On the metric regularity of affine optimal control problems

M. Quincampoix, T. Scarinci, V.M. Veliov

Research Report 2018-08

November 2018

ISSN 2521-313X

Operations Research and Control Systems
Institute of Statistics and Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Wiedner Hauptstraße 8 / E105-4
1040 Vienna, Austria
E-mail: orcos@tuwien.ac.at

On the metric regularity of affine optimal control problems *

M. Quincampoix[†], T. Scarinci[‡], V.M. Veliov[§]

Abstract

The paper establishes properties of the type of (strong) metric regularity of the set-valued map associated with the system of necessary optimality conditions for optimal control problems that are affine with respect to the control (shortly, *affine problems*). It is shown that for such problems it is reasonable to extend the standard notions of metric regularity by involving two metrics in the image space of the map. This is done by introducing (following an earlier paper by the first and the third author) the concept of (strong) *bi-metric* regularity in a general space setting. Lyusternik-Graves-type theorems are proved for (strongly) bi-metrically regular maps, which claim stability of these regularity properties with respect to "appropriately small" perturbations. Based on that, it is proved that in the case of a map associated with affine optimal control problems, the strong bi-metric regularity is invariant with respect to linearization. This result is complemented with a sufficient condition for strong bi-metric regularity for linear-quadratic affine optimal control problems, which applies to the "linearization" of a nonlinear affine problem. Thus the same conditions are also sufficient for strong bi-metric regularity in the nonlinear affine problem.

Key words: optimal control, affine problems, metric regularity, solution stability, perturbed control problems.

AMS subject classifications: 49J30, 49K15, 49K40.

1 Introduction

Properties of set-valued mappings known under the general name "metric regularity" play an important role in the theory of optimization, and in optimal control, in particular. In optimization, in general, the regularity properties of the set-valued mappings appearing in the corresponding necessary optimality conditions are of special interest; such properties provide the basis for analysis of stability of the solutions with respect to perturbations and approximations. In the optimal control context, the regularity theory has been developed mainly for problems satisfying

*This research is supported by the Austrian Science Foundation (FWF) under grant No P31400-N32. The second author is supported by the Doctoral Program "Vienna Graduate School on Computational Optimization" funded by the Austrian Science Fund (FWF), project No W1260-N35.

[†]Laboratoire de Mathématiques de Bretagne Atlantique, unité CNRS UMR6205, Université de Brest, France, marc.quincampoix@univ-brest.fr.

[‡]Dept. of Statistics and Operations Research, University of Vienna, Austria, teresa.scarinci@univie.ac.at.

[§]Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Austria, vladimir.veliov@tuwien.ac.at.

certain *coercivity conditions* (see e.g. [5] and the recent paper [6]). The goal of the present paper is to contribute to the metric regularity theory for optimal control problems that are affine with respect to the control (further to be called *affine problems*). This requires some conceptual enhancement of the general regularity theory, as explained below.

Metric regularity

We recall two basic definitions (see e.g. [4, Section 3G] and [10]) for a set-valued map $\Phi : Y \rightrightarrows Z$, where (Y, d_Y) and (Z, d_Z) are metric spaces. Let $I\!B_Y(y; a)$ and $I\!B_Z(z; a)$ be the closed balls with radius $a > 0$ centered at $y \in Y$ and $z \in Z$, respectively. Moreover, let $\operatorname{gph}\Phi := \{(y, z) \in Y \times Z : z \in \Phi(y)\}$ denotes the graph of Φ, and $\Phi^{-1} : Z \rightrightarrows Y$ be its inverse defined as $\Phi^{-1}(z) := \{y \in Y : z \in \Phi(y)\}$.

The map $\Phi : Y \rightrightarrows Z$ is *metrically regular* (MR) at $\hat{y} \in Y$ for $\hat{z} \in Z$ with constants $a > 0$, $b > 0$ and $\kappa \geq 0$ if $(\hat{y}, \hat{z}) \in \operatorname{gph}(\Phi)$ and for all $z, z' \in I\!B_Z(\hat{z}; b)$ it holds that

$$e_Y(\Phi^{-1}(z) \cap I\!B_Y(\hat{y}; a), \Phi^{-1}(z')) \leq \kappa d_Z(z, z'),$$

where $e_Y(A, B) := \sup_{a \in A} d_Y(a, B) = \sup_{a \in A} \inf_{b \in B} d_Y(a, b)$ denotes the excess of $A \subset Y$ beyond $B \subset Y$.

If Φ is (MR) at $\hat{y}$ for $\hat{z}$ and $\Phi^{-1}(z) \cap I\!B(\hat{y}; a)$ is a singleton for every $z \in I\!B_Z(\hat{z}; b)$, then Φ is called strongly metrically regular (SMR) at $\hat{y}$ for $\hat{z}$. In this case the mapping $I\!B_Z(\hat{z}; b) \ni z \mapsto \Phi^{-1}(z) \cap I\!B(\hat{y}; a)$ is single-valued and Lipschitz continuous with Lipschitz constant κ.

We refer to [4, 10] for several equivalent definitions of MR; see also [9] and [10, Chapter 2.7] and the references therein for extensions of the notion to Hölder and nonlinear metric regularity.

It is of crucial importance that the MR and SMR are invariant with respect to functional single-valued perturbations of the mapping Φ, provided that these perturbations are sufficiently small in an appropriate sense. This type of results are known as Lyusternik-Graves-type and Robinson-type theorems (see [4, 10] for a detailed historical account). We recall one theorem of this type concerning SMR, which is close to the extension we develop in the present paper. It is a part of [4, Theorem 5G.3][1]:

Theorem 1.1. *Let Y and Z be Banach spaces (so that d_Y and d_Z are induced by the norms) and let $\operatorname{gph}\Phi$ be closed. Let a, b and κ be positive scalars such that Φ is strongly metrically regular at $\hat{y}$ for $\hat{z}$ with constants a, b and κ. Let $\mu > 0$ be such that $\kappa\mu < 1$ and let $\kappa' > \kappa/(1 - \kappa\mu)$. Then for every positive a' and b' such that*

$$a' \leq a/2, \quad 2\mu a' + 2b' \leq b \quad \text{and} \quad 2\kappa' b' \leq a'$$

and for every function $\varphi : Y \to Z$ satisfying

$$d_Z(\varphi(\hat{y}), 0) \leq b' \quad \text{and} \quad d_Z(\varphi(y), \varphi(y')) \leq \mu\, d_Y(y, y') \quad \text{for every } y, y' \in I\!B_Y(\hat{y}; 2a'),$$

the mapping $z \mapsto (\varphi + \Phi)^{-1}(z) \cap I\!B_Y(\hat{y}; a')$ is a Lipschitz continuous function on $I\!B_Z(\hat{z}; b')$ with Lipschitz constant κ'.

[1] See Errata and Addenda at https://sites.google.com/site/adontchev/

An important consequence of this theorem is that SMR of a mapping $\varphi + \Phi$ with a Fréchet differentiable function $\varphi : Y \to Z$ is equivalent to SMR of the partial linearization $\varphi(\hat{y}) + \varphi(\hat{y})(y-\hat{y})+\Phi(y)$. This result was proved (in a finite-dimensional space setting) by S. Robinson in his seminal paper [14] in the case of $\Phi(y) = N_{\mathcal{U}}(y)$, where $\mathcal{U}$ is a closed convex set and $N_{\mathcal{U}}(y)$ is the normal cone to $\mathcal{U}$ at y. Thus the result concerns a variational inequality of the form (the variable y is changed to u for further convenience)

$$\varphi(u) + N_{\mathcal{U}}(u) \ni 0. \tag{1.1}$$

Control-affine problems and lack of metric regularity

Let us introduce the affine optimal control problem that will be in the focus of the present paper:

$$\begin{array}{ll} \text{minimize} & \ell(x(T)) \\ \text{subject to} & \dot{x}(t) = f(t,x(t)) + g(t,x(t))u(t), \quad \text{a.e. on } [0,T], \\ & u(t) \in U := [0,1]^m, \\ & x(0) = x_0. \end{array} \tag{$\mathcal{P}$}$$

Here $x(t) \in \mathbb{R}^n$ is the state at time t, $u(t) \in \mathbb{R}^m$ is the control, $\ell : \mathbb{R}^n \to \mathbb{R}$, $f : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$, $g : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are given functions, the time horizon $[0,T]$ and the initial condition x_0 are fixed and given.

It is well known that the Pontryagin maximum (here minimum) principle involves a variational inequality of the type (1.1), where, however, the convex set $\mathcal{U}$ is a subset of the space $L^\infty(0,T)$, namely, $\mathcal{U} = \{u \in L^\infty : u(t) \in U \text{ for a.e. } t \in [0,T]\}$. Here the function $\varphi(u)$ is the derivative with respect to u of the Hamiltonian associated with problem $(\mathcal{P})$ (the details are described in Section 3). It has the special form $\varphi(u)(t) = \hat{\sigma}(t)$, where $\hat{\sigma}$ is the so-called *switching function* associated with the problem and a reference optimal solution. Observe that the function $\hat{\sigma}$ is not directly dependent on the control u due to the affine dependence of differential equation in $(\mathcal{P})$ on u. This is a substantial difference with coercive problems, and it creates a trouble for establishing MR or SMR of the mapping in (1.1). In order to explain this trouble and the reason for which we extend the notions of MR and SMR, we consider the following simple example.

Let $U = [-1,1]$, $\hat{\sigma}(t) = -1+t$, $t \in [0,2]$. The variational inequality (1.1) takes the point-wise form $\hat{\sigma}(t) + N_U(u(t)) \ni 0$, and this inclusion has the unique solution (modulo a set of measure zero)

$$\hat{u}(t) = \begin{cases} 1 & \text{if } t \in [0,1), \\ -1 & \text{if } t \in (1,2]. \end{cases}$$

The (control) functions $u : [0,2] \to U$ belong to L^∞, but in order to obtain some kind of regularity of the mapping $u(\cdot) \mapsto \hat{\sigma}(\cdot) + N_U(u(\cdot))$ one has to consider the L^1-metric in this space.

Now, consider the disturbed inclusions

$$\sigma(t) + N_U(u(t)) \ni 0, \qquad \sigma'(t) + N_U(u(t)) \ni 0, \tag{1.2}$$

with two "special" continuous disturbances, σ and σ' of $\hat{\sigma}$, such that where $\|\sigma - \hat{\sigma}\|_{L^\infty} = \|\sigma' - \hat{\sigma}\|_{L^\infty} = b$ (b is a fixed positive number) but $\|\sigma - \sigma'\|_{L^\infty} \le \varepsilon$ where $\varepsilon > 0$ is arbitrarily small. The definition of these disturbances is clearly presented on Figure 1.1.

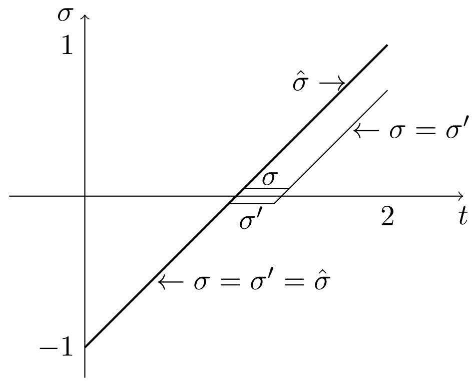

Figure 1.1: Here $\|\hat\sigma - \sigma\|_{L^\infty} = \|\hat\sigma - \sigma'\|_{L^\infty} = b$, and $\|\sigma - \sigma'\|_{L^\infty} = \varepsilon << b/4$. The number b equals the length of the three "short" horizontal segments, ε is the distance between the lower and the upper one.

The corresponding solutions u and u' of (1.2) are unique and $\|u-u'\|_1 \geq b$, since $u(t)-u'(t) = 2$ on a set of measure at least $b/2$ (if ε is small enough). If the mapping $\hat\sigma + N_U(u)$ were MR with constants a, b, κ at $\hat u$ for $\hat\sigma$ in the spaces $Y = L^1$, $Z = C$, then the inequalities

$$b \leq \|u - u'\|_{L^1} \leq \kappa\|\sigma - \sigma'\|_{L^\infty} \leq \kappa\varepsilon$$

had to be fulfilled. This is a contradiction, since $\varepsilon > 0$ can be chosen arbitrarily small, interdependently on b.

