A CHARACTERISTIC METHOD FOR FULLY CONVEX BOLZA PROBLEMS OVER ARCS OF BOUNDED VARIATION*

CRISTOPHER HERMOSILLA† AND PETER WOLENSKI‡

Abstract. The aim of this paper is to study the value function of a Fully Convex Bolza problem with state constraints and under no coercivity assumptions. This requires that the state trajectories be of bounded variation rather than merely absolutely continuous. Our approach is based on the duality theory of classical convex analysis, and we establish a Fenchel-Young type equality between the value function of the Bolza problem and a suitable value function associated with its dual problem. The main result we present in this paper is a characteristic method that describes the evolution of the subgradients of the associated value functions. A few simple examples are provided to demonstrate the theory.

Key words. Fully convex Bolza problems, Convex value functions, Nonsmooth Hamiltonian systems, Extended method of characteristics, Impulsive systems.

AMS subject classifications. 49N15, 49N25, 49K15

1. Introduction. This paper is concerned with a duality approach to study problems of Bolza type and the corresponding value functions associated with them. Our emphasis is on Fully Convex Bolza (FCB) problems in which the Lagrangian (running cost) is an extended real-valued function jointly convex in the state and velocity variables, and whose end-point cost is also an extended real-valued convex function.

Rockafellar set the foundations of a duality theory for this type of problems in the 1970s with a series of papers [18, 21, 23, 24, 26]. In this approach, an FCB problem is paired with a dual problem which has the same structure as the primal. A sample result from the theory is that, under reasonable conditions, solutions of the adjoint equation for the primal problem are optimal solutions of the dual problem; see for instance [18, Section 10]. This in particular leads to optimality conditions, which in this case are also sufficient, and can be given in terms of a Hamiltonian system; cf. [18, 26, 28, 17]. Let us also mention that convex problems of Bolza type have been widely investigated in several contexts. For example [31, 32, 29, 8, 7, 10, 9, 12] are concerned with the study of the value function of FCB problems and its relation with the Hamilton-Jacobi equation. Some attention has also been put on Hamiltonian systems associated with FCB problems as for instance in [20, 11, 15].

It is worth mentioning that extended real-valued Lagrangians allow to implicitly encode in FCB problems dynamical and state constraints, and this in turn leads to formulating optimal control problems as FCB problems; we refer to [18, Section 4] for precise examples. Most notably, Linear Quadratic optimal control problems can be formulated as FCB problems (see [27, 13, 16]). This duality theory provides a framework to treat Linear Quadratic optimal control problems with state constraints and without coercivity assumptions in the control variable; this was pointed out in [16]. We also mention that this theory provides a framework to study convex optimal

*Submitted to the editors DATE.

Funding: C. Hermosilla was supported by CONICYT-Chile through FONDECYT grant number 3170485.

†Departamento de Matemática, Universidad Técnica Federico Santa María, Avenida España 1680, Valparaíso, Chile (cristopher.hermosill@usm.cl).

‡Department of Mathematics, Louisiana State University, Baton Rouge, LA70803-4918, USA (pwolens@lsu.edu).

control problems with mixed constraints; see [18, Example 2].

A global *characteristic method* for FCB problems was introduced by Rockafellar and the second author in [31, Theorem 2.4]. The characteristic method describes the time-evolution of the subgradients of the value functions through following trajectories of a Hamiltonian system. The cited result holds true for FCB problems under standard hypotheses where the Lagrangian is coercive and there are no state constraints implicitly encoded (see [31, Assumption (A) and Section 3]). This means the minimization in both the primal and dual problems takes places over the space of absolutely continuous arcs. The main purpose of the current work is to extend these results to FCB problems with state constraints and with a lack of coercivity.

The lack of coercivity is equivalent to the dual problem having nontrivial state constraints, and by symmetry, the converse is true as well. Hence (non)coercivity assumptions and (non)trivial state constraints are dual concepts in FCB problems. In optimal control, when state constraints are involved, it is expected that the adjoint arc will have jumps whenever the constraint is active, and so it is natural that adjoint arcs of Bounded Variation (**BV**) be considered; see for instance [23, 14, 33].

The works [23, 26, 17, 16] on FCB problems cited earlier are set in an impulsive framework. Notably, [26] provides the foundations of FCB problems for impulsive systems and introduces the fundamental tools (the so-called *fundamental kernels* presented in subsection 5.1) we use to prove our main results. A less restrictive assumption on the time-dependence was introduced in [17] and proved the lower semicontinuity of the integral functional (see (2.2)) of the **BV**-extension of FCB problems.

The fact that no coercivity assumptions are made in the current paper leads to Bolza problems whose trajectories are of bounded variation rather than merely absolutely continuous, and consequently, to a new notion of Hamiltonian system suitable for arcs of bounded variation. The definition of Hamiltonian system considered here closely resembles the one introduced in [26] and revisited in [17], however, our definition (see Definition 3.7) does not impose conditions on jumps at the initial and terminal times. This feature will be seen inherent to the way an FCB problem and its dual counterpart are extended to minimizing over arcs of bounded variation.

The **BV**-extension of the value functions of the primal and dual problems follows the ideas introduced in [26], and requires two different extensions. Namely, one that is finite only in the closure of the state constraint, and a second one that can be finite beyond that set. The latter corresponds to the case when, from an initial state, trajectories are allowed to jump instantaneously at the initial time into the closure of the state constraint, or at the terminal time can jump out of the state constraint.

The duality approach taken in the paper provides an intermediate result that has an interest by itself. We prove (see Theorem 3.6) that, for any of the impulsive extensions of the FCB problem, a Fenchel-Young type equality between the value function of the primal problem and a value function associated with its dual problem holds; this result can be interpreted as a strong duality theorem for FCB problems. A similar result for Linear Quadratic optimal control problems was reported in [16].

The development in this work will have time-dependent Lagrangians, and as such, this paper extends [31, Theorem 2.4] and [31, Theorem 5.1] to this context. The assumptions on time-dependence are essentially the same considered in [26], and it has not been our purpose here to go deeper into that issue. However, extensions to less restrictive hypotheses, such as the ones done in [17], might be possible.

1.1. Notation and essentials. Throughout this paper, $|\cdot|$ is the Euclidean norm and $a \cdot b$ stands for the Euclidean inner product of $a, b \in \mathbb{R}^n$. The indicator and support functions of $\mathbf{S} \subseteq \mathbb{R}^n$ are denoted by $\delta_{\mathbf{S}}$ and $\sigma_{\mathbf{S}}$, respectively. The normal cone to $\mathbf{S}$ at $x \in \mathbf{S}$ is

$$\mathcal{N}_{\mathbf{S}}(x) := \{z \in \mathbb{R}^n \mid \ z \cdot (s - x) \leq 0, \ \forall s \in \mathbf{S}\}.$$

Suppose $\varphi : \mathbb{X} \to \mathbb{R} \cup \{\pm\infty\}$ is a function with $\mathbb{X}$ being a topological vector space. The effective domain of φ is the set $\mathrm{dom}(\varphi) := \{x \in \mathbb{X} \mid \ \varphi(x) < +\infty\}$. Then φ is called *proper* if $\mathrm{dom}(\varphi) \neq \emptyset$ and $\varphi(x) > -\infty$ for all $x \in \mathbb{X}$; *convex* if $\mathrm{epi}(\varphi) := \{(x, r) \in \mathbb{X} \times \mathbb{R} \mid \ \varphi(x) \leq r\}$ is a convex set, and *lower semicontinuous (l.s.c. for short)* if $\mathrm{epi}(\varphi)$ is a closed set.

Suppose $\mathbb{X}$ is in duality with another topological vector space $\mathbb{Y}$ via a bilinear mapping $\langle \cdot, \cdot \rangle : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$; if $\mathbb{X} = \mathbb{R}^n$, then it is in duality with itself via the Euclidean inner product. The (Legendre-Fenchel) conjugate of $\varphi : \mathbb{X} \to \mathbb{R} \cup \{+\infty\}$ is the mapping $\varphi^* : \mathbb{Y} \to \mathbb{R} \cup \{\pm\infty\}$ defined via $\varphi^*(y) := \sup\{\langle x, y\rangle - \varphi(x) \mid \ x \in \mathbb{X}\}$ and its subdifferential at $x \in \mathrm{dom}(\varphi)$ is the set

$$\partial\varphi(x) := \{y \in \mathbb{Y} \mid \varphi(x) + \langle y, z - x\rangle \leq \varphi(z), \ \forall z \in \mathbb{X}\}.$$

These mathematical objects are related via the Fenchel-Young equality:

$$y \in \partial\varphi(x) \quad \Longleftrightarrow \quad \varphi(x) + \varphi^*(y) = \langle x, y\rangle. \tag{1.1}$$

If φ is convex proper and l.s.c., so is φ^*. Moreover $\varphi^{**} = (\varphi^*)^*$ agrees with φ and $y \in \partial\varphi(x)$ if and only if $x \in \partial\varphi^*(y)$.

A function $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is called concave-convex if $h_y(\cdot) = -h(\cdot, y)$ and $h_x(\cdot) = h(x, \cdot)$ are convex functions. The (concave-convex) subdifferential of h is

$$\partial h(x, y) := [-\partial h_y(x)] \times \partial h_x(y).$$

We suppose $T > 0$ is fixed. In our setting an arc is a function $x : [0, T] \to \mathbb{R}^n$. The space of continuous, absolutely continuous and bounded variation arcs are denoted by $\mathbf{C}$, $\mathbf{AC}$ and $\mathbf{BV}$, respectively. If $x \in \mathbf{BV}$ then $x(t^-)$ and $x(t^+)$ stand for the left and right limits of x at t, with the convention that $x(0) = x(0^-)$ and $x(T) = x(T^+)$. Furthermore, when $t = \tau \in [0, T]$ is understood from the context as the initial time, then we also use the convention $x(\tau^-) = x(\tau)$.

We write $\mathbf{L}^1$ for the (equivalence class of) Lebesgue integrable real-valued functions defined on $[0, T]$ endowed with the usual norm $\|\cdot\|_{\mathbf{L}^1}$. The product space $(\mathbf{L}^1)^n$ is denoted by $\mathbf{L}^1_n$ and the equivalence class of measurable $\mathbb{R}^n$-valued functions that are essentially bounded on $[0, T]$ is denoted by space $\mathbf{L}^\infty_n$, both spaces are endowed with their standard norms.

Given a measure μ on $[0, T]$, we say that a property holds $d\mu$-a.e. on $A \subseteq [0, T]$ if there is a Borel-measurable set $B \subseteq A$ such that $\mu(A \setminus B) = 0$ and the property holds for any $t \in B$. If μ is the Lebesgue measure we denote $d\mu(t) = dt$ and we simply say that the property holds a.e. on A.

A set-valued map $t \mapsto S(t) \subseteq \mathbb{R}^n$ defined on an interval $[a, b] \subseteq \mathbb{R}$ is said to be measurable if $\{t \in [a, b] \mid S(t) \cap O \neq \emptyset\}$ is a measurable set of $[a, b]$ for any open set $O \subseteq \mathbb{R}^n$. A set-valued map is also said to be upper semicontinuous (u.s.c. for short) if for any $t_0 \in [a, b]$ and any neighborhood O of $S(t_0)$ there is $r > 0$ such that $S(t) \subseteq O$ for any $t \in [a, b]$ such that $|t - t_0| < r$.

2. FCB problems. Given a Lagrangian $L : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and a terminal cost $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, consider the Bolza problem:

$$\text{(P)} \quad \text{Minimize} \int_0^T L(t, x(t), \dot{x}(t))dt + g(x(T)) \text{ over all } x \in \mathbf{AC} \text{ such that } x(0) = \xi_0.$$

Our task in this paper is to study an extension of the Bolza problem (P) to minimization over arcs of bounded variation and their corresponding associated value functions. We also want to explore duality relationships and pose the discussion in a fully convex context, meaning that we are mainly concerned with the case where $(x, v) \mapsto L(t, x, v)$ and $a \mapsto g(a)$ are proper convex l.s.c. functions. In order for the integral cost to be well-defined, we also impose conditions so that $t \mapsto L(t, x(t), \dot{x}(t))$ is measurable for any arc $x \in \mathbf{AC}$, and this will be done through the *epigraphical mapping* associated with the Lagrangian:

$$t \mapsto S_L(t) := \{(x, v, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \mid L(t, x, v) \leq z\}.$$

To be more precise, the paper will operate under the following hypothesis:

Hypothesis 2.1. (Standing convexity assumptions)

(i) g is a proper convex l.s.c. function.

(ii) $t \mapsto S_L(t)$ is measurable on $[0, T]$ with nonempty convex closed images.

(iii) $\exists \alpha \in \mathbf{L}^1, \beta \in \mathbf{L}^1_n$ and $\gamma \in \mathbf{L}^\infty_n$ such that $L(t, x, v) \geq \alpha(t) + \beta(t) \cdot x + \gamma(t) \cdot v$.

The fact that $S_L(t)$ is nonempty convex and closed is equivalent to $(x, v) \mapsto L(t, x, v)$ being proper, convex and l.s.c. for any $t \in [0, T]$ fixed. Also, the fact that $t \mapsto S_L(t)$ is measurable implies that $t \mapsto L(t, x, v)$ is measurable for any $x, v \in \mathbb{R}^n$ fixed. Moreover, $t \mapsto L(t, x(t), \dot{x}(t))$ is measurable for any arc $x \in \mathbf{AC}$; see for instance [30, Proposition 14.28]. Furthermore, by Hypothesis 2.1 we have that

$$\int_0^T L(t, x(t), \dot{x}(t))dt + g(x(T)) \geq -\|\alpha\|_{\mathbf{L}^1} - \|\beta\|_{\mathbf{L}^1_n} \|x\|_\infty - \|\gamma\|_{\mathbf{L}^\infty_n} \|\dot{x}\|_{\mathbf{L}^1_n}, \qquad \forall x \in \mathbf{AC}.$$

Thus, the integral cost is well-defined, although it may be $+\infty$ but never $-\infty$. Let us point out that in the literature, functions satisfying property (*ii*) in Hypothesis 2.1 are known as *(convex) normal integrands*. Also, in the autonomous case, that is, when L doesn't depend on the time variable, conditions (*ii*) and (*iii*) in Hypothesis 2.1 are equivalent to require L to be proper, convex and l.s.c..

