \mathbf{u_f}(\theta, t_1, u_l, \cdot)$ is a surjection from Θ_σ to the set $\mathcal{E}_\sigma(t_1, u_l)$.*

In the sequel we will keep all the hypotheses of Theorem 2 in addition to the preliminary assumptions $(\mathcal{PA})$.

4 Equivalent formulations of problems $(\mathrm{OSVBC})_\sigma$ and $(\mathrm{PSVBC})_\sigma$

Consider, for each $(\theta, t_1, u_l) \in \Theta_\sigma \times \mathcal{T} \times \mathcal{U}_l \subset \mathbb{R}^p \times \mathbb{R} \times L_2^{m_l}([t_0, T])$, the function $F(\theta, t_1, u_l, \cdot) : \mathcal{U}_f(t_1, u_l) \to \mathbb{R}$ defined by

$$\forall \mathbf{u_f} \in \mathcal{U}_f(t_1, u_l) \qquad F(\theta, t_1, u_l, \mathbf{u_f}) := \sum_{i=1}^p \theta_i \tilde{J}_i(t_1, u_l, \mathbf{u_f}),$$

where $\mathcal{U}_f(t_1, u_l)$ and $\tilde{J}_i$ are given respectively in (11) and (9).

Note that problem $(OSVBC)_\sigma$ can be written equivalently as an optimistic semivectorial bilevel optimization problem

$$(OSVB)_\sigma \qquad \min_{(t_1,u_l)\in\mathcal{T}\times\mathcal{U}_l} \min_{\mathbf{u_f}\in\mathcal{E}_\sigma(t_1,u_l)} \tilde{J}_l(t_1,u_l,\mathbf{u_f}).$$

According to Theorem 2, for each $(\theta,t_1,u_l)\in\Theta_\sigma\times\mathcal{T}\times\mathcal{U}_l$, there exists a unique minimizer $\mathbf{u_f}(\theta,t_1,u_l)\in\mathcal{U}_f(t_1,u_l)$ of $F(\theta,t_1,u_l,\cdot)$ over $\mathcal{U}_f(t_1,u_l)$. According to Corollary 1, for each $(t_1,u_l)\in\mathcal{T}\times\mathcal{U}_l$, we have

$$\mathcal{E}_\sigma(t_1,u_l) = \bigcup_{\theta\in\Theta_\sigma} \{\mathbf{u_f}(\theta,t_1,u_l,\cdot)\}. \tag{13}$$

Then we obviously have the following.

Proposition 2. *(see [17] Problem $(OSVB)_\sigma$ is equivalent to the problem*

$$\min_{(t_1,u_l)\in\mathcal{T}\times\mathcal{U}_l} \min_{\theta\in\Theta_\sigma} \tilde{J}_l(t_1,u_l,\mathbf{u_f}(\theta,t_1,u_l,\cdot))$$

Thus, the optimistic semivectorial problem $(OSVB)_\sigma$, can be rewritten as an optimistic bilevel optimization problem (also called strong Stackelberg problem)

$$(OB)_\sigma \qquad \begin{cases} \displaystyle\min_{(t_1,u_l)\in\mathcal{T}\times\mathcal{U}_l} \min_{\theta\in\Theta_\sigma} \tilde{J}_l(t_1,u_l,\mathbf{u_f}(\theta,t_1,u_l,\cdot)) \\[1em] \text{where } \mathbf{u_f}(\theta,t_1,u_l) \text{ is the unique minimizer to the problem} \\[1em] (S)_{(\theta,t_1,u_l)}: \quad \displaystyle\min_{\mathbf{u_f}\in\mathcal{U}_f(t_1,u_l)} F(\theta,t_1,u_l,\mathbf{u_f}). \end{cases}$$

Here the *upper and lower level are given by scalar optimization problems and the lower level admits a unique solution.*

In the same way the pessimistic semivectorial problem can be rewritten as a pessimistic bilevel optimization problem (leading to a so-called weak Stackelberg problem, see [19] where this terminology was introduced).

Proposition 3. *(see [17]) Problem $(PSVBC)_\sigma$ is equivalent to the problem*

$$\min_{(t_1,u_l)\in\mathcal{T}\times\mathcal{U}_l} \sup_{\theta\in\Theta_\sigma} \tilde{J}_l(t_1,u_l,\mathbf{u_f}(\theta,t_1,u_l,\cdot))$$

Finally, we can rewrite that problem as

$$(PB)_\sigma \qquad \begin{cases} \displaystyle\min_{(t_1,u_l)\in\mathcal{T}\times\mathcal{U}_l} \sup_{\theta\in\Theta_\sigma} \tilde{J}_l(t_1,u_l,\mathbf{u_f}(\theta,t_1,u_l,\cdot)) \\[1em] \text{where } \mathbf{u_f}(\theta,t_1,u_l) \text{ is the unique minimizer of the problem} \\[1em] (S)_{(\theta,t_1,u_l)}: \quad \displaystyle\min_{\mathbf{u_f}\in\mathcal{U}_f(t_1,u_l)} F(\theta,t_1,u_l,\mathbf{u_f}). \end{cases}$$

5 Necessary and Sufficient Conditions for the Scalarized Lower Level Problem

Let $(t_1, u_l) \in \mathcal{T} \times \mathcal{U}_l$, and $\theta = (\theta_1, \ldots, \theta_p) \in \Theta_\sigma$ be given. The scalarized problem $(\mathrm{S})_{(\theta, t_1, u_l)}$ can be written as

$$\min_{(\mathbf{u_f}, x) \in L_2^{m_f}([t_0, T]) \times H_1^n([t_0, t_1])} \left[\sum_{i=1}^p \theta_i \psi_i(x(t_1)) + \int_{t_0}^{t_1} \left(\sum_{i=1}^p \theta_i f_i(t, u_l(t), \mathbf{u_f}(t), x(t)) \right) dt \right] \quad \text{s.t.}$$

$$\begin{aligned}
\mathbf{u_f}(t) &\in \mathbf{U_f} \text{ a.e. on } [t_0, T], \quad \mathbf{u_f}(t) = 0 \text{ a.e. on } [t_1, T], \\
\dot{x}(t) &= A(t)\, x(t) + B_l(t)\, u_l(t) + \mathbf{B_f}(t)\mathbf{u_f}(t) \text{ a.e. on } [t_0, t_1] \\
x(t_0) &= x_0 \\
x(t_1) &\in \mathcal{F}.
\end{aligned}$$

Let $H : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^{m_l} \times \mathbb{R}^{m_f} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be the Hamilton-Pontryagin function associated to this control defined by

$$H(t, u_l, \mathbf{u_f}, x, \lambda_0, \lambda) = \lambda^T \left(A(t)x + B_l(t)u_l + \mathbf{B_f}(t)\mathbf{u_f} \right) - \lambda_0 \sum_{i=1}^p \theta_i f_i(t, u_l, \mathbf{u_f}, x).$$

Let $\lambda(\cdot) = (\lambda_1(\cdot), \ldots, \lambda_n(\cdot)) \in W_{1,\infty}^n([t_0, t_1])$ be the adjoint function, where $W_{1,\infty}^n([t_0, t_1])$ is the Banach space of absolutely continuous functions from $[t_0, t_1]$ to $\mathbb{R}^n$ having derivative in the Banach space $L_\infty^n([t_0, t_1])$ of essentially bounded measurable functions (see e.g.[21] for details).

Since we use L_2 controls, and the Pontryagin maximum principle usually uses controls in L_∞, we will consider two particular situations in order to be able to get necessary and sufficient conditions for problem $(\mathrm{S})_{(\theta, t_1, u_l)}$, as stated below.

5.1 The case when $\mathbf{U_f}$ is bounded and $\mathcal{U}_l \subset L_\infty^{m_l}([t_0, T]) \cap L_2^{m_l}([t_0, T])$

In this subsection we assume the set $\mathbf{U_f}$ is bounded (and closed, convex with non-empty interior), and the leader's controls are essentially bounded, i.e. $\mathcal{U}_l \subset L_\infty^{m_l}([t_0, T]) \cap L_2^{m_l}([t_0, T])$. Also, suppose the target set $\mathcal{F} = \{x \in \mathbb{R}^n \,|\, Gx = a\}$, where the matrix $G \in \mathbb{R}^{k \times n}$, and $a \in \mathbb{R}^k$ are given. Moreover we assume that $\mathrm{rank}(G) = k$. However the results presented in this subsection are also valid when $\mathcal{F} = \mathbb{R}^n$ by taking $G = 0$, $a = 0$.

We obtain the following.

Theorem 3. Necessary conditions. *Let $(\mathbf{u_{f*}}, x_*) \in L_2^{m_f}([t_0, T]) \times H_1^n([t_0, t_1])$ be an optimal control process for problem $(\mathrm{S})_{(\theta, t_1, u_l)}$. Then there exist $\lambda(\cdot) \in W_{1,\infty}^n([t_0, t_1])$, a nonnegative real number λ_0 and a vector $v \in \mathbb{R}^k$ with $(\lambda(\cdot), \lambda_0, v) \neq 0$ such that*

$$\dot{\lambda}^T(t) = -\lambda^T(t) A(t) + \lambda_0 \sum_{i=1}^{p} \theta_i \frac{\partial f_i}{\partial x}(t, u_l(t), \mathbf{u_{f*}}(t), x_*(t)), \quad a.e. \ on \ [t_0, t_1] \quad (14)$$

$$\lambda^T(t_1) = -\lambda_0 \sum_{i=1}^{p} \theta_i \frac{\partial \psi_i}{\partial x}(x_*(t_1)) + v^T G, \quad (15)$$

and, for almost all $t \in [t_0, t_1]$

$$H(t, u_l(t), \mathbf{u_{f*}}(t), x_*(t), \lambda_0, \lambda(t)) = \max_{\mathbf{v_f} \in \mathbf{U_f}} H(t, u_l(t), \mathbf{v_f}, x_*(t), \lambda_0, \lambda(t)). \quad (16)$$

Moreover, if the linearized system

$$\dot{x}(t) = A(t) x(t) + \mathbf{B_f}(t)\mathbf{u_f}(t) \quad a.e. \ on \ [t_0, t_1] \quad (17)$$
$$x(t_0) = 0 \quad (18)$$

is controllable[(i)], then we can take above $\lambda_0 = 1$.

Sufficient conditions. *Let* $(x_*, \mathbf{u_{f*}}) \in H_1^n([t_0, t_1]) \times L_2^{m_f}([t_0, T])$ *verifying (2-5). If there exist* $\lambda(\cdot) \in W_{1,\infty}^n([t_0, t_1])$ *and* $v \in \mathbb{R}^k$ *such that (14-16) are verified with* $\lambda_0 = 1$, *then* $(x_*, \mathbf{u_{f*}})$ *is an optimal control process for problem* $(S)_{(\theta, t_1, u_l)}$.

Proof. Since $\mathbf{U_f}$ is bounded, $\{\mathbf{u_f}(\cdot) \in L_2^{m_f}([t_0, T])| \mathbf{u_f}(t) \in \mathbf{U_f}\} \subset L_\infty^{m_f}([t_0, T])$. For the same reason $u_l(\cdot) \in L_\infty^{m_l}([t_0, t_1])$. Thus we have $\mathbf{u_{f*}} \in L_\infty^{m_f}([t_0, T])$, hence $x_* \in W_{1,\infty}^n([t_0, t_1])$ and $\lambda(\cdot) \in W_{1,\infty}^n([t_0, t_1])$. Therefor we can apply [39, Theorem 5.19] to obtain the first part (necessary conditions). Note that [39, Theorem 5.19] is stated for autonomous systems, but the same proof apply for non-autonomous systems.

For the second part (sufficiency conditions) we can use [39, Theorem 5.22] which also holds for non autonomous systems with the same proof. $\square$

REMARK 3. Since $\mathbf{U_f}$ is convex and closed and H is concave w.r.t. $\mathbf{u_f}$, relation (16) can equivalently be written as a variational inequality

$$\forall \mathbf{v_f} \in \mathbf{U_f} \quad \left(\lambda^T(t)\mathbf{B_f}(t) - \lambda_0 \sum_{i=1}^{p} \theta_i \frac{\partial f_i}{\partial \mathbf{u_f}}(t, u_l(t), \mathbf{u_{f*}}(t), x_*(t))\right)(\mathbf{v_f} - \mathbf{u_{f*}}(t)) \leq 0 \quad a.e. \ on \ [t_0, t_1]$$

Finally, we can conclude the following.