On the other hand, if we consider only disturbances from the space $Z = W^{1,\infty}$, one can easily verify that for any $\sigma, \sigma' \in W^{1,\infty}$ with $\|\sigma - \hat\sigma\|_{W^{1,\infty}}$, $\|\sigma' - \hat\sigma\|_{W^{1,\infty}} \leq b := 1/2$ the corresponding solutions u and u' of (1.2) are again unique and

$$\|u - u'\|_{L^1} \leq 4\|\sigma - \sigma'\|_{W^{1,\infty}}.$$

This means that the mapping $\hat\sigma(\cdot) + N_U(u(\cdot))$ is SMR at $\hat u$ for zero in the space settings L^1 for u and $W^{1,\infty}$ for the disturbances. This, however is not a satisfactory result. In fact, one can easily prove that the inequality

$$\|u - u'\|_{L^1} \leq 4\|\sigma - \sigma'\|_{L^\infty} \tag{1.3}$$

holds, provided that $\|\sigma - \hat\sigma\|_{W^{1,\infty}} \leq 1/2$ and $\|\sigma' - \hat\sigma\|_{W^{1,\infty}} \leq 1/2$.

As a recapitulation, we formulate the above observations in the following way. The mapping $\hat\sigma + N_U(u)$ is not MR in the space setting $Y = L^1$, $Z = L^\infty$ but is SMR in the spaces $Y = L^1$, $Z = W^{1,\infty}$ with constants $a = +\infty$, $b = 1/2$ and $\kappa = 4$. However, this regularity property is too weak, because although the disturbances have to be close to $\hat\sigma$ in the $W^{1,\infty}$-norm, the Lipschitz property in the definition of SMR holds in the L^∞-norm in Z (see 1.3). Therefore , in order to grasp in a right way the situation, we have to use two norms in the space $W^{1,\infty}$: the $W^{1,\infty}$-norm for the neighborhood $I\!B_Z(\hat\sigma)$ and the L^∞-norm for the Lipschitz property.

Aims and content of the paper

For the reasons explained in the subsection above, in this paper we introduce in an abstract space setting the notions of *bi-Metric Regularity* (bi-MR) and *Strong bi-Metric Regularity* (Sbi-MR). This done in Section 2 (the Sbi-MR) and in Section 5 (the bi-MR), where we also prove Lyusternik-Graves-type theorems for these notions of regularity. We mention that the notion of Sbi-MR was first introduced in [13] and utilized in the context of linear optimal control problems. The Lyusternik-Graves-type theorem for Sbi-MR maps proved there is substantially more restrictive than the one obtained later in [12] and presented in Section 2.

The Lyusternik-Graves-type theorems claim stability of the respective regularity properties under perturbations that are small in an appropriate sense. Thanks to that, in Section 3 we prove that the Sbi-MR of the map appearing in the Pontryagin necessary optimality conditions for problem $\mathcal{P}$ is invariant with respect to linearization. Thus, any sufficient conditions for Sbi-MR in the linearized version of problem $\mathcal{P}$ are also sufficient for Sbi-MR of the map associated with problem $\mathcal{P}$ itself. For this reason, in Section 3 we also prove sufficient conditions for Sbi-MR for linear-quadratic affine optimal control problems that arise, in particular, as linearization of problem $\mathcal{P}$. These conditions extend the ones in [12] by weaker requirements about the data, which is essential if the linear-quadratic problem results from linearization. The obtained sufficient conditions are, nevertheless, restrictive since they require a purely bang-bang structure of the optimal control and additional properties, which exclude the case of singular arcs. Here we mention that similar conditions have been involved in various contexts (including sufficiency of the Pontryagin conditions and error analysis of approximation methods) in several papers out of which we mention [1, 7, 11, 8, 2, 3].

As another application of our version of the Lyusternik-Graves theorem applied to affine optimal control problems, we prove in Section 4 a result about stability of the Sbi-MR for affine problems with disturbances, extending to the non-linear case results from [7] and [13].

2 Strong bi-metric regularity: abstract setting and a Lyusternik-Graves-type theorem

Let (Y, d_Y), (Z, d_Z), $(\tilde{Z}, d_{\tilde{Z}})$ be metric spaces, with $\tilde{Z} \subset Z$ and $d_Z \le d_{\tilde{Z}}$ on $\tilde{Z}$. Denote by $I\!\!B_Y(\hat{y}; a)$, $I\!\!B_Z(\hat{z}; b)$ and $I\!\!B_{\tilde{Z}}(\hat{z}; b)$ the closed balls in the metric spaces (Y, d_Y), (Z, d_Z) and $(\tilde{Z}, d_{\tilde{Z}})$ with radius $a > 0$ or $b > 0$ centered at $\hat{y}$ and $\hat{z}$, respectively.

Given a set-valued map (or multifunction) $\Phi : Y \rightrightarrows Z$, $\operatorname{gph} \Phi := \{(y, z) \in Y \times Z : \; z \in \Phi(y)\}$ denotes the graph of Φ. Its inverse, $\Phi^{-1} : Z \rightrightarrows Y$, is the multifunction defined as $\Phi^{-1}(z) := \{y \in Y : z \in \Phi(y)\}$.

As explained in the introduction, in order to grasp affine optimal control problems we involve two norms in the definition of strong metric regularity. The following definition is a modification of [13, Definition 2.1] (see [12]).

Definition 2.1. The map $\Phi : Y \rightrightarrows Z$ is *strongly bi-metrically regular* (Sbi-MR) at $\hat{y} \in Y$ for $\hat{z} \in \tilde{Z}$ with constants $\kappa \ge 0$, $a > 0$ and $b > 0$ if $(\hat{y}, \hat{z}) \in \operatorname{graph}(\Phi)$ and the following properties are fulfilled:

(i) the mapping $I\!\!B_{\tilde{Z}}(\hat{z}; b) \ni z \mapsto \Phi^{-1}(z) \cap I\!\!B_Y(\hat{y}; a)$ is single-valued;

(ii) for all $z, z' \in I\!B_{\tilde Z}(\hat z; b)$

$$d_Y(\Phi^{-1}(z) \cap I\!B_Y(\hat y; a), \Phi^{-1}(z') \cap I\!B_Y(\hat y; a)) \le \kappa d_Z(z, z'). \tag{2.1}$$

It is important to notice that in this definition only 'disturbances" z, z' belonging to the smaller space $\tilde Z$ are allowed, and they have to be sufficiently close to reference point $\hat z$ in the stronger metric $d_{\tilde Z}$. However, the Lipschitz property (2.1) holds with the (weaker) metric d_Z. This is the difference with the standard definition of strong metric regularity (see e.g. [4, Section 3G] and [10]), where a single metric in Z is involved. Clearly, Sbi-MR implies strong metric regularity of the set-valued mapping $y \mapsto \Phi(y) \cap \tilde Z$ at $\hat y$ for $\hat z$ in the metric $d_{\tilde Z}$, but then the Lipschitz property is also in the metric $d_{\tilde Z}$ (rather than d_Z), which can be a much weaker property. This is the case, in particular, when dealing with mappings that arise in optimal control problems for affine systems.

The next result resembles the main features of the Lyusternik-Graves-type Theorem 2.1 in [13] but under substantially weaker requirements, as explained below. It resembles Theorem 4.2 in [12].

Theorem 2.2. *Let the metric space Y be complete, let Z be a linear space and $\tilde Z$ a subspace of Z. Let the two metrics, d_Z and $d_{\tilde Z}$, respectively in Z and $\tilde Z$ be shift-invariant and $d_Z \le d_{\tilde Z}$ on $\tilde Z$. Let the set-valued map $\Phi : Y \rightrightarrows Z$ be strongly bi-metrically regular at $\hat y \in Y$ for $\hat z \in \tilde Z$ with constants κ, a, b. Let $\mu > 0$ and κ' be such that $\kappa\mu < 1$ and $\kappa' \ge \kappa/(1 - \kappa\mu)$. Then for every positive constants a', b', and γ satisfying*

$$a' \le a, \quad b' + \gamma \le b, \quad \kappa b' \le (1 - \kappa\mu)a', \tag{2.2}$$

and for every function $\varphi : Y \to \tilde Z$ such that

$$d_{\tilde Z}(\varphi(\hat y), \varphi(y)) \le \gamma \quad \forall y \in I\!B_Y(\hat y; a'), \tag{2.3}$$

and

$$d_Z(\varphi(y), \varphi(y')) \le \mu d_Y(y, y') \quad \forall y, y' \in I\!B_Y(\hat y; a'), \tag{2.4}$$

the mapping $I\!B_{\tilde Z}(\hat z + \varphi(\hat y); b') \ni z \mapsto (\varphi + \Phi)^{-1}(z) \cap I\!B_Y(\hat y; a')$ is single-valued and Lipschitz continuous with constant κ' with respect to the metric d_Z.

The main improvement in the above theorem, compared with [13, Theorem 2.1], is that the Lipschitz property (2.4) is required in [13, Theorem 2.1] to be fulfilled in the stronger metric $d_{\tilde Z}$, while here it is required in d_Z. This improvement makes the theorem usable in the context of the present paper, in particular.

Corollary 2.3. *Let the assumptions of Theorem 2.2 be fulfilled with $\hat z = 0$, and let $\varphi : Y \to \tilde Z$ be such that the mapping $I\!B_{\tilde Z}(\varphi(\hat y); b') \ni z \mapsto (\varphi + \Phi)^{-1}(z) \cap I\!B_Y(\hat y; a')$ is single-valued and Lipschitz continuous with constant κ' with respect to the metric d_Z, where $a', b', \kappa' > 0$. Let, in addition, the following inequalities be fulfilled:*

$$d_{\tilde Z}(\varphi(\hat y), 0) \le b'/2, \qquad \kappa' d_Z(\varphi(\hat y), 0) \le a'/2.$$

Then

(i) the inclusion $0 \in \varphi(y) + \Phi(y)$ *has a unique solution* y^* *in* $I\!B(\hat{y}; a')$, *and*

$$d_Y(y^*, \hat{y}) \le \kappa' d_Z(\varphi(\hat{y}), 0); \tag{2.5}$$

(ii) the mapping $\varphi + \Phi$ *is strongly bi-metrically regular at* y^* *for zero.*

Before starting the proof we mention that the parameters a'', b'', κ'' of Sbi-MR of $\varphi + \Phi$ at y^* for zero can explicitly defined as any numbers satisfying the inequalities

$$a'' \le a'/2, \quad \kappa' b'' \le a'', \quad b'' \le b'/2, \quad \kappa'' \ge \kappa'.$$

Proof. Since $0 \in I\!B_{\tilde{Z}}(\varphi(\hat{y}); b')$, the solution y^* exists and is unique in $I\!B(\hat{y}; a')$. Moreover, (2.5) is fulfilled since $\hat{y} \in (\varphi + \Phi)^{-1}(\varphi(\hat{y}))$. To prove (ii) we take arbitrarily $z_i \in I\!B_{\tilde{Z}}(0; b'')$, $i = 1, 2$. We have

$$d_{\tilde{Z}}(z_i, \varphi(\hat{y})) \le d_{\tilde{Z}}(z_i, 0) + d_{\tilde{Z}}(0, \varphi(\hat{y})) \le b'' + \frac{b'}{2} \le b'.$$

Then $y_i \in (\varphi + \Phi)^{-1}(z_i) \cap I\!B_Y(\hat{y}; a')$ exist, and $d_Y(y_1, y_2) \le \kappa' d_Z(z_1, z_2)$. Moreover, $y_i \in I\!B_Y(y^*; a'')$ since

$$d_Y(y_i, y^*) \le \kappa' d_Z(z_i, 0) \le \kappa' b'' \le a''.$$

Finally, y_i is the unique element of $(\varphi + \Phi)^{-1}(z_i) \cap I\!B_Y(y^*; a'')$ since

$$I\!B_Y(y^*; a'') \subset I\!B_Y(\hat{y}; d_Y(y^*, \hat{y}) + a'') \subset I\!B_Y(\hat{y}; \kappa' d_Z(0, \varphi(\hat{y})) + a'') \subset I\!B_Y(\hat{y}; a'/2 + a'/2) = I\!B_Y(\hat{y}; a')$$

and $(\varphi + \Phi)^{-1}(z_i) \cap I\!B_Y(\hat{y}; a')$ is a singleton. Q.E.D.