2.1. Constraints encoded in FCB formulations. By allowing L to take infinite values, we are handling implicitly constraints over the state of system $x(t)$ and its velocity $\dot{x}(t)$. Indeed, these constraints are determined by the set-valued maps

$$\mathbf{X}(t) := \{x \in \mathbb{R}^n \mid \exists v \in \mathbb{R}^n, \; L(t, x, v) \in \mathbb{R}\} \quad \text{and} \quad \Gamma_L(t, x) := \{v \in \mathbb{R}^n \mid \; L(t, x, v) \in \mathbb{R}\}.$$

In a similar way, the fact that g can take infinite values implies that the set

$$\mathbf{D}_g := \text{dom}(g) = \{a \in \mathbb{R}^n \mid \; g(a) < +\infty\},$$

can be understood as a terminal constraint. In other words, this means that the minimization in (P) occurs and is finite only when the initial state of the system $x(0) = \xi_0$ is brought to the target $\mathbf{D}_g$ at time $t = T$. Explicitly, any feasible arc x of the Bolza problem (P) satisfies the state and velocity constraints, and reaches the terminal constraint set:

$$x(t) \in \mathbf{X}(t) \quad \text{and} \quad \dot{x}(t) \in \Gamma_L(t, x(t)), \quad \text{for a.e. } t \in [0, T], \quad \text{and} \quad x(T) \in \mathbf{D}_g.$$

The fully convex framework we are concerned with yields naturally to a meaningful duality theory that can be obtained via the standard approach with perturbation functions (cf. [25] or [34, Section 2.6]). Indeed, following the theory developed in [18, 21, 31], one can define a dual problem to (P) as the FCB problem

$$\text{(D)}\quad \text{minimize} \int_0^T K(t, y(t), \dot{y}(t))dt + f(y(T)) \text{ over all } y \in \mathbf{AC} \text{ so that } y(0) = -\eta_0.$$

Here the Lagrangian $K : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and the terminal cost $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ are

$$K(t, y, w) := \sup_{x,v \in \mathbb{R}^n} \{x \cdot w + v \cdot y - L(t, x, v)\} \quad \text{and} \quad f(b) := g^*(-b).$$

By symmetry, this formulation considers implicitly constraints over the state of the dual system. Indeed, the corresponding dynamics and state constraints are

$$\Gamma_K(t, y) := \{w \in \mathbb{R}^n \mid K(t, y, w) \in \mathbb{R}\} \quad \text{and} \quad \mathbf{Y}(t) := \{y \in \mathbb{R}^n \mid \Gamma_K(t, y) \neq \emptyset\},$$

and the terminal constraint is

$$\mathbf{D}_f := \text{dom}(f) = \{b \in \mathbb{R}^n \mid f(b) < +\infty\}.$$

Likewise for the primal problem, any feasible arc of the dual Bolza problem (D) respects a state and dynamical constraint, and satisfies the terminal constraint:

$$y(t) \in \mathbf{Y}(t) \quad \text{and} \quad \dot{y}(t) \in \Gamma_K(t, y(t)), \qquad \text{for a.e. } t \in [0,T] \quad \text{and} \quad y(T) \in \mathbf{D}_f.$$

The primal and dual state constraints are related to the dual and primal Lagrangian through the following relations (e.g. [19, Theorem 13.3])

$$r_L(t, d) = \sigma_{\overline{\mathbf{Y}(t)}}(d) \quad \text{and} \quad r_K(t, d) = \sigma_{\overline{\mathbf{X}(t)}}(d), \quad \forall d \in \mathbb{R}^n, \ \forall t \in [0,T], \tag{2.1}$$

where $r_\Lambda : [0,T] \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ denotes the recession function of the mapping $p \mapsto \Lambda(t, z, p)$, that is,

$$r_\Lambda(t, d) := \lim_{s \to +\infty} \frac{\Lambda(t, z, p + sd) - \Lambda(t, z, p)}{s}.$$

The functions r_L and r_K will play an active role in the sequel. As we will discuss next, the assumptions we are considering in this paper do not ensure the existence of solutions to (P) nor to (D) in the space $\mathbf{AC}$. As a matter of fact, minimizing sequences may converge to discontinuous arcs; see the examples in section 7. This means that the velocities of feasible states need to be understood as measures rather than functions, and so the functionals to be minimized must be extended to a $\mathbf{BV}$ setting. An extension has been proposed by Rockafellar in [26], where the cost associated with the jumps in the state are penalized by the recession functions r_L and r_K in the primal and dual problem, respectively. Let us point out that the recession functions r_L and r_K also allow to establish the following principle: *if the primal problem has no state constraints, solutions to the dual problem are necessarily absolutely continuous arcs.* This in particular allows to recover the usual case studied in [31, 32, 7, 8, 9, 10, 11, 12] where jumps and impulses are not present.

2.2. Extended FCB problems. It has been proven (cf. [18, Example 12]) that, for FCB problems, a solution of the dual problem is actually the costate trajectory that appears in Pontryagin's maximum principle. In optimal control, when state constraints are involved, it is expected that the adjoint arc will have jumps whenever the constraint is active; see for instance [23, 14, 33]. This fact naturally leads to the dual problem minimizing over $\mathbf{BV}$ rather than $\mathbf{AC}$. The philosophy of convex analysis is that symmetry between primal and dual problems should be adhered to, and hence the primal problem should be extended to minimizing over $\mathbf{BV}$ arcs as well. For these reasons, and following the ideas outlined in [26, 15, 16], an extension of the problems (P) and (D) to ones with minimization over $\mathbf{BV}$ is made in the following manner. First of all, we fix a singular (regular) measure (w.r.t. the Lebesgue measure) μ on $[0,T]$. Any $z \in \mathbf{BV}$ induces a Borel measure $dz(t)$ and has Lebesgue decomposition of form

$$dz(t) = \dot{z}(t)dt + \pi_z(t)d\mu(t),$$

where $\dot{z}(t)$ and $\pi_z(t)$ are the densities associated with the absolutely continuous and singular part of the measure $dz(t)$ (w.r.t. the Lebesgue measure).

There are two main issues to be taken into account when extending the functional of an FCB problem to a $\mathbf{BV}$ setting. The first one is concerned with the restrictions over the end-points of the admissible arcs and the second one with the cost associated with the jumps in the state.

As pointed out in [16], the first issue will force us to consider two different value functions associated with the same problem. The functionals to be minimized in (2.5) and (2.7) defined below are the same, however in (2.5) the end-points of the admissible arcs are required to satisfy the state constraints (2.3) while in (2.7) they are not.

The second issue has already been addressed by Rockafellar in [26] inspired by the theory developed in [22] for convex integral functionals; see [22, Theorem 5] for more details. In our framework this means that the extended FCB problem is concerned with the minimization of the functional

$$x \mapsto \int_0^T L(t, x(t), \dot{x}(t))dt + \int_0^T r_L(t, \pi_x(t))d\mu(t) + g(x(T)).$$

Note that this functional has a particular structure, it penalizes the absolutely continuous and singular part of the measure $dx(t)$ in a somehow independent way.

The extra integral part associated with the extended problem requires additional assumptions in order to be well defined. In particular, $t \mapsto r_L(t, \pi_x(t))$ has to be at least a measurable function. The following assumption alleviates this issue.

Hypothesis 2.2. The multifunctions $t \mapsto \overline{\mathbf{X}(t)}$ and $t \mapsto \overline{\mathbf{Y}(t)}$ are u.s.c. on $[0,T]$.

Note that this condition is trivial for the autonomous case, but in the non-autonomous case it essentially requires the state constraints processes to have bounded images; see for instance [1, Theorem 1.1.2]. Furthermore, since we are concerned with the value functions of the primal Bolza problem, we need to parametrize the FCB problem at hand w.r.t. its initial data. For this reason, for a given Lagrangian $\Lambda : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, we introduce for each $\tau \in [0,T]$ the functional $J_\tau^\Lambda : \mathbf{BV} \to \mathbb{R} \cup \{+\infty\}$ defined via

$$J_\tau^\Lambda(z) := \int_\tau^T \Lambda(t, z(t), \dot{z}(t))dt + \int_\tau^T r_\Lambda(t, \pi_z(t))d\mu(t), \qquad \forall z \in \mathbf{BV}. \tag{2.2}$$

In the light of (2.1), Hypothesis 2.2 ensures that the mappings $t \mapsto r_L(t, \pi_x(t))$ and $t \mapsto r_K(t, \pi_y(t))$ are measurable (e.g. [30, Example 14.51]) and so the functionals J_τ^L and J_τ^K are well-defined

2.3. Value functions of extended FCB problems. The first type of value functions we introduce are those that satisfy the state constraints up to the initial and final times. In other words, these value functions don't allow to jump at the initial or final time from outside the state constraints sets into them. For this reason primal and dual trajectories must satisfy respectively

$$x(\tau) \in \overline{\mathbf{X}(\tau)}, \quad \text{and} \quad x(T) \in \overline{\mathbf{X}(T)}, \tag{2.3}$$

$$y(\tau) \in \overline{\mathbf{Y}(\tau)}, \quad \text{and} \quad y(T) \in \overline{\mathbf{Y}(T)}. \tag{2.4}$$

Hence, the value function associated with the extended primal problem is

$$\mathbf{V}_\tau(\xi) := \inf_{x \in \mathbf{BV}} \left\{ J_\tau^L(x) + g(x(T)) \middle| \; x \text{ satisfies (2.3) and} x(\tau) = \xi \right\} \tag{2.5}$$

In a similar way, we consider the extended dual value function given by

$$\mathbf{W}_\tau(\eta) := \inf_{y \in \mathbf{BV}} \left\{ J_\tau^K(y) + f(y(T)) \middle| \; y \text{ satisfies (2.4) and } y(\tau) = -\eta \right\} \tag{2.6}$$

Note that in the light of (2.1), these definitions imply that when no state constraints are present in the dual problem ($\mathbf{Y}(t) = \mathbb{R}^n$ for any $t \in [0, T]$), then the optimal value of the Bolza problem (P) can be recovered through the value function (2.5), and furthermore, a feasible trajectory for (2.5) (if available) is forced to have a null singular part because $\sigma_{\mathbb{R}^n} = \delta_{\{0\}}$, and hence, the minimization is carried out in **AC** rather than in **BV**. From an optimization point of view, the fact that the dual problem has no state constraint means that the cost of the primal problem is coercive w.r.t. the velocity; we refer to the discussion in [31, Section 3]. Similar comments apply to the dual problem in the case when the primal has no state constraints ($\mathbf{X}(t) = \mathbb{R}^n$ for any $t \in [0, T]$).

The explicit constraints (2.3) and (2.4) are rather natural for problems over continuous arcs with well-behaved state constraints processes. However, since we are considering discontinuous trajectories we are not forced to impose them and it may be beneficial to disregard them. As a matter of fact, if these additional constraints are not considered, other value functions are obtained, which also play an interesting role in the theory we develop here as we will see shortly. The primal and dual value functions that may violate (2.3) or (2.4) are simply given by

$$\mathbb{V}_\tau(\xi) \; := \inf_{x \in \mathbf{BV}} \left\{ J_\tau^L(x) + g(x(T)) \middle| \; x(\tau) = \xi \right\}, \tag{2.7}$$

$$\mathbb{W}_\tau(\eta) := \inf_{y \in \mathbf{BV}} \left\{ J_\tau^K(y) + f(y(T)) \middle| \; y(\tau) = -\eta \right\}, \tag{2.8}$$

By definition of the value functions in (2.5)-(2.8), the following relations hold:

$$\mathbb{V}_\tau(\xi) \leq \mathbf{V}_\tau(\xi) \quad \text{and} \quad \mathbb{W}_\tau(\eta) \leq \mathbf{W}_\tau(\eta), \qquad \forall \tau \in [0, T], \; \forall \eta, \xi \in \mathbb{R}^n.$$

The importance of these value functions will be clear from the discussion in the next part, where we show that $\mathbb{V}_\tau$ must be paired with $\mathbf{W}_\tau$ whereas $\mathbf{V}_\tau$ must be with $\mathbb{W}_\tau$; see Theorem 3.6. This also is demonstrated with the examples in section 7.

3. Main results. In this section we present the main contribution of this paper.

3.1. Weak and strong duality. Under the assumptions we posed so far we have that both value functions $\mathbb{V}_\tau$ and $\mathbb{W}_\tau$ satisfy, thanks to the definition of the Legendre-Fenchel conjugate, a weak duality relation with the value functions $\mathbf{W}_\tau$ and $\mathbf{V}_\tau$, respectively.

PROPOSITION 3.1. *Under Hypotheses 2.1 and 2.2, for any $\tau \in [0,T]$ we have that the functionals J_τ^L and J_τ^K are convex and well-defined on $\mathbf{BV}$ and, with the convention $+\infty - \infty = -\infty + \infty = +\infty$, we also have*

$$J_\tau^L(x) + J_\tau^K(y) \geq x(T)\cdot y(T) - x(\tau)\cdot y(\tau), \qquad \forall x, y \in \mathbf{BV}$$

as long as x satisfies (2.3) or y satisfies (2.4). Moreover, we also have

$$\mathbf{V}_\tau(\xi) + \mathbb{W}_\tau(\eta) \geq \xi\cdot\eta \quad \textit{and} \quad \mathbb{V}_\tau(\xi) + \mathbf{W}_\tau(\eta) \geq \xi\cdot\eta, \qquad \forall \xi, \eta \in \mathbb{R}^n.$$

Proof. The first part of the statement comes from [26, Theorem 1]. For the second part, the one regarding the value functions, we note that it is enough to check the inequalities for $\xi, \eta \in \mathbb{R}^n$ such that $\mathbf{V}_\tau(\xi), \mathbf{W}_\tau(\eta) < +\infty$, otherwise the conclusion is straightforward (considering the convention $+\infty \pm \infty = +\infty$). Note that we are not assuming a priori that $\mathbf{V}_\tau(\xi), \mathbf{W}_\tau(\eta) > -\infty$. Let us focus on the first inequality, the other one follows by symmetry.

Let $x, y \in \mathbf{BV}$ be any feasible arc for the optimization problem defined in (2.5) and (2.8), respectively. In other words, $x(\tau) = \xi$, $y(\tau) = -\eta$, x satisfies (2.3) and

$$J_\tau^L(x) + g(x(T)) \in \mathbb{R} \quad \text{and} \quad J_\tau^K(y) + f(y(T)) \in \mathbb{R}.$$

Since $f(b) = g^*(-b)$, the definition of the conjugate of g leads to

$$g(x(T)) + f(y(T)) \geq -x(T)\cdot y(T).$$

It follows that

$$J_\tau^L(x) + g(x(T)) + J_\tau^K(y) + f(y(T)) \geq \xi\cdot\eta.$$

Taking infimum over $x, y \in \mathbf{BV}$ we get the first inequality. In particular, we must have that $\mathbf{V}_\tau(\xi) > -\infty$ and $\mathbb{W}_\tau(\eta) > -\infty$, and thus $\eta \in \mathrm{dom}(\mathbb{W}_\tau)$ and $\xi \in \mathrm{dom}(\mathbf{V}_\tau)$. ∎

In order to ensure the existence of optimal solutions we might impose the following qualification assumption, which can be interpreted as a Slater condition:

Hypothesis 3.2. There are $\bar{x} \in \mathbf{AC}$ and $\bar{y} \in \mathbf{AC}$ such that
(i) $\bar{x}(T) \in \mathrm{int}(\mathbf{D}_g)$, $J_\tau^L(\bar{x}) \in \mathbb{R}$ and $\bar{x}(\tau) \in \mathrm{int}(\mathbf{X}(\tau))$, $\forall \tau \in [0,T]$.
(ii) $\bar{y}(T) \in \mathrm{int}(\mathbf{D}_f)$, $J_\tau^K(\bar{y}) \in \mathbb{R}$ and $\bar{y}(\tau) \in \mathrm{int}(\mathbf{Y}(\tau))$, $\forall \tau \in [0,T]$.