Corollary 2. *Let* $(t_1, u_l) \in \mathcal{U}_l$, *and let* $\theta \in \Theta_\sigma$. *Assume that the linearized system (17-18) is controllable. Let* $\mathbf{u_f} \in L_2^{m_f}([t_0, T])$. *Then* $\mathbf{u_f}(\cdot) = \mathbf{u_f}(\theta, t_1, u_l, \cdot)$ *(i.e.* $\mathbf{u_f}$ *is the unique optimal control for problem* $S_{(\theta, t_1, u_l)}$ *presented in Theorem 2) if, and only if, there exists* $(x(\cdot), \lambda(\cdot), v) \in H_1^n([t_0, t_1]) \times W_{1,\infty}^n([t_0, t_1]) \times \mathbb{R}^k$ *such that*

[(i)]If A and $\mathbf{B_f}$ do not depend on t, it is well known that this system is controllable if, and only if, rank $(\mathbf{B_f}, A\mathbf{B_f}, A^2\mathbf{B_f}, \ldots, A^{n-1}\mathbf{B_f}) = n$

$$\mathbf{u_f}(t) \ \in \ \mathbf{U_f} \ a.e. \ on \ [t_0, T], \quad \mathbf{u_f}(t) = 0 \ a.e. \ on \ [t_1, T] \tag{19}$$

$$\dot{x}(t) \ = \ A(t)\,x(t) + B_l(t)\,u_l(t) + \mathbf{B_f}(t)\mathbf{u_f}(t) \quad a.e. \ on \ [t_0, t_1] \tag{20}$$

$$x(t_0) \ = \ x_0 \tag{21}$$

$$Gx(t_1) \ = \ a \tag{22}$$

$$\dot{\lambda}^T(t) \ = \ -\lambda^T(t)\,A(t) + \sum_{i=1}^{p} \theta_i \frac{\partial f_i}{\partial x}(t, u_l(t), \mathbf{u_{f*}}(t), x_*(t)) \quad a.e. \ on \ [t_0, t_1] \tag{23}$$

$$\lambda^T(t_1) \ = \ -\sum_{i=1}^{p} \theta_i \frac{\partial \psi_i}{\partial x}(x_*(t_1)) + v^T G, \tag{24}$$

and, for almost all $t \in [t_0, t_1]$,

$$\forall \mathbf{v_f} \in \mathbf{U_f} \qquad \left(\lambda^T(t)\mathbf{B_f}(t) - \sum_{i=1}^{p} \theta_i \frac{\partial f_i}{\partial \mathbf{u_f}}(t, u_l(t), \mathbf{u_{f*}}(t), x_*(t)) \right)(\mathbf{v_f} - \mathbf{u_{f*}}(t)) \leq 0. \tag{25}$$

5.2 The case $\mathbf{U_f} = \mathbb{R}^{m_f}$: the followers problem is linear-quadratic; explicit expressions of $\mathbf{u_f}(\theta, t_1, u_l, \cdot)$ and $x_{(t_1, u_l, \mathbf{u_f}(\theta, t_1, u_l, \cdot))}$

In this subsection we consider the case when $\mathbf{U_f} = \mathbb{R}^{m_f}$, $\mathcal{U}_l$ is an arbitrary closed, convex set with non empty interior in $L_2^{m_l}([t_0, T])$, and the endpoint is free, i.e. the target set $\mathcal{F} = \mathbb{R}^n$. The objectives of the followers are quadratic, i.e. for $i = 1, \ldots, p$, and $(t, u_l, \mathbf{u_f}, x) \in [t_0, T] \times \mathbb{R}^n \times \mathbb{R}^{m_l} \times \mathbb{R}^{m_f}$

$$f_i(t, u_l, \mathbf{u_f}, x) = x^T Q_i(t)x + \mathbf{u_f}^T R_i(t)\mathbf{u_f},$$

where $Q_i(\cdot) : [t_0, T] \to \mathbb{R}^{n \times n}$ and $R_i(\cdot) : [t_0, T] \to \mathbb{R}^{m_f \times m_f}$ are continuous positive semidefinite matrix valued functions.

Also

$$\psi_i(x) = x^T Q_i^f x,$$

where Q_i^f is a symmetric semidefinite positive matrix.

Moreover we make the following assumption

$$(HLQP)_\sigma : \quad \begin{cases} \forall (i,t) \in \{1, \ldots, p\} \times [t_0, T] \quad R_i(t) > 0 \quad \text{if } \sigma = we \\[2mm] (\exists i \in \{1, \ldots, p\})\,(\forall t \in [t_0, T]) \quad R_i(t) > 0 \quad \text{if } \sigma = pe. \end{cases}$$

Note that this particular choice of f_i and ψ_i agrees with all the assumptions $(\mathcal{PA})$.

Let us denote

$$Q(\theta, \cdot) = \sum_{i=1}^{p} \theta_i Q_i(\cdot); \quad R(\theta, \cdot) = \sum_{i=1}^{p} \theta_i R_i(\cdot); \quad Q^f(\theta) = \sum_{i=1}^{p} \theta_i Q_i^f.$$

Thus, the scalarized problem $(S)_{(\theta, t_1, u_l)}$ becomes the linear quadratic problem

11

$$\text{(LQP)} \quad \begin{cases} \min \left(x(t_1)^T Q^f(\theta) x(t_1) + \int_{t_0}^{t_1} (x(t)^T Q(\theta,t) x(t) + \mathbf{u_f}(t)^T R(\theta,t) \mathbf{u_f}(t)) dt \right) & \text{s.t.} \\ \dot{x}(t) = A(t) x(t) + \mathbf{B_f}(t) \mathbf{u_f}(t) + B_l(t) u_l(t) & \text{a.e. on } [t_0, t_1], \\ x(t_0) = x_0. \end{cases}$$

We have the following result which is probably known also for L_2 controls, but we will present a proof for the sake of completeness.

Theorem 4. *Let $(x_*(\cdot), \mathbf{u_{f*}}(\cdot)) \in H_1^n([t_0, t_1]) \times L_2^{m_f}([t_0, t_1])$ verify the differential system and the initial condition for problem (LQP). Then the control process $(x_*(\cdot), \mathbf{u_{f*}}(\cdot))$ is optimal for problem (LQP) if, and only if, there exists a function $\lambda(\cdot) \in H_1^n([t_0, t_1])$ such that*

$$
\begin{align}
\dot{\lambda}^T(t) &= -\lambda^T(t) A(t) - x_*^T(t) Q(\theta, t) \quad \text{a.e. on } [t_0, t_1] \tag{26} \\
\lambda^T(t_1) &= x_*^T(t_1) Q^f(\theta) \tag{27} \\
\mathbf{u_{f*}}(t) &= -R^{-1}(\theta, t) \mathbf{B_f}^T(t) \lambda(t) \quad \text{a.e. on } [t_0, t_1]. \tag{28}
\end{align}
$$

Proof. Assume that $\lambda(\cdot) \in H_1^n([t_0, t_1])$ verifies (26-28). Let $(x, \mathbf{u_f}) \in H_1^n([t_0, t_1]) \times L_2^{m_f}([t_0, t_1])$ verify the differential system and the initial condition for problem (LQP). We have for almost all $t \in [t_0, t_1]$

$$
\begin{align*}
\frac{d}{dt} \left(\lambda^T(t)(x(t) - x_*(t)) \right) &= \dot{\lambda}^T(t)(x(t) - x_*(t)) + \lambda^T(t)(\dot{x}(t) - \dot{x}_*(t)) \\
&= -(\lambda^T(t) A(t) + x_*^T(t) Q(\theta, t))(x(t) - x_*(t)) \\
&\quad + \lambda^T(t) \left(A(t)(x(t) - x_*(t)) + \mathbf{B_f}(t)(\mathbf{u_f}(t) - \mathbf{u_{f*}}(t)) \right) \\
&= -x_*^T(t) Q(\theta, t)(x(t) - x_*(t)) - \mathbf{u_{f*}}^T(t) R(\theta, t)(\mathbf{u_f}(t) - \mathbf{u_{f*}}(t))
\end{align*}
$$

With the initial condition for $x(\cdot), x_*(\cdot)$ and final condition for $\lambda(\cdot)$ we get by integration

$$
x_*^T(t_1) Q^f(\theta)(x(t_1) - x_*(t_1)) = -\int_{t_0}^{t_1} \left(x_*^T(t) Q(\theta, t)(x(t) - x_*(t)) + \mathbf{u_{f*}}^T(t) R(\theta, t)(\mathbf{u_f}(t) - \mathbf{u_{f*}}(t)) \right) dt.
\tag{29}
$$

Denote

$$
J(x(\cdot), \mathbf{u_f}(\cdot)) = \left(x(t_1)^T Q^f(\theta) x(t_1) + \int_{t_0}^{t_1} (x(t)^T Q(\theta, t) x(t) + \mathbf{u_f}(t)^T R(\theta, t) \mathbf{u_f}(t)) dt \right).
$$

Using the fact that for any symmetric positive semidefinite matrix P, for all vectors v, v_* we obviously have

$$
v^T P v - v_*^T P v_* \geq 2 v_*^T P(v - v_*)
$$

we get

$$
\begin{align*}
J(x(\cdot), \mathbf{u_f}(\cdot)) - J(x_*(\cdot), \mathbf{u_{f*}}(\cdot)) \geq 2 \Big[& x_*^T(t_1) Q^f(\theta)(x(t_1) - x_*(t_1)) \\
& + \int_{t_0}^{t_1} \left(x_*^T(t) Q(\theta, t)(x(t) - x_*(t)) + \mathbf{u_{f*}}^T(t) R(\theta, t)(\mathbf{u_f}(t) - \mathbf{u_{f*}}(t)) \right) dt \Big]
\end{align*}
$$

From (29) the last expression is zero, hence $J(x(\cdot), u(\cdot)) - J(x_*(\cdot), \mathbf{u_{f*}}(\cdot)) \geq 0$. Thus $(x_*(\cdot), \mathbf{u_{f*}}(\cdot))$ is an optimal control process for problem (SQP).

Conversely, let $(x_*(\cdot), \mathbf{u_{f*}}(\cdot)) \in H_1^n([t_0, t_1]) \times L_2^{m_f}([t_0, t_1])$ be a solution of (LQP) (which exists and is unique according to Theorem 2). Let $\lambda(\cdot) \in H_1^n([t_0, t_1])$ be the solution of the linear system (26) verifying the final condition (27). For any $\mathbf{u_f}(\cdot) \in L_2^{m_f}([t_0, t_1])$, denoting by $x(\cdot)$ the corresponding solution of the differential system and the initial condition for problem (LQP), we have (using a similar calculus as before)

$$\lambda^T(t_1)(x(t_1) - x_*(t_1)) = -\int_{t_0}^{t_1} \left(x_*^T(t)Q(\theta, t)(x(t) - x_*(t)) + \lambda^T(t)\mathbf{B_f}(t)(\mathbf{u_f}(t) - \mathbf{u_{f*}}(t)) \right) dt.$$

On the other, using the fact that the directional derivative of J at the optimal point $(x_*(\cdot), \mathbf{u_{f*}}(\cdot))$ in the direction $(x(\cdot), \mathbf{u_f}(\cdot)) - (x_*(\cdot), \mathbf{u_{f*}}(\cdot))$ is positive we have

$$x_*^T(t_1)Q^f(\theta)(x(t_1) - x_*(t_1)) + \int_{t_0}^{t_1} (x_*^T(t)Q(\theta, t)(x(t) - x_*(t)) + \mathbf{u_{f*}}^T(t)R(\theta, t)(\mathbf{u_f}(t) - \mathbf{u_{f*}}(t))) dt \geq 0$$

Finally we obtain

$$\int_{t_0}^{t_1} (\lambda^T(t)\mathbf{B_f}(t) - \mathbf{u_{f*}}^T(t)R(\theta, t))(\mathbf{u_f}(t) - \mathbf{u_{f*}}(t))) dt \leq 0.$$

Since $\mathbf{u_f}(\cdot)$ can be arbitrarily chosen in $L_2^{m_f}([t_0, t_1])$, we obtain that (28) is satisfied. $\qquad\square$

Next we will show that, in the linear-quadratic case, it is possible to compute explicitly the optimal control and state as a function of the parameters θ, t_1, u_l by means of a $2n \times 2n$ resolvent matrix of a linear differential system based on data. This fact will allow us to find explicit optimality conditions for our bilevel problems.