3 Strong bi-metric regularity of affine problems

To begin with, we introduce the basic notation which will be used throughout the paper. The euclidean norm in $\mathbb{R}^n$ is denoted by $|\cdot|$ and $I\!B(x; a)$ stays for the closed ball in $\mathbb{R}^n$ centered at $x \in \mathbb{R}^n$ with radius $a \ge 0$. As usual, $L^1([0,T], \mathbb{R}^n)$ and $L^\infty([0,T], \mathbb{R}^n)$ are the spaces of all measurable absolutely integrable, respectively essentially bounded, functions from $[0,T]$ to $\mathbb{R}^n$; these spaces are equipped with the standard Lebesgue norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ and sometimes they will be shortened by L^1 and L^∞, respectively. For $k \in \{1, \infty\}$, $W^{1,k}([0,T], \mathbb{R}^n)$ is the space of absolutely continuous functions x mapping $[0,T]$ to $\mathbb{R}^n$ such that the derivative x has a finite L^k-norm; moreover, $W^{1,k}_{x_0}([0,T], \mathbb{R}^n) := \{x \in W^{1,k}([0,T], \mathbb{R}^n) : x(0) = x_0\}$. These spaces are equipped with the standard corresponding Sobolev norm, $\|\cdot\|_{1,k}$, and possibly shortened by $W^{1,k}$ or $W^{1,k}_{x_0}$. For a given matrix M, $M^\top$ and $\|M\|$ denote the transpose and operator norm of M, respectively.

3.1 The problem and the concept of strong bi-metric regularity

We consider the affine problem $(\mathcal{P})$ on the following assumptions.

Assumption (A1). (Smoothness of the data) The functions f, g and ℓ, and their partial derivatives f_x, g_x, ℓ_x, f_{xt}, g_{xt}, f_{xx}, g_{xx}, ℓ_{xx} are continuous. Moreover, f_{xx}, g_{xx} and ℓ_{xx} are locally Lipschitz continuous in x uniformly in t.

The set of admissible controls, $\mathcal{U}$, consists of all Lebesgue measurable functions $u : [0,T] \to U$. The system of necessary optimality conditions provided by the Pontryagin minimum principle reads as follows:

$$\begin{aligned}
&0 = \dot{x}(t) - f(t,x(t)) - g(t,x(t))u(t),\\
&0 = \dot{p}(t) + \left(f(t,x(t)) + g(t,x(t))u(t)\right)_x^\top p(t),\\
&0 \in g(t,x(t))^\top p(t) + N_U(u(t)),\\
&0 = p(T) - \nabla\ell(x(T)),
\end{aligned} \tag{PMP}$$

where $\nabla\ell = \ell_x^\top$ is the gradient of ℓ and $N_U(u)$ is the normal cone to U at u defined in the usual way:

$$N_U(u) := \begin{cases} \emptyset & \text{if } u \notin U \\ \{l \in \mathbb{R}^m : \langle l, v-u\rangle \le 0 \ \forall v \in U\} & \text{if } u \in U. \end{cases}$$

System (PMP) will be reformulated as an abstract generalized equation, for which appropriate spaces for (x,p,u) and the image of the right-hand side in (PMP) should be chosen. As shown in [12] for the case of linear-quadratic problems, it is appropriate (in the context of Sbi-MR of the set-valued mapping associated with (PMP)) to consider (x,p,u) as elements of the space $W^{1,1}_{x_0}([0,T],\mathbb{R}^n) \times W^{1,1}([0,T],\mathbb{R}^n) \times \mathcal{U}$, where $\mathcal{U}$ is endowed with the metric induced by the L_1-norm (note that $\mathcal{U}$ is a complete metric space with this metric).

However, this choice of the space above is too redundant: it contains elements (x,p,u) with arbitrarily large norms (even with infinite L^∞-norms of $\dot{x}$ and $\dot{p}$), which are irrelevant to the considered optimal control problem. To avoid this (troublesome) redundancy we first make the following natural assumption.

Assumption (A2). There exists a number $M > 0$ such that for every $u \in \mathcal{U}$ the corresponding solutions $x \in W^{1,1}_{x_0}$ and $p \in W^{1,1}$ (with $p(T) = \nabla\ell(x(T))$) of the differential equations in (PMP) exist and satisfy $x(t), p(t), \dot{x}(t), \dot{p}(t) \in \mathbb{B}(0;M)$.

Then the following space setting is relevant to the problem into consideration:

$$Y := \{(x,p,u) \in W^{1,1}_{x_0}([0,T],\mathbb{R}^n) \times W^{1,1}([0,T],\mathbb{R}^n) \times \mathcal{U} :\ x(t), p(t), \dot{x}(t), \dot{p}(t) \in \mathbb{B}(0;2M)\}.$$

The set Y will be considered as a metric space where the distance between two elements $y_1 = (x_1,p_1,u_1)$ and $y_2 = (x_2,p_2,u_2)$ of Y is defined as

$$d_Y(y_1,y_2) = \|y_1 - y_2\|_Y := \|x_1 - x_2\|_{1,1} + \|p_1 - p_2\|_{1,1} + \|u_1 - u_2\|_1.$$

It is easy to verify that the metric space Y is complete.

In addition, we define the spaces

$$
\begin{aligned}
Z &:= L^\infty([0,T],\mathbb{R}^n) \times L^\infty([0,T],\mathbb{R}^n) \times L^\infty([0,T],\mathbb{R}^m) \times \mathbb{R}^n, \\
\tilde{Z} &:= L^\infty([0,T],\mathbb{R}^n) \times L^\infty([0,T],\mathbb{R}^n) \times W^{1,\infty}([0,T],\mathbb{R}^m) \times \mathbb{R}^n,
\end{aligned}
$$

endowed, respectively, with the following two norms:

$$
\begin{aligned}
\|(\xi,\pi,\rho,\nu)\|_Z &:= \|\xi\|_1 + \|\pi\|_1 + \|\rho\|_\infty + |\nu|, \\
\|(\xi,\pi,\rho,\nu)\|_\sim &:= \max\{1,T\}\,(\|\xi\|_\infty + \|\pi\|_\infty) + \|\rho\|_{1,\infty} + |\nu|.
\end{aligned}
$$

The distances introduced by these norms are denoted by d_Z and $d_{\tilde{Z}}$, respectively. Clearly, $\|z\|_Z \le \|z\|_\sim$ for every $z \in \tilde{Z}$ (the multiplier $\max\{1,T\}$ is just to ensure this inequality). Similarly as in Section 2, the closed balls in the spaces (Y, d_Y), $(Z, \|\cdot\|_Z)$, $(\tilde{Z}, \|\cdot\|_\sim)$ will be denoted by $I\!\!B_Y$, $I\!\!B_Z$ and $I\!\!B_{\tilde{Z}}$, respectively (with specified center and radius as arguments).

Now, we may recast system (PMP) as the generalized equation

$$
\psi(y) + F(y) \ni 0, \tag{3.1}
$$

where $y := (x,p,u) \in Y$, and the function $\psi : Y \to Z$ and the set-valued map $F : Y \rightrightarrows Z$ are defined as

$$
\psi(y)(t) = \begin{pmatrix} \dot{x}(t) - f(t,x(t)) - g(t,x(t))u(t) \\ \dot{p}(t) + (f(t,x(t)) + g(t,x(t))u(t))_x^\top p(t) \\ g(t,x(t))^\top p(t) \\ p(T) - \nabla \ell(x(T)) \end{pmatrix}, \quad F(y) := \begin{pmatrix} 0 \\ 0 \\ N_{\mathcal{U}}(u) \\ 0 \end{pmatrix}. \tag{3.2}
$$

Here,

$$
N_{\mathcal{U}}(u) := \begin{cases} \emptyset & \text{if } u \notin \mathcal{U} \\ \{v \in L^\infty :\ v(t) \in N_U(u(t)) \text{ for a.e. } t \in [0,T]\} & \text{if } u \in \mathcal{U}, \end{cases}
$$

is the normal cone to the set $\mathcal{U}$ at u in the space L^1.

Due to the definition of the spaces Y and Z, it is obvious that ψ and F map Y into Z.

In the above setting the Sbi-MR property for system (PMP) associated with problem $(\mathcal{P})$ translates as follows.

Property Sbi-MR of the mapping $\psi + F$ in (3.1) at a solution $\hat{y}$: There exist positive numbers a, b and κ such that the mapping $I\!\!B_{\tilde{Z}}(0;b) \ni z \mapsto (\psi+F)^{-1}(z) \cap I\!\!B_Y(\hat{y};a)$ is single-valued and

$$
d_Y\big((\psi+F)^{-1}(z') \cap B_Y(\hat{y};a),\, (\psi+F)^{-1}(z) \cap B_Y(\hat{y};a)\big) \le \kappa d_Z(z,z') \tag{3.3}
$$

for every $z, z' \in I\!\!B_{\tilde{Z}}(0;b)$.

3.2 Invariance of the strong bi-metric regularity with respect to linearization

Let $\hat{y} = (\hat{x}, \hat{p}, \hat{u})$ be a (from now on fixed) solution of the optimality conditions (PMP). According to Assumption (A2) we have $\hat{y} \in Y$, and $\hat{y}$ solves (3.1). We shall formally introduce a partial linearization of (3.1) in the usual way, introducing first the matrices

$$A(t) := (f(t, \hat{x}(t)) + g(t, \hat{x}(t))\hat{u}(t))_x\,, \tag{3.4}$$

$$B(t) := g(t, \hat{x}(t)), \tag{3.5}$$

$$W(t) := \Big((f(t, \hat{x}(t)) + g(t, \hat{x}(t))\hat{u}(t))_x^\top \, \hat{p}(t) \Big)_x\,, \tag{3.6}$$

$$S(t) := \Big(g(t, \hat{x}(t))^\top \hat{p}(t) \Big)_x^\top\,, \tag{3.7}$$

$$K := \ell_{xx}(\hat{x}(T)). \tag{3.8}$$

Then we introduce the mapping $\Gamma(\hat{y}) : Y - \hat{y} \to Z$ acting on any $\Delta y := (\Delta x, \Delta p, \Delta u) := (x - \hat{x}, p - \hat{p}, u - \hat{u}) \in Y - \hat{y}$ as follows:

$$\Gamma(\hat{y})\Delta y := \begin{pmatrix} (\Delta x)^{\cdot} - A\Delta x - B\Delta u \\ (\Delta p)^{\cdot} + A^\top \Delta p + W\Delta x + S\Delta u \\ B^\top \Delta p + S^\top \Delta x \\ \Delta p(T) - K\Delta x(T) \end{pmatrix}. \tag{3.9}$$

Notice that the function $y \mapsto \Gamma(\hat{y})(y - \hat{y})$ maps Y to Z. Then we have the identity

$$\psi(y) = \psi(\hat{y}) + \Gamma(\hat{y})(y - \hat{y}) + \varphi(y), \qquad \forall\, y \in Y,$$

where the mapping $\varphi : Y \to \tilde{Z}$ is defined (suppressing the argument t) as

$$\varphi(y) = \begin{pmatrix} -f(x) - g(x)u + f(\hat{x}) + g(\hat{x})\hat{u} + A\Delta x + B\Delta u \\ (f(x) + g(x)u)_x^\top \, p - (f(\hat{x}) + g(\hat{x})\hat{u})_x^\top \, \hat{p} - A^\top \Delta p - W\Delta x - S\Delta u \\ g(x)^\top p - g(\hat{x})^\top \hat{p} - B^\top \Delta p - S^\top \Delta x \\ -\nabla \ell(x(T)) + \nabla \ell(\hat{x}(T)) + K\Delta x(T) \end{pmatrix}. \tag{3.10}$$

Thus the generalized equation (3.1), which is satisfied by $\hat{y}$, can be equivalently reformulated as

$$\psi(\hat{y}) + \Gamma(\hat{y})(y - \hat{y}) + \varphi(y) + F(y) \ni 0. \tag{3.11}$$

This representation enables us to use the Lyusternik-Graves-type Theorem 2.2 and to deduce the following theorem, whose proof is given in Appendix.