Let us point out that this assumption implies that the primal and dual problems are feasible. Even more, we have that $\tau \mapsto \mathbf{V}_\tau(\bar{x}(\tau))$ and $\tau \mapsto \mathbf{W}_\tau(\bar{y}(\tau))$ are bounded mappings on $[0,T]$ with

$$\bar{x}(\tau) \in \mathrm{dom}(\mathbf{V}_\tau) \quad \text{and} \quad \bar{y}(\tau) \in \mathrm{dom}(\mathbf{W}_\tau), \qquad \forall \tau \in [0,T],$$

which means in particular, that for any $\tau \in [0,T]$ the value functions $\mathbf{V}_\tau$, $\mathbb{V}_\tau$, $\mathbf{W}_\tau$ and $\mathbb{W}_\tau$ are all proper maps (due to Proposition 3.1). Furthermore, Hypothesis 3.2 implies that $\mathrm{int}(\mathbf{X}(\tau))$ and $\mathrm{int}(\mathbf{Y}(\tau))$ are nonempty sets for any $\tau \in [0,T]$.

Remark 3.3. The assumptions in our case require that the domain of the primal and dual end-points costs have non-empty interior, while in [31, 32] they do not. By applying the perturbation technique, the lack of coercivity on the primal and dual Lagrangians must be compensated in some way on the end-point cost.

On the other hand, in order to have a strong duality between the value functions, further regularity over the data is required. In particular, the state constraint processes need to be continuous multifunctions. For this reason we consider the following:

Hypothesis 3.4. For any $\tau \in [0,T]$, $\xi \in \operatorname{int}(\mathbf{X}(\tau))$ and $\eta \in \operatorname{int}(\mathbf{Y}(\tau))$ the mapping $t \mapsto H(t,\xi,\eta)$ is integrable on an open subinterval of $[0,T]$ containing τ.

Here $H : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is the Hamiltonian associated with the primal problem (P), that is

$$H(t,x,y) := \sup_{v \in \mathbb{R}^n} \{y \cdot v - L(t,x,v)\}.$$

Remark 3.5. Note that Hypothesis 3.4 holds immediately in the autonomous case, because $H(t,\xi,\eta)$ finite and constant for any $t \in [0,T]$. The fact Hypothesis 3.4 together with the other hypotheses imply that $t \mapsto \overline{\mathbf{X}(t)}$ and $t \mapsto \overline{\mathbf{Y}(t)}$ are actually continuous multifunctions was shown in [26, Proposition 5].

We are now in a position to state a strong duality result between the value functions defined earlier. This result links the values functions and their conjugates.

THEOREM 3.6. *Under Hypotheses 2.1, 2.2, 3.2, and 3.4, for any $\tau \in [0,T]$ we have that $\mathbf{V}_\tau$, $\mathbb{V}_\tau$, $\mathbf{W}_\tau$ and $\mathbb{W}_\tau$ are convex proper and l.s.c. functions on $\mathbb{R}^n$. Moreover, the pairs $(\mathbf{V}_\tau, \mathbb{W}_\tau)$ and $(\mathbb{V}_\tau, \mathbf{W}_\tau)$ are conjugate to each other, that is,*

$$\mathbf{V}_\tau^* = \mathbb{W}_\tau, \quad \mathbf{V}_\tau = \mathbb{W}_\tau^*, \quad \mathbb{V}_\tau^* = \mathbf{W}_\tau \quad \text{and} \quad \mathbb{V}_\tau = \mathbf{W}_\tau^*.$$

In order to prove Theorem 3.6, a few more concepts and results need to be presented. For these reasons, we postpone the details of the proof to section 6. Examples that demonstrate this result and the next one (the characteristic method, Theorem 3.8) are available in section 7.

3.2. Generalized Characteristic method. We now turn our attention into the most important result of this paper, *a characteristic method for impulsive FCB problems* that describes the evolution of the subgradients of the value functions.

We begin by revisiting the notion of Hamiltonian system for impulsive problems. Definition 3.7 has been inspired by the *extended Hamiltonian conditions* introduced in [26]; for a general account on Hamiltonian conditions in optimal control we refer to [33, Sections 7.8 and 10.7]. The novelty of Definition 3.7 is that it weakens the conditions at the initial and final times, a fact that is fundamental for being able to deal with the two types of extended value functions we have introduced in section 2.

DEFINITION 3.7. *A pair of* **BV** *arcs (x,y) is said to be a* Hamiltonian trajectory *on $[t_0,t_1] \subseteq [0,T]$ provided that*

a) $(-\dot{y}(t), \dot{x}(t)) \in \partial H_t(x(t), y(t))$ *for a.e.* $t \in [t_0,t_1]$, *where* $H_t = H(t,\cdot,\cdot)$.
b) $x(t^+) \in \overline{\mathbf{X}(t)}$ *and* $y(t^+) \in \overline{\mathbf{Y}(t)}$ *for any* $t \in [t_0,t_1)$.
c) $x(t^-) \in \overline{\mathbf{X}(t)}$ *and* $y(t^-) \in \overline{\mathbf{Y}(t)}$ *for any* $t \in (t_0,t_1]$.
d) $\pi_x(t) \in \mathcal{N}_{\overline{\mathbf{Y}(t)}}(y(t^+)) \cap \mathcal{N}_{\overline{\mathbf{Y}(t)}}(y(t^-))$ $d\mu$*-a.e.* $t \in (t_0,t_1)$.
e) $\pi_y(t) \in \mathcal{N}_{\overline{\mathbf{X}(t)}}(x(t^+)) \cap \mathcal{N}_{\overline{\mathbf{X}(t)}}(x(t^-))$ $d\mu$*-a.e.* $t \in (t_0,t_1)$.

The characteristic method and main result of this paper is the following.

THEOREM 3.8. *Under Hypotheses 2.1, 2.2, 3.2, and 3.4, for any $\tau \in [0,T]$ and $\xi, \eta \in \mathbb{R}^n$ we have that $\eta \in \partial \mathbf{V}_\tau(\xi)$ if and only if there are $x, y \in \mathbf{BV}$ such that*
a) *(x,y) is a Hamiltonian trajectory on $[\tau, T]$.*
b) *$(x(\tau), y(\tau)) = (\xi, -\eta)$ and $-y(T) \in \partial g(x(T))$.*
c) *$\xi \in \overline{\mathbf{X}(\tau)}$, $x(T) \in \overline{\mathbf{X}(T)}$.*
d) *$\pi_y(\tau) \in \mathcal{N}_{\overline{\mathbf{X}(\tau)}}(\xi)$ and $\pi_y(T) \in \mathcal{N}_{\overline{\mathbf{X}(T)}}(x(T))$.*
e) *$\pi_x(\tau) \in \mathcal{N}_{\overline{\mathbf{Y}(\tau)}}(y(\tau^+))$ and $\pi_x(t) \in \mathcal{N}_{\overline{\mathbf{Y}(T)}}(y(T^-))$.*

The details of Theorem 3.8's proof are postponed to section 6.

Remark 3.9. It is worth mentioning that Theorem 3.8 shows that if the value function of an extended FCB is not differentiable (i.e. $\partial \mathbf{V}_\tau(\xi)$ is not a singleton) then the dual problem could have more than one solution (infinitely many actually). This is for instance the case when Items a to e in Theorem 3.8 define a well-posed measure differential inclusion in **BV**. In the same setting, one could establish a well-know principle in convex optimization that asserts that uniqueness of minimizers of the dual problem is equivalent to the value function of the primal problem being differentiable. The well-posedness of the dynamical system determined by Items a and b has been studied in [20, 24] in the standard setting (no state constraints). To the best of our knowledge, similar results (as in [20, 24]) on the well-posedness of the Hamiltonian systems for **BV** arcs has not been investigated in the literature.

The symmetry between the primal and dual problems leads to a similar result for the value function $\mathbb{V}_\tau$, which can be obtained by applying Theorem 3.8 to the value function $\mathbf{W}_\tau$ instead of $\mathbf{V}_\tau$ and using the fact that $\eta \in \partial \mathbb{V}_\tau(\xi)$ if and only if $\xi \in \partial \mathbf{W}_\tau(\eta)$ (Thanks to Theorem 3.6).

THEOREM 3.10. *Under Hypotheses 2.1, 2.2, 3.2, and 3.4, for any $\tau \in [0,T]$ and $\xi, \eta \in \mathbb{R}^n$ we have that $\eta \in \partial \mathbb{V}_\tau(\xi)$ if and only if there are $x, y \in \mathbf{BV}$ such that*
a) *(x,y) is a Hamiltonian trajectory on $[\tau, T]$.*
b) *$(x(\tau), y(\tau)) = (\xi, -\eta)$ and $-y(T) \in \partial g(x(T))$.*
c) *$-\eta \in \overline{\mathbf{Y}(\tau)}$, $y(T) \in \overline{\mathbf{Y}(T)}$.*
d) *$\pi_x(\tau) \in \mathcal{N}_{\overline{\mathbf{Y}(\tau)}}(-\eta)$ and $\pi_x(T) \in \mathcal{N}_{\overline{\mathbf{Y}(T)}}(y(T))$.*
e) *$\pi_y(\tau) \in \mathcal{N}_{\overline{\mathbf{X}(\tau)}}(x(\tau^+))$ and $\pi_y(t) \in \mathcal{N}_{\overline{\mathbf{X}(T)}}(x(T^-))$.*

Remark 3.11. As pointed out earlier, Theorem 3.8 and Theorem 3.10 extend [31, Theorem 2,4] in two ways. First of all, it covers problems with non-autonomous Lagrangians. Secondly and most important, it deals with problems with state constraints. In particular, under coercivity assumptions on the primal and dual Lagrangians the statements agree. This is because the coercive case means that no state constraints are present in the primal or dual problems, and so all the normal cones are reduced to the trivial cone that contains only the zero vector.

4. The Linear Quadratic case. Let us now provide some explicit formulas for the Linear Quadratic (LQ) problem, and show how this type of problems fits in our fully convex setting. Along this section we assume that $\mathbf{X} \subseteq \mathbb{R}^n$ and $\mathbf{U} \subseteq \mathbb{R}^m$ are given sets and P_T is an $n \times n$ matrix. Also, for any $t \in [0,T]$ given we consider that A_t, Q_t are $n \times n$ matrices, B_t is an $n \times m$ matrix and R_t is an $m \times m$ matrix. The set $\mathbf{U}_\infty$ stands for the recession cone of the set $\mathbf{U}$.

The LQ problem is the following: Given $\tau \in [0,T]$ and $\xi \in \mathbf{X}$,

$$
\text{(LQ)} \quad \begin{cases} \text{Minimize} & \frac{1}{2}\left[\int_\tau^T [(Q_t x(t))\cdot x(t) + (R_t u(t))\cdot u(t)]dt + (P_T x(T))\cdot x(T)\right] \\ \text{over all} & x \in \mathbf{AC} \text{ and } u \in \mathbf{L}^1_m(dt) \text{ with } x(\tau) = \xi \\ \text{such that} & \dot{x}(t) = A_t x(t) + B_t u(t), \quad \text{for a.e. } t \in [\tau, T] \\ & u(t) \in \mathbf{U}, \text{ for a.e.} t \in [\tau, T] \\ & x(t) \in \mathbf{X}, \ \forall t \in [\tau, T]. \end{cases}
$$

To formulate the problem as an FCB problem, the end-point cost $g : \mathbb{R}^n \to \mathbb{R}$ is taken as $g(a) := \frac{1}{2}(P_T a)$, and the Lagrangian $L : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined via

$$
L(t,x,v) = \frac{1}{2}(Q_t x)\cdot x + \delta_{\mathbf{X}}(x) + \inf_{u \in \mathbf{U}} \left\{ \frac{1}{2}(R_t u)\cdot u \mid v = A_t x + B_t u \right\}
$$

4.1. Basic Assumptions. In order to ensure that the Lagrangian is l.s.c. we assume the following qualification condition:

Hypothesis 4.1. For any $t \in [0,T]$ we have $\ker(B_t) \cap \ker(R_t) \cap \mathbf{U}_\infty = \{0\}$.

Since the recession function of $u \mapsto \frac{1}{2}(R_t u)\cdot u + \delta_{\mathbf{U}}(u)$ is the indicator function of $\ker(R_t) \cap \mathbf{U}_\infty$, we have that **??** implies the qualification condition on [3, Corollary 3.5.7], and so if R_t is positive semi-definite and $\mathbf{U}$ nonempty convex and closed, then the infimum in the Lagrangian is attained and it is an l.s.c. function. Therefore, in order to fulfill the standing assumption in this work, we assume that

Hypothesis 4.2. (Standing assumptions for the LQ problem)
- (i) $\mathbf{X}$ and $\mathbf{U}$ are convex closed nonempty sets.
- (ii) For any $t \in [0,T]$ the matrices Q_t and P_T are symmetric positive definite.
- (iii) For any $t \in [0,T]$ the matrix R_t is a symmetric positive semi-definite matrix.
- (iv) The mappings $t \mapsto A_t$, $t \mapsto Q_t$, $t \mapsto B_t$ and $t \mapsto R_t$ are measurable from $[0,T]$ into the corresponding matrix space with the Borel $\sigma-$Algebra.

It is not difficult to see that this guarantees that the end-point cost satisfies Hypothesis 2.1(i) and, in the light of the preceding discussion, that the epigraphical mapping $t \mapsto S_L(t)$ has nonempty convex closed images for any $t \in [0,T]$. To see that the epigraphical mapping S_L is measurable we note that the set-valued map

$$
t \mapsto C(t) := \{(x,u,v) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \mid x \in \mathbf{X},\ u \in \mathbf{U},\ v = A_t x + B_t u\}
$$

is measurable on $[0,T]$ (cf. [30, Theorem 14.36]), and so by definition its indicator function is a normal integrand. Thus, the epigraphical mapping of the function

$$
(t,x,u,v) \mapsto \frac{1}{2}(Q_t x)\cdot x + \frac{1}{2}(R_t u)\cdot u + \delta_{C(t)}(x,u,v)
$$

is measurable on $[0,T]$ because the first two summand are normal integrands (cf. [30, Example 14.29]) and the sum of normal integrands is also a normal integrand ([30, Proposition 14.44]). Hence, the Lagrangian is a normal integrand because it is l.s.c. and it is the marginal function of a normal integrand ([30, Proposition 14.47]). We have shown then that Hypotheses 4.1 and 4.2 lead to Hypothesis 2.1.