Recall that $\mathbf{u_f}(\theta, t_1, u_l, \cdot)$ denotes the unique optimal control of the scalarized problem $(S)_{(\theta, t_1, u_l)}$. The corresponding unique state and adjoint state (verifying Theorem 4) will be denoted by $x(\theta, t_1, u_l, \cdot)$ and $\lambda(\theta, t_1, u_l, \cdot)$.

To be more precise, the functions $x(\theta, t_1, u_l, \cdot)$ and $\lambda(\theta, t_1, u_l, \cdot)$ verify the following boundary linear problem

$$\frac{\partial x}{\partial t}(\theta, t_1, u_l, t) = A(t)x(\theta, t_1, u_l, t) - \mathbf{B_f}(t)R^{-1}(\theta, t)\mathbf{B_f}(t)^T\lambda(\theta, t_1, u_l, t) \tag{30}$$
$$+ B_l(t)u_l(t) \quad \text{a.e. on } [t_0, t_1]$$
$$\frac{\partial \lambda}{\partial t}(\theta, t_1, u_l, t) = -A(t)^T\lambda(\theta, t_1, u_l, t) - Q(\theta, t)x(\theta, t_1, u_l, t) \quad \text{a.e. on } [t_0, t_1] \tag{31}$$
$$x(\theta, t_1, u_l, t_0) = x_0 \tag{32}$$
$$\lambda(\theta, t_1, u_l, t_1) = Q^f(\theta)x(\theta, t_1, u_l, t_1) \tag{33}$$

and

$$\mathbf{u_f}(\theta, t_1, u_l, t) = -R^{-1}(\theta, t)\mathbf{B_f}^T(t)\lambda(\theta, t_1, u_l, t) \quad \text{a.e. on } [t_0, t_1]. \tag{34}$$

Given $t_1 \in \mathcal{T}$ and $\theta \in \Theta_\sigma$, consider the matrix valued function $P(\theta, t_1, \cdot) : [t_0, t_1] \to \mathbb{R}^{n \times n}$ which, under our hypotheses about matrices $Q^f(\theta)$, $Q(\theta, t)$, $R(\theta, t)$, is the unique continuously differentiable solution (see e.g. [1]) of the Riccati matrix differential equation (RMDE) on $[t_0, t_1]$:

$$\frac{\partial P}{\partial t}(\theta, t_1, t) = -A(t)^T P(\theta, t_1, t) - P(\theta, t_1, t)A(t) - Q(\theta, t) + P(\theta, t_1, t)\mathbf{B_f}(t)R(\theta, t)^{-1}\mathbf{B_f}(t)^T P(\theta, t_1, t)$$

satisfying the final time condition

$$P(\theta, t_1, t_1) = Q^f(\theta). \tag{35}$$

Moreover, $P(\theta, t_1, t)$ is a symmetric positive definite matrix for each t.

Following [18] we can express P in terms of a resolvent matrix depending directly on data. Thus consider for all $(\theta, t) \in \Theta_\sigma \times [t_0, t_1]$ the $2n \times 2n$ matrix which defines the linear system (30, 31)

$$L(\theta, t) = \begin{pmatrix} A(t) & -\mathbf{B_f}(t)R^{-1}(\theta, t)\mathbf{B_f}^T(t) \\ -Q(\theta, t) & -A^T(t) \end{pmatrix}$$

The proof of the following result can be found in [18].

Proposition 4. *Let $\Psi(\theta, \cdot, \cdot)$ be the resolvent (or state transition) matrix associated to the linear differential system defined by $L(\theta, t)$, i.e., for each $s \in [t_0, T]$, $\Psi(\theta, \cdot, s)$ satisfies the Cauchy problem:*

$$\frac{\partial \Psi}{\partial t}(\theta, t, s) = L(\theta, t)\Psi(\theta, t, s), \quad t \in [t_0, T], \quad \Psi(\theta, s, s) = I_{2n}.$$

Let us divide the matrix $\Psi(\theta, t, s)$ into four $n \times n$ blocks

$$\Psi(\theta, t, s) = \begin{pmatrix} \Psi_{11}(\theta, t, s) & \Psi_{12}(\theta, t, s) \\ \Psi_{21}(\theta, t, s) & \Psi_{22}(\theta, t, s) \end{pmatrix}.$$

Then, for all $t \in [t_0, t_1]$, the matrix $[\Psi_{11}(\theta, t, t_1) + \Psi_{12}(\theta, t, t_1)Q^f(\theta)]$ is invertible and

$$P(\theta, t_1, t) = \left[\Psi_{21}(\theta, t, t_1) + \Psi_{22}(\theta, t, t_1)Q^f(\theta)\right]\left[\Psi_{11}(\theta, t, t_1) + \Psi_{12}(\theta, t, t_1)Q^f(\theta)\right]^{-1}. \tag{36}$$

Next, let us denote by $\xi(\theta, t_1, u_l, \cdot) \in H_1^n([t_0, t_1])$ the unique solution of the following linear Cauchy problem

$$\frac{\partial \xi}{\partial t}(\theta, t_1, u_l, t) = \left(-A(t)^T + P(\theta, t_1, t)\mathbf{B_f}(t)R^{-1}(\theta, t)\mathbf{B_f}(t)\right)\xi(\theta, t_1, u_l, t) \tag{37}$$
$$-P(\theta, t_1, t)B_l(t)u_l(t) \quad \text{a.e. on } [t_0, t_1]$$
$$\xi(\theta, t_1, u_l, t_1) = 0. \tag{38}$$

Lemma 2. *For all $t \in [t_0, t_1]$ we have*

$$\lambda(\theta, t_1, u_l, t) = P(\theta, t_1, t)x(\theta, t_1, u_l, t) + \xi(\theta, t_1, u_l, t). \tag{39}$$

14

Proof. Computing the derivative $\frac{\partial}{\partial t}\Big(\lambda(\theta, t_1, u_l, t) - P(\theta, t_1, t)x(\theta, t_1, u_l, t) - \xi(\theta, t_1, u_l, t)\Big)$, and then, using (30-33), (RMDE), (35), (37) and (38), the result follows easily. $\square$

Denote by $\Xi(\theta, t_1, \cdot, \cdot)$ the resolvent matrix associated to (37), i.e. for all $(\theta, t_1, s) \in \Theta_\sigma \times \mathcal{T} \times [t_0, T]$

$$\frac{\partial \Xi}{\partial t}(\theta, t_1, t, s) = \big(- A(t)^T + P(\theta, t_1, t)\mathbf{B_f}(t)R^{-1}(\theta, t)\mathbf{B_f}(t)\big)\Xi(\theta, t_1, t, s), \quad t \in [t_0, T] \quad (40)$$

$$\Xi(\theta, t_1, s, s) = I_n. \quad (41)$$

Based on this we are able to solve the boundary problem (30-33) in terms of data.

Corollary 3. *For all $(\theta, t_1, u_l) \in \Theta_\sigma \times \mathcal{T} \times L_2^{m_l}([t_0, T])$, and for all $t \in [t_0, t_1]$ we have*

$$\begin{pmatrix} x(\theta, t_1, u_l, t) \\ \lambda(\theta, t_1, u_l, t) \end{pmatrix} = \Psi(\theta, t, t_0) \begin{pmatrix} x_0 \\ P(\theta, t_1, t_0)x_0 + \xi(\theta, t_1, u_l, t_0) \end{pmatrix} + \int_{t_0}^{t} \Psi(\theta, t, s) \begin{pmatrix} B_l(s)u_l(s) \\ 0 \end{pmatrix} ds,$$

where

$$\xi(\theta, t_1, u_l, t_0) = \int_{t_0}^{t_1} \Xi(\theta, t_1, t_0, s)P(\theta, t_1, s)B_l(s)u_l(s)ds.$$

REMARK 4. The right hand side member in the formulas giving $x(\theta, t_1, u_l, t)$ and $\lambda(\theta, t_1, u_l, t)$ in Corollary 3 is defined for all $(t_1, t) \in]t_0, T[\times[t_0, T]$ (and not only for $(t_1, t) \in \mathcal{T} \times [t_0, t_1]$), and for all θ belonging to an open convex set Ω with $\Theta_\sigma \subseteq \Omega$. Indeed, the formulas in Corollary 3 have a meaning as long as $R(\theta, t) > 0$.

When $\sigma = pe$, by $(\text{HLQP})_{pe}$ it is obvious that we can take $\Omega = \mathbb{R}_{++}^p$.

When $\sigma = we$, the continuous function $[t_0, T] \times \mathbb{R}^{m_f} \ni (t, \mathbf{u_f}) \mapsto \mathbf{u_f}^T R_i(t)\mathbf{u_f}$ attains its minimum value, say α_i, on the compact set $[t_0, T] \times \mathbb{S}$, where $\mathbb{S}$ is the unit sphere in $\mathbb{R}^{m_f}$, $i = 1, \ldots, p$. According to $(\text{HLQP})_{we}$ we have $\alpha_i > 0$ for all i. Then, it is easy to see that we can take

$$\Omega = \{\theta \in \mathbb{R}^p | \sum_{i=1}^{p} \theta_i\alpha_i > 0\}.$$

We will extend the functions $x(\cdot, \cdot, \cdot, \cdot)$ and $\lambda(\cdot, \cdot, \cdot, \cdot)$ based on these formulas as continuous functions from $\Omega \times]t_0, T[\times L_2^{m_l}([t_0, T]) \times [t_0, T]$ to $\mathbb{R}^n$. Moreover, based on (34), we will extend also the function $\mathbf{u_f}(\cdot, \cdot, \cdot, \cdot)$ as a continuous function from $\Omega \times]t_0, T[\times L_2^{m_l}([t_0, T]) \times [t_0, T]$ to $\mathbb{R}^{m_f}$. These extensions are necessary further in order to obtain optimality conditions for the upper level.

Using the differentiability with respect to parameters of a differential equation and some straightforward computation we have the following.

Proposition 5. *The resolvent $\Psi(\cdot, \cdot, \cdot)$ is continuously differentiable on $\Omega \times [t_0, T] \times [t_0, T]$. We have the following formulas for all $(\theta, t, s) \in \Omega \times [t_0, T] \times [t_0, T]$ and $i = 1, \ldots, p$.*

15

$$\frac{\partial \Psi}{\partial \theta_i}(\theta, t, s) = \int_s^t \Psi(\theta, t, \tau) \frac{\partial L}{\partial \theta_i}(\theta, \tau) \Psi(\theta, \tau, s) d\tau, \quad where \tag{42}$$

$$\frac{\partial L}{\partial \theta_i}(\theta, t) = \begin{pmatrix} 0 & \mathbf{B_f}(t) R^{-1}(\theta, t) R_i(t) R^{-1}(\theta, t) \mathbf{B_f}(t)^T \\ -Q_i(t) & 0 \end{pmatrix} \tag{43}$$

$$\frac{\partial \Psi}{\partial s}(\theta, t, s) = -\Psi(\theta, t, s) L(\theta, s). \tag{44}$$

By (36) and the previous Proposition we obtain immediately the following.