Theorem 3.1. *Let Assumptions (A1) and (A2) be fulfilled, and let the mappings ψ, F and Γ be defined as in (3.2) and (3.9). Then strong bi-metric regularity of the mapping $y \mapsto \psi(y) + F(y)$ at $\hat{y}$ for 0 (in the above defined spaces and metrics) is equivalent to the strong bi-metric regularity of the mapping $y \mapsto \psi(\hat{y}) + \Gamma(\hat{y})(y - \hat{y}) + F(y)$ at $\hat{y}$ for 0.*

We mention that the partly linearized generalized equation

$$\psi(\hat{y}) + \Gamma(\hat{y})(y - \hat{y}) + F(y) \ni 0 \tag{3.12}$$

represents the Pontryagin system for the optimal control problem

$$
\begin{aligned}
&\min \Big\{ J(x,u) := \tfrac{1}{2} x(T)^\top K x(T) + q_T^\top x(T) \\
&\qquad\qquad + \textstyle\int_0^T \Big(\tfrac{1}{2} x(t)^\top W(t) x(t) + x(t)^\top S(t) u(t) + q(t)^\top x(t) + e(t)^\top u(t) \Big)\, \mathrm{d}t \Big\}, \\
&\dot{x}(t) = A(t)x(t) + B(t)u(t) + d(t), \quad x(0) = x_0, \quad t \in [0,T], \\
&u(t) \in U := [0,1]^m,
\end{aligned}
\tag{LP}
$$

where the matrices appearing in this problem are defined in (3.4)-(3.8) and the vectors $d(t)$, $q(t)$, $e(t)$ and q_T, resulting from the expression $\psi(\hat{y}) - \Gamma(\hat{y})\hat{y}$ in (3.12), are given by (we skip the argument t):

$$
d := f(\hat{x}) - A\hat{x}, \qquad q := -W\hat{x} - S\hat{u}, \tag{3.13}
$$

$$
e := -S^\top \hat{x}, \qquad q_T := \hat{p}(T) - K\hat{x}(T). \tag{3.14}
$$

3.3 Sufficient conditions for strong bi-metric regularity of affine problems

According to Theorem 3.1, in order to obtain SbiMR of the Pontryagin system (PMP) (or equivalently, of generalized equation (3.1)), associated with the affine problem ($\mathcal{P}$) and a given reference point $\hat{y} = (\hat{x}, \hat{p}, \hat{u}) \in Y$, it is enough to obtain the same property for the partially linearized system (3.12), which has the same solution $\hat{y}$. As mentioned in the end of the previous session, the latter represents the Pontryagin system for the optimal control problem (LP).

For this reason, in the present section we consider the linear-quadratic affine problem (LP), now for general data A, B, W, S, d, e, q, not necessarily in the form of those resulting from linearization. The generalized equation (Pontryagin system) associated with this problem is

$$
L(y) + F(y) \ni 0, \tag{3.15}
$$

where

$$
L(y) = \begin{pmatrix} \dot{x} - Ax - Bu - d \\ \dot{p} + A^\top p + Wx + Su + q \\ B^\top p + S^\top x + e \\ p(T) - Kx(T) \end{pmatrix}
$$

and F is defined as in (3.2). The so-called switching function for a reference solution y of (3.15) is defined as

$$
\sigma(t) = B(t)^\top p(t) + S(t)^\top x(t) + e(t).
$$

Below we fix a reference solution $\hat{y} = (\hat{x}, \hat{p}, \hat{u}) \in Y$ of (3.15).

Assumption (B1). The functions A, W, d, q are piecewise continuous, B, S and e are Lipschitz continuous with piecewise continuous first derivatives. The matrices $W(t)$ and $S^\top(t)B(t)$ are symmetric for every $t \in [0,T]$.

We clarify, that here "piecewise continuity" of a function ω on $[0,T]$ means that there exist finite number of points $t_0 = 0 < t_1 \ldots, t_k = T$ such that on each interval (t_i, t_{i+1}) the function ω is continuous, with a finite right limit at t_i and left limit at t_{i+1}, $i = 0, \ldots, k-1$.

Assumption (B2). It holds that

$$z(T)^\top K z(T) + \int_0^T \Big(z(t)^\top W(t) z(t) + 2z(t)^\top S(t) v(t) \Big) \,\mathrm{d}t \ge 0$$

for any pair $(z, v) \in W^{1,\infty} \times L^\infty$ solving

$$\begin{aligned} \dot z(t) &= A(t)z(t) + B(t)v(t) \text{ a.e. on } [0,T], \\ v(t) &\in [-1,1], \\ z(0) &= 0. \end{aligned}$$

Assumption (B3). For the switching function $\hat\sigma(t) := B(t)^\top \hat p(t) + S(t)^\top \hat x(t) + e(t)$ there exist real numbers $\alpha, \tau > 0$ such that for all $j \in \{1, \dots, m\}$ and $s \in [0,T]$ with $\hat\sigma_j(s) = 0$ (the j-th component of $\hat\sigma$) we have

$$|\hat\sigma_j(t)| \ge \alpha |t - s| \qquad \forall t \in [s-\tau, s+\tau] \cap [0,T].$$

Remark 3.2. When the data of the problem (LP) come from linearization, as in the previous subsection, then the inclusion (3.15) is identical with (3.12). Moreover, due to the definition of e in (3.14), the switching function corresponding to the reference solution $\hat y$ in this case is

$$\hat\sigma(t) = B(t)^\top \hat p(t) + S(t)^\top \hat x(t) + e(t) = g(t, \hat x(t))^\top \hat p(t).$$

Thus, $\hat\sigma$ is Lipschitz continuous and coincides with the switching function of the nonlinear problem (see the third relation in (PMP) evaluated at $(\hat x,\, \hat p)$). Moreover, the requirements in the first sentence in Assumption (B1) are fulfilled if the reference optimal control $\hat u$ in the non-linear problem $(\mathcal{P})$, hence also in the linearized problem (LP), is piecewise continuous.

Theorem 3.3. (Sufficient conditions for Sbi-MR for linear-quadratic problems.) *Let Assumptions (B1)–(B3) be fulfilled for generalized equation (3.15) associated with problem (LP) and its reference solution $\hat y = (\hat x, \hat p, \hat u)$. Then the mapping $Y \ni y \mapsto L(y) + F(y) \subset Z$ is strongly bi-metrically regular at $\hat y$ for* 0.

A similar result is proved in [12, Theorem 4.5], where conditions (B2), (B3) and (B1') below.

Assumption (B1'). The functions A, W, d, e, q are continuous, B and S have continuous first derivatives. The matrices $W(t)$ and $S^\top(t)B(t)$ are symmetric for every $t \in [0,T]$.

However, this result is not applicable if this generalized equation results from linearization, as (3.12). This is because assumption (B1') is not fulfilled here, in general, due to the possible discontinuity of A, ... , e, and the derivatives of B and S (notice that these functions may depend on t through $\hat u$, which may be discontinuous). Therefore, we need to reconsider Theorem 4.5 in [12] and prove the result in Theorem 3.3, where Assumption (B1') has been replaced by the new Assumption (B1).

Proof. The proof of Theorem 4.5 in [12], which claims the same result, but with (B1') instead of (B1), is based on Theorem 3.3 and Proposition 4.3 in [12]. Theorem 3.3 is proved in that paper on the weaker assumptions (B1)–(B3) (in the notations of [12] these are (A1)–(A3)). The only reason for which assumption (B1') is needed in that paper (instead of (B1)) is the proof of Proposition 4.3. Therefore, below we formulate and prove an improved version of Proposition 4.3 in [12] which uses (B1)–(B3) only. This will complete the proof of the present theorem. Q.E.D.

Following [12, Section 3], we introduce the following notations for a continuous function $\sigma : [0,T] \mapsto \mathbb{R}^m$ and $\delta > 0$:

$$I_j(\sigma,\delta) := \bigcup_{s\in[0,T]:\,\sigma_j(s)=0} (s-\delta, s+\delta)\cap[0,T], \qquad l_{\min}(\sigma,\delta) := \min_{1\le j\le m}\ \min_{t\in[0,T]\setminus I_j(\sigma,\delta)} |\sigma_j(t)|.$$

Clearly, $l_{\min}(\sigma,\delta) > 0$ for any $\delta > 0$.

Proposition 3.4. *(Stability of Assumption (B3).) Let assumptions (B1)–(B3) be fulfilled for the linearized problem (LP). Then there exist positive constants $\tilde{b}$, $\tilde{\alpha}$, $\tilde{\tau}$ and $\tilde{m}_0$ such that for every $z = (\xi,\pi,\rho,\nu) \in Z$ with $\|z\|_\sim \le \tilde{b}$ and for any triple $y = (x,p,u) \in Y$ solving the disturbed version $L(y) + F(y) \ni z$ of (3.15), the corresponding switching function $\sigma(t) = B(t)^\top p(t) + S(t)^\top x(t) + e(t) - \rho(t)$ satisfies Assumption (B3) with constants $\tilde{\alpha}$ and $\tilde{\tau}$ replacing α and τ, respectively, and with $l_{\min}(\sigma,\widetilde{\tau}) \ge \tilde{m}_0$.*

Proof. Let α and τ be the constants appearing in Assumption (B3), and let $j \in \{1,\ldots,m\}$ be arbitrarily fixed. Further, we consider only disturbances $z \in Z$ satisfying $\|z\|_\sim \le 1$.

First, observe that for all $t \in [0,T]$ it holds that

$$|\sigma_j(t) - \hat{\sigma}_j(t)| \le \left|\Big(B(t)^\top(p(t)-\hat{p}(t)) + S(t)^\top(x(t)-\hat{x}(t))\Big)_j\right| + |\rho_j(t)|.$$

Now we use Theorem 3.3 in [12] (applied with $b = 1$) to obtain that there is a constant c_1 such that

$$|\sigma_j(t) - \hat{\sigma}_j(t)| \le c_1\|z\|$$

for all $j = 1,\ldots,m$, $t \in [0,T]$, and $z \in Z$ with $\|z\|_\sim \le 1$. Hence,

$$|\sigma_j(t)| \ge |\hat{\sigma}_j(t)| - c_1\|z\|, \quad t \in [0,T], \quad j \in \{1,\ldots,m\}. \tag{3.16}$$

Consider (skipping the argument t) the derivative

$$\begin{aligned}
\dot{\hat{\sigma}}_j &= \Big[\dot{B}^\top\hat{p} + B^\top\dot{\hat{p}} + \dot{S}^\top\hat{x} + S^\top\dot{\hat{x}} + \dot{e}\Big]_j \\
&= \Big[\dot{B}^\top\hat{p} + B^\top(-A^\top\hat{p} - W\hat{x} - S\hat{u} - q) + \dot{S}^\top\hat{x} + S^\top(A\hat{x} + B\hat{u} + d) + \dot{e}\Big]_j \\
&= \Big[\dot{B}^\top\hat{p} + B^\top(-A^\top\hat{p} - W\hat{x} - q) + \dot{S}^\top\hat{x} + S^\top(A\hat{x} + d) + \dot{e}\Big]_j,
\end{aligned} \tag{3.17}$$

where in the last inequality we use the symmetry of $B^\top S$. Thus $\dot{\hat{\sigma}}$ is piecewise continuous.

Let us fix a component j of $\hat{\sigma}$. Observe that due to (B3) σ_j has a finite number of zeros in $[0,T]$ (not more than $T/2\tau$). Then there is a number $\tau_0 \in (0,\tau)$ such that for every zero s of $\hat{\sigma}_j$ the function $\dot{\hat{\sigma}}_j$ is continuous in each of the intervals $[s-\tau_0, s)$ and $(s, s+\tau_0]$ and the values

$$\dot{\hat{\sigma}}_j(s;t) := \begin{cases} \lim_{\xi \to s_-} \dot{\hat{\sigma}}_j(\xi) & \text{if } t < s, \\ \lim_{\xi \to s_+} \dot{\hat{\sigma}}_j(\xi) & \text{if } t \geq s \end{cases}$$

are well defined.