4.2. Extended problem and corresponding dual state constraints. Note that [3, Corollary 3.5.7] implies, thanks to Hypotheses 4.1 and 4.2, that

$$r_L(t,d) = \begin{cases} 0 & \text{if } \exists \theta \in \ker(R_t) \cap \mathbf{U}_\infty, \text{ such that} B\theta = d, \\ +\infty & \text{otherwise.} \end{cases}$$

This means that extended LQ problem in the **BV** setting can be written as follows:

$$\begin{cases} \text{Minimize} & \frac{1}{2}\left[\int_\tau^T [(Q_t x(t))\cdot x(t) + (R_t u(t))\cdot u(t)]dt + (P_T x(T))\cdot x(T)\right] \\ \text{over all} & x \in \mathbf{AC} \text{ and } u \in \mathbf{L}^1_m(dt) \text{ with } x(\tau) = \xi \\ \text{such that} & \dot{x}(t) = A_t x(t) + B_t u(t), \quad \text{for a.e. } t \in [\tau, T] \\ & \pi_x(t) = B\theta(t), \quad \text{for } d\mu\text{-a.e. } t \in [\tau, T] \\ & u(t) \in \mathbf{U}, \text{ for a.e.} t \in [\tau, T] \\ & \theta(t) \in \ker(R_t) \cap \mathbf{U}_\infty, \quad \text{for } d\mu\text{-a.e. } t \in [\tau, T] \\ & x(t) \in \mathbf{X}, \ \forall t \in (\tau, T). \end{cases}$$

The novelty in this extended version of the LQ problem is the singular part, which is determined by a control $t \mapsto \theta(t)$ via a dynamical constraint. This control also satisfies the given input constraints. The dual Lagrangian is

$$K(t,y,w) = \sup_{x\in\mathbf{X}} \left\{ (w + A_t^\top y)\cdot x - \frac{1}{2}(Q_t x)\cdot x \right\} + \sup_{u\in\mathbf{U}} \left\{ (B_t^\top y)\cdot u - \frac{1}{2}(R_t u)\cdot u \right\}.$$

Since Q_t is assumed to be positive definite, the dual state constraint is determined only by the second supremum in the definition above. Moreover, by [3, Theorem 2.5.4], we have that

$$\overline{\mathbf{Y}(t)} = \left\{ y \in \mathbb{R}^n \mid (B_t^\top y)\cdot d \le 0, \quad \forall d \in \ker(R_t)\cap \mathbf{U}_\infty \right\},$$

that is, it is the set of $y \in \mathbb{R}^n$ such that $B^\top y$ belongs to the (negative) polar cone to $\ker(R_t)\cap\mathbf{U}_\infty$. In particular, a sufficient condition for $\overline{\mathbf{Y}(t)}$ to have non empty interior is that $\ker(R_t)\cap\mathbf{U}_\infty$ is pointed for any $t \in [0,T]$. Also, the structure of $\overline{\mathbf{Y}(t)}$ implies that it is either $\{0\}$ or unbounded. As we have discussed earlier, this could be rather restrictive for Hypothesis 2.2. However, if R_t is positive definite for any $t \in [0,T]$ or $\mathbf{U}$ is compact, then $\mathbf{Y}(t) = \mathbb{R}^n$ and Hypothesis 2.2 holds immediately. The general non-autonomous case is still an open question.

4.3. Hamiltonian system. The Hamiltonian of the problem is given by

$$H(t,x,y) = g_t\left(B_t^\top y\right) + (A_t x)\cdot y - \frac{1}{2}(Q_t x)\cdot x - \delta_{\mathbf{X}}(x), \qquad \forall t \in [0,T], \ \forall x, y \in \mathbb{R}^n,$$

where $g_t(z) = (h_t + \delta_{\mathbf{U}})^* = (h_t^* \square \sigma_{\mathbf{U}})^{**}$ and $h_t(u) = \frac{1}{2}(R_t u)\cdot u$; here $\varphi_1 \square \varphi_2$ stands for the inf-convolution of φ_1 and φ_2. We claim that $g_t(z) = h_t^* \square \sigma_{\mathbf{U}}$ and that the infimum is attained. Indeed, note that by standard results in Convex Analysis (see for instance [3, Theorem 2.5.4]) we have that $(h_t^*)_\infty = \sigma_{\mathrm{dom}(h_t)} = \sigma_{\mathbb{R}^n} = \delta_{\{0\}}$ and $(\sigma_{\mathbf{U}})_\infty = \sigma_{\mathbf{U}}$. Hence, by [3, Corrollary 3.5.8] we have that $h_t^* \square \sigma_{\mathbf{U}}$ is proper convex l.s.c. and agrees with g_t.

Note that $u \in \partial g_t(z)$ if and only if $z \in R_t u + \mathcal{N}_{\mathbf{U}}(u)$, that is

$$\partial g_t(z) = \{u \in \mathbf{U} \mid z \in R_t u + \mathcal{N}_{\mathbf{U}}(u)\}.$$

Moreover, if $\mathbf{Y}(t)$ has nonempty interior for any $t \in [0,T]$, we have that

$$\partial_x H(t,x,y) = A_t^\top y - Q_t x - \mathcal{N}_{\mathbf{X}}(x) \quad \text{and} \quad \partial_y H(t,x,y) = B_t \partial g_t \left(B_t^\top y\right) + A_t x.$$

Consequently, a Hamiltonian trajectory on $[0,T]$ must satisfy a.e. the system of inclusions

$$\dot{x}(t) \in A_t x(t) + B_t \partial g_t \left(B_t^\top y(t)\right) \quad \text{and} \quad \dot{y}(t) \in -A_t^\top y(t) + Q_t x(t) + \mathcal{N}_{\mathbf{X}}(x(t)).$$

Note that in the classical case that $\mathbf{X} = \mathbb{R}^n$, $\mathbf{U} = \mathbb{R}^m$ and R_t is non singular, we recover the well-known system of equations

$$\dot{x}(t) = A_t x(t) + B_t R_t^{-1} B_t^\top y(t) \quad \text{and} \quad \dot{y}(t) = -A_t^\top y(t) + Q_t x(t).$$

5. Minimizers and optimality conditions. In this section we discuss about the existence of optimal solutions. To begin with, we state some intermediate lemmas and for sake of the exposition we revisit the notion of *fundamental kernel* studied by Rockafellar in [26], and introduced as such in [31] for **AC** arcs.

5.1. Fundamental Kernel. For given $\tau \in [0,T]$ and $a, b \in \mathbb{R}^n$ the fundamental kernel associated with a Lagrangian $\Lambda : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is

$$E_\tau^\Lambda(a,b) := \inf_{z \in \mathbf{BV}} \left\{ J_\tau^\Lambda(z) \;\middle|\; z(\tau) = a,\ z(T) = b \right\}$$

where, we recall that the functional J_τ^Λ has been defined in (2.2).

The fundamental kernel plays a key role in the analysis because it allows us to rewrite a Bolza problem in simple terms. Essentially, it encloses the information associated with the running cost. For example, it follows that

$$\text{(5.1)}\quad \mathbb{V}_\tau(\xi) = \inf_{a \in \mathbb{R}^n} \left\{ E_\tau^L(\xi,a) + g(a) \right\} \quad \text{and} \quad \mathbb{W}_\tau(\eta) = \inf_{b \in \mathbb{R}^n} \left\{ E_\tau^K(-\eta,b) + f(b) \right\}$$

Furthermore, it also can be checked that

$$\mathbf{V}_\tau(\xi) = \inf_{a \in \overline{\mathbf{X}(T)}} \left\{ E_\tau^L(\xi,a) + g(a) \right\} \quad \text{if } \xi \in \overline{\mathbf{X}(\tau)} \tag{5.2}$$

$$\mathbf{W}_\tau(\eta) = \inf_{b \in \overline{\mathbf{Y}(T)}} \left\{ E_\tau^K(-\eta,b) + f(b) \right\} \quad \text{if } \eta \in \overline{\mathbf{Y}(\tau)} \tag{5.3}$$

LEMMA 5.1 ([26, Theorem 3 and 3']). *Assume Hypotheses 2.1, 2.2, 3.2, and 3.4.*

I - *For any $a, b \in \mathbb{R}^n$ we have that*

$$E_\tau^L(a,b) = \sup_{y \in \mathbf{BV}} \left\{ y(T)\cdot b - y(\tau)\cdot a - J_\tau^K(y) \;\middle|\; y(\tau) \in \overline{\mathbf{Y}(\tau)},\ y(T) \in \overline{\mathbf{Y}(T)} \right\},$$

$$E_\tau^K(a,b) = \sup_{x \in \mathbf{BV}} \left\{ x(T)\cdot b - x(\tau)\cdot a - J_\tau^L(x) \;\middle|\; x(\tau) \in \overline{\mathbf{X}(\tau)},\ x(T) \in \overline{\mathbf{X}(T)} \right\}.$$

II - *For any $a \in \overline{\mathbf{X}(\tau)}$ and $b \in \overline{\mathbf{X}(T)}$ we have that*

$$E_\tau^L(a,b) = \sup_{y \in \mathbf{BV}} \left\{ y(T)\cdot b - y(\tau)\cdot a - J_\tau^K(y) \right\}$$

III - *For any $a \in \overline{\mathbf{Y}(\tau)}$ and $b \in \overline{\mathbf{Y}(T)}$ we have that*

$$E_\tau^K(a,b) = \sup_{x \in \mathbf{BV}} \left\{ x(T)\cdot b - x(\tau)\cdot a - J_\tau^L(x) \right\}$$

Fundamental kernels are families of optimization problems parametrized by given end-points. As such, solutions to them can be characterized by means of a Hamiltonian system. The following definitions is closely related to our definition of Hamiltonian trajectory. We emphasize that the main difference between Definition 5.2 and Definition 3.7, is that some conditions in Definition 3.7 are not necessarily satisfied at the initial or final times, whereas in the next definition they are.

Definition 5.2. *We say that a* **BV** *arc* x *is an extremal for the Lagrangian* L *if there is another* **BV** *arc* y*, called coextremal, such that*
a) $(-\dot{y}(t), \dot{x}(t)) \in \partial H_t(x(t), y(t))$ *for a.e.* $t \in [\tau, T]$*, where* $H_t = H(t, \cdot, \cdot)$.
b) $x(t^+) \in \overline{\mathbf{X}(t)}$ *and* $y(t^+) \in \overline{\mathbf{Y}(t)}$ *for any* $t \in [\tau, T]$.
c) $x(t^-) \in \overline{\mathbf{X}(t)}$ *and* $y(t^-) \in \overline{\mathbf{Y}(t)}$ *for any* $t \in [\tau, T]$.
d) $\pi_x(t) \in \mathcal{N}_{\overline{\mathbf{Y}(t)}}(y(t^+)) \cap \mathcal{N}_{\overline{\mathbf{Y}(t)}}(y(t^-))$ $d\mu$*-a.e.* $t \in [\tau, T]$.
e) $\pi_y(t) \in \mathcal{N}_{\overline{\mathbf{X}(t)}}(x(t^+)) \cap \mathcal{N}_{\overline{\mathbf{X}(t)}}(x(t^-))$ $d\mu$*-a.e.* $t \in [\tau, T]$.

We are now in a position to recall a duality theorem for the optimization problem involved in the definition of the fundamental kernel E^L_τ and a suitable dual problem.

Lemma 5.3 ([26, Theorem 2]). *Assume that Hypotheses 2.1 and 2.2 hold. Let* $\tau \in [0, T]$ *be fixed and* $x, y \in \mathbf{BV}$*. Then, the following are equivalent*
a) x *is an extremal for the Lagrangian* L *with coextremal* y.
b) *The following assertions hold*
  i) $E^L_\tau(x(\tau), x(T)) = J^L_\tau(x)$ *and* $E^K_\tau(y(\tau), y(T)) = J^K_\tau(y)$.
  ii) $E^L_\tau(x(\tau), x(T)) + E^K_\tau(y(\tau), y(T)) = x(T) \cdot y(T) - x(\tau) \cdot y(\tau)$.
  iii) x *satisfies* (2.3) *and* y *satisfies* (2.4).

Remark 5.4. Note that a pair (x, y) with $x \in \mathbf{BV}$ being an extremal for L and $y \in \mathbf{BV}$ being a corresponding coextremal is a Hamiltonian trajectory in terms of Definition 3.7, but the converse is not true. Our definition of a Hamiltonian trajectory can be explained in the light of Lemma 5.3. Note that if we had defined a Hamiltonian trajectory as exactly a pair extremal-coextremal, we would have got that the end-points of the primal and dual trajectory must satisfy the state constraints. Nevertheless, since $\mathbf{V}_\tau$ is conjugate with $\mathbb{W}_\tau$ (not with $\mathbf{W}_\tau$), we have to allow to the possible optimal dual arc to violate the state constraints at its end-points.

5.2. Existence of minimizers. We now continue by stating an existence of solution theorem. Let us point out that this result doesn't rely directly on a compactness argument, but instead on the Fermat's rule.

Before continuing let us discuss about the lower semicontinuity of the functional J^L_τ on the weak-$\star$ topology of **BV**. Here we use the identifications $\mathbf{BV} \cong \mathbb{R}^n \times \mathbf{C}^*$ and the fact that **BV** can be put in duality with $\mathbb{R}^n \times \mathbf{C}$ via the bilinear mapping

$$\langle x, (a, p)\rangle := x(\tau)\cdot a + \int_\tau^T p(t)\cdot\dot{x}(t)dt + \int_\tau^T p(t)\cdot\pi_x(t)d\mu(t), \quad x \in \mathbf{BV},\ (a, p) \in \mathbb{R}^n \times \mathbf{C}.$$

Lemma 5.5. *Assume Hypotheses 2.1, 2.2, 3.2, and 3.4. Then, for any* $\tau \in [0, T]$ *fixed, the functional* J^L_τ *is l.s.c. on the weak-*$\star$ *topology of* **BV**.

The proof of this lemma follows essentially the same arguments presented by Pennanen and Perkkiö to prove the last part of [17, Theorem 2.1], so we skip it; the time-dependence hypotheses in [17] are weaker than ours. The underlying idea is that the functional J^L_τ can be seen as the conjugate of a convex integral functional defined on the space **C** as in [22, Theorem 5]. This in particular, implies that the functional is l.s.c. for the weak-$\star$ topology of $\mathbb{R}^n \times \mathbf{L}^\infty_n$, and a posterior, l.s.c. for the weak-$\star$ topology of $\mathbf{BV}$.

Thanks to Lemma 5.5 we get the next result about the existence of minimizers.