Proposition 6. *The matrix valued function $P(\cdot, \cdot, \cdot)$ is continuously differentiable on $\Omega \times [t_0, T] \times [t_0, T]$, and verifies the following formulas.*

$$\frac{\partial P}{\partial \theta_i}(\theta, t_1, t) = \left[\frac{\partial \Psi_{21}}{\partial \theta_i}(\theta, t, t_1) + \frac{\partial \Psi_{22}}{\partial \theta_i}(\theta, t, t_1) Q^f(\theta) + \Psi_{22}(\theta, t, t_1) Q_i^f \right] \left[\Psi_{11}(\theta, t, t_1) + \Psi_{12}(\theta, t, t_1) Q^f(\theta) \right]^{-1}$$
$$\text{ï?‘}\frac{1}{2} \quad - \left[\Psi_{21}(\theta, t, t_1) + \Psi_{22}(\theta, t, t_1) Q^f(\theta) \right] \left[\Psi_{11}(\theta, t, t_1) + \Psi_{12}(\theta, t, t_1) Q^f(\theta) \right]^{-1} \tag{45}$$
$$\text{ï?‘}\frac{1}{2} \quad \left[\frac{\partial \Psi_{11}}{\partial \theta_i}(\theta, t, t_1) + \frac{\partial \Psi_{12}}{\partial \theta_i}(\theta, t, t_1) Q^f(\theta) + \Psi_{12}(\theta, t, t_1) Q_i^f \right] \left[\Psi_{11}(\theta, t, t_1) + \Psi_{12}(\theta, t, t_1) Q^f(\theta) \right]^{-1}$$

and

$$\frac{\partial \Psi}{\partial \theta_i}(\theta, t, s) = \begin{pmatrix} \frac{\partial \Psi_{11}}{\partial \theta_i}(\theta, t, s) & \frac{\partial \Psi_{12}}{\partial \theta_i}(\theta, t, s) \\ \frac{\partial \Psi_{21}}{\partial \theta_i}(\theta, t, s) & \frac{\partial \Psi_{22}}{\partial \theta_i}(\theta, t, s) \end{pmatrix}$$

Using an analog calculus we obtain

$$\frac{\partial P}{\partial t_1}(\theta, t_1, t) = \left[\frac{\partial \Psi_{21}}{\partial t_1}(\theta, t, t_1) + \frac{\partial \Psi_{22}}{\partial t_1}(\theta, t, t_1) Q^f(\theta) \right] \left[\Psi_{11}(\theta, t, t_1) + \Psi_{12}(\theta, t, t_1) Q^f(\theta) \right]^{-1}$$
$$- \left[\Psi_{21}(\theta, t, t_1) + \Psi_{22}(\theta, t, t_1) Q^f(\theta) \right] \left[\Psi_{11}(\theta, t, t_1) + \Psi_{12}(\theta, t, t_1) Q^f(\theta) \right]^{-1} \tag{46}$$
$$\left[\frac{\partial \Psi_{11}}{\partial t_1}(\theta, t, t_1) + \frac{\partial \Psi_{12}}{\partial t_1}(\theta, t, t_1) Q^f(\theta) \right] \text{ï?‘}\frac{1}{2} \left[\Psi_{11}(\theta, t, t_1) + \Psi_{12}(\theta, t, t_1) Q^f(\theta) \right]^{-1}$$

The computation of $\frac{\partial \Psi_{ij}}{\partial t_1}(\theta, t, t_1)$ can be obtained using (44)

$$\begin{pmatrix} \frac{\partial \Psi_{11}}{\partial t_1}(\theta, t, t_1) & \frac{\partial \Psi_{12}}{\partial t_1}(\theta, t, t_1) \\ \text{ï?‘}\frac{1}{2}\frac{\partial \Psi_{21}}{\partial t_1}(\theta, t, t_1) & \frac{\partial \Psi_{22}}{\partial t_1}(\theta, t, t_1) \end{pmatrix} = - \begin{pmatrix} \Psi_{11}(\theta, t, t_1) & \Psi_{12}(\theta, t, t_1) \\ \text{ï?‘}\frac{1}{2} & \\ \Psi_{21}(\theta, t, t_1) & \Psi_{22}(\theta, t, t_1) \end{pmatrix} L(\theta, t_1) \tag{47}$$

16

Proposition 7. *The resolvent* $\Xi(\cdot, \cdot, \cdot, \cdot)$ *is continuously differentiable on* $\Omega \times [t_0, T] \times [t_0, T]$, *and, denoting*

$$\mathcal{A}(\theta, t_1, t) := -A(t)^T + P(\theta, t_1, t)\mathbf{B_f}(t)R^{-1}(\theta, t)\mathbf{B_f}(t), \tag{48}$$

we have

$$\frac{\partial \Xi}{\partial \theta_i}(\theta, t_1, t, s) = \int_s^t \Xi(\theta, t_1, t, \tau)\frac{\partial \mathcal{A}}{\partial \theta_i}(\theta, t_1, \tau)\Xi(\theta, t_1, \tau, s)d\tau, \tag{49}$$

$$\frac{\partial \Xi}{\partial t_1}(\theta, t_1, t, s) = \int_s^t \Xi(\theta, t_1, t, \tau)\frac{\partial \mathcal{A}}{\partial t_1}(\theta, t_1, \tau)\Xi(\theta, t_1, \tau, s)d\tau, \tag{50}$$

$$\frac{\partial \Xi}{\partial s}(\theta, t_1, t, s) = -\Xi(\theta, t_1, t, s)\mathcal{A}(\theta, t_1, s). \tag{51}$$

The computation of the partial derivatives of $\mathcal{A}(\theta, t_1, t)$ *can be obtained using (36), Proposition 6, and the obvious formulas*

$$\frac{\partial}{\partial \theta_i}R^{-1}(\theta, t) = -R^{-1}(\theta, t)R_i(t)R^{-1}(\theta, t)$$

Proposition 8. *For all* $(\theta, t_1) \in \Omega \times]t_0, T[$, *the maps* $u_l \mapsto x(\theta, t_1, u_l, \cdot)$, $u_l \mapsto \lambda(\theta, t_1, u_l, \cdot)$, *respectively* $u_l \mapsto \mathbf{u_f}(\theta, t_1, u_l, \cdot)$ *are affine and continuous from* $L_2^{m_l}([t_0, T])$ *to* $H_1^n([t_0, t_1])$, *respectively from* $L_2^{m_l}([t_0, T])$ *to* $L_2^{m_f}([t_0, T])$. *Therefore they are continuously Fréchet differentiable on* $L_2^{m_l}([t_0, T])$ *and, for any* $u_l \in L_2^{m_l}([t_0, t_1])$, *their Fréchet differentials (which are linear continuous maps from* $L_2^{m_l}([t_0, T])$ *to* $H_1^n([t_0, t_1])$, *and respectively from* $L_2^{m_l}([t_0, T])$ *to* $L_2^{m_f}([t_0, T])$) *verify for all* $h \in L_2^{m_l}([t_0, T])$, *and for all* $t \in [t_0, t_1]$

$$\frac{\partial}{\partial u_l}x(\theta, t_1, u_l, t) \cdot h = \Psi_{12}(\theta, t, t_0)\int_{t_0}^{t_1} \Xi(\theta, t_1, t_0, s)P(\theta, t_1, s)B_l(s)h(s)ds \tag{52}$$

$$+ \int_{t_0}^t \Psi_{11}(\theta, t, s)B_l(s)h(s)ds$$

$$\text{ï?‘}\frac{1}{2}\frac{\partial}{\partial u_l}\lambda(\theta, t_1, u_l, t) \cdot h = \Psi_{22}(\theta, t, t_0)\int_{t_0}^{t_1} \Xi(\theta, t_1, t_0, s)P(\theta, t_1, s)B_l(s)h(s)ds \tag{53}$$

$$\text{ï?‘}\frac{1}{2} + \int_{t_0}^t \Psi_{21}(\theta, t, s)B_l(s)h(s)ds$$

$$\frac{\partial}{\partial u_l}\mathbf{u_f}(\theta, t_1, u_l, t) \cdot h = -R^{-1}(\theta, t)\mathbf{B_f}(t)^T\frac{\partial}{\partial u_l}\lambda(\theta, t_1, u_l, t) \cdot h \tag{54}$$

Proof. It is easy to see from Corollary 3 and (30-31) that the maps $u_l \mapsto x(\theta, t_1, u_l, \cdot)$ and $u_l \mapsto \lambda(\theta, t_1, u_l, \cdot)$ are affine and continuous from $L_2^{m_l}([t_0, T])$ to $H_1^n([t_0, t_1])$, hence (52) and (53) hold. Then, by (34) we obtain that the map $u_l \mapsto \mathbf{u_f}(\theta, t_1, u_l, \cdot)$ from $L_2^{m_l}([t_0, T])$ to $L_2^{m_f}([t_0, T])$ is affine and continuous and we get (54). $\square$

Theorem 5. *(REGULARITY OF* $\mathbf{u_f}(\cdot, \cdot, \cdot, \cdot)$ *AND* $x(\cdot, \cdot, \cdot, \cdot)$*)*

1. *The functions* $\mathbf{u_f}(\cdot, \cdot, \cdot, \cdot) : \Omega \times]t_0, T[\times L_2^{m_l}([t_0, T]) \times [t_0, T] \rightarrow \mathbb{R}^{m_f}$ *and* $x(\cdot, \cdot, \cdot, \cdot) : \Omega \times]t_0, T[\times L_2^{m_l}([t_0, T]) \times [t_0, T] \rightarrow \mathbb{R}^n$ *are continuous.*

2. *The function $(\theta, t_1, u_l) \mapsto \mathbf{u_f}(\theta, t_1, u_l, \cdot)$ from $\Omega \times]t_0, T[\times L_2^{m_l}([t_0, T])$ to $L_2^{m_f}([t_0, T])$ is continuous as well as the function $(\theta, t_1, u_l) \mapsto x(\theta, t_1, u_l, \cdot)$ from $\Omega \times]t_0, T[\times L_2^{m_l}([t_0, T])$ to $L_2^n([t_0, T])$.*

3. *For each fixed $(\bar{\theta}, \bar{t}_1, \bar{u}_l) \in \Omega \times]t_0, T[\times L_2^{m_l}([t_0, T])$,*

   - *The function $\theta \mapsto \mathbf{u_f}(\theta, \bar{t}_1, \bar{u}_l, \cdot)$ from Ω to $L_2^{m_f}([t_0, T])$, and the function[ii] $\theta \mapsto x(\theta, \bar{t}_1, \bar{u}_l, \cdot)$ from Ω to $L_2^n([t_0, T])$ are continuously Fréchet differentiable on Ω.*

   - *The function $u_l \mapsto \mathbf{u_f}(\bar{\theta}, \bar{t}_1, u_l, \cdot)$ from $L_2^{m_l}([t_0, T])$ to $L_2^{m_f}([t_0, T])$, and the function $u_l \mapsto x(\bar{\theta}, \bar{t}_1, u_l, \cdot)$ from $L_2^{m_l}([t_0, T])$ to $H_1^n([t_0, T])$ are continuously Fréchet differentiable.*

   - *The functions $t_1 \mapsto \mathbf{u_f}(\bar{\theta}, t_1, \bar{u}_l, \cdot)$ from $]t_0, T[$ to $L_2^{m_f}([t_0, T])$ and $t_1 \mapsto x(\bar{\theta}, t_1, \bar{u}_l, \cdot)$ from $]t_0, T[$ to $L_2^n([t_0, T])$ are a.e. differentiable on $]t_0, T[$, and for almost all $t_1 \in]t_0, T[$, $\dfrac{\partial \mathbf{u_f}}{\partial t_1}(\bar{\theta}, \bar{t}_1, \bar{u}_l, \cdot) \in L_2^{m_f}([t_0, T])$ and $\dfrac{\partial x}{\partial t_1}(\bar{\theta}, \bar{t}_1, \bar{u}_l, \cdot) \in L_2^n([t_0, T])$.*

     Moreover, for each $t_1 \in]t_0, T[$ such that $\bar{u}_l$ is continue[iii] at t_1, these functions are differentiable in t_1.

4. *The functions $\mathbf{u_f}(\cdot, \cdot, \cdot, \cdot)$, $x(\cdot, \cdot, \cdot, \cdot)$ and their partial derivatives can be explicitly represented as functions of data (supposing we are able to compute the resolvent matrices Ψ and Ξ).*

Proof. By Corollary 3, Remark 4 and Propositions 5-8, we obtain the points 1 and 4.

To prove the point 2 we will use the fact that, by Corollary 3, we can write

$$x(\theta, t_1, u_l, t) = \alpha(\theta, t_1, t) + \int_{t_0}^{T} X(\theta, t_1, t, s) u_l(s) ds,$$

where

$$\alpha(\theta, t_1, t) = \big(\Psi_{11}(\theta, t, t_0) + \Psi_{12} P(\theta, t_1, t_0)\big) x_0$$

and $X(\theta, t_1, t, s)$ is described later in relations (61) and (63). Obviously $\alpha : \Omega \times]t_0, T[\times [t_0, T] \to \mathbb{R}^n$ is a continuous function, and for each $s \in [t_0, T]$, $X(\cdot, \cdot, \cdot, s)$ is continuous on $\Omega \times]t_0, T[\times [t_0, T] \to \mathbb{R}^{n \times m_l}$, and, for each $(\theta, t_1, t) \in \Omega \times]t_0, T[\times [t_0, T]$, $X(\theta, t_1, t, \cdot) \in L_2^{n \times m_l}([t_0, T])$.