Let us fix an arbitrary zero $\hat{s}$ of $\hat{\sigma}_j$. Choose $\tau_1 \in (0,\tau_0)$ so small that $|\dot{\hat{\sigma}}_j(\theta) - \dot{\hat{\sigma}}_j(\hat{s};\theta)| \leq \alpha/4$ for every $\theta \in [\hat{s}-\tau_1, \hat{s}+\tau_1] \cap [0,T] \setminus \{\hat{s}\}$.

Using Assumption (B3) we obtain that for an arbitrary $t \in (\hat{s}-\tau_1, \hat{s}+\tau_1) \cap [0,T]$

$$\begin{aligned} \alpha|t-\hat{s}| &\leq |\hat{\sigma}_j(t) - \hat{\sigma}_j(\hat{s})| = \Big| \int_{\hat{s}}^t \dot{\hat{\sigma}}_j(\theta)\, d\theta \Big| \leq \Big| \int_{\hat{s}}^t \dot{\hat{\sigma}}_j(\hat{s};\theta)\, d\theta \Big| + \int_{\hat{s}}^t |\dot{\hat{\sigma}}_j(\theta) - \dot{\hat{\sigma}}_j(\hat{s};\theta)| d\theta \\ &\leq |t-\hat{s}||\dot{\hat{\sigma}}_j(\hat{s};t)| + |t-\hat{s}|\frac{\alpha}{4}. \end{aligned}$$

Hence, $|\dot{\hat{\sigma}}_j(\hat{s};t)| \geq 3\alpha/4$ for any zero $\hat{s}$ of $\hat{\sigma}_j$, $j = 1,\ldots,m$ and any $t \in [0,T]$ (we remind that $\dot{\hat{\sigma}}_j(\hat{s};t)$ takes at most two values: one on the left of $\hat{s}$ and one on the right).

The equality (3.17) holds also for σ_j (where $(\hat{x},\hat{p})$ is replaced with (x,p)), with the additional term $\left[B^\top \pi + S^\top \xi - \dot{\rho}\right]_j$ in the right-hand side. Then using Assumption (B1) and Theorem 3.3 in [12] we estimate

$$\|\dot{\sigma}_j - \dot{\hat{\sigma}}_j\|_\infty \leq c_2(\|y\| + \|\xi\|_\infty + \|\pi\|_\infty + \|\dot{\rho}\|_\infty) \leq c_3\|z\|_\sim, \tag{3.18}$$

where c_2 and c_3 are independent of j and $z \in \widetilde{\mathcal{Y}}$, $\|z\|_\sim \leq 1$.

Define $\widetilde{\tau} := \tau_1/2$ and choose the number $\widetilde{b} > 0$ in such a way that

$$c_1\widetilde{b} \leq \min\left\{ \frac{l_{\min}(\hat{\sigma}, \widetilde{\tau}/2)}{2}, \frac{\alpha\widetilde{\tau}}{4} \right\} \quad \text{and} \quad 4c_3\widetilde{b} \leq \alpha, \quad \widetilde{b} \leq 1, \tag{3.19}$$

and let $\|z\|_\sim \leq \widetilde{b}$. Since from (3.16) and the first inequality in (3.19) we have that for $t \in [0,T] \setminus I_j(\hat{\sigma}, \widetilde{\tau}/2)$

$$|\sigma_j(t)| \geq |\hat{\sigma}_j(t)| - c_1\|z\| \geq l_{\min}(\hat{\sigma}, \widetilde{\tau}/2) - c_1\widetilde{b} \geq \frac{l_{\min}(\hat{\sigma}, \widetilde{\tau}/2)}{2} > 0,$$

we obtain that any zero s of σ_j is contained in $I_j(\hat{\sigma}, \widetilde{\tau}/2)$. Thus $s \in (\hat{s} - \widetilde{\tau}/2, \hat{s} + \widetilde{\tau}/2) \cap [0,T]$ for some zero $\hat{s}$ of $\hat{\sigma}_j$.

Now take an arbitrary $t \in (s-\widetilde{\tau}, s+\widetilde{\tau}) \cap [0,T]$. Then t, $s \in (\hat{s}-\tau_1, \hat{s}+\tau_1) \cap [0,T]$ and using (3.18) and the second inequality in (3.19) we obtain that

$$\begin{aligned} |\sigma_j(t)| &= \left| \int_s^t \dot{\sigma}_j(\theta)\, d\theta \right| = \left| \int_s^t \left[\dot{\hat{\sigma}}_j(\hat{s};\theta) + (\dot{\hat{\sigma}}_j(\theta) - \dot{\hat{\sigma}}_j(\hat{s};\theta)) + (\dot{\sigma}_j(\theta) - \dot{\hat{\sigma}}_j(\theta)) \right] d\theta \right| \\ &\geq |\dot{\hat{\sigma}}_j(\hat{s};t)||t-s| - \frac{\alpha}{4}|t-s| - c_3\|z\|_\sim |t-s| \\ &\geq \frac{3\alpha}{4}|t-s| - \frac{\alpha}{4}|t-s| - \frac{\alpha}{4}|t-s| = \frac{\alpha}{4}|t-s|. \end{aligned}$$

Thus (B3) holds for σ with constants $\widetilde{\alpha} = \alpha/4$ and $\widetilde{\tau}$.

Further, for $t \in I(\hat{\sigma}, \widetilde{\tau}) \setminus I(\sigma, \widetilde{\tau})$ we have

$$|\sigma_j(t)| \geq \alpha|t - \hat{s}| - c_1\|z\| \geq \alpha|t - s| - \alpha|s - \hat{s}| - c_1\|z\| \geq \frac{\alpha\widetilde{\tau}}{4}$$

for some zeros $\hat{s}$ and s of $\hat{\sigma}$ and σ respectively. So if we set $m_0 := \min\{\frac{\alpha\widetilde{\tau}}{4}, l_{\min}(\hat{\sigma}, \widetilde{\tau})\}$ then $l_{\min}(\sigma, \widetilde{\tau}) \geq m_0$. Q.E.D.

Theorem 3.1, Theorem 3.3 and Remark 3.2 imply the following corollary.

Corollary 3.5. (Sufficient conditions for Sbi-MR for affine-control problems.) *Let Assumption (A1) be fulfilled for the nonlinear problem ($\mathcal{P}$), and let $\hat{y} = (\hat{x}, \hat{p}, \hat{u})$ be a solution of the associated generalized equation (Pontryagin system) (3.1). Moreover, let Assumptions (B2) and (B3) be fulfilled by the linearized problem (LP). Then, the mapping $Y \ni y \mapsto \psi(y) + F(y) \subset Z$ associated to the generalized equation (Pontryagin system) (PMP) of Problem ($\mathcal{P}$) is strongly bi-metrically regular at $\hat{y}$ for zero.*

We complete this section by a more familiar representation of the condition for symmetry of the matrix $S^\top B$ in Assumption (B1) when B and S result from linearization of ($\mathcal{P}$) and, thus, are given by (3.5) and (3.7).

We denote by $g_i(t, x) \in \mathbb{R}^n$, $i = 1, \cdots, m$, the i-column of the matrix $g(t, x) \in \mathbb{R}^{n \times m}$ and by $[g_i, g_j]$, $i, j = 1, \dots, m$ – the Lie bracket of g_i and g_j as functions of x, namely,

$$[g_i, g_j](t, x) := \frac{\partial g_j}{\partial x}(t, x) g_i(t, x) - \frac{\partial g_i}{\partial x}(t, x) g_j(t, x),$$

where $\frac{\partial g_i}{\partial x}(t, x)$ denotes the Jacobian matrix of the vector field $g_i(t, \cdot) : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, \dots, m$.

Below we shall prove that the linearization of problem ($\mathcal{P}$) fulfills the condition $S^\top(t)B(t) = B^\top(t)S(t)$ whenever the vector fields $g_i(t, \cdot)$ "commute". This condition is trivially satisfied when $g(t, x) \equiv g(t)$ is independent of x or when the control is one-dimensional ($m = 1$).

Lemma 3.6. *If*

$$[g_i, g_k](t, x) \mid_{x = \hat{x}(t)} = 0$$

for any $i \neq k$, $i, k = 1, \dots, m$ and $t \in [0, T]$, then the assumption $S^\top(t)B(t) = B^\top(t)S(t)$ for all $t \in [0, T]$ is fulfilled by the linearization of (3.12).

Proof. The condition $S^\top(t)B(t) = B^\top(t)S(t)$ for the linearization of ($\mathcal{P}$) reads as follows:

$$\Big(g(t, x)^\top \hat{p}(t)\Big)_x g(t, x) = g(t, x)^\top \Big(g(t, x)^\top \hat{p}(t)\Big)_x^\top, \tag{3.20}$$

evaluated at $x = \hat{x}(t)$. Let us first note that

$$\Big(g(t, x)^\top \hat{p}(t)\Big)_x = \begin{pmatrix} \frac{\partial g_1}{\partial x}(t, x)^\top \hat{p}(t) \\ \vdots \\ \frac{\partial g_m}{\partial x}(t, x)^\top \hat{p}(t) \end{pmatrix}.$$

Thus, the left-hand side of (3.20) is

$$\left(g(t,x)^\top \hat{p}(t)\right)_x g(t,x) = \begin{pmatrix} \left(\frac{\partial g_1}{\partial x}(t,x)^\top \hat{p}(t)\right)^\top g_1(t,x) & \dots & \left(\frac{\partial g_1}{\partial x}(t,x)^\top \hat{p}(t)\right)^\top g_m(t,x) \\ \vdots & & \vdots \\ \left(\frac{\partial g_m}{\partial x}(t,x)^\top \hat{p}(t)\right)^\top g_1(t,x) & \dots & \left(\frac{\partial g_m}{\partial x}(t,x)^\top \hat{p}(t)\right)^\top g_m(t,x) \end{pmatrix},$$

while the right-hand side is

$$g(t,x)^\top \left(g(t,x)^\top \hat{p}(t)\right)_x^\top = \begin{pmatrix} g_1^\top(t,x) \left(\frac{\partial g_1}{\partial x}(t,x)^\top \hat{p}(t)\right) & \dots & g_1(t,x)^\top \left(\frac{\partial g_m}{\partial x}(t,x)^\top \hat{p}(t)\right) \\ \vdots & & \vdots \\ g_m(t,x) \left(\frac{\partial g_1}{\partial x}(t,x)^\top \hat{p}(t)\right) & \dots & g_m(t,x)^\top \left(\frac{\partial g_m}{\partial x}(t,x)^\top \hat{p}(t)\right) \end{pmatrix}.$$

Thus, condition (3.20) is verified whenever

$$\frac{\partial g_j}{\partial x}(t,x) g_i(t,x) - \frac{\partial g_i}{\partial x}(t,x) g_j(t,x) = 0,$$

for all $i,j = 1,\dots,m$ and $t \in [0,T]$, at $x = \hat{x}(t)$; namely, $[g_i, g_j](t,\hat{x}(t)) = 0$ for all $i,j = 1,\dots,m$ and $t \in [0,T]$. Q.E.D.

4 Stability of strong bi-metric regularity for affine-control systems under perturbations

In this subsection, we will apply the Lyusternik-Graves-type Theorem 2.2 to prove that the strong bi-metric regularity of the mapping in the Pontryagin system for problem $(\mathcal{P})$ is stable when the problem is affected by some "small" nonlinear disturbances. We consider, therefore, the following perturbed version of $(\mathcal{P})$:

$$\begin{aligned} &\text{minimize} && \ell(x(T)) + b(x(T)) \\ &\text{subject to} && \dot{x}(t) = f(t,x(t)) + g(t,x(t))u(t) + h(t,x(t)) + \beta(t,x(t))u(t), \quad \text{a.e. on } [0,T], \\ &&& u(t) \in U := [0,1]^m, \\ &&& x(0) = x_0, \end{aligned} \tag{4.1}$$

where $b : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^n$ and $\beta : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are suitably smooth and "small" disturbances, in a sense which will be specified below. The system of necessary condition of (4.1) is

$$\begin{aligned} &0 = \dot{x}(t) - f(t,x(t)) - g(t,x(t))u(t) - h(t,x(t)) - \beta(t,x(t))u(t), \\ &0 = \dot{p}(t) + (f(t,(x(t)) + g(t,x(t))u(t) + h(t,x(t)) + \beta(t,x(t))u(t))_x^\top p(t), \\ &0 \in (g(t,x(t)) + \beta(t,x(t)))^\top p(t) + N_U(u(t)), \\ &0 = p(T) - \nabla\ell(x(T)) - \nabla b(x(T)). \end{aligned} \tag{4.2}$$

In the theorem below we prove that the strong bi-metric regularity of $(\mathcal{P})$ is not destroyed by the disturbance (b,h,β), provided that the latter is sufficiently "small". We recall the similar results have been proved in [12, 13] for linear and for linear-quadratic control problems, respectively.