PROPOSITION 5.6. *Assume Hypotheses 2.1, 2.2, 3.2, and 3.4 hold.*

I - *If $\mathbb{V}_\tau(\xi) \in \mathbb{R}$ for some $\xi \in \mathbb{R}^n$, then there is $\tilde{x} \in \mathbf{BV}$ with $\tilde{x}(\tau) = \xi$ such that*

$$\mathbb{V}_\tau(\xi) = J^L_\tau(\tilde{x}) + g(\tilde{x}(T)).$$

Moreover, if in addition $\xi \in \overline{\mathbf{X}(\tau)}$ and $\mathbf{V}_\tau(\xi) \in \mathbb{R}$, then there is $x \in \mathbf{BV}$ with $x(\tau) = \xi$ and $x(T) \in \overline{\mathbf{X}(T)}$ such that

$$\mathbf{V}_\tau(\xi) = J^L_\tau(x) + g(x(T)).$$

II - *If $\mathbb{W}_\tau(\eta) \in \mathbb{R}$ for some $\eta \in \mathbb{R}^n$, then there is $\tilde{y} \in \mathbf{BV}$ with $\tilde{y}(\tau) = -\eta$ so that*

$$\mathbb{W}_\tau(\eta) = J^K_\tau(\tilde{y}) + f(\tilde{y}(T)).$$

Moreover, if in addition $\eta \in \overline{\mathbf{Y}(\tau)}$ and $\mathbf{W}_\tau(\eta) \in \mathbb{R}$, then there is $y \in \mathbf{BV}$ with $y(\tau) = -\eta$ and $y(T) \in \overline{\mathbf{Y}(T)}$ such that

$$\mathbf{W}_\tau(\xi) = J^K_\tau(y) + f(y(T)).$$

Proof. By symmetry, we just need to focus on the primal value functions. Note that by (2.7), the value function $\mathbb{V}_\tau$ can be written as a marginal function:

$$\mathbb{V}_\tau(\xi) = \inf_{x \in \mathbf{BV}} \Psi(x, \xi), \quad \forall \xi \in \mathbb{R}^n,$$

where $\Psi : \mathbf{BV} \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is given by

$$\Psi(x, a) := J^L_\tau(x) + g(x(T)) + \delta_{\{0\}}(x(\tau) - a), \quad \forall x \in \mathbf{BV},\ \forall a \in \mathbb{R}^n.$$

Thanks to Lemma 5.5, it's not difficult to see that, $(x, a) \mapsto \Psi(x, a)$ is convex and l.s.c. for the weak-$\star$ topology on $\mathbf{BV}$. Furthermore, the fact that $\mathbb{V}_\tau(\xi) \in \mathbb{R}$ implies that Ψ is also proper with $(\check{x}, \xi) \in \mathrm{dom}(\Psi)$ for some $\check{x} \in \mathbf{BV}$. Hence, to establish that the infimum is attained, it is enough to check the Fermat's rule, that is, $0 \in \partial \Psi_\xi(\tilde{x})$ for some $\tilde{x} \in \mathbf{BV}$, where $\Psi_\xi(\cdot) = \Psi(\cdot, \xi)$. Let us begin by pointing out that Ψ_ξ agrees with the conjugate of the mapping $\varphi_\xi : \mathbb{R}^n \times \mathbf{C} \to \mathbb{R} \cup \{\pm\infty\}$ defined via

$$(a, p) \mapsto \varphi_\xi(a, p) := \inf_{\eta \in \mathbb{R}^n} \{\Psi^*((a, p), \eta) - \xi \cdot \eta\}. \tag{5.4}$$

If φ_ξ given by (5.4) is bounded on a neighborhood of the origin of $\mathbb{R}^n \times \mathbf{C}$, then we have that either φ_ξ is identically $-\infty$ or it is continuous around the origin; see for instance [34, Theorem 2.2.9]. Nevertheless, as we will see shortly $\varphi_\xi(0) \geq -\mathbb{V}_\tau(\xi)$, and so φ_ξ must be continuous at the origin and consequently $\partial \varphi_\xi(0) \neq \emptyset$; see for example [34, Theorem 2.4.9]. The fact that is equivalent (via the Fenchel-Young equality (1.1)) to the Fermat's rule provides the desired result. We divide the rest of the proof into several steps:

1. Let us first check that φ_ξ bounded below at the origin. Since, $\mathbb{V}_\tau(\xi) \in \mathbb{R}$ for any $\varepsilon > 0$ there is $x_\varepsilon \in \mathbf{BV}$ with $x_\varepsilon(\tau) = \xi$ such that

$$\mathbb{V}_\tau(\xi) \geq J^L_\tau(x_\varepsilon) + g(x_\varepsilon(T)) - \varepsilon,$$

and thus

$$\Psi^*(0,\eta) \geq \xi\cdot\eta - J^L_\tau(x_\varepsilon) - g(x_\varepsilon(T)) \geq \xi\cdot\eta - \mathbb{V}_\tau(\xi) - \varepsilon.$$

From where we get that $\varphi_\xi(0) > \infty$, and as a matter of fact

$$\varphi_\xi(0) \geq -\mathbb{V}_\tau(\xi) - \varepsilon, \quad \forall \varepsilon > 0.$$

2. For any $p \in \mathbf{C}$ we set $L_p(t,x,v) := L(t,x,v) - p(t)\cdot v$ and the corresponding dual counterpart $K_p(t,y,w) := K(t,y+p(t),w)$. Note that

$$r_{L_p}(t,d) = r_L(t,d) - p(t)\cdot d \quad \text{and} \quad r_{K_p}(t,d) = r_K(t,d)$$

and so it can be verified that

$$\Psi^*((a,p),\eta) = \sup_{x\in\mathbf{BV}} x(\tau)\cdot(a+\eta) - J^{L_p}_\tau(x) - g(x(T)).$$

It is rather clear Hypotheses 2.1, 2.2, and 3.4 hold true when replacing L with L_p and K with K_p. Furthermore, Hypothesis 3.2(i) holds too when interchanging L by L_p. Let us next show that Hypothesis 3.2(ii) holds as well for K_p instead of K when p lies on a neighborhood of the origin of $\mathbf{C}$.

3. Following the arguments presented in [26, Page 183-184], we claim that there are $\varepsilon > 0$, $\alpha \in \mathbf{L}^1$ and $F : [\tau,T]\times\mathbb{R}^n \to \mathbb{R}^n$ measurable in the first variable and Lipschitz continuous in the second one, that in addition satisfy

$$|y - \bar{y}(t)| < \varepsilon \quad \Rightarrow \quad K(t,y,F(t,y)) \leq \alpha(t), \quad \text{for a.e. } t\in[\tau,T],$$

where $\bar{y} \in \mathbf{AC}$ is the arc given by part (ii) in Hypothesis 3.2. Since $\bar{y}(t) \in \operatorname{int}(\mathbf{Y}(t))$ for any $t \in [\tau,T]$, we may assume without loss of generality that $\mathbb{B}(\bar{y}(t),\varepsilon) \subseteq \operatorname{int}(\mathbf{Y}(t))$ for any $t\in[\tau,T]$. Moreover, due to the fact that F is measurable w.r.t. the variable t and Lipschitz continuous w.r.t. the variable y, the Carathéodory Theorem [6, Theorem 2.1,1] implies that for any $p \in \mathbf{C}$ there is a unique $y_p \in \mathbf{AC}$ solution of the (backward) Cauchy problem

$$\dot{y}(t) = F(t,y(t)+p(t)), \quad \forall t\in[\tau,T], \quad \text{with } y(T) = \bar{y}(T).$$

Let us point out that $\bar{y}$ is the (unique) solution for the case $p = 0$ and also that, thanks to the Gronwall's Lemma [5, Proposition 4.1.4], the operator $p \mapsto y_p$ is continuous from $\mathbf{C}$ into itself, which means that for some $\delta > 0$ we have that

$$\|p\|_\infty < \delta \quad \Rightarrow \quad \|y_p + p - \bar{y}\|_\infty < \varepsilon.$$

In particular, for any $p\in\mathbf{C}$ with $\|p\|_\infty < \delta$ it follows that

$$K_p(t,y_p(t),\dot{y}_p(t)) \leq \alpha(t), \quad \text{for a.e.} t\in[\tau,T].$$

Hence $J^{K_p}_\tau(y_p) \in \mathbb{R}$ and $y_p(T) \in \mathbf{D}_f$. Moreover, if we call $\mathbf{Y}_p(t)$ the state constraints induced by K_p at time t, we note that

$$y \in \mathbf{Y}_p(t) \quad \Longleftrightarrow y + p(t) \in \mathbf{Y}(t).$$

Therefore, since $\mathbb{B}(\bar{y}(t),\varepsilon) \subseteq \text{int}(\mathbf{Y}(t))$ for any $t \in [\tau, T]$, we have that $y_p(t) \in \text{int}(\mathbf{Y}_p(t))$ for any $t \in [\tau, T]$. Thus, part (ii) in Hypothesis 3.2 holds as well when replacing K with K_p and if $\|p\|_\infty$ is small enough.

4. On the other hand, since $g(b) = f^*(-b)$ we obtain

$$\Psi^*((a,p),\eta) = \sup_{x\in\mathbf{BV}} \inf_{b\in\mathbb{R}^n} x(\tau)\cdot(a+\eta) + x(T)\cdot b - J_\tau^{L_p}(x) + f(b).$$

It follows then that

$$\varphi_\xi(a,p) = \inf_{\eta\in\mathbb{R}^n} \sup_{x\in\mathbf{BV}} \inf_{b\in\mathbb{R}^n} x(\tau)\cdot(a+\eta) + x(T)\cdot b - \xi\cdot\eta - J_\tau^{L_p}(x) + f(b)$$
$$\le \sup_{x\in\mathbf{BV}} \left\{ x(T)\cdot\bar{y}(T) - x(\tau)\cdot y_p(\tau) + -J_\tau^{L_p}(x)\right\} + \xi\cdot(y_p(\tau)+a) + f(\bar{y}(T)),$$

where $\bar{y} \in \mathbf{AC}$ is the arc given by Hypothesis 3.2 and y_p the one given by the previous step (the trajectory that makes Hypothesis 3.2 to hold for K_p). Moreover, since $y_p(\tau) \in \text{int}(\mathbf{Y}_p(\tau))$ and $y_p(T) = \bar{y}(T) \in \mathbf{Y}_p(T)$, one gets that

$$\varphi_\xi(a,p) \le E_\tau^{K_p}(y_p(\tau), y_p(T)) + f(\bar{y}(T)) + \xi\cdot(y_p(\tau)+a)$$

where the last step is a consequence of Item III in Lemma 5.1 applied with L_p and K_p; we have already seen that Hypotheses 2.1, 2.2, 3.2, and 3.4 hold true if one replaces L and K with L_p and K_p, respectively. Note also that $E_\tau^{K_p}(y_p(\tau), y_p(T)) \le \|\alpha\|_{\mathbf{L}^1}$ and so, φ_ξ is bounded above on a neighborhood of the origin, which implies that the infimum in the definition of $\mathbb{V}_\tau(\xi)$ is attained at some $\tilde{x} \in \mathbf{BV}$.

5. The argument for the case $\mathbf{V}_\tau(\xi) \in \mathbb{R}$ is similar to the above; the difference is the application of Lemma 5.1(Item III) should be replaced by Lemma 5.1(Item I). We omit these details. ∎

6. Proof of main results. In this section we provide the arguments that prove Theorem 3.6 and Theorem 3.8.

6.1. Proof of Theorem 3.6.

Proof. Let $\tau \in [0, T]$ be given. Note that the fact that the value functions are proper is a direct consequence of Hypothesis 3.2. Consequently, we only need to prove that the pairs $(\mathbf{V}_\tau, \mathbb{W}_\tau)$ and $(\mathbb{V}_\tau, \mathbf{W}_\tau)$ are conjugate to each other. Indeed, the convex and l.s.c. character of the value functions is implied by the fact that they, as being the conjugate of another function, can be written as the supremum of a family of affine continuous functions.

Let us focus on $\mathbb{W}_\tau = (\mathbf{V}_\tau)^*$. Note that by symmetry, a proof for $\mathbb{V}_\tau = (\mathbf{W}_\tau)^*$ would have the same estructure, hence we skip it to avoid repetition. We begin by pointing out that (5.1) and combined with Item I in Lemma 5.1 lead to the following relation

$$\mathbb{W}_\tau(\eta) = \inf_{b\in\mathbb{R}^n} \sup_{x\in\mathbf{BV}} f(b) - \Psi_\eta(b,x),$$

where $\Psi_\eta : \mathbb{R}^n \times \mathbf{BV} \to \mathbb{R}\cup\{+\infty\}$ is defined via

$$\Psi_\eta(b,x) := \begin{cases} J_\tau^L(x) - x(\tau)\cdot\eta - x(T)\cdot b & \text{if } x(\tau) \in \overline{\mathbf{X}(\tau)},\ x(T) \in \overline{\mathbf{X}(T)}, \\ +\infty & \text{otherwise.} \end{cases}$$

Note that $b \mapsto f(b) - \Psi_\eta(b,x)$ is convex and l.s.c. for any $x \in \mathrm{dom}(J_\tau^L)$ with $x(\tau) \in \overline{\mathbf{X}(\tau)}$ and $x(T) \in \overline{\mathbf{X}(T)}$. Note as well that the lower levels sets of $\varphi(b) := f(b) - \Psi_\eta(b,\bar{x})$ are compact subsets of $\mathbb{R}^n$, where $\bar{x}$ is given by Hypothesis 3.2. Indeed, note that by definition

$$\varphi^*(a) = g(\bar{x}(T) - a) + J_\tau^L(\bar{x}) - x(\tau)\cdot\eta, \qquad \forall a \in \mathbb{R}^n.$$

Since $\bar{x}(T) \in \mathrm{int}(\mathbf{D}_g)$ we have that $0 \in \mathrm{int}(\mathrm{dom}(\varphi^*))$ and so, by the Moreau's Theorem ([3, Proposition 3.1.3]) we have that φ has compact lower level sets.