We obtain easily that the function $(\theta, t_1) \mapsto \alpha(\theta, t_1, \cdot)$ is continuous from $\Omega \times]t_0, T[$ to $\mathcal{C}([t_0, T]; \mathbb{R}^n)$, where $\mathcal{C}([t_0, T]; \mathbb{R}^n)$ is the Banach space of continuous functions on $[t_0, T]$ with values in $\mathbb{R}^n$ endowed with the uniform convergence norm.

Since the embedding $\mathcal{C}([t_0, T]; \mathbb{R}^n) \subset L_2^n([t_0, T])$ is continuous, we obtain that the function $(\theta, t_1) \mapsto \alpha(\theta, t_1, \cdot)$ is continuous from $\Omega \times]t_0, T[$ to $L_2^n([t_0, T])$.

Also, using Lebesgue's dominated convergence theorem, we obtain easily that the function $(\theta, t_1, t) \mapsto X(\theta, t_1, t, \cdot)$ is continuous from $\Omega \times]t_0, T[\times [t_0, T]$ to $L_2^{n \times m_l}([t_0, T])$. Denoting $y(\theta, t_1, u_l, t) = \int_{t_0}^{T} X(\theta, t_1, t, s) u_l(s) ds$, and writing

$$y(\theta', t_1', u_l', t) - y(\theta, t_1, u_l, t) = \big(y(\theta', t_1', u_l', t) - y(\theta', t_1', u_l, t)\big) + \big(y(\theta', t_1', u_l, t) - y(\theta, t_1, u_l, t)\big),$$

[ii] Note that the embedding $H_1^n([t_0, T]) \subset L_2^n([t_0, T])$ is continue.

[iii] In the sense that there exists a function $\tilde{u}_l$ continue at t_1 and $\bar{u}_l(t) = \tilde{u}_l(t)$ a.e. on $[t_0, T]$. Note that by Lusin's theorem, we can find measurable sets of arbitrarily small positive measure, and such functions $\tilde{u}_l$ which are continuous on the complement of those sets.

we obtain that

$$|y(\theta', t_1', u_l', t) - y(\theta, t_1, u_l, t)| \leq \|X(\theta', t_1', t, \cdot)\|_2 \cdot \|u_l' - u_l\|_2 + \|X(\theta', t_1', t, \cdot) - X(\theta, t_1, t, \cdot)\|_2 \cdot \|u_l\|_2$$

which finally prove the continuity of the function $(\theta, t_1, u_l) \mapsto x(\theta, t_1, u_l, \cdot)$ from $\Omega \times]t_0, T[\times L_2^{m_l}([t_0, T])$ to $L_2^n([t_0, T])$.

With similar arguments we can prove the continuity of the function $(\theta, t_1, u_l) \mapsto \mathbf{u_f}(\theta, t_1, u_l, \cdot)$ from $\Omega \times]t_0, T[\times L_2^{m_l}([t_0, T])$ to $L_2^{m_f}([t_0, T])$, and the point 3. $\qquad \square$

6 Optimality conditions for the upper level, i.e. for problems $(OB)_\sigma$ and $(PB)_\sigma$

In this section we will restrain to the case considered in Subsection 5.2. Moreover we will suppose that $\mathcal{U}_l$ is the closed ball

$$\mathcal{U}_l = \{u_l \in L_2^{m_l}([t_0, T]) \,|\, \|u_l\|_2 \leq R\}, \tag{55}$$

where R is a strictly positive real.

6.1 The optimistic bilevel problem

We begin with some preliminary results in order to obtain an existence result when $\mathbf{U_f}$ is not assumed to be bounded so we cannot apply the results obtained in [17]. We could adapt the proof given in [17] but we will give direct proofs for the sake of completeness.

Lemma 3. *Let X and Y be arbitrary sets and let $J : X \times Y \to \mathbb{R} \cup \{+\infty\}$ such that, for each $x \in X$, the set argmin $J(x, \cdot)$ is nonempty. Then the problems*

$$\min_{(x,y) \in X \times Y} J(x, y) \tag{56}$$

and

$$\min_{x \in X} \min_{y \in Y} J(x, y) \tag{57}$$

are equivalent, i.e. problem (56) is solvable if and only if problem (57) is solvable. In this case the solution sets coincide as well as the minimal values.

Proof. Let $(\hat{x}, \hat{y}) \in X \times Y$ be a solution for problem (56), i.e. $(\hat{x}, \hat{y}) \in$ argmin $J(\cdot, \cdot)$. Then, for each $x \in X$, we have obviously $J(\hat{x}, \hat{y}) = \min_{y \in Y} J(\hat{x}, y) \leq \min_{y \in Y} J(x, y)$, hence $J(\hat{x}, \hat{y}) = \min_{x \in X} \min_{y \in Y} J(x, y)$, and $(\hat{x}, \hat{y})$ is a solution for problem (57).

Conversely, let $(\bar{x}, \bar{y})$ be a solution for problem (57). This means that, for all $x \in X$ and $y' \in$ argmin $J(x, \cdot)$, we have we have $J(\bar{x}, \bar{y}) \leq J(x, y') = \min_{y \in Y} J(x, y)$, hence for all $(x, y) \in X \times Y$, we have $J(\bar{x}, \bar{y}) \leq J(x, y)$. Therefore $(\bar{x}, \bar{y})$ is a solution for problem (56). $\qquad \square$

Lemma 4. *Let $X = X' \times X''$ where X' is a compact metric space, X'' is a closed bounded convex set in a reflexive Banach space $\mathcal{X}''$, and let Y be a compact metric space. Let $J :*

$X \times Y \to \mathbb{R} \cup \{+\infty\}$ *be a lower semicontinuous function on the topological product space* $X' \times (X'', s) \times Y$, *where s denotes the topology on X'' induced by the strong topology of $\mathcal{X}''$.* *Suppose that $J(x', \cdot, y)$ is convex for each fixed $(x', y) \in X' \times Y$.*

Then the hypotheses of Lemma 3 are fulfilled, and argmin $J(\cdot, \cdot, \cdot) \neq \emptyset$.

Proof. 1. From Banach-Alaoglu-Kakutani theorem, X'' is compact for the weak topology of $\mathcal{X}''$ denoted w. Thus $X \times Y = (X' \times X'') \times Y$ is compact in the topological product space $[X' \times (\mathcal{X}'', w)] \times Y$. Let us show that J is sequentially lower semicontinuous on $[X' \times (X'', w_{X''})] \times Y$, where $w_{X''}$ stands for the topology on X'' induced by the weak topology of $\mathcal{X}''$. Indeed, for any real α, let us denote

$$SL_\alpha = \{(x', x'', y) \in X' \times X'' \times Y | J(x', x'', y) \leq \alpha\}.$$

Since J is lower semicontinuous on $X' \times (X'', s) \times Y$ we have that SL_α is closed in $X' \times (X'', s) \times Y$. Consider now a sequence $((x'_k, x''_k, y_k))_k$ in SL_α convergent to some (x', x'', y) in $X' \times (\mathcal{X}'', w) \times Y$. Since (x''_k) converges weakly to x'', by Mazur's lemma [32, page 6], there is a sequence $(\bar{x}''_k)$ converging to x'' in (X'', s) such that, for any k, $\bar{x}''_k$ is a convex combination of x''_k's. Then, by the convexity of X'' and of $J(x'_k, \cdot, y_k)$, we have $\bar{x}''_k \in X''$ and

$$J(x'_k, \bar{x}''_k, y_k) \leq J(x'_k, x''_k, y_k) \leq \alpha.$$

Thus $(x'_k, \bar{x}''_k, y_k) \in SL_\alpha$ and $(x'_k, \bar{x}''_k, y_k)$ converges to (x', x'', y) in $X' \times (X'', s) \times Y$, hence $(x', x'', y) \in SL_\alpha$. Therefore SL_α is sequentially closed in $X' \times (\mathcal{X}'', w) \times Y$, hence J is sequentially lower semicontinuous on $X' \times (\mathcal{X}'', w) \times Y$. Finally, by Weierstrass' theorem we obtain that argmin $J(\cdot, \cdot, \cdot) \neq \emptyset$.

Let now $x = (x', x'') \in X = X' \times X''$ be fixed. Since Y is compact and $J(x, \cdot)$ is lower semicontinuous on Y, we obtain from Weierstrass' theorem that argmin $J(x, \cdot) \neq \emptyset$. $\square$

Let $\hat{J}_l : \Omega \times]t_0, T[\times \mathcal{U}_l \to \mathbb{R} \cup \{+\infty\}$ be defined by

$$\hat{J}_l(\theta, t_1, u_l) := \tilde{J}_l(t_1, u_l, \mathbf{u_f}(\theta, t_1, u_l, \cdot)) = J_l(t_1, u_l, \mathbf{u_f}(\theta, t_1, u_l, \cdot), x(\theta, t_1, u_l, \cdot)). \qquad (58)$$

Theorem 6. *In addition to hypotheses $(\mathcal{PA})$ we suppose that, for each $t \in [t_0, T]$, $f_l(t, \cdot, \cdot, \cdot)$ is a convex function.*

Moreover we suppose the following hypothesis

$$(Hf) \quad \begin{cases} \text{there is some } \alpha \in L_\infty([t_0, T]) \text{ and some real constant } \beta \text{ such that,} \\ \text{for almost all } t \in [t_0, T], \text{ and for all } (u_l, \mathbf{u_f}, x) \in \mathbb{R}^{m_l} \times \mathbb{R}^{m_f} \times \mathbb{R}^n, \\ \left| \nabla_{(u_l, \mathbf{u_f}, x)} f_l(t, u_l, \mathbf{u_f}, x) \right| \leq \alpha(t) + \beta |(u_l, \mathbf{u_f}, x)|. \end{cases} \quad (59)$$

Then problem $(OB)_{we}$ has at least one solution and it is equivalent to the problem

$$(P_l) \qquad \min_{(\theta, t_1, u_l) \in \Theta_{we} \times \mathcal{T} \times \mathcal{U}_l} \hat{J}_l(\theta, t_1, u_l).$$

Proof. We will show that all the hypotheses of Lemma 4 are fulfilled (denoting $X' = \mathcal{T}$, $X'' = \mathcal{U}_l$, $Y = \Theta_{we}$, $\mathcal{X}'' = L_2^{m_l}([t_0, T])$, $x' = t_1$, $x'' = u_l$, $y = \theta$, $J(x', x'', y) = \hat{J}_l(\theta, t_1, u_l))$, and then the conclusion follows from Lemma 3.

$\mathcal{U}_l$ is (strongly) closed, bounded and convex in $L_2^{m_l}([t_0, T])$, $\mathcal{T}$ and Θ_{we} are compact. For fixed $(t_1, \theta) \in \mathcal{T} \times \Theta_{we}$, the function $\hat{J}_l(\theta, \cdot, t_1)$ is convex since, for any $t \in [t_0, T]$, the function $f_l(t, \cdot, \cdot, \cdot)$ is convex, and $u_l \mapsto \mathbf{u_f}(\theta, t_1, u_l, \cdot)$, $u_l \mapsto x(\theta, t_1, u_l, \cdot)$ are affine functions by Proposition 8.

To finish the proof it is sufficient to show that $\hat{J}_l$ is lower semicontinuous on $\Theta_{we} \times \mathcal{T} \times \mathcal{U}_l$, where $\mathcal{U}_l$ is endowed with the topology induced by the strong topology of $L_2^{m_l}([t_0, T])$. Let $(\theta^k, t_1^k, u_l^k)_k$ be a sequence in $\Theta_{we} \times \mathcal{T} \times \mathcal{U}_l$ which converges (strongly) to an element $(\bar{\theta}, \bar{t}_1, \bar{u}_l)$. Since $\Theta_{we} \times \mathcal{T} \times \mathcal{U}_l$ is closed we have $(\bar{\theta}, \bar{t}_1, \bar{u}_l) \in \Theta_{we} \times \mathcal{T} \times \mathcal{U}_l$.