Theorem 4.1. *Assume that* $(\mathcal{P})$ *has a local optimal solution* $\hat{y} = (\hat{x}, \hat{p}, \hat{u})$ *and that the mapping* $\psi + F : Y \rightrightarrows Z$ *(as introduced in* (3.1)*) is strongly bi-metrically regular at* $(\hat{x}, \hat{p}, \hat{u})$ *for* 0*. Moreover, assume that* $(\mathcal{P})$ *fulfills Assumption (A1) and (A2).*

Then, there exist positive real numbers ϵ_0*,* δ *and* c *such that the following conclusion is true. For any positive number* $\epsilon < \epsilon_0$*, and any continuously differentiable functions* h*,* β *and* b*, if*

(i) the functions h*,* β*,* h_x*,* β_x*,* h_t*,* β_t *are bounded by* ϵ *in* $[0,T] \times \mathbb{B}(0; 2M)$*. Moreover,* ∇b *is bounded by* ϵ *in* $\mathbb{B}(0; 2M)$*;*

(ii) the functions h*,* β*,* h_x*,* β_x *are Lipschitz continuous with respect to* x *with constant* ϵ *in* $[0,T] \times \mathbb{B}(0; 2M)$*.* ∇b *is Lipschitz continuous with respect to* x *with constant* ϵ *in* $\mathbb{B}(0; 2M)$*,*

then

(a) system (4.2) *has a unique solution* $y^* := (x^*, p^*, u^*)$ *in the* δ*-neighborhood of* $\hat{y}$ *and*

$$d_Y(y^*, \hat{y}) \leq c\epsilon, \tag{4.3}$$

(b) the mapping in system (4.2) *is strongly bi-metrically regular at the solution* (x^*, p^*, u^*) *for zero.*

Proof. First of all, we observe that system (4.2) can be shortly recast as

$$\varphi(y) + \psi(y) + F(y) \ni 0 \tag{4.4}$$

where ψ and F are introduced in (3.2) and $\varphi : Y \to \tilde{Z}$ is defined as follows:

$$\varphi(y)(t) = \begin{pmatrix} -h(t, x(t)) - \beta(t, x(t))u(t) \\ (h(t, x(t)) + \beta(t, x(t))u(t))_x^\top p(t) \\ \beta(t, x(t))^\top p(t) \\ -\nabla b(x(T)) \end{pmatrix}. \tag{4.5}$$

Fix $\epsilon_0 > 0$ such that for any admissible control u the solution x of the controlled ODE in (4.1) exists on the time interval $[0, T]$ and takes values in $\mathbb{B}(0; 2M)$.

Now, we shall apply Theorem 3.1 for the mappings $\Phi = \psi + F$. Let κ, a, b be the constant of strongly bi-metric regularity for the mapping $\psi + F$ at $(\hat{y}, 0)$. Then, let μ, κ', a', b', and γ the constants introduced in Theorem 2.2. We will verify that φ fulfils the conditions (2.3) and (2.4). Let us start with (2.3). For any $y \in \mathbb{B}_Y(\hat{y}; a')$,

$$d_{\tilde{Z}}(\varphi(\hat{y}), \varphi(y)) \leq d_{\tilde{Z}}(\varphi(\hat{y}), 0) + d_{\tilde{Z}}(\varphi(y), 0).$$

Let us estimate the second term in the right-hand side.

$$\begin{aligned} d_{\tilde{Z}}(\varphi(y), 0) = \max\{1, T\}&\Big(\|h(\cdot, x(\cdot)) + \beta(\cdot, x(\cdot))u(\cdot)\|_\infty + \|(h(\cdot, x(\cdot)) + \beta(\cdot, x(\cdot))u(\cdot))_x^\top p(\cdot)\|_\infty\Big) \\ &+ \|\beta(\cdot, x(\cdot))^\top p(\cdot)\|_{1,\infty} + |b_x(x(T))| \\ \leq \max\{1, T\}&\Big(\|h(\cdot, x(\cdot)) + \beta(\cdot, x(\cdot))u(\cdot)\|_\infty + \|(h(\cdot, x(\cdot)) + \beta(\cdot, x(\cdot)u(\cdot))_x^\top p(\cdot)\|_\infty\Big) \\ &+ \|\beta(\cdot, x(\cdot))^\top p(\cdot)\|_\infty\Big) + \|\beta_t(\cdot, x(\cdot))^\top p(\cdot) + (\beta(\cdot, x(\cdot))\dot{x}(\cdot))_x^\top p(\cdot) \\ &\quad + \beta(\cdot, x(\cdot))^\top \dot{p}(\cdot)\|_\infty + |b_x(x(T))|. \end{aligned}$$

From the conditions in (i) and the facts that $\|u\|_\infty \le m$, $\|p\|_\infty \le 2M$, $\|\dot{x}\|_\infty \le 2M$, $\|\dot{p}\|_\infty \le 2M$ it follows then that

$$d_{\tilde{Z}}(\varphi(y), 0) \le \max\{1, T\}(\varepsilon(1+m) + 2\varepsilon(1+m)M + 2\varepsilon M + \varepsilon(2M + 4M^2 + 2M) + \varepsilon) := C_0\varepsilon. \quad (4.6)$$

Since a similar estimate can be deduced as well for $d_{\tilde{Z}}(\varphi(\hat{y}), 0)$, it holds that

$$\tilde{d}_Z(\varphi(\hat{y}), \varphi(y)) \le 2C_0\varepsilon = C_1\epsilon,$$

where C_1 (as also C_0) is a constant that can be chosen as dependent only on M, m and T.

Now, let us check (2.4). By (i), (ii), $\|u - u'\|_\infty \le m$ we obtain that

$$\begin{aligned}
d_Z(\varphi(y), \varphi(y')) =& \|h(\cdot, x(\cdot)) + \beta(\cdot, x(\cdot))u(\cdot) - h(\cdot, x'(\cdot)) - \beta(\cdot, x'(\cdot))u'(\cdot)\|_1 \\
&+ \| \left(h(\cdot, x(\cdot)) + \beta(\cdot, x(\cdot))u(\cdot)\right)_x^\top p(\cdot) - \left(h(\cdot, x'(\cdot)) + \beta(\cdot, x'(\cdot))u'(\cdot)\right)_x^\top p'(\cdot)\|_1 \\
&+ \|B^\top(\cdot, x(\cdot))p(\cdot) - B^\top(\cdot, x'(\cdot))p'(\cdot)\|_\infty + |b_x(x(T)) - b_x(x'(T))| \\
\le& \varepsilon\|x - x'\|_1 + m\varepsilon\|x - x'\|_1 + \varepsilon\|u - u'\|_1 \\
&+ 2\varepsilon M(1+m)\|x - x'\|_1 + \varepsilon(1+m)\|p - p'\|_1 + 2\varepsilon M\|u - u'\|_1 \\
&+ \varepsilon\|p - p'\|_\infty + 2\varepsilon M\|x - x'\|_\infty + \varepsilon|x(T) - x'(T)|.
\end{aligned}$$

We now can use the facts that $\|p - p'\|_\infty \le \max\{1, T^{-1}\}\|p - p'\|_{1,1}$ and $\|x - x'\|_\infty \le \|x - x'\|_{1,1}$ for any $x, x' \in W^{1,1}_{x_0}([0,T])$ and $p, p' \in W^{1,1}([0,T])$, to conclude that

$$d_Z(\varphi(y), \varphi(y')) \le \varepsilon C_2\, d_Y(y, y'),$$

where C_2 is a constant depending only on M, m and T.

Choosing $\epsilon_0 > 0$ such that $C_1\epsilon_0 \le \gamma$, $C_2\epsilon_0 \le \mu$, we can apply Theorem 2.2 and deduce that the mapping $\tilde{I\!B}_Z(\varphi(\hat{y}); b') \ni z \mapsto (\varphi + \psi + F)^{-1}(z) \cap I\!B_Y(\hat{y}; a')$ is single-valued and Lipschitz continuous with respect to the metric d_Z with constant κ'. Now we utilize Corollary 2.3, which is possible since $d_{\tilde{Z}}(\varphi(\hat{y}), 0) \le C_0\varepsilon$ (see (4.6) applied for the particular case $y = \hat{y}$) and $\kappa' d_Z(\varphi(\hat{y}), 0) \le C_0\varepsilon$. For that we have to chose $\varepsilon > 0$ so small that $C_0\varepsilon < b'/2$ and $\kappa' C_0\varepsilon < a'/2$. Then Corollary 2.3 yields the claims of the theorem. Q.E.D.

5 A Lyusternik-Graves-type theorem for bi-metrically regular maps

In this section we introduce the notion of *bi-Metric Regularity* (bi-MR) as an extension of the widely used notion of *metric regularity*, [4, 10]. Similarly as for the Sbi-MR property, the bi-MR may be a relevant property in the context of generalized equations containing variational inequalities of the form of (1.1) where φ is not strictly monotone.

Below we use the notation introduced in Section 2.

Definition 5.1. The map $\Phi : Y \rightrightarrows Z$ is *bi-metrically regular* (bi-MR) at $\hat{y} \in Y$ for $\hat{z} \in \tilde{Z}$ with constants $\kappa \ge 0$, $a > 0$ and $b > 0$ if $(\hat{y}, \hat{z}) \in \text{graph}(\Phi)$, and for all $z, z' \in I\!B_{\tilde{Z}}(\hat{z}; b)$ it holds that

$$e_Y(\Phi^{-1}(z) \cap I\!B_Y(\hat{y}; a), \Phi^{-1}(z')) \le \kappa d_Z(z, z'), \quad (5.1)$$

where, as in the introduction, $e_Y(A, B)$ is the excess of $A \subset Y$ beyond $B \subset Y$.

Observe that the only difference with the definition of MR given in the introduction is that z and z' should be sufficiently small in the metric $d_{\tilde{Z}}$ instead of d_Z.

Below, we prove a Lyusternik-Graves-type theorem for bi-MR maps. The proof follows a similar idea as the proofs of Lyusternik-Graves-type theorems for metrically regular maps by utilization of a contraction mapping theorem, but needs modifications due to the involvement of two metrics.

Theorem 5.2. *Let the metric space Y be complete, let Z be a linear space and $\tilde{Z} \subset Z$ be a subspace. Moreover, let d_Z and $d_{\tilde{Z}}$ be shift-invariant metrics in, Z and $\tilde{Z}$, respectively, and $d_Z \le d_{\tilde{Z}}$ on $\tilde{Z}$. Assume that $\Phi : Y \rightrightarrows Z$ is a set-valued map which is bi-metrically regular at $\hat{y} \in Y$ for $\hat{z} \in \tilde{Z}$ with constants κ, a, b. Also assume that the set* $\operatorname{gph}(\Phi) \cap (I\!\!B_Y(\hat{y}; a) \times I\!\!B_Z(\hat{z}; b))$ *is closed.*

Let $\mu > 0$ be such that $\kappa\mu < 1$. Let the numbers κ', a', b' and γ satisfy the relations

$$\kappa' = \frac{\kappa}{1 - \kappa\mu}, \quad 2a' \le a, \quad b' + \gamma \le b, \quad 2\kappa' b' \le a'.$$

Then for every function $\varphi : Y \to \tilde{Z}$ such that

$$d_{\tilde{Z}}(\varphi(\hat{y}), \varphi(y)) \le \gamma \quad \forall y \in I\!\!B_Y(\hat{y}; 2a'), \tag{5.2}$$

and

$$d_Z(\varphi(y), \varphi(y')) \le \mu d_Y(y, y') \quad \forall y, y' \in I\!\!B_Y(\hat{y}; 2a'), \tag{5.3}$$

the mapping $\varphi + \Phi$ is bi-metrically regular at $\hat{y}$ for $\hat{z} + \varphi(\hat{y})$ with constants κ', a' and b'.