Furthermore, $x \mapsto \Psi(b,x)$ is convex for any $b \in \mathbf{D}_f$. Note that the subset of $\mathrm{dom}(J_\tau^L)$ that satisfies the additional conditions $x(\tau) \in \overline{\mathbf{X}(\tau)}$ and $x(T) \in \overline{\mathbf{X}(T)}$ is convex and nonempty thanks to Hypotheses 2.1 and 3.2. Hence, in the light of the Lopsided Minimax Theorem [2, Theorem 6.2.7] we have that

$$\mathbb{W}_\tau(\eta) = -\inf_{x\in\mathbf{BV}} \sup_{b\in\mathbb{R}^n} \Psi(b,x) - f(b).$$

But since $f(b) = g^*(-b)$, given $x \in \mathbf{BV}$, we have that

$$\sup_{b\in\mathbb{R}^n} x(T)\cdot(-b) - f(b) = g^{**}(x(T)) = g(x(T)).$$

This implies that

$$\begin{aligned}\mathbb{W}_\tau(\eta) &= -\inf_{x\in\mathbf{BV}} \left\{ J_\tau^L(x) - x(\tau)\cdot\eta + g(x(T)) \,\middle|\, x(\tau)\in\overline{\mathbf{X}(\tau)},\ x(T)\in\overline{\mathbf{X}(T)} \right\}\\ &= -\inf_{\xi\in\overline{\mathbf{X}(\tau)}} \left\{ -\xi\cdot\eta + \inf_{x\in\mathbf{BV}} \left[J_\tau^L(x) + g(x(T)) + \delta_{\overline{\mathbf{X}(T)}}(x(T)) \,\middle|\, x(\tau)=\xi \right] \right\}\\ &= \sup_{\xi\in\overline{\mathbf{X}}(\tau)} \{\xi\cdot\eta - \mathbf{V}_\tau(\xi)\} = (\mathbf{V}_\tau)^*(\eta)\end{aligned}$$

On the other hand, using similar arguments, (5.3) and Item III in Lemma 5.1, we can show that for any $\eta \in \overline{\mathbf{Y}(\tau)}$, we have $\mathbf{W}_\tau(\eta) = \sup_{\xi\in\mathbb{R}^n} \xi\cdot\eta - \psi_\tau(\xi)$ where

$$\psi_\tau(\xi) = \inf_{x\in\mathbf{BV}} \left\{ J_\tau^L(x) + \left(f + \delta_{\overline{\mathbf{Y}(T)}}\right)^*(-x(T)) \,\middle|\, x(\tau) = \xi \right\}.$$

Therefore, $\mathbf{W}_\tau$ is convex l.s.c., and since we know that it is proper, we get

$$\mathbf{W}_\tau(\eta) = (\mathbf{W}_\tau)^{**}(\eta) = (\mathbb{V}_\tau)^*(\eta).$$

Symmetric arguments (but using Item II in Lemma 5.1) allow to show that $\mathbf{V}_\tau$ is l.s.c. and convex, and since it is also proper , we obtain

$$\mathbf{V}_\tau(\xi) = (\mathbf{V}_\tau)^{**}(\xi) = (\mathbb{W}_\tau)^*(\xi).$$

This completes the proof. ∎

6.2. Proof of Theorem 3.8. The proof is divided in two parts.

6.2.1. Necessity part.

Proof. Let $\eta \in \partial \mathbf{V}_\tau(\xi)$, then in particular $\xi \in \mathrm{dom}(\mathbf{V}_\tau) \cap \overline{\mathbf{X}(\tau)}$, and so, by Proposition 5.6 there is an optimal trajectory that realizes $\mathbf{V}_\tau(\xi)$, that is, $x \in \mathbf{BV}$ such that $x(\tau) = \xi$ and

$$\mathbf{V}_\tau(\xi) = J^L_\tau(x) + g(x(T)) \quad \text{and} \quad x(T) \in \overline{\mathbf{X}(T)}.$$

On the other hand, $\eta \in \partial \mathbf{V}_\tau(\xi)$ is equivalent to $\xi \in \partial \mathbb{W}_\tau(\eta)$ because $\mathbf{V}$ and $\mathbb{W}$ are conjugate to each other. Hence, $\eta \in \mathrm{dom}(\mathbb{W}_\tau)$, and thus Proposition 5.6 implies there is an optimal trajectory that realizes $\mathbb{W}_\tau(\eta)$, that is, $y \in \mathbf{BV}$ such that $y(\tau) = -\eta$ and $\mathbb{W}_\tau(\eta) = J^K_\tau(y) + f(y(T))$. Furthermore, by the Fenchel-Young equality (1.1) we have

$$\xi \cdot \eta = \mathbf{V}_\tau(\xi) + \mathbb{W}_\tau(\eta) = J^L_\tau(x) + J^K_\tau(y) + f(y(T)) + g(x(T)).$$

Rearranging the terms we get

$$0 = \left(J^L_\tau(x) + J^K_\tau(y) - x(T) \cdot y(T) - \xi \cdot \eta\right) + \left(f(y(T)) + g(x(T)) + x(T) \cdot y(T)\right).$$

Now, since both terms in parenthesis at the right-hand side are nonnegative, we actually get that

$$J^L_\tau(x) + J^K_\tau(y) = x(T) \cdot y(T) + \xi \cdot \eta \tag{6.1}$$

$$f(y(T)) + g(x(T)) = -x(T) \cdot y(T) \tag{6.2}$$

Let us point out that (6.2) is equivalent to $-y(T) \in \partial g(x(T))$. Let us show now that (x, y) is a Hamiltonian trajectory on $[\tau, T]$. Note that (6.1) and Lemma 5.1 imply

$$E^L_\tau(\xi, x(T)) = J^L_\tau(x) \quad \text{and} \qquad E^K_\tau(-\eta, y(T)) = J^K_\tau(y).$$

Let $\hat{y}$ be the $\mathbf{BV}$ arc that agrees with y on (τ, T), and that is right continuous at $t = \tau$ and left continuous at $t = T$. In particular, $\hat{y}(\tau) = y(\tau^+)$ and $\hat{y}(T) = y(T^-)$. Note that $\hat{y}(\tau) \in \overline{\mathbf{Y}(\tau)}$ and $\hat{y}(T) \in \overline{\mathbf{Y}(T)}$ because of Hypothesis 2.2. Moreover

$$E^K_\tau(-\eta, y(T)) = J^K_\tau(\hat{y}) + r_K(\tau, \hat{y}(\tau) + \eta) + r_K(T, y(T) - \hat{y}(T)).$$

By [26, Proposition 2] it follows that $E^K_\tau(\hat{y}(\tau), \hat{y}(T)) = J^K_\tau(\hat{y})$.
Furthermore, from (6.1) and the preceding identities we get that

$$\begin{gathered} E^L_\tau(\xi, x(T)) + E^K_\tau(\hat{y}(\tau), \hat{y}(T)) = \\ x(T) \cdot y(T) - r_K(T, y(T) - \hat{y}(T)) + \xi \cdot \eta - r_K(\tau, \hat{y}(\tau) + \eta) \le x(T) \cdot \hat{y}(T) - \xi \cdot \hat{y}(\tau) \end{gathered}$$

where the inequality comes from the fact that $r_K(t, d) = \sigma_{\overline{\mathbf{X}(t)}}(d)$ for any $t \in [\tau, T]$, and because $\xi \in \overline{\mathbf{X}(\tau)}$ and $x(T) \in \overline{\mathbf{X}(T)}$. It turns out then, thanks to Proposition 3.1, that the inequality is actually an equality. Hence, by Lemma 5.3 we have that x is an extremal with $\hat{y}$ being a coextremal. But since $\hat{y}(t) = y(t)$ for $t \in (\tau, T)$, we obtain that the pair (x, y) is a Hamiltonian trajectory on $[\tau, T]$. Given that

$$x(T) \cdot y(T) - r_K(T, y(T) - \hat{y}(T)) + \xi \cdot \eta - r_K(\tau, \hat{y}(\tau) + \eta) = x(T) \cdot \hat{y}(T) - \xi \cdot \hat{y}(\tau) \tag{6.3}$$

we must also have that

$$x(T)\cdot(y(T)-\hat{y}(T)) = r_K(T,y(T)-\hat{y}(T)) \quad \text{and} \quad \xi\cdot(\eta+\hat{y}(\tau)) = r_K(\tau,\hat{y}(\tau)+\eta).$$

Hence, for any $a\in\overline{\mathbf{X}(T)}$ we have that

$$\pi_y(T)\cdot(a-x(T)) = (y(T)-\hat{y}(T))\cdot(a-x(T)) = (y(T)-\hat{y}(T))\cdot a - r_K(T,y(T)-\hat{y}(T)),$$

$$\pi_y(\tau)\cdot(a-x(\tau)) = (y(\tau)-\hat{y}(\tau))\cdot(a-x(\tau)) = (y(\tau)-\hat{y}(\tau))\cdot a - r_K(\tau,y(\tau)-\hat{y}(\tau)).$$

But, since $r_K(T,d) = \sigma_{\overline{\mathbf{X}(T)}}(d)$, we conclude, by taking $a\in\overline{\mathbf{X}(T)}$ or $a\in\overline{\mathbf{X}(\tau)}$ when appropriate, that

$$\pi_y(T)\in\mathcal{N}_{\overline{\mathbf{X}(T)}}(x(T)) \quad \text{and} \quad \pi_y(\tau)\in\mathcal{N}_{\overline{\mathbf{X}(\tau)}}(x(\tau)) \tag{6.4}$$

and the conclusion follows. ∎

6.2.2. Sufficiency part.

Proof. This part is rather direct from what we have done before, however, we provide the details for sake of completeness. First of all, the fact that (x,y) is a Hamiltonian trajectory implies that x is an extremal with $\hat{y}$ being a coextremal. Note that there are some conditions that are trivially verified because $\pi_{\hat{y}}(\tau)=\pi_{\hat{y}}(T)=0$. Thus, by Lemma 5.3 we have $E^L_\tau(\xi,x(T)) = J^L_\tau(x)$, $E^K_\tau(\hat{y}(\tau),\hat{y}(T)) = J^K_\tau(\hat{y})$ and

$$J^L_\tau(x)+J^K_\tau(\hat{y}) = x(T)\cdot\hat{y}(T)-\xi\cdot\hat{y}(\tau).$$

On the other hand, is it not difficult to see that (6.4) is actually equivalent to (6.3), which leads to

$$J^L_\tau(x)+J^K_\tau(y) = x(T)\cdot y(T)+\xi\cdot\eta.$$

Using the transversality condition we also have that

$$\mathbf{V}_\tau(\eta)+\mathbb{W}_\tau(\eta)\le J^L_\tau(x)+J^K_\tau(y)+g(x(T))+g^*(-y(T)) = \xi\cdot\eta.$$

From here we get that the inequality holds with equality instead, and so the Fenchel-Young equality (1.1) is obtained, which means that $\eta\in\partial\mathbf{V}_\tau(\xi)$. ∎

7. Some examples. In this section we illustrate the results we have provided in the paper with some examples.

7.1. Example 1. Let us begin by illustrating Theorem 3.6 and Theorem 3.8 through a 1D example. Let

$$L(t,x,v) = \begin{cases} 0 & \text{if } x\ge 0,\ v\le 0,\\ +\infty & \text{otherwise,}\end{cases} \quad \text{and} \quad g(a)=\frac{1}{2}(a+1)^2.$$

Simple calculations yield to $K(t,y,w)=L(t,y,w)$, $f(b)=g(b)-\frac{1}{2}$ and

$$r_L(t,d)=r_K(t,d)=\begin{cases} 0 & \text{if } d\le 0,\\ +\infty & \text{otherwise.}\end{cases}$$

That is, primal and dual trajectory can only approach to zero from the right, and jumps are only allowed to the left. It follows then that

$$\mathbf{V}_\tau(\xi) = \begin{cases} \frac{1}{2} & \text{if } \xi \geq 0, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad \mathbb{V}_\tau(\xi) = \begin{cases} 0 & \text{if } \xi \geq 0, \\ +\infty & \text{otherwise,} \end{cases}$$

where the infima are attained respectively at

$$x_1(t) = \frac{T-t}{T-\tau}\xi \ \text{ if } \ t \in [\tau, T] \quad \text{and} \quad x_2(t) = \begin{cases} \xi & \text{if } t \in [\tau, T), \\ -1 & \text{if } t = T. \end{cases}$$

By symmetry, the dual value functions are

$$\mathbf{W}_\tau(\eta) = \begin{cases} 0 & \text{if } \eta \leq 0, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad \mathbb{W}_\tau(\eta) = \begin{cases} -\frac{1}{2} & \text{if } \eta \leq 0, \\ +\infty & \text{otherwise,} \end{cases}$$

with the infima being attained respectively at

$$y_1(t) = -\frac{T-t}{T-\tau}\eta \ \text{ if } \ t \in [\tau, T] \quad \text{and} \quad y_2(t) = \begin{cases} -\eta & \text{if } t \in [\tau, T), \\ -1 & \text{if } t = T. \end{cases}$$

We see then that $\mathbf{V}_\tau(\xi) + \mathbb{W}_\tau(\eta) = \xi \cdot \eta$ and $\mathbb{V}_\tau(\xi) + \mathbf{W}_\tau(\eta) = \xi \cdot \eta$ if either $\xi = 0$ or $\eta = 0$, which implies that the functions are conjugate to one another, as claimed in Theorem 3.6. Note that in general $\mathbf{V}_\tau(\xi) + \mathbf{W}_\tau(\eta)$ is either $\frac{1}{2}$ or $+\infty$. Furthermore, $\mathbb{V}_\tau(\xi) + \mathbb{W}_\tau(\eta)$ is either $-\frac{1}{2}$ or $+\infty$. Hence, the pairs $(\mathbf{V}_\tau, \mathbf{W}_\tau)$ and $(\mathbb{V}_\tau, \mathbb{W}_\tau)$ are not conjugate to each other.

The Hamiltonian of the problem is

$$H(x, y) = \begin{cases} 0 & \text{if } y \geq 0, \\ +\infty & \text{if } y < 0, \end{cases} - \begin{cases} 0 & \text{if } x \geq 0, \\ +\infty & \text{if } x < 0. \end{cases}$$

In particular, $\mathbf{X}(t) = \mathbf{Y}(t) = [0, +\infty)$ and

$$\partial H(x, y) = \begin{cases} \{(0,0)\} & \text{if } x, y > 0, \\ [0, +\infty) \times \{0\} & \text{if } x = 0, \ y > 0, \\ \{0\} \times (-\infty, 0] & \text{if } x > 0, \ y = 0. \\ [0, +\infty) \times (-\infty, 0] & \text{if } x = y = 0. \end{cases}$$

Hence, it's not difficult to see that if $\xi > 0$

$$x(t) = \frac{T-t}{T-\tau}\xi \ \text{ if } \ t \in [\tau, T] \quad \text{and} \quad y(t) = \begin{cases} 0 & \text{if } t \in [\tau, T), \\ -1 & \text{if } t = T, \end{cases}$$

defines a Hamiltonian trajectory on $[\tau, T]$, with

$$\pi_x(t) = 0 \quad \text{and} \quad \pi_y(t) = y(t), \qquad \forall t \in [\tau, T].$$

Note that $y(T) \notin \overline{\mathbf{Y}(T)}$, but $x(\tau) \in \overline{\mathbf{X}(\tau)}$ and $x(T) \in \overline{\mathbf{X}(T)}$. Moreover, x and y are optimal solutions for the value functions $\mathbf{V}_\tau$ and $\mathbb{W}_\tau$, respectively.