We obtain from Lemma 1, Theorem 5 and (58) that, for each fixed $t_1 \in \mathcal{T}$, the function $\hat{J}_l(\cdot, t_1, \cdot)$ is lower semicontinuous. On the other hand we have

$$\hat{J}_l(\theta^k, t_1^k, u_l^k) = \hat{J}_l(\theta^k, \bar{t}_1, u_l^k) + (\hat{J}_l(\theta^k, t_1^k, u_l^k) - \hat{J}_l(\theta^k, \bar{t}_1, u_l^k)),$$

and the term $(\hat{J}_l(\theta^k, t_1^k, u_l^k) - \hat{J}_l(\theta^k, \bar{t}_1, u_l^k))$ tends to 0 as $k \to +\infty$. Indeed,

$$\hat{J}_l(\theta^k, t_1^k, u_l^k) - \hat{J}_l(\theta^k, \bar{t}_1, u_l^k) = \int_{t_0}^{t_1^k} f_l(t, u_l^k(t), \mathbf{u_f}(\theta^k, t_1^k, u_l^k, t), x(\theta^k, t_1^k, u_l^k, t))dt -$$
$$\int_{t_0}^{\bar{t}_1} f_l(t, u_l^k(t), \mathbf{u_f}(\theta^k, \bar{t}_1, u_l^k, t), x(\theta^k, \bar{t}_1, u_l^k, t))dt, \quad (60)$$

Since the sequence (u_l^k) is bounded in $L_2^{m_l}([t_0, T])$, by (Hf) and Theorem 5 there is a constant $M > 0$, such that, for all $k \in \mathbb{N}$ and almost all $t \in [t_0, T]$,

$$|f_l(t, u_l^k(t), \mathbf{u_f}(\theta^k, t_1^k, u_l^k, t), x(\theta^k, t_1^k, u_l^k, t))| \leq M,$$

and

$$|f_l(t, u_l^k(t), \mathbf{u_f}(\theta^k, \bar{t}_1, u_l^k, t), x(\theta^k, \bar{t}_1, u_l^k, t))| \leq M.$$

Finally, let us show that both integrals in (60) have the same limit as $k \to +\infty$, which is $\int_{t_0}^{\bar{t}_1} f_l(t, \bar{u}_l(t), \mathbf{u_f}(\bar{\theta}, \bar{t}_1, \bar{u}_l, t), x(\bar{\theta}, \bar{t}_1, \bar{u}_l, t))dt$. To do this it is sufficient to prove that these convergences hold for a subsequence. Since (u_l^k) converges in $L_2^{m_l}([t_0, T])$, there exists a subsequence $(u_l^{k'})_{k'}$, such that $(u_l^{k'}(t))_{k'}$ converges to $\bar{u}_l(t)$ a.e. on $[t_0, T]$. Then, we can apply Lebesgue's dominated convergence theorem to obtain the last claim.

Therefore, using the fact that for each $t_1 \in \mathcal{T}$ the function $\hat{J}_l(\cdot, t_1, \cdot)$ is lower semicontinuous, we obtain

$$\lim_{k \to +\infty} \hat{J}_l(\theta^k, t_1^k, u_l^k) = \lim_{k \to +\infty} \hat{J}_l(\theta^k, \bar{t}_1, u_l^k) \geq \hat{J}_l(\bar{\theta}, \bar{t}_1, \bar{u}_l).$$

$\square$

We denote $(f_l)'_{u_l}(\cdot, \cdot, \cdot, \cdot) : [t_0, T] \times \mathbb{R}^{m_l} \times \mathbb{R}^{m_f} \times \mathbb{R}^n \to \mathbb{R}^{m_l}$, $(f_l)'_{\mathbf{u_f}}(\cdot, \cdot, \cdot, \cdot) : [t_0, T] \times \mathbb{R}^{m_l} \times \mathbb{R}^{m_f} \times \mathbb{R}^n \to \mathbb{R}^{m_f}$, $(f_l)'_x(\cdot, \cdot, \cdot, \cdot) : [t_0, T] \times \mathbb{R}^{m_l} \times \mathbb{R}^{m_f} \times \mathbb{R}^n \to \mathbb{R}^n$ the partial derivatives of f_l with respect to the variables located on the second, third and fourth position respectively.

Also, let us denote for all $(\theta, t_1, t, s) \in \Omega \times]t_0, T[\times [t_0, T] \times [t_0, T]$,

$$X(\theta, t_1, t, s) = \Big[\chi_{[t_0, t_1]}(s) \Psi_{12}(\theta, t, t_0) \Xi(\theta, t_1, t_0, s) P(\theta, t_1, s) \tag{61}$$

$$+ \chi_{[t_0, t]}(s) \Psi_{11}(\theta, t, s) \Big] B_l(s)$$

$$Y(\theta, t_1, t, s) = -R^{-1}(\theta, t) \mathbf{B_f}(t)^T \Big[\chi_{[t_0, t_1]}(s) \Psi_{22}(\theta, t, t_0) \Xi(\theta, t_1, t_0, s) P(\theta, t_1, s) \tag{62}$$

$$\text{ï?¿}\frac{1}{2} \qquad\qquad + \chi_{[t_0, t]}(s) \Psi_{21}(\theta, t, s) \Big] B_l(s),$$

where $\chi_{[t_0, t]} : [t_0, T] \to \mathbb{R}$ is the characteristic function

$$\chi_{[t_0, t]}(s) = \begin{cases} 1 & \text{if } s \in [t_0, t] \\ 0 & \text{otherwise.} \end{cases} \tag{63}$$

Thus, formulas (52), (54) become

$$\frac{\partial}{\partial u_l} x(\theta, t_1, u_l, \cdot) \cdot h = \int_{t_0}^{T} X(\theta, t_1, \cdot, s) h(s) ds \tag{64}$$

$$\frac{\partial}{\partial u_l} \mathbf{u_f}(\theta, t_1, u_l, \cdot) \cdot h = \int_{t_0}^{T} Y(\theta, t_1, \cdot, s) h(s) ds. \tag{65}$$

Next result is necessary to ensure the differentiability of $\hat{J}_l$.

Lemma 5. *Suppose that f_l satisfies the hypothesis (Hf) given in Theorem 6, in addition to the hypothesis $(\mathcal{PA})$. Then, for each fixed $t_1 \in]t_0, T[$, the functional $\hat{J}_l(\cdot, t_1, \cdot) : \Omega \times L_2^{m_l}([t_0, T]) \to \mathbb{R}$ is well defined and continuously Fréchet differentiable. Its partial derivatives with respect to θ_i, $i = 1, \ldots, p$ are given by*

$$\frac{\partial \hat{J}_l}{\partial \theta_i}(\theta, t_1, u_l) = \int_{t_0}^{t_1} (f_l)'_{\mathbf{u_f}}(t, u_l(t), \mathbf{u_f}(\theta, t_1, u_l, t), x(\theta, t_1, u_l, t))^T \frac{\partial \mathbf{u_f}}{\partial \theta_i}(\theta, t_1, u_l, t) dt \tag{66}$$

$$+ \int_{t_0}^{t_1} (f_l)'_x(t, u_l(t), \mathbf{u_f}(\theta, t_1, u_l, t), x(\theta, t_1, u_l, t))^T \frac{\partial x}{\partial \theta_i}(\theta, t_1, u_l, t) dt.$$

Its partial Fréchet gradient with respect to u_l at (θ, t_1, u_l) is given, for almost all $s \in [t_0, t_1]$, by[iv]

$$\nabla_{u_l} \hat{J}_l(\theta, t_1, u_l)(s) = (f_l)'_{u_l}(s, u_l(s), \mathbf{u_f}(\theta, t_1, u_l, s), x(\theta, t_1, u_l, s)) \tag{67}$$

$$+ \int_{t_0}^{T} L^T(\theta, t_1, t, s)(f_l)'_{\mathbf{u_f}}(t, u_l(t), \mathbf{u_f}(\theta, t_1, u_l, t), x(\theta, t_1, u_l, t)) dt$$

$$+ \int_{t_0}^{T} X^T(\theta, t_1, t, s)(f_l)'_x(t, u_l(t), \mathbf{u_f}(\theta, t_1, u_l, t), x(\theta, t_1, u_l, t)) dt$$

[iv] We identify the Hilbert space $L_2^{m_l}([t_0, T])$ with its dual according to Riesz-Fréchet theorem, hence $\nabla_{u_l} \hat{J}_l(\theta, t_1, u_l) \in L_2^{m_l}([t_0, T])$ (see e.g. [7, page 38]).

Moreover, for each fixed $(\theta, u_l) \in \Omega \times L_2^{m_l}([t_0, T])$, the function $\hat{J}_l(\theta, \cdot, u_l) \in H_1([t_0, T])$, and for almost all $t_1 \in]t_0, T[$, its derivative is given by

$$\frac{\partial \hat{J}_l}{\partial t_1}(\theta, t_1, u_l) = f_l(t_1, u_l(t_1), \mathbf{u_f}(\theta, t_1, u_l, t_1), x(\theta, t_1, u_l, t_1)) \tag{68}$$

$$+ \int_{t_0}^{t_1} (f_l)'_{\mathbf{u_f}}(t, u_l(t), \mathbf{u_f}(\theta, t_1, u_l, t), x(\theta, t_1, u_l, t))^T \frac{\partial \mathbf{u_f}}{\partial t_1}(\theta, t_1, u_l, t) dt$$

$$+ \int_{t_0}^{t_1} (f_l)'_x(t, u_l(t), \mathbf{u_f}(\theta, t_1, u_l, t), x(\theta, t_1, u_l, t))^T \frac{\partial x}{\partial t_1}(\theta, t_1, u_l, t) dt$$

In particular, at each point t_1 such that u_l is continuous at t_1 (see footnote (iii)), the real valued function $t \mapsto \hat{J}_l(\theta, t, u_l)$ is differentiable.

Proof. By [4, Example 2, page 20] we have that the functional $J_l(t_1, \cdot, \cdot, \cdot) : L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T]) \times H_1^n([t_0, T]) \to \mathbb{R}$ is well defined and is continuously Fréchet differentiable for each fixed $t_1 \in]t_0, T[$. Moreover, its partial derivatives satisfy for all $(t_1, u_l, \mathbf{u_f}, x) \in]t_0, T[\times L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T]) \times H_1^n([t_0, T])$, the following equations

$$\frac{\partial J_l}{\partial u_l}(t_1, u_l, \mathbf{u_f}, x) \cdot v = \int_{t_0}^{t_1} (f_l)'_{u_l}(t, u_l(t), \mathbf{u_f}(t), x(t))^T v(t) dt \quad \forall v \in L_2^{m_l}([t_0, T])$$

$$\frac{\partial J_l}{\partial \mathbf{u_f}}(t_1, u_l, \mathbf{u_f}, x) \cdot w = \int_{t_0}^{t_1} (f_l)'_{\mathbf{u_f}}(t, u_l(t), \mathbf{u_f}(t), x(t))^T w(t) dt \quad \forall w \in L_2^{m_f}([t_0, T])$$

$$\frac{\partial J_l}{\partial x}(t_1, u_l, \mathbf{u_f}, x) \cdot z = \int_{t_0}^{t_1} (f_l)'_x(t, u_l(t), \mathbf{u_f}(t), x(t))^T z(t) dt \quad \forall z \in H_1^n([t_0, T]),$$

Also, for each fixed $(u_l, \mathbf{u_f}, x) \in L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T]) \times H_1^n([t_0, T])$ and for almost all $t_1 \in]t_0, T]$,

$$\frac{\partial J_l}{\partial t_1}(t_1, u_l, \mathbf{u_f}, x) = f_l(t_1, u_l(t_1), \mathbf{u_f}(t_1), x(t_1))$$