Proof. We have to prove that

$$e_Y((\varphi + \Phi)^{-1}(z') \cap B_Y(\hat{y}; a')), (\varphi + \Phi)^{-1}(z)) \le \kappa' d_Z(z, z') \qquad \forall\, z, z' \in I\!\!B_{\tilde{Z}}(\hat{z} + \varphi(\hat{y}); b').$$

For that we take two arbitrary points $z, z' \in I\!\!B_{\tilde{Z}}(\hat{z} + \varphi(\hat{y}); b')$ and any $y' \in (\varphi + \Phi)^{-1}(z') \cap I\!\!B_Y(\hat{y}; a')$. The theorem will be proved if we find

$$y \in (\varphi + \Phi)^{-1}(z) \text{ such that } d_Y(y, y') \le \kappa' d_Z(z, z'). \tag{5.4}$$

Similarly as in the proof of [4, Theorem 5G.3], we shall use the following theorem, adapted to our notations.

Theorem 5.3. (Theorem 5E.2, [4].) *Let $\mathcal{Y}$ be a complete metric space with a metric ρ, let $G : \mathcal{Y} \rightrightarrows \mathcal{Y}$ be a set-valued mapping and $y' \in \mathcal{Y}$ be arbitrarily fixed. Assume that there are numbers $p > 0$ and $\lambda \in (0, 1)$ such that*

(a) the set $\operatorname{gph}(G) \cap (I\!\!B_{\mathcal{Y}}(y'; p) \times I\!\!B_Z(y'; p))$ *is closed;*

(b) $\rho(y', G(y')) \le p(1 - \lambda)$;

(c) $e_{\mathcal{Y}}(G(v_1) \cap I\!\!B_{\mathcal{Y}}(y'; p), G(v_2)) \le \lambda\rho(v_1, v_2)$ for every $v_1, v_2 \in I\!\!B_{\mathcal{Y}}(y'; p)$.

Then there exists $y \in I\!\!B_{\mathcal{Y}}(y'; p)$ such that $y \in G(y)$.

We shall apply this theorem with $\mathcal{Y} = I\!B_Y(\hat{y}; a)$, $\rho = d_Y$, $p = \kappa' d_Z(z, z')$, $\lambda = \kappa\mu$ and the mapping G defined as

$$G(y) := \Phi^{-1}(z - \varphi(y)) \cap I\!B_Y(\hat{y}; a).$$

We shall use the following facts:

$$I\!B_Y(y'; p) \subset I\!B_Y(\hat{y}; 2a'), \tag{5.5}$$

$$d_{\tilde{Z}}(z - \varphi(v), \hat{z}) \leq b \qquad \forall v \in I\!B_Y(y'; p). \tag{5.6}$$

The first one follows from $I\!B_Y(y'; p) \subset I\!B_Y(\hat{y}; p + d_Y(y', \hat{y}))$ and

$$p + d_Y(y', \hat{y}) \leq \kappa' d_Z(z, z') + a' \leq 2\kappa' b' + a' \leq 2a'.$$

For the second we estimate

$$d_{\tilde{Z}}(z - \varphi(v), \hat{z}) \leq d_{\tilde{Z}}(z, \hat{z} + \varphi(\hat{y})) + d_{\tilde{Z}}(\varphi(\hat{y}), \varphi(v)) \leq b' + \gamma \leq b,$$

where we use (5.5) to verify that $v \in I\!B_Y(\hat{y}; 2a')$ and then use (5.3).

Now, let us check the assumptions of Theorem 5.3 starting with (a). Let $(y_k, w_k) \in \mathrm{gph}\,(G) \cap (I\!B_Y(y'; p) \times I\!B_Y(y'; p))$ and (y_k, w_k) converges to (y_0, w_0). Clearly, $(y_0, w_0) \in (I\!B_Y(y'; p) \times I\!B_Y(y'; p))$. Moreover, $w_k \in \Phi^{-1}(z - \varphi(y_k)) \cap I\!B_Y(\hat{y}; a)$, hence, $z - \varphi(y_k) \in \Phi(w_k)$. Due to (5.3), φ is continuous in $I\!B_Y(\hat{y}; 2a')$ and $y_k \in I\!B_Y(y'; p) \subset I\!B_Y(\hat{y}; 2a')$. Thus the sequence $\varphi(y_k)$ converges to $\varphi(y_0)$. Moreover, using (5.6) we have $(y_k, z - \varphi(y_k)) \subset I\!B_Y(\hat{y}; 2a') \times I\!B_Z(\hat{z}; b)) \subset I\!B_Y(\hat{y}; a) \times I\!B_Z(\hat{z}; b)$. The set $\mathrm{gph}\,(\Phi) \cap (I\!B_Y(\hat{y}; a) \times I\!B_Z(\hat{z}; b))$ is assumed closed, hence $z - \varphi(y_0) \in \Phi(w_0)$ and $(y_0, w_0) \in \mathrm{gph}\,(G)$.

In order to prove (b) we use the equivalence of the properties $y' \in (\varphi + \Phi)^{-1}(z')$ and $y' \in \Phi^{-1}(z' - \varphi(y'))$, and the inclusion $y' \in I\!B_Y(\hat{y}; a') \subset I\!B_Y(\hat{y}; a)$ to estimate

$$\begin{aligned} d_Y(y', G(y')) &= d_Y(y', \Phi^{-1}(z - \varphi(y')) \cap I\!B_Y(\hat{y}; a)) = d_Y(y', \Phi^{-1}(z - \varphi(y'))) \\ &\leq e_Y(\Phi^{-1}(z' - \varphi(y')) \cap I\!B_Y(\hat{y}; a), \Phi^{-1}(z - \varphi(y'))). \end{aligned}$$

Due to (5.6), which holds also with z' at the place of z, and the bi-MR property of Φ we obtain that

$$d_Y(y', G(y')) \leq \kappa d_Z(z, z') = \kappa'(1 - \kappa\mu)\, d_Z(z, z') = p(1 - \lambda).$$

Let us verify condition (c) in Theorem 5.3. Using again (5.5) and (5.6) with $v = v_i$ we have

$$\begin{aligned} e_{\mathcal{Y}}(G(v_1) \cap I\!B_{\mathcal{Y}}(y'; p), G(v_2)) &= \\ e_{\mathcal{Y}}\Big(\Phi^{-1}(z - \varphi(v_1)) \cap I\!B_Y(\hat{y}; a) \cap I\!B_{\mathcal{Y}}(y'; p), \Phi^{-1}(z - \varphi(v_2)) \cap I\!B_Y(\hat{y}; a)\Big) \\ &\leq e_{\mathcal{Y}}(\Phi^{-1}(z - \varphi(v_1)) \cap I\!B_{\mathcal{Y}}(y'; p), \Phi^{-1}(z - \varphi(v_2))) \\ &\leq \kappa d_Z(\varphi(v_1), \varphi(v_2)) \leq \kappa\mu\, d_Y(v_1, v_2) = \lambda\, d_Y(v_1, v_2). \end{aligned}$$

Here we used condition (5.3) thanks to the inequality $d_Y(v_i, \hat{y}) \leq 2a'$.

Theorem 5.3 gives us existence of y as required in (5.4). Q.E.D.

It is trivial to construct (academic) examples where b-MR is present but not Sbi-MR. Typically this situation appears if there are redundant state and control variables that create a continual variety of optimal solutions.

Appendix: Prof of Theorem 3.1

We denote by $L(f)$, $L(g)$, $L(f_x)$, $L(g_x)$, $L(g_t)$, $L(f_{xx})$, $L(g_{xx})$, $L(\ell_x)$ the Lipschitz constants of the functions in the parentheses with respect to $x \in I\!B(0;2M)$. Set $\bar{d} := \max_{u\in U} |u| = \sqrt{m}$. Below c_1, c_2, ... will be constants depending only on the above Lipschitz constants and the numbers $\bar{d}$, T and

$$D := \sup_{t\in[0,T]} \max\Big\{\|A(t)\|,\ \|S(t)\|,\ \|W(t)\|,\ \sup_{|x|\le M} |f_x(t,x)|,\ \sup_{|x|\le M} |g_x(t,x)|,\ \sup_{|x|\le M} |\ell_{xx}(x)|\Big\}.$$

We suppose that the mapping $\psi + F$ is strongly bi-metrically regular at $\hat{y} = (\hat{x}, \hat{p}, \hat{u})$ for 0 with constant (κ, a, b). Let the numbers μ, κ', γ, a' and b' be chosen as in Theorem 2.2 applied with $\Phi = \psi + F$, $\hat{y}$, and $\hat{z} = (0,0,0,0)$. Let φ be as in (3.10). Observe that if the conclusion of Theorem 2.2 holds for some positive a' and b' (together with μ, κ' and γ), then it also holds for any other choice of smaller positive a' and b', provided that the inequality $\kappa b' \le (1-\kappa\mu)a'$ is still fulfilled. Hence, without lost of generality, we can assume that $a' < M$.

1. First we prove that inequality (2.3) holds for every $y \in I\!B_Y(\hat{y}; a')$, which reads as $d_{\tilde{Z}}(\varphi(y), 0) \le \gamma$, since $\varphi(\hat{y}) = 0$. Notice that, by using the definition of B and S and after few simplifications, $\varphi(y)$ can be equivalently recast as

$$\varphi(y) = \begin{pmatrix} f(\hat{x}) - f(x) + (g(\hat{x}) - g(x))u + A\Delta x \\ f(x)_x^\top p - f(\hat{x})_x^\top \hat{p} + ((g(x)^\top p)_x - (g(\hat{x})^\top \hat{p})_x)^\top u - A^\top \Delta p - W\Delta x \\ (g(x) - g(\hat{x}))^\top p - S^\top \Delta x \\ -\nabla \ell(x(T)) + \nabla \ell(\hat{x}(T)) + K\Delta x(T) \end{pmatrix}. \tag{5.7}$$

We shall estimate each of the components $\varphi_1, \dots, \varphi_4$ of φ separately. By the Lipschitz continuity of f and g in $I\!B(0;2M)$, we obtain from (5.7) that

$$|\varphi_1(y)(t)| \le \left(L(f) + \bar{d}L(g) + \sup_{t\in[0,T]} \|A(t)\| \right) |\Delta x(t)|.$$

Hence,

$$\|\varphi_1(y)\|_\infty \le c_1 \|\Delta x\|_\infty \le c_1 \|\dot{x} - \dot{\hat{x}}\|_1 \le c_1 \|\Delta x\|_{1,1} \le c_1 a', \qquad \forall y \in I\!B_Y(\hat{y}; a'). \tag{5.8}$$

Similarly we can obtain by (5.7) the following estimation for the second component of $\varphi(y)$:

$$\begin{aligned} |\varphi_2(y)(t)| \ \le\ & \left(ML(f_x) + \bar{d}ML(g_x) + \bar{d}M \sup_{t\in[0,T], |x|\le M} |g_x| \right) |\Delta x(t)| + \sup_{t\in[0,T], |x|\le M} |f_x|\, |\Delta p(t)| \\ & + \sup_{t\in[0,T]} \|A(t)\|\, |\Delta p(t)| + \sup_{t\in[0,T]} \|W(t)\|\, |\Delta x(t)|. \end{aligned}$$

Hence,

$$\|\varphi_2(y)\|_\infty \le c_2 (\|\Delta x\|_\infty + \|\Delta p\|_\infty), \qquad \forall y \in I\!B_Y(\hat{y}; a').$$

Moreover, for any $\Delta p \in W^{1,1}([0,T])$ it hods that

$$\|\Delta p\|_\infty \le \max\{1, T^{-1}\} \|\Delta p\|_{1,1}.$$

Then, for an appropriate c_3 we have

$$\|\varphi_2(y)\|_\infty \leq c_3 a', \qquad \forall y \in I\!B_Y(\hat{y}; a'). \tag{5.9}$$