7.2. Example 2. Let us now consider a non coercive problem on $\mathbb{R}^2$. We set

$$L(x,v) = \frac{1}{2}\left(x_1^2 + x_2^2 + v_1^2\right) + \delta_{\mathbb{R}_+}(v_2) \quad \text{and} \quad g(a) = \frac{1}{2}\left(a_1^2 + a_2^2\right), \quad \forall x, v, a \in \mathbb{R}^2.$$

The Lagrangian above corresponds to a non coercive LQ problem with dynamical constraints. The related FCB problem is:

$$\begin{cases} \text{Minimize} & \frac{1}{2}\left[\int_\tau^T \left[x_1^2(t) + x_2^2(t) + \dot{x}_1^2(t)\right] dt + x_1^2(T) + x_2^2(T)\right] \\ \text{over all} & x \in \mathbf{AC} \quad \text{such that } \dot{x}_2(t) \geq 0, \text{ for a.e. } t \in [\tau, T], \quad x(\tau) = \xi. \end{cases} \tag{7.1}$$

Let $\Gamma = \mathbb{R} \times \{0\}$ and $\mathbf{\Omega}$ be the (Euclidean) open lower half-plane in $\mathbb{R}^2$. Let $\tilde{x} \in \mathbf{AC}$ be arbitrary such that $\tilde{x}(\tau) = (\xi_1, 0)$ for some $\xi_1 \in \mathbb{R}$. Note that if $\xi = (\xi_1, \xi_2) \in \mathbf{\Omega}$ (in particular $\xi_2 < 0$), then for any positive integer k large enough the arc given by

$$x^k(t) = \begin{cases} (\xi_1, \xi_2 + (t-\tau)k) & \text{if } \tau \leq t \leq \tau - \frac{\xi_2}{k}, \\ \tilde{x}\left(t + \frac{\xi_2}{k}\right) & \text{if } \tau - \frac{\xi_2}{k} < t \leq T, \end{cases}$$

reaches $(\xi_1, 0)$ from $\xi = (\xi_1, \xi_2)$ within time $\Delta t = \frac{-\xi_2}{k}$. The running cost associated with this $\mathbf{AC}$ arc on the time interval $\left[\tau, \tau - \frac{\xi_2}{k}\right]$ is just the one corresponding to the state (the cost related to the velocity is zero), and it is given by

$$\int_\tau^{\tau - \frac{\xi_2}{k}} \left[|x_1^k(t)|^2 + |x_2^k(t)|^2\right] dt = \frac{1}{6k}\left(-3\xi_1^2\xi_2 - \xi_2^3\right) \to 0 \quad \text{as } k \to +\infty$$

Notice that x^k converges pointwise to the $\mathbf{BV}$ arc

$$x(t) = \begin{cases} \xi & \text{if } t = \tau, \\ \tilde{x}(t) & \text{if } \tau < t \leq T. \end{cases}$$

This fact, combined with the estimate for the running cost associated with x^k on the time interval $\left[\tau, \tau - \frac{\xi_2}{k}\right]$, implies that $\mathbf{V}_\tau(\xi_1, \xi_2) = \mathbf{V}_\tau(\xi_1, 0)$ for any $(\xi_1, \xi_2) \in \mathbf{\Omega}$. In particular, this means that an impulsive trajectory, optimal for the extended FCB problem associated with (7.1), must jump from any point in $\mathbf{\Omega}$ into Γ instantaneously at time $t = \tau$ because the cost of doing so is zero.

On the other hand, note that for any $\bar{\xi}_2 \geq 0$ fixed, any affine subspace of $\mathbb{R}^2$ having the form $\left\{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \geq \bar{\xi}_2\right\}$ is (strongly) invariant w.r.t. the dynamical system associated with (7.1). This implies then, because of the quadratic cost over the state, that the function

$$\xi_2 \mapsto \mathbf{V}_\tau(\xi_1, \xi_2)$$

is increasing. This in turn implies that for computing $\mathbf{V}_\tau$ on $\mathbb{R}^2 \setminus \mathbf{\Omega}$, the state component x_2 can be set as ξ_2 and so, we can reduce the problem to solve the following 1D Linear-Quadratic problem:

$$\begin{cases} \text{Minimize} & \frac{1}{2}\int_\tau^T \left[z^2(t) + \dot{z}^2(t)\right] dt + \frac{1}{2}z^2(T) \\ \text{over all} & z \in \mathbf{AC} \text{ such that} z(\tau) = \xi_1. \end{cases} \tag{7.2}$$

If $\vartheta_\tau(\xi_1)$ stands for the value function of the problem (7.2), we have the relation

$$\mathbf{V}_\tau(\xi_1,\xi_2) = \vartheta_\tau(\xi_1) + \frac{1}{2}(T+1-\tau)\max\{0,\xi_2\}^2, \qquad \forall(\xi_1,\xi_2)\in\mathbb{R}^2.$$

Problem (7.2) is a rather classical one, and it is not difficult to see that it has a unique solution. As a matter of fact, its optimal solution and the associated value function are

$$z(t) = \xi_1 e^{\tau-t}, \quad \forall t\in[\tau,T] \quad \text{and} \quad \vartheta_\tau(\xi_1) = \frac{1}{2}\xi_1^2.$$

Therefore, a suitable optimal trajectory for the extended problem to **BV** associated with (7.1) that starts at $\xi=(\xi_1,\xi_2)$ at time $t=\tau$ has the following form

$$x^*(t) = \begin{cases} (\xi_1,\xi_2) & \text{if } t=\tau,\\ (\xi_1 e^{\tau-t}, \max\{0,\xi_2\}) & \text{if } t\in(\tau,T]. \end{cases}$$

7.2.1. Dual problem. Let us now take a look at the dual problem. First of all let us note that, after some calculations, the dual Lagrangian and dual endpoint cost have the form

$$K(y,w) = \frac{1}{2}\left(y_1^2+w_1^2+w_2^2\right) + \delta_{\mathbb{R}_-}(y_2) \quad \text{and} \quad f(b) = \frac{1}{2}\left(b_1^2+b_2^2\right).$$

Therefore, the dual problem has a state constraint, namely $\mathbf{Y}=\mathbb{R}\times\mathbb{R}_-$. Consequently, the dual problem is given by

$$\text{(7.3)} \quad \begin{cases} \text{Minimize} & \frac{1}{2}\left[\int_\tau^T \left[y_1^2(t)+\dot{y}_1^2(t)+\dot{y}_2^2(t)\right]dt + y_1^2(T)+y_2^2(T)\right]\\ \text{over all} & y\in\mathbf{AC} \text{ so that} y_2(t)\le 0 \text{ for any } t\in[\tau,T] \text{ and } y(\tau)=-\eta. \end{cases}$$

Notice that the problem can be separated into two, one that depends exclusively on the first variable and another on the second. In particular, when no state constraints are present these two problems are:

$$\text{(7.4)} \quad \begin{cases} \text{Minimize} & \frac{1}{2}\left[\int_\tau^T \left[y_1^2(t)+\dot{y}_1^2(t)\right]dt + y_1^2(T)\right]\\ \text{over all} & y\in\mathbf{AC} \text{ such that } y_1(\tau)=-\eta_1. \end{cases}$$

and

$$\text{(7.5)} \quad \begin{cases} \text{Minimize} & \frac{1}{2}\left[\int_\tau^T \dot{y}_2^2(t)dt + y_2^2(T)\right]\\ \text{over all} & y\in\mathbf{AC} \text{ such that} y_2(\tau)=-\eta_2. \end{cases}$$

We have already seen that the minimizer and associated value function to (7.4) is

$$y_1(t) = -\eta_1 e^{\tau-t}, \quad \forall t\in[\tau,T] \quad \text{and} \quad \omega_\tau^1(\eta_1) = \frac{1}{2}\eta_1^2.$$

On the other hand, by the Euler-Lagrange equation and the transversality condition (see [4, Chapter 14]) we get a twice continuously differentiable solution of the problem

(7.5), which satisfy

$$\ddot{y}_2(t) = 0 \quad \forall t \in [\tau, T] \quad \text{and} \quad \dot{y}_2(T) + y_2(T) = 0.$$

Therefore, the optimal solution and associated value function to (7.5) is

$$y_2(t) = -\eta_2 \frac{T+1-t}{T+1-\tau}, \quad \forall t \in [\tau, T] \quad \text{and} \quad \omega_\tau^2(\eta_2) = \frac{\eta_2^2}{2(T+1-\tau)}.$$

Then, the optimal solution for (7.3) is given by

$$y^*(t) = \left(-\eta_1 e^{\tau - t}, -\eta_2 \frac{T+1-t}{T+1-\tau} \right) \qquad \text{whenever} \eta_2 \geq 0$$

and the associated value function is

$$\mathbf{W}_\tau(\eta) = \begin{cases} \frac{1}{2}\eta_1^2 + \frac{\eta_2^2}{2(T+1-\tau)} & \text{if } \eta_2 \geq 0 \\ +\infty & \text{otherwise} \end{cases}$$

It is not difficult to see that $\mathbf{W}_\tau^* = \mathbf{V}_\tau$ and $\mathbf{V}_\tau^* = \mathbf{W}_\tau$. Moreover, since $\mathbf{V}_\tau$ is continuously differentiable, we also have that

$$\eta \in \partial \mathbf{V}_\tau(\xi) \iff \eta_1 = \xi_1 \quad \text{and} \quad \eta_2 = (T+1-\tau)\max\{0, \xi_2\}. \tag{7.6}$$

7.2.2. Application of Theorem 3.8. Let us now see what can be obtained from Theorem 3.8. Note that the Hamiltonian of the problem is

$$H(x,y) = \frac{1}{2}\left(y_1^2 - x_1^2 - x_2^2\right) + \delta_{\mathbb{R}_-}(y_2), \qquad \forall x, y \in \mathbb{R}^2.$$

Thus, according to Definition 3.7, a Hamiltonian trajectory on $[\tau, T]$ is a pair (x, y) of **BV** arcs that satisfies

$$\dot{y}_1 = x_1, \quad \dot{x}_1 = y_1, \qquad \text{for a.e. } t \in [\tau, T], \tag{7.7}$$

and

$$\dot{y}_2 = x_2, \quad \dot{x}_2 \in \begin{cases} \{0\} & \text{if } y_2(t) < 0, \\ [0, +\infty) & \text{if } y_2(t) = 0, \end{cases} \qquad \text{for a.e. } t \in [\tau, T], \tag{7.8}$$

together with $y_2(t) \leq 0$ for any $t \in [\tau, T)$ and

$$\pi_{x_2}(t) \in \begin{cases} \{0\} & \text{if } y_2(t) < 0, \\ [0, +\infty) & \text{if } y_2(t) = 0, \end{cases} \qquad \text{for } d\mu\text{-a.e. } t \in [\tau, T].$$

The fact that $\mathbf{X}(t) = \mathbb{R}^2$ combined with (d) in Theorem 3.8, makes all the other conditions trivial and also leads to claim that $y \in \mathbf{AC}$. Furthermore, the transversality condition implies that

$$x(T) + y(T) = 0. \tag{7.9}$$

By (7.7) and the initial time $t = \tau$, we have that

$$x_1(t) = \xi_1 \cosh(t-\tau) - \eta_1 \sinh(t-\tau) \quad \text{and} \quad y_1(t) = -\eta_1 \cosh(t-\tau) + \xi_1 \sinh(t-\tau).$$

Thanks to the transversality condition (7.9), we can actually conclude that we must have $\xi_1 = \eta_1$. On the other hand, suppose that $\eta_2 > 0$. Then since $y_2 \in \mathbf{AC}$, there is $t_0 \in (\tau, T]$ for which $y_2(t) < 0$ for any $t \in [\tau, t_0)$. This in turn, combined with (7.8), implies that $x_2(t) = \xi_2$ for any $t \in [\tau, t_0)$ and

$$y_2(t) = -\eta_2 + \xi_2(t - \tau), \quad \forall t \in [\tau, t_0).$$

Suppose that $y_2(t_0) = 0$. Then, necessarily we must have $\xi_2 > 0$. But, due to the fact that $y_2 \in \mathbf{AC}$ and

$$dx_2(t) = \dot{x}_2(t)dt + \pi_{x_2}(t)d\mu(t) \geq 0,$$

The equation $\dot{y}_2(t) = x_2(t) > \xi_2 > 0$ for a.e. $t \in [\tau, T]$, and so y_2 does not satisfy the state constraint after the time $t = t_0$, which yields to a contradiction, so we can assume that $t_0 = T$. Moreover, if $y_2(T) = 0$, by the transversality condition we must have that $x_2(T) = 0$ as well. This in turn implies that $\pi_{x_2}(T) = -\xi_2 < 0$ which is not possible. So, we must have that if $\eta_2 > 0$, then $y_2(t) < 0$ and $x_2(t) = \xi_2$ for any $t \in [\tau, T]$. In this case, the transversality condition (7.9) implies that $\eta_2 = (T+1-\tau)\xi_2$.

Finally, let us consider the case $\eta_2 = 0$. Using the same argument as before, we must have that $y_2(t) = \eta_2 = 0$ for any $t \in [\tau, T]$. Furthermore, the transversality condition (7.9) implies that $x_2(T) = 0$ and (7.8) that $x_2(t) = 0$ for a.e. $t \in [\tau, T]$. But, since $dx_2(t) \geq 0$ the only option for x_2 to be discontinuous is at time $t = \tau$ and this can only happen if $\xi_2 \leq 0$.

The analysis we have shown confirms the expressions that have been found in (7.6), which corroborates in turn Theorem 3.8. This also shows how Theorem 3.8 can be used to determine optimal solutions to FCB in the context of **BV** arcs.

It is worth mentioning that the example we have just studied demonstrates that the uniqueness of minimizers to the dual problem is equivalent to the value function of the primal problem to be differentiable as pointed out in Remark 3.9.

7.3. Example 3. We now study a similar problem as in the previous example, but with a state constraint on the primal problem, that is, we are concerned now with the LQ problem with dynamical and state constraints:

$$\begin{cases} \text{Minimize} & \dfrac{1}{2}\left[\displaystyle\int_\tau^T \left[x_1^2(t) + x_2^2(t) + \dot{x}_1^2(t)\right] dt + x_1^2(T) + x_2^2(T)\right] \\ \text{over all} & x \in \mathbf{AC} \text{ such that } \quad x(\tau) = \xi, \\ & x_2(t) \geq -1, \quad \text{for any } t \in [\tau, T], \\ & \dot{x}_2(t) \geq 0, \quad \text{for a.e. } t \in [\tau, T]. \end{cases}$$

Under these circumstances two value functions arise. One that is finite only on the state constraint, and another that is finite beyond that set. According to our notation and based on the analysis exposed in Example 2, we have for any $(\xi_1, \xi_2) \in \mathbb{R}^2$

$$\mathbf{V}_\tau(\xi) = \begin{cases} \mathbb{V}_\tau(\xi) & \text{if } \xi_2 \geq -1, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad \mathbb{V}_\tau(\xi) = \frac{1}{2}\xi_1^2 + \frac{1}{2}(T + 1 - \tau)\max\{0, \xi_2\}^2.$$

This is because the optimal solutions have the same form as in Example 2, it only changes the domain where it is defined.