Let us identify, using Riesz-Fréchet theorem, the Hilbert spaces $L_2^{m_l}([t_0, T])$, $L_2^{m_f}([t_0, T])$ and $L_2^n([t_0, T])$ with their duals, and do not identify $H_1^n([t_0, T])$ with its dual $H_1^n([t_0, T])^*$. Based on the fact that (see [21, pp. 81-82] for details)

$$H_1^n([t_0, T]) \subset L_2^n([t_0, T]) \equiv L_2^n([t_0, T])^* \subset H_1^n([t_0, T])^*$$

and both embeddings are continuous and dense, and the duality product between $H_1^n([t_0, T])$ and $H_1^n([t_0, T])^*$ coincide with the inner product in $L_2^n([t_0, T])$ on $H_1^n([t_0, T]) \times L_2^n([t_0, T])$, we have that the Fréchet gradients $\nabla_{u_l} J_l(t_1, u_l, \mathbf{u_f}, x) \in L_2^{m_l}([t_0, T])$, $\nabla_{\mathbf{u_f}} J_l(t_1, u_l, \mathbf{u_f}, x) \in$

$L_2^{m_f}([t_0, T])$ and $\nabla_x J_l(t_1, u_l, \mathbf{u_f}, x) \in L_2^n([t_0, T])$ are given for almost all $t \in [t_0, T]$ by

$$\nabla_{u_l} J_l(t_1, u_l, \mathbf{u_f}, x)(t) = \begin{cases} (f_l)'_{u_l}(t, u_l(t), \mathbf{u_f}(t), x(t)) & \text{if } t \in [t_0, t_1] \\ \\ 0 & \text{if } t \in]t_1, T] \end{cases}$$

$$\nabla_{\mathbf{u_f}} J_l(t_1, u_l, \mathbf{u_f}, x)(t) = \begin{cases} (f_l)'_{\mathbf{u_f}}(t, u_l(t), \mathbf{u_f}(t), x(t)) & \text{if } t \in [t_0, t_1] \\ \\ 0 & \text{if } t \in]t_1, T] \end{cases}$$

$$\nabla_x J_l(t_1, u_l, \mathbf{u_f}, x)(t) = \begin{cases} (f_l)'_x(t, u_l(t), \mathbf{u_f}(t), x(t)) & \text{if } t \in [t_0, t_1] \\ \\ 0 & \text{if } t \in]t_1, T] \end{cases}$$

Now, using the chain rule in (58), we obtain immediately (66) and (68), and also

$$\nabla_{u_l} \hat{J}_l(\theta, t_1, u_l)(t) = (f_l)'_{u_l}(t, u_l(t), \mathbf{u_f}(\theta, t_1, u_l, t), x(\theta, t_1, u_l, t)) \tag{69}$$
$$+ \left(\frac{\partial}{\partial u_l} \mathbf{u_f}(\theta, t_1, u_l, \cdot) \right)^* (f_l)'_{\mathbf{u_f}}(t, u_l(t), \mathbf{u_f}(\theta, t_1, u_l, t), x(\theta, t_1, u_l, t))$$
$$+ \left(\frac{\partial}{\partial u_l} x(\theta, t_1, u_l, \cdot) \right)^* (f_l)'_x(t, u_l(t), \mathbf{u_f}(\theta, t_1, u_l, t), x(\theta, t_1, u_l, t)),$$

and, for almost all $t \in]t_1, T]$, $\nabla_{u_l} \hat{J}_l(\theta, t_1, u_l)(t) = 0$, where M^* stands for the adjoint operator of a linear continuous operator M between two Hilbert spaces.

Fix $(\theta, t_1, u_l) \in \Omega \times]t_0, T[\times L_2^{m_l}([t_0, T])$. Since the embedding $H_1^n([t_0, T]) \subset L_2^n([t_0, T])$ is continuous, we can consider the partial Fréchet derivative $\dfrac{\partial}{\partial u_l} x(\theta, t_t, u_l, \cdot)$ as a linear continuous operator from $L_2^{m_l}([t_0, T])$ to $L_2^n([t_0, T])$. Denote $\langle \cdot, \cdot \rangle_n$ the inner product in $L_2^n([t_0, T])$. For all $h \in L_2^{m_l}([t_0, T])$, $k \in L_2^n([t_0, T])$ we have

$$\langle \frac{\partial}{\partial u_l} x(\theta, t_t, u_l, \cdot) h, k \rangle_n = \int_{t_0}^T k^T(t) \left(\int_{t_0}^T X(\theta, t_1, t, s) h(s) ds \right) dt$$
$$= \int_{t_0}^T h^T(s) \left(\int_{t_0}^T X^T(\theta, t_1, t, s) k(t) dt \right) ds$$
$$= \langle h, \left(\frac{\partial}{\partial u_l} x(\theta, t_t, u_l, \cdot) \right)^* k \rangle_{m_l},$$

hence

$$\left(\frac{\partial}{\partial u_l} x(\theta, t_t, u_l, \cdot) \right)^* \cdot k = \int_{t_0}^T X^T(\theta, t_1, t, \cdot) k(t) dt. \tag{70}$$

In the same way we get for all $k \in L_2^{m_f}([t_0, T])$,

$$\left(\frac{\partial}{\partial u_l} \mathbf{u_f}(\theta, t_t, u_l, \cdot) \right)^* \cdot k = \int_{t_0}^T Y^T(\theta, t_1, t, \cdot) k(t) dt. \tag{71}$$

24

Finally (67) follows from (69).

$\square$

Theorem 7. (FIRST ORDER NECESSARY CONDITIONS WHEN THE FINAL TIME IS FIXED, I.E. $\mathcal{T} = \{t_1\}$) *Suppose that $\mathcal{T} = \{t_1\}$, and f_l satisfies hypotheses ($\mathcal{PA}$), (Hf), and $f_l(t, \cdot, \cdot, \cdot)$ is convex for all $t \in [t_0, T]$.*

Let $(\bar{\theta}, \bar{u}_l) \in \Theta_{we} \times \mathcal{U}_l$ solve $(OB)_{we}$. Then there are nonnegative real numbers $\mu, l_1, \ldots, l_p$ and a real number ν such that

$$\nabla_{u_l} \hat{J}_l(t_1, \bar{\theta}, \bar{u}_l)(t) + \mu \bar{u}_l(t) = 0 \qquad \text{a.e. on } [t_0, T], \tag{72}$$

$$\frac{\partial \hat{J}_l}{\partial \theta_i}(t_1, \bar{\theta}, \bar{u}_l) - l_i + \nu = 0, \qquad i = 1, \ldots, p, \tag{73}$$

$$\mu(\|\bar{u}_l\|_2 - R) = 0, \tag{74}$$

$$l_i \bar{\theta}_i = 0 \qquad i = 1, \ldots, p, \tag{75}$$

and of course

$$\sum_{i=1}^{p} \bar{\theta}_i = 1. \tag{76}$$

$$\|\bar{u}_l\|_2 \le R, \quad \bar{\theta}_i \ge 0 \qquad i = 1, \ldots, p. \tag{77}$$

REMARK 5. According to (67), equation (72) is a Fredholm integral equation in the unknown $\bar{u}_l$ (linear if $f_l(t, \cdot, \cdot, \cdot)$ is quadratic, case which satisfies hypothesis (Hf)), depending on $2p+1$ parameters (μ and $\bar{\theta}_i$). Assuming that we are able to solve this integral equation, (73-76) represent an nonlinear system with $2p + 2$ equations and $2p + 2$ unknowns μ, ν, θ_i, l_i.

A similar remark applies to the next theorem.

Theorem 8. (FIRST ORDER NECESSARY CONDITIONS WHEN THE FINAL TIME $t_1 \in \mathcal{T} = [\underline{t}, \bar{t}] \subset]t_0, T[$) *Suppose that f_l satisfies hypotheses ($\mathcal{PA}$), (Hf) and $f_l(t, \cdot, \cdot, \cdot)$ is convex for all $t \in [t_0, T]$.*

Let $(\bar{t}_1, \bar{\theta}, \bar{u}_l) \in \mathcal{T} \times \Theta_{we} \times \mathcal{U}_l$ solve $(OB)_{we}$. Suppose that $\bar{u}_l$ is continuous at $\bar{t}_1$ (see footnote (iii)). Then there are nonnegative real numbers $\mu, l_1, \ldots, l_p, l_{p+1}, l_{p+2}$ and a real number ν such that

$$\nabla_{u_l} \hat{J}_l(\bar{t}_1, \bar{\theta}, \bar{u}_l)(t) + \mu \bar{u}_l(t) = 0 \qquad \text{a.e. on } [t_0, T], \tag{78}$$

$$\frac{\partial \hat{J}_l}{\partial \theta_i}(\bar{t}_1, \bar{\theta}, \bar{u}_l) - l_i + \nu = 0, \qquad i = 1, \ldots, p, \tag{79}$$

$$\frac{\partial \hat{J}_l}{\partial t_1}(\bar{t}_1, \bar{\theta}, \bar{u}_l) - l_{p+1} + l_{p+2} = 0, \tag{80}$$

$$\mu(\|\bar{u}_l\|_2 - R) = 0, \tag{81}$$

$$l_i \bar{\theta}_i = 0 \qquad i = 1, \ldots, p, \tag{82}$$

$$l_{p+1}(\bar{t}_1 - \underline{t}) = 0, \tag{83}$$

$$l_{p+2}(\bar{t} - \bar{t}_1) = 0, \tag{84}$$

25

and of course

$$\sum_{i=1}^{p} \bar{\theta}_i = 1. \tag{85}$$

$$\|\bar{u}_l\|_2 \leq R, \quad \bar{\theta}_i \geq 0 \qquad i = 1, \dots, p. \tag{86}$$

The proof of Theorems 7 and 8 is a direct application of the generalized Lagrange multiplier rule under Kurcyusz-Robinson-Zowe regularity condition (see [39, Theorem 5.3]), and is based on Theorem 6 and on Lemma 5.

6.2 The pessimistic bilevel problem

In this section we assume that $f_l(t, \cdot, \cdot, \cdot)$ is quadratic, i.e. for all $(t, u_l, \mathbf{u_f}, x) \in [t_0, T] \times \mathbb{R}^{m_l} \times \mathbb{R}^{m_f} \times \mathbb{R}^n$,

$$f_l(t, u_l, \mathbf{u_f}, x) = u_l^T S_l(t) u_l + \mathbf{u_f}^T R_l(t) \mathbf{u_f} + x^T Q_l(t) x, \tag{87}$$

where $S_l(\cdot), R_l(\cdot), Q_l(\cdot)$ are continuous symmetric matrix valued functions. Note that this function satisfies hypotheses $(\mathcal{PA})$ and (Hf).

According to [4, Example 3, page 14] the functional $J_l(t_1, \cdot, \cdot, \cdot) : L_2^{m_l}([t_0, T]) \times L_2^{m_f}([t_0, T]) \times H_1^n([t_0, T]) \times$ is well defined and continuous. Therefore, by Theorem 5, the functional $\hat{J}_l(\cdot, \cdot, \cdot)$ has finite values and is continuous on $\Theta_{we} \times \mathcal{T} \times \mathcal{U}_l$.

Moreover, since Θ_{we} is compact, the pessimistic problem $(PB)_{we}$ can be written as

$$\min_{(t_1, u_l) \in \mathcal{T} \times \mathcal{U}_l} \max_{\theta \in \Theta_{we}} \hat{J}_l(\theta, t_1, u_l)$$

Theorem 9. (FIRST ORDER NECESSARY CONDITIONS WHEN THE FINAL TIME IS FIXED, I.E. $\mathcal{T} = \{t_1\}$) *Suppose that* $\mathcal{T} = \{t_1\}$.

Let $(\bar{\theta}, \bar{u}_l) \in \Theta_{we} \times \mathcal{U}_l$ *solve* $(PB)_{we}$. *Then there are nonnegative real numbers* $\mu, l_1, \dots, l_p$ *and a real number* ν *such that*

$$\nabla_{u_l} \hat{J}_l(t_1, \bar{\theta}, \bar{u}_l)(t) + \mu \bar{u}_l(t) = 0 \qquad a.e. \ on \ [t_0, T], \tag{88}$$

$$\frac{\partial \hat{J}_l}{\partial \theta_i}(t_1, \bar{\theta}, \bar{u}_l) + l_i + \nu = 0, \qquad i = 1, \dots, p, \tag{89}$$

$$\mu(\|\bar{u}_l\|_2 - R) = 0, \tag{90}$$

$$l_i \bar{\theta}_i = 0 \qquad i = 1, \dots, p, \tag{91}$$

and of course

$$\sum_{i=1}^{p} \bar{\theta}_i = 1. \tag{92}$$

$$\|\bar{u}_l\|_2 \leq R, \quad \bar{\theta}_i \geq 0 \qquad i = 1, \dots, p. \tag{93}$$

Proof. We have that $\bar{\theta}$ is a maximizer of $\hat{J}_l(\cdot, t_1, \bar{u}_l)$ over Θ_{we}. By Karush-Kuhn-Tucker theorem, since on Θ_{we} the linear independence of gradients of active constraints holds (hence Mangasarian-Fromowitz regularity condition holds), and based on Lemma 5, we obtain that there are nonnegative reals $l_1, \ldots, l_p$, and a real ν such that (89), (91) hold, and of course (92, 93).