Now we shall estimate the $W^{1,\infty}$-norm of $\varphi_3(y)$. We first observe that

$$| \varphi_3(y)(t) | \leq (ML(g) + \sup_{t\in[0,T]} \|S(t)\|) \mid \Delta x(t) \mid . \tag{5.10}$$

The derivative of $\varphi_3(y)$ can be calculated as follows:

$$\begin{aligned}
\frac{d}{dt}\varphi_3(y)(t) &= (g_t(t,x(t)) - g_t(t,\hat{x}(t)))^\top p(t) + \Big(g(t,x(t))^\top p(t)\Big)_x \dot{x}(t) - \Big(g(t,\hat{x}(t))^\top p(t)\Big)_x \dot{\hat{x}}(t) \\
&\quad + (g(t,x(t)) - g(t,\hat{x}(t)))^\top \dot{p}(t) - \dot{S}(t)^\top \Delta x(t) - S(t)^\top (\dot{x}(t) - \dot{\hat{x}}(t)) \\
&= (g_t(t,x(t)) - g_t(t,\hat{x}(t)))^\top p(t) + (g(t,x(t)) - g(t,\hat{x}(t)))^\top \dot{p}(t) - \dot{S}(t)^\top \Delta x(t) \\
&\quad - \Big(g(t,\hat{x}(t))^\top \Delta p(t)\Big)_x \dot{\hat{x}}(t) + (g(t,x(t))p(t) - g(t,\hat{x}(t))\hat{p}(t))_x \dot{x}(t).
\end{aligned}$$

Then it is easy to deduce that

$$\begin{aligned}
\Big|\frac{d}{dt}\varphi_3(y)(t)\Big| &\leq \Big(ML(g_t) + ML(g) + \sup_{t\in[0,T]} \|\dot{S}(t)\|\Big)|\Delta x(t)| + M \sup_{t\in[0,T],|x|\leq M} \|g_x\| \, |\Delta p(t)| \\
&+ \; M\Big(ML(g)|\Delta x(t)| + \sup_{t\in[0,T],|x|\leq M} \|g_x\| \, |\Delta p(t)|\Big).
\end{aligned} \tag{5.11}$$

By (5.10) and (5.11) there exist constants c_4 and c_5 such that

$$\|\varphi_3(y)\|_{1,\infty} \leq c_4(\|\Delta x\|_\infty + \|\Delta p\|_\infty) \leq c_5 a', \qquad \forall y \in I\!B_Y(\hat{y}; a'). \tag{5.12}$$

Finally, we observe that

$$|\varphi_4(y)(T)| \leq \Big(L(\ell_x) + \sup_{|x|\leq M} \|\ell_{xx}\|\Big)|\Delta x(T)| \leq c_6 a', \qquad \forall y \in I\!B_Y(\hat{y}; a'), \tag{5.13}$$

for some $c_6 > 0$. By (5.8), (5.9), (5.12) and (5.13) we deduce that there exists a constant c_7 such that

$$d_{\tilde{Z}}(\varphi(y), 0) \leq c_7 a', \qquad \forall y \in I\!B_Y(\hat{y}; a'). \tag{5.14}$$

2. We now prove the inequality (2.4). Let $y = (x, p, u)$ and $y' = (x', p', u')$ be arbitrary triples in $I\!B_Y(\hat{y}; a')$. Below $y - y'$ will be shortened by $\delta y := (\delta x, \delta p, \delta u)$. First we observe that

$$\varphi(y) - \varphi(y') = \begin{pmatrix} f(x') + g(x')u' + A\delta x + B\delta u - f(x) - g(x)u \\ (f(x) + g(x)u)_x^\top p - (f(x') + g(x')u')_x^\top p' - A^\top \delta p - W\delta x - S\delta u \\ g(x)^\top p - g(x')^\top p' - B^\top \delta p - S^\top \delta x \\ -\nabla \ell(x(T)) + \nabla \ell(x'(T)) + K\delta x(T) \end{pmatrix}.$$

We shall analyze $d_Z(\varphi(y), \varphi(y'))$ working component by component.

We represent (still suppressing the variable t in the right-hand side)

$$\begin{aligned}
\varphi_1(y)(t) - \varphi_1(y')(t) &= f(x') + g(x')u - (f(x) + g(x)u) + A\delta x - g(x')\delta u + B\delta u \\
&= f(x') + g(x')u - (f(x) + g(x)u) + (f(\hat{x}) + g(\hat{x})\hat{u})_x\delta x - g(x')\delta u + g(\hat{x})\delta u \\
&= \int_0^1 \Big[\frac{d}{ds}\big(f(\xi(s)) + g(\xi(s))u\big) + (f(\hat{x}) + g(\hat{x})\hat{u})_x\delta x\Big] ds + (g(\hat{x}) - g(x'))\delta u \\
&= \int_0^1 \Big[-\big(f(\xi(s)) + g(\xi(s))u\big)_x + (f(\hat{x}) + g(\hat{x})\hat{u})_x\Big] ds\, \delta x + (g(\hat{x}) - g(x'))\delta u,
\end{aligned}$$

where we have denoted $\xi(s) := x + s(x' - x)$. Now we observe that $|\xi(s) - \hat{x}| \le 2a'$, $|x' - \hat{x}| \le a'$, and $|u - \hat{u}| \le a'$. Then we estimate

$$\begin{aligned}
|\varphi_1(y)(t) - \varphi_1(y')(t)| \;\le\; & \int_0^1 \big[L(f_x)|\xi(s) - \hat{x}| + L(g_x)\bar{d}|\xi(s) - \hat{x}| + L(g_x)|u - \hat{u}|\big]\, ds\, |\delta x| \\
& + L(g)|x' - \hat{x}|\,|\delta u| \le c_8 a'(|\delta x(t)| + |\delta u(t)|).
\end{aligned}$$

Since this estimation holds for any $t \in [0, T]$, we obtain that

$$\|\varphi_1(y) - \varphi_1(y')\|_1 \le c_9 a' d_Y(y, y'). \tag{5.15}$$

For the second component of $\varphi(y) - \varphi(y')$ we use the representation

$$\begin{aligned}
\varphi_2(y)(t) - \varphi_2(y')(t) &= \Big((f(x) + g(x)u')_x^\top - (f(x') + g(x')u')_x^\top\Big)\, p' - W\delta x \\
&\quad + \Big((f(x) + g(x)u')_x^\top - A\Big)\, \delta p + \Big((g(x)^\top p)_x^\top u - S\Big)\, \delta u \\
&= \Big((f(x) + g(x)u')_x^\top - (f(x') + g(x')u')_x^\top\Big)\, p' - \Big((f(\hat{x}) + g(\hat{x})\hat{u})_x^\top\, \hat{p}\Big)_x \delta x \\
&\quad + \Big((f(x) + g(x)u')_x^\top - (f(\hat{x}) + g(\hat{x})\hat{u})_x\Big)\, \delta p \\
&\quad + \Big((g(x)^\top p)_x^\top u - \Big(g(\hat{x})^\top \hat{p}\Big)_x^\top\Big)\, \delta u,
\end{aligned}$$

where we have used the identity $(g(x)u)_x^\top p = (g(x)^\top p)_x^\top u$. The two terms in the right-hand side can be estimated using the Lipschitz continuity of the involved functions. The remaining term in the right-hand side can be treated similarly as the "bad" term in $\varphi_1(y) - \varphi_1(y')$. Thus, skipping the routine calculation, we obtain a similar estimation

$$\|\varphi_2(y) - \varphi_2(y')\|_1 \le c_{10} a' d_Y(y, y'). \tag{5.16}$$

Analogously, one can represent

$$\begin{aligned}
\varphi_3(y)(t) - \varphi_3(y')(t) &= (g(x) - B)^\top \delta p + (g(x) - g(x'))^\top p' - S^\top \delta x) \\
&= (g(x) - g(\hat{x}))^\top \delta p + \Big((g(x) - g(x'))^\top p' - (g(\hat{x})\hat{p})_x^\top \delta x\Big).
\end{aligned}$$

Similarly as above this yields

$$\|\varphi_3(y) - \varphi_3(y')\|_\infty \le c_{11} a' d_Y(y, y'). \tag{5.17}$$

Finally, since ℓ has Lipschitz continuous derivative, there exists $\bar{x}$ in the interval $[x(T), x'(T)]$ such that

$$\begin{aligned} |\varphi_4(y)(T) - \varphi(y')(T)| &= \ell_{xx}(\bar{x}(T))\delta x(T) - \ell_{xx}(\hat{x}(T))\delta x(T)| \leq L(l_{xx})2a'|\delta x(T)| \\ &\leq c_{12}a' d_Y(y, y'). \end{aligned} \tag{5.18}$$

From (5.15), (5.16), (5.17) and (5.18) we finally conclude that there exists a constant c_{13} such that for all $y, y' \in I\!B_Y(\hat{y}; a')$

$$d_Z(\varphi(y), \varphi(y')) \leq c_{13}a' d_Y(y, y'). \tag{5.19}$$

3. The inequalities (5.14) and (5.19) for $y, y' \in I\!B_Y(\hat{y}; a')$ show that the relations (2.3) and (2.4) in Theorem 2.2 are fulfilled, provided that

$$c_7 a' \leq \gamma, \qquad c_{13}a' \leq \mu. \tag{5.20}$$

As explained in the beginning of the proof, we can assume that the conditions in (5.20) are satisfied, taking smaller values of a' and b' if necessary. The conclusion of the theorem then follows by applying Theorem 2.2 to the map $\psi + F$ and recalling that the generalized equation $\psi(y) + F(y) \ni 0$ can be equivalently recast as in (3.11).

References

[1] Agrachev A.A., Stefani G., Zezza P.: Strong optimality for a bang-bang trajectory. SIAM J. Control Optim., 41(4), 991–1014, 2003.

[2] Alt W., Baier R., Gerdts M., Lempio F.: Approximation of Linear Control Problems with Bang-Bang Solutions. Optimization. 62(1), 9–32,2013.

[3] Alt W., Schneider C., Seydenschwanz M.: Regularization and implicit Euler discretization of linear-quadratic optimal control problems with bang-bang solutions. Appl. Math. and Comp. 287-288, 104–124, 2016.

[4] Dontchev A.L., Rockafellar R.T.: Implicit Functions and Solution Mappings: A View from Variational Analysis. Second edition. Springer, New York , 2014.

[5] Dontchev A.L., W. W. Hager W.W.: Lipschitzian stability in nonlinear control and optimization. SIAM J. Control Optim., 31, 569–603, 1993.

[6] Dontchev A.L., Krastanov M.I., Veliov V.M.: On the existence of Lipschitz continuous optimal feedback control in ODE control problems. Submitted. Available as Research Report 2018-04, ORCOS, TU Wien, 2018.

[7] Felgenhauer U.: On Stability of Bang-Bang Type Controls. SIAM J. Control Optim., 41(6), 1843–1867, 2003.

[8] Felgenhauer U., Poggiolini L., Stefani G.: Optimality and stability result for bang-bang optimal controls with simple and double switch behaviour. Control and Cybernetics, 38, 1305–1325, 2009.

[9] Frankowska H., Quincampoix M.: Hölder metric regularity of set-valued maps. Math. Program. 132, 333-354, 2012.

[10] Ioffe A.D.: Variational Analysis of Regular Mappings: Theory and Applications. Springer, 2917.

[11] Maurer H. and Osmolovskii N.P.: Second Order Sufficient Conditions for Time-Optimal Bang-Bang Control SIAM J. Control Optim., 42(6), 2239–2263, 2004.

[12] Preininger J, Scarinci T., Veliov V.M: Metric regularity properties in bang-bang type linear-quadratic optimal control problems. Set-Valued and Variational Analysis, https://doi.org/10.1007/s11228-018-0488-1.

[13] Quincampoix M., Veliov V.M.: Metric Regularity and Stability of Optimal Control Problems for Linear Systems. SIAM J. Control Optim., 51(5), 4118–4137, 2013.

[14] Robinson S.M.: Strongly regular generalized equations. Math. Oper. Res., 5, 43–62, 1980.