Let us now study the dual problem. The Lagrangian in this case takes the form

$$K(y,w)=\frac{1}{2}w_1^2+\begin{cases}\frac{1}{2}w_2^2 & \text{if } w_2\geq -1,\\ -w_2-\frac{1}{2} & \text{otherwise,}\end{cases}+\begin{cases}\frac{1}{2}y_1^2 & \text{if } y_2\leq 0,\\ +\infty & \text{otherwise.}\end{cases}$$

Therefore, the dual problem has a state constraints as well, namely $\mathbf{Y}=\mathbb{R}\times\mathbb{R}_-$ and is non coercive. Consequently, the dual problem is given by

$$\begin{cases}\text{Minimize} & \frac{1}{2}\left[\int_\tau^T\left[y_1^2(t)+\dot{y}_1^2(t)+\varphi(\dot{y}_2(t))\right]dt+y_1^2(T)+y_2^2(T)\right]\\ \text{over all} & y\in\mathbf{AC}\ \text{ such that} y_2(t)\leq 0 \text{ for any } t\in[\tau,T] \text{ and } y(\tau)=-\eta.\end{cases}$$

where $\varphi(w)=\frac{1}{2}w^2$ if $w\geq -1$ and $\varphi(w)=-w-\frac{1}{2}$ otherwise. Since the problem associated with the value function $\mathbf{W}_\tau$ is the one that respects the state constraint at any time, we have that the optimal solution in this case agrees with the one found in example 2, that is,

$$\mathbf{W}_\tau(\eta)=\begin{cases}\frac{1}{2}\eta_1^2+\frac{\eta_2^2}{2(T+1-\tau)} & \text{if } \eta_2\geq 0,\\ +\infty & \text{otherwise.}\end{cases}$$

On the other hand, when considering the value function that allows jump at the initial and final times we have that optimal solution can start outside the state constraint and jump at time $t=\tau$ to the set $\mathbf{Y}$. Thus, it is not difficult to see that the following is an optimal solution for problem $\mathbb{W}_\tau(\eta)$

$$y^*(t)=\left(-\eta_1e^{\tau-t},-\max\{\eta_2,0\}\frac{T+1-t}{T+1-\tau}\right),\qquad \forall t\in(\tau,T].$$

The associated cost for jumping from outside the state constraint coincides with the length of the jump. Therefore,

$$\mathbb{W}_\tau(\eta)=\begin{cases}\frac{1}{2}\eta_1^2+\frac{\eta_2^2}{2(T+1-\tau)} & \text{if } \eta_2\geq 0,\\ \frac{1}{2}\eta_1^2-\eta_2 & \text{otherwise.}\end{cases}$$

We have already seen that $\mathbb{V}_\tau$ and $\mathbf{W}_\tau$ are conjugate to each other, and that Theorem 3.8 is verified in this case. Let us now take a look at $\mathbf{V}_\tau$ and $\mathbb{W}_\tau$. It can be checked that $\mathbf{V}_\tau$ and $\mathbb{W}_\tau$ are conjugate to each other, a fact that confirms Theorem 3.6.

7.3.1. Application of Theorem 3.8. Let us now apply Theorem 3.8 to these data. Note that the Hamiltonian of the problem is

$$H(x,y)=\frac{1}{2}\left(y_1^2-x_1^2-x_2^2\right)+\delta_{\mathbb{R}_-}(y_2)-\delta_{\mathbb{R}_+}(x_2+1),\qquad \forall x,y\in\mathbb{R}^2.$$

In this case we have assumed, for sake of simplicity, the convention that $+\infty-\infty=+\infty$. However, in strict sense we may have to work with an equivalent class of convex-concave functions; see the discussion in [19, Chapter 33-34].

According to Definition 3.7, a Hamiltonian trajectory on $[\tau,T]$ is a pair (x,y) of **BV** arcs that satisfies

$$\dot{y}_1(t)=x_1(t),\quad \dot{x}_1(t)=y_1(t),\qquad \text{for a.e. } t\in[\tau,T], \tag{7.10}$$

and for a.e. $t \in [\tau, T]$

$$(7.11)\quad \dot{y}_2(t) \in \begin{cases} \{x_2(t)\} & \text{if } x_2(t) > -1, \\ (-\infty, x_2(t)] & \text{if } x_2(t) = -1, \end{cases} \quad \dot{x}_2(t) \in \begin{cases} \{0\} & \text{if } y_2(t) < 0, \\ [0, +\infty) & \text{if } y_2(t) = 0, \end{cases}$$

together with $y_2(t^+) \leq 0$ and $x_2(t^+) \geq -1$ for any $t \in [\tau, T)$. Also we have $y_2(t^-) \leq 0$ and $x_2(t^-) \geq -1$ for any $t \in (\tau, T]$. The singular parts must verify $d\mu$-a.e. on (τ, T)

$$(7.12)\qquad \pi_{x_2}(t) \in \begin{cases} \{0\} & \text{if } y_2(t^-) < 0 \text{ or } y_2(t^+) < 0, \\ [0, +\infty) & \text{if } y_2(t^+) = y_2(t^-) = 0, \end{cases}$$

$$(7.13)\qquad \pi_{y_2}(t) \in \begin{cases} \{0\} & \text{if } x_2(t^-) > -1 \text{ or } x_2(t^+) > -1, \\ (-\infty, 0] & \text{if } x_2(t^+) = x_2(t^-) = -1. \end{cases}$$

The preceding conditions (7.12) and (7.13) imply that, at least on the interval (τ, T) the optimal trajectories cannot jump at the same time. Furthermore, Theorem 3.8 implies that $\xi_2 \geq -1$ as well as the transversality condition (7.9). It in addition imposes the following condition on the singular parts

$$(7.14)\quad \pi_{x_2}(\tau) \in \begin{cases} \{0\} & \text{if } y_2(\tau^+) < 0, \\ [0, +\infty) & \text{if } y_2(\tau^+) = 0, \end{cases} \quad \pi_{x_2}(T) \in \begin{cases} \{0\} & \text{if } y_2(T^-) < 0, \\ [0, +\infty) & \text{if } y_2(T^-) = 0, \end{cases}$$

$$(7.15)\quad \pi_{y_2}(\tau) \in \begin{cases} \{0\} & \text{if } \xi_2 > -1, \\ (-\infty, 0] & \text{if } \xi_2 = -1, \end{cases} \quad \pi_{y_2}(T) \in \begin{cases} \{0\} & \text{if } x(T) > -1, \\ (-\infty, 0] & \text{if } x(T) = -1. \end{cases}$$

By (7.10) and the initial condition at time $t = \tau$, we have that

$$x_1(t) = \xi_1 \cosh(t - \tau) - \eta_1 \sinh(t - \tau) \quad \text{and} \quad y_1(t) = -\eta_1 \cosh(t - \tau) + \xi_1 \sinh(t - \tau).$$

Thanks to the transversality condition (7.9), we can actually conclude that we must have $\xi_1 = \eta_1$. By (7.11), (7.12) and (7.14) we get that $dx_2 \geq 0$. Therefore, if $\xi_2 > -1$ then $x_2(t) > -1$ for any $t \in [\tau, T]$. Consequently, (7.11) implies that $\dot{y}_2 = x_2$ for a.e. $t \in [\tau, T]$. But under these circumstances the singular part of y_2 is zero because of (7.13) and (7.15), and so $y_2 \in \mathbf{AC}$, which leads to the same solution as in Example 2.

So, the only interesting case remaining to analyze is when $\xi_2 = -1$. Note that $x_2(t)$ cannot be constantly equal to -1 on $[\tau, T]$. Indeed, if this was the case, the transversality condition would imply that $y_2(T) = 1$. But, since $y(T^-) \leq 0$ because of the state constraint, we would get that $\pi_{y_2}(T) = 1$, which contradicts (7.15). So, since $dx_2 \geq 0$, for some $t_0 \in [\tau, T]$, an optimal solution must have the form

$$x_2(t) \in \begin{cases} \{-1\} & \text{if } t \in [\tau, t_0) \\ (-1, +\infty) & \text{otherwise.} \end{cases}$$

Now, if $t_0 = T$, we must have that $y_2(T^-) = 0$ because of (7.14). But, since $x(T) > -1$ we have by (7.15) that $y_2(T) = 0$. Note that by (7.11), (7.13) and (7.15) we have that $dy_2(t) \leq 0$, and so by the state constraint we should have $y(t) = 0$ for any $t \in (\tau, T]$. However, this is not possible because of (7.13) we have $\dot{y}_2(t) \leq -1$ a.e. on $[\tau, T]$. Therefore, we can assume that $t_0 < T$.

On the other hand, note that $dy_2(t) \leq 0$ on $[\tau, t_0)$ and so, if $\eta_2 > 0$ we get that $y_2(t) \leq -\eta_2 < 0$ on $[\tau, t_0)$. But since $\pi_{y_2}(t) \leq 0$ for $d\mu$-a.e. $t \in [\tau, T]$ we must have that $y_2(t) < 0$ for any $t \in [\tau, T]$, and so $\pi_{x_2}(t) = 0$ for $d\mu$-a.e. $t \in [\tau, T]$, which means that $x_2(T) = -1$. However, the transversality condition (7.9) leads to $-1 = \xi_2 = x_2(T) = -y_2(T) > 0$. Therefore, the case $\eta_2 > 0$ is not possible.

Suppose now that $\eta_2 \leq 0$ and $t_0 < T$. Then on $(t_0, T]$ the trajectory y_2 is absolutely continuous (because of the analysis above). This means that either (i) $y(t) = 0$ for any $t \in (t_0, T]$ or (ii) $y(t) < 0$ for any $t \in (t_0, T]$. In the case (ii), we can conclude that x_2 is constant on $[\tau, T]$, because its singular part is zero on that interval and the Hamiltonian equation (7.11) implies that $\dot{x}_2 = 0$ a.e. on $[\tau, T]$. However, because of the transversality condition (7.9) we get that $-1 = \xi_2 = x_2(T) = -y_2(T) > 0$. Therefore, the only possible option is (i). On the other hand, the condition over the singular part implies that x_2 as well as y_2 can only jump at $t = t_0$, but we have discussed that this is only possible at $t = \tau$ or $t = T$; recall that we have already ruled out the case $t_0 = T$. Thus we can conclude that

$$x_2(t) = \begin{cases} -1 & \text{if } t = \tau \\ 0 & \text{if } t \in (\tau, T] \end{cases} \quad \text{and} \quad y_2(t) = \begin{cases} -\eta_2 & \text{if } t = \tau \\ 0 & \text{if } t \in (\tau, T]. \end{cases}$$

Note that $\partial \mathbf{V}_\tau(\xi_1, -1) = \{\xi_1\} \times (-\infty, 0]$, which confirms what we have found above; the only constraints over η_2 is that it must be non positive.

Acknowledgments. We thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions that have helped to improve the presentation of the manuscript.

REFERENCES

[1] J.-P. Aubin and A. Cellina, *Differential Inclusions*, Springer-Verlag, 1984.

[2] J.-P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, John Wiley & Sons Inc., 1984.

[3] A. Auslender and M. Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer-Verlag, 2003.

[4] F. Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*, Springer, 2013.

[5] F. Clarke, Y. Ledyaev, R. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, 1998.

[6] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, Tata McGraw-Hill Education, 1955.

[7] R. Goebel, *Regularity of the optimal feedback and the value function in convex problems of optimal control*, Set-Valued Analysis, 12 (2004), pp. 127–145.

[8] R. Goebel, *Stationary Hamilton–Jacobi equations for convex control problems: uniqueness and duality of solutions*, in Optimal control, stabilization and nonsmooth analysis, Springer, 2004, pp. 313–322.

[9] R. Goebel, *Convex optimal control problems with smooth Hamiltonians*, SIAM Journal on Control and Optimization, 43 (2005), pp. 1787–1811.

[10] R. Goebel, *Duality and uniqueness of convex solutions to stationary Hamilton–Jacobi equations*, Transactions of the American Mathematical Society, 357 (2005), pp. 2187–2203.

[11] R. Goebel and R. T. Rockafellar, *Generalized conjugacy in Hamilton-Jacobi theory for fully convex Lagrangians*, Journal of Convex Analysis, 9 (2002), pp. 463–473.

[12] R. Goebel and R. T. Rockafellar, *Linear-convex control and duality*, in Geometric Control And Nonsmooth Analysis: In Honor of the 73rd Birthday of H Hermes and of the 71st Birthday of RT Rockafellar, World Scientific, 2008, pp. 280–299.

[13] R. Goebel and M. Subbotin, *Continuous time linear quadratic regulator with control constraints via convex duality*, IEEE Transactions on Automatic Control, 52 (2007), pp. 886–892.

[14] R. F. Hartl, S. P. Sethi, and R. G. Vickson, *A survey of the maximum principles for optimal control problems with state constraints*, SIAM Review, 37 (1995), pp. 181–218.

[15] C. Hermosilla and P. R. Wolenski, *Self-dual approximations to fully convex impulsive systems*, in Proceedings of the 55th IEEE CDC, 2016, pp. 6198–6203.

[16] C. Hermosilla and P. R. Wolenski, *Constrained and impulsive linear quadratic control problems*, in Proceeding of thye 20th World Congress of the IFAC, vol. 50, 2017, pp. 1637–1642.
[17] T. Pennanen and A.-P. Perkkiö, *Duality in convex problems of Bolza over functions of bounded variation*, SIAM Journal on Control and Optimization, 52 (2014), pp. 1481–1498.
[18] R. T. Rockafellar, *Conjugate convex functions in optimal control and the calculus of variations*, Journal of Mathematical Analysis and Applications, 32 (1970), pp. 174–222.
[19] R. T. Rockafellar, *Convex Analysis*, vol. 28, Princeton University Press, 1970.
[20] R. T. Rockafellar, *Generalized Hamiltonian equations for convex problems of Lagrange*, Pacific Journal of Mathematics, 33 (1970), pp. 411–427.
[21] R. T. Rockafellar, *Existence and duality theorems for convex problems of Bolza*, Transactions of the American Mathematical Society, 159 (1971).
[22] R. T. Rockafellar, *Integrals which are convex functionals II*, Pacific Journal of Mathematics, 39 (1971), pp. 439–469.
[23] R. T. Rockafellar, *State constraints in convex control problems of Bolza*, SIAM Journal on Control and Optimization, 10 (1972), pp. 691–715.
[24] R. T. Rockafellar, *Saddle points of Hamiltonian systems in convex problems of Lagrange*, Journal of Optimization Theory and Applications, 12 (1973), pp. 367–390.
[25] R. T. Rockafellar, *Conjugate duality and optimization*, vol. 16 of CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 1974.
[26] R. T. Rockafellar, *Dual problems of Lagrange for arcs of bounded variation*, in Calculus of variations and control theory (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1975; dedicated to Laurence Chisholm Young on the occasion of his 70th birthday), Academic Press, New York, 1976, pp. 155–192.
[27] R. T. Rockafellar, *Linear-quadratic programming and optimal control*, SIAM Journal on Control and Optimization, 25 (1987), pp. 781–814.
[28] R. T. Rockafellar, *Hamiltonian trajectories and duality in the optimal control of linear systems with convex costs*, SIAM Journal on Control and Optimization, 27 (1989), pp. 1007–1025.
[29] R. T. Rockafellar, *Hamilton-Jacobi theory and parametric analysis in fully convex problems of optimal control*, Journal of Global Optimization, 28 (2004), pp. 419–431.
[30] R. T. Rockafellar and R. Wets, *Variational Analysis*, Springer-Verlag, 2009.
[31] R. T. Rockafellar and P. R. Wolenski, *Convexity in Hamilton-Jacobi theory I: Dynamics and duality*, SIAM Journal on Control and Optimization, 32 (2000), pp. 442–470.
[32] R. T. Rockafellar and P. R. Wolenski, *Convexity in Hamilton-Jacobi theory II: Envelope representations*, SIAM Journal on Control and Optimization, 39 (2000), pp. 1351–1372.
[33] R. Vinter, *Optimal Control*, Springer, 2010.
[34] C. Zalinescu, *Convex Analysis in General Vector Spaces*, World Scientific, 2002.