Moreover, $\bar{u}_l$ is a minimizer of $\hat{J}_l(\bar{\theta}, t_1, \cdot)$ over the ball $\mathcal{U}_l$. By the generalized Lagrange multiplier rule under Kurcyusz-Robinson-Zowe regularity condition (see [39, Theorem 5.3]), and based on Lemma 5, we obtain (88), (90). $\qquad\square$

Theorem 10. (FIRST ORDER NECESSARY CONDITIONS WHEN THE FINAL TIME $t_1 \in \mathcal{T} = [\underline{t}, \bar{t}] \subset]t_0, T[$) *Let $(\bar{t}_1, \bar{\theta}, \bar{u}_l) \in \mathcal{T} \times \Theta_{we} \times \mathcal{U}_l$ solve $(PB)_{we}$. Suppose that $\bar{u}_l$ is continuous at $\bar{t}_1$ (see footnote (iii)). Then there are nonnegative real numbers $\mu, l_1, \ldots, l_p, l_{p+1}, l_{p+2}$ and a real number ν such that*

$$\nabla_{u_l} \hat{J}_l(\bar{t}_1, \bar{\theta}, \bar{u}_l)(t) + \mu \bar{u}_l(t) = 0 \qquad a.e. \ on \ [t_0, T], \tag{94}$$

$$\frac{\partial \hat{J}_l}{\partial \theta_i}(\bar{t}_1, \bar{\theta}, \bar{u}_l) + l_i + \nu = 0, \qquad i = 1, \ldots, p, \tag{95}$$

$$\frac{\partial \hat{J}_l}{\partial t_1}(\bar{t}_1, \bar{\theta}, \bar{u}_l) - l_{p+1} + l_{p+2} = 0, \tag{96}$$

$$\mu(\|\bar{u}_l\|_2 - R) = 0, \tag{97}$$

$$l_i \bar{\theta}_i = 0 \qquad i = 1, \ldots, p, \tag{98}$$

$$l_{p+1}(\bar{t}_1 - \underline{t}) = 0, \tag{99}$$

$$l_{p+2}(\bar{t} - \bar{t}_1) = 0, \tag{100}$$

and of course

$$\sum_{i=1}^{p} \bar{\theta}_i = 1. \tag{101}$$

$$\|\bar{u}_l\|_2 \leq R, \quad \bar{\theta}_i \geq 0 \qquad i = 1, \ldots, p. \tag{102}$$

The proof is identical to the proof of Theorem 9.

REMARK 6. A similar comment as in Remark 5 can be done for the last two theorem. Moreover, in this case the computation of the partial derivatives and gradients in Lemma 5 is simplified since, by (87), we have

$$(f_l)'_{u_l}(t, u_l, \mathbf{u_f}, x) = 2u_l^T S_l(t), \quad (f_l)'_{\mathbf{u_f}}(t, u_l, \mathbf{u_f}, x) = 2\mathbf{u_f}^T R_l(t), \quad (f_l)'_x(t, u_l, \mathbf{u_f}, x) = x^T Q_l(t).$$

References

[1] H. ABOU-KANDIL, G. FREILING, V. IONESCU AND G. JANK, Matrix Riccati Equations in Control and Systems Theory, Birkhauser Verlag, Basel, 2003.

[2] V. M. ALEXEEV, V. M. TIKHOMIROV, AND S. V. FOMIN, Optimal Control, Plenum, New York, 1987.

[3] Z. Ankhili, A. Mansouri, An exact penalty on bilevel programs with linear vector optimization lower level, Eur. J. Oper. Res. 197, 36-41 (2009)

[4] J.-P. Aubin and I. Ekeland, Applied Nonlinear Analysis, John Wiley & Sons, New York, 1984.

[5] T.Basar and G.J.Olsder, Dynamic Noncooperative Game Theory, Academic Press, London/New York, January 1995. (Revised 2nd edition of 1982 book with the same title.)

[6] A.Bagchi, Stackelberg Differential Games in Economic Models, Lecture Notes in Control and Information Sciences, Vol. 64, Springer-Verlag, Berlin, (1984).

[7] H.H. Bauschke and P.L. Combettes, Convex Analysis and Monotone Operetor Theory in Hilbert Spaces, CMS Books in Mathematics, Springer, 2010.

[8] H.P. Benson, Optimization over the efficient set, J. Math. Anal.Appl., 98, (1984) pp. 562-580.

[9] H.P. Benson, A finite, non-adjacent extreme point search algorithm for optimization over the efficient set, J. Optim. Theory Appl., 73, (1992), pp. 47-64.

[10] S. Bolintinéanu, Optimality conditions for Minimization over the (Weakly or Properly) Efficient Set, J. Math. Anal. Appl., Vol 173, No. 2, (1993) pp. 523-541.

[11] S. Bolintinéanu, Minimization of a quasi-concave function over an efficient set, Math. Program., 61, (1993) pp. 89-110.

[12] S. Bolintinéanu, Necessary conditions for nonlinear suboptimization over the weakly-efficient set, J. Optim. Theory Appl., 78, (1993) pp. 579-598.

[13] S. Bolintinéanu and M. El Maghri, Pénalisation dans l'optimisation sur l'ensemble faiblement efficient, RAIRO Oper. Res., 31, No. 3, (1997) pp 295-310.

[14] H. Bonnel, Optimality Conditions for the Semivectorial Bilevel OptimizationProblem, Pacific Journal of Optimization, Vol. 2, No. 3, (2006) pp. 447-468.

[15] H. Bonnel and C.Y. Kaya, Optimization Over the Efficient Set in Multiobjective Convex Optimal Control Problems, J. Optim. Theory Appl., Vol. 147, No. 1., (2010), pp. 93-112.

[16] H. Bonnel, J. Morgan, Semivectorial Bilevel Optimization Problem: Penalty Approach, Journal of Optimization Theory and Applications, Vol. 131, No. 3, (2006) pp. 365-382.

[17] H. Bonnel, J. Morgan, Semivectorial Bilevel Convex Optimal Control Problems: Existence Results (submitted)

[18] H. Bonnel and N.S. Pham, Nonsmooth Optimization over the (Weakly or Properly) Pareto Set of a Linear-quadratic Multi-objective Control Problem: Explicit Optimality Conditions, Journal of Industrial and Management Optimization, Vol. 7, No. 4, (2011), pp. 789-809.

[19] M. Breton, A. Alj and A. Haurie, Sequential Stackelberg Equilibrium in Two-person Games, Journal of Optimization Theory and Applications, Vol. 59, (1988) pp. 71-97.

[20] J. Borwein, Proper efficient points for maximizations with respect to cones, SIAM J. Control Optimization 15 (1977), no. 1, 57-63.

[21] H. Brezis, Analyse fonctionnelle : théorie et applications, Dunod, Paris, 1999.

[22] H. Calvete and C. Galï¿½½, On linear bilevel problems with multiple objectives at the lower level, Omega 39 (2011) 33-40. Elsevier.

[23] T. Chen and J.B. Cruz, Jr., Stackelberg solution for two person games with biased information patterns, I.E.E.E. Trans.on Automatic Control, A-C-17, (1972), pp. 791-798.

[24] B. Colson, P. Marcotte and G. Savard, An overview of bilevel optimization, Ann. Op. Research, Vol. 153, (2007), 235-256.

[25] B.D. Craven, Aspects of multicriteria optimization, Recent Prospects in Mathematical Programming, Gordon and Breach, Philadelphia, 1991.

[26] J. P. Dauer, Optimization over the efficient set using an active constraint approach, Z. Oper. Res. 35 (1991) pp. 185-195.

[27] J.P. Dauer and T.A. Fosnaugh, Optimization over the efficient set, J. Global Optim., 7, (1995) pp. 261-277.

[28] S. Dempe, Foundations of Bilevel Programming, Kluwer Academic Publishers, Dordrecht, 2002.

[29] S. Dempe, Annotated Bibliography on Bilevel Programming and Mathematical Programs with Equilibrium Constraints Optimization, 52 (2003) pp. 333-359.

[30] S. Dempe, N. Gadhi, A.B. Zemkoho, New optimality conditions for the semivectorial bilevel optimization problem, (submitted)

[31] G. Eichfelder, Multiobjective bilevel optimization, Math. Program., Ser. A (2010) 123:419-449

[32] I. Ekeland, R. Témam, Convex Analysis and Variational Problems, Classics in Applied Mathematics, SIAM, 1999.

[33] J. Fülöp, A cutting plane algorithm for linear optimization over the efficient set, Generalized Convexity, Lecture notes in Economics and Mathematical System 405, Springer-Verlag, Berlin, (1994), pp. 374-385.

[34] A. M. Geoffrion, Proper efficiency and the theory of vector maximization. J. Math. Anal. Appl. 22 1968 618-630.

[35] A.Haurie, A Historical Perspective on Cooperative Differential Games, Annals of the International Society of Dynamic Games, 6, Part I, (2001), pp. 19-29.

[36] R. Horst and N.V. Thoai, Maximizing a concave function over the efficient or weakly-efficient set, European J. Oper. Res., 117, 1999, 239-252.

[37] R. Horst, N.V. Thoai, Y. Yamamoto and D. Zenke, On Optimization over the Efficient Set in Linear Multicriteria Programming, J. Optim. Theory Appl., vol. 134, (2007), pp. 433-443.

[38] J. Jahn, Vector Optimization, Springer-Verlag, Berlin, Heidelberg, New York, 2004.

[39] J. Jahn, Introduction to the Theory of Nonlinear Optimization, Springer-Verlag, Berlin, Heidelberg, New York, 2007.

[40] M. B. Lignola and J. Morgan, Topological Existence and Stability for Stackelberg Problems, J. Optim. Theory Appl. 84, (1995) pp. 575-596.

[41] M. B. Lignola and J. Morgan, Stability of regularized bilevel programming problems, J. Optim. Theory Appl. 93, (1997) pp. 575-596.

[42] P. Loridan and J. Morgan, New results on approximate solutions in two-level optimization, Optimization 20, (1989), pp. 819-836.

[43] P. Loridan and J. Morgan, A theoretical approximation scheme for Stackelberg problems, J. Optim. Theory Appl. 61, (1989), 95-110.

[44] P. Loridan and J. Morgan, Approximation of the Stackelberg Problem and Applications in Control Theory, Proceedings of the Vth I.F.A.C. Workshop on "Control Applications of Non Linear Programming and Optimization" (Capri, 1985), Pergamon Press (1986).

[45] J. Philip, Algorithms for the vector maximization problem, Math. Program., (1972) pp. 207-229.

[46] J. Morgan, Constrained well-posed two-level optimization problems, in: Nonsmooth Optimization and related topics (F. Clarke, V. F. Dem'yanov, and F. Giannessi, eds.), Ettore Majorana International Sciences Series, Plenum Press, New York (1989), pp. 307-326.

[47] J. Morgan, Existence for Hierarchical Differential Non-zero Sum Games with Coupled Constraints, Workshop of the International Society of Game Theory and Applications, Sils-Maria, Switzerland, (1997).

[48] M. Simaan and J.B. Cruz, Jr., On the Stackelberg Strategy in Nonzero-Sum Games, Journal of Optimization Theory and Applications, Vol. 11, No. 5, (1973), pp. 533-555.

[49] H. Von Stackelberg, The Theory of the Market Economy, Oxford University Press, Oxford, England, (1952).

[50] Y. Yamamoto, Optimization over the efficient set : overview, J. Global Optim., 22 (2002), pp. 285-317.

[51] Y.Zheng and Z. Wan, A solution method for semivectorial bilevel programming problem via penalty method, J. Appl. Math. Comput. 37 (2011), 207-219.