TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

On the accuracy of the model predictive control method

G. Angelov, A. Domínguez Corella, V.M. Veliov

Research Report 2021-05

November 2021

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 / E105-04
1040 Vienna, Austria
E-mail: orcos@tuwien.ac.at

On the accuracy of the model predictive control method*

Georgi Angelov† Alberto Domínguez Corella‡ Vladimir M. Veliov§

Abstract

The paper investigates the accuracy of the Model Predictive Control (MPC) method for finding on-line approximate optimal feedback control for Lagrange type problems on a fixed finite horizon. The predictions for the dynamics, the state measurements, and the solution of the auxiliary open-loop control problems that appear at every step of the MPC method may be inaccurate. The main result provides an error estimate of the MPC-generated solution compared with the optimal open-loop solution of the "ideal" problem, where all predictions and measurements are exact. The technique of proving the estimate involves an extension of the notion of strong metric sub-regularity of set-valued maps and utilization of a specific new metric in the control space, which makes the proof non-standard. The result is specialized for two problem classes: coercive problems, and affine problems.

Keywords: optimal control, Lagrange problem, model predictive control, metric sub-regularity

MSC Classification: 93B45, 49M99, 49J40, 47J20

1 Introduction

Model Predictive Control (MPC) is a powerful method for approximate on-line feedback control, widely used in industrial applications and recently in digital engine control and microelectronics, see, e.g., [18, 22, 25]. On the other hand, the rigorous mathematical theory investigating the scope of validity and the efficiency of the MPC method under appropriate assumptions is still underdeveloped. The present paper contributes to this theory by investigating the accuracy of a version of the MPC method applied to a finite horizon optimal control problem (the so-called *economic MPC* with shrinking horizon).

To set the stage, we consider the following optimal control problem, further denoted by $\mathscr{P}_p(0,x_0)$:

$$\min_{u\in\mathscr{U}} \left\{ J_p(u) := \int_0^T g(p(t),x(t),u(t))\,\mathrm{d}t \right\}, \tag{1}$$

subject to

$$\dot{x}(t) = f(p(t),x(t),u(t)) \quad x(0) = x_0. \tag{2}$$

*This research is supported by the Austrian Science Foundation (FWF) under grant No I4571.

†Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Austria, georgi.angelov@tuwien.ac.at

‡Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Austria, alberto.corella@tuwien.ac.at

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

Here the state vector $x(t)$ belongs to $\mathbb{R}^n$ and the control function $u(\cdot)$ belongs to the set $\mathscr{U}$ of all Lebesgue measurable functions $u : [0,T] \to U$, where $U \subset \mathbb{R}^m$. The function p represents an uncertain time-dependent parameter which is known to belong to a set Π of bounded Lebesgue measurable functions $p : [0,T] \to \mathbb{R}^l$. Correspondingly, f and g are defined on $\mathbb{R}^l \times \mathbb{R}^n \times \mathbb{R}^m$ with values in $\mathbb{R}^n$ and $\mathbb{R}$, respectively. The initial state $x_0 \in \mathbb{R}^n$ and the final time $T > 0$ are fixed.

A version of the MPC method (called further MPC algorithm) applied to the above problem is described in detail in Subsection 3.1. Here we briefly present the main result given in Theorem 3.1, Subsection 3.2. It is assumed that for some particular parameter function, $\hat{p}$, equation (2) represents a real system, therefore problem (1)–(2) with $p = \hat{p}$ (that is, problem $\mathscr{P}_{\hat{p}}(0,x_0)$) is called *reference problem*. However, $\hat{p}$ and the initial state x_0 are not assumed to be exactly known. The MPC algorithm generates a control function by solving a sequence of auxiliary open-loop optimal control problems. Given a mesh $0 = t_0 < t_1 < \ldots < t_N = T$, the auxiliary problem at the k-th step is of the same kind as (1)–(2), but on the shorter time-interval $[t_k, T]$. A prediction p for the parameter function on $[t_k, T]$ is given, and the initial state at t_k is obtained by measuring the real systems state at time t_k. Both the prediction and the measurement may be inaccurate. In addition, the auxiliary problem at the k-th step is only approximately solved, which is another source of error. The approximate optimal control in the k-th auxiliary problem is only applied to the ‘’real” system on the interval $[t_k, t_{k+1}]$, then the procedure is further repeated on the next interval, resulting at the end in what is called the MPC- generated control.

The main result in the paper gives an estimate of the difference between the MPC-generated control and the optimal open-loop control for the reference problem (the latter corresponding to the “ideal” scenario where the prediction, the measurement, and the solution of the auxiliary problems are all exact). A remarkable feature of the estimation is that the overall error of the MPC-generate control depends on the *average* of the errors appearing at the steps of the algorithm, thus occasional relatively large errors in the prediction or measurement do not substantially damage the MPC-generated control. Another interesting feature of the overall error is that for some classes of problems it depends linearly on the averaged errors appearing at the steps of the method, while for other classes, the estimate of the overall error depends on the square root of the averaged errors (and this estimate is sharp). We mention that an estimate of the difference between the MPC-generated control and the realization of the optimal *feedback* control in the reference problem is obtained [11]. In [10], this result is extended to a comparison with the optimal open-loop control in the reference problem, as in the present paper. However, in both quoted papers the results are obtained within a much more restrictive framework: a single prediction is used which does not change from step to step, the results only apply to the Euler discretization of the auxiliary problems, and most importantly, the reference optimal control problem is assumed *coercive* (see Subsection 4.1 for the notion). In fact, the main goal of the present paper is to extend the results about the accuracy of the MPC method beyond the coercive case, especially for affine control problems.

The main result (the error estimate in Theorem 3.1) is obtained under general assumptions; the most demanding one is the requirement that the map associated with the first order optimality conditions (called *optimality map*) for the reference problem is *strongly metrically sub-regular* in an appropriate space setting. In Subsection 2.1, we extend the abstract notion of strong metric sub-regularity of a set-valued map by involving two metrics in the domain of the map. In Subsection 2.2, we define the optimality map and the space setting. In the control space (which is a projection of the domain of the optimality map) we introduce a specific new metric, which is of key importance for the analysis of the MPC method and may be useful in other contexts.

The proof of Theorem 3.1 (given in Appendix) is non-trivial and substantially differs from all proofs of error estimates in optimal control that the authors know.

Section 4 presents or recalls sufficient conditions for the extended strong metric sub-regularity of the optimality map for two non-intersecting classes of problems: for *coercive problems* and for *affine problems*. The paper concludes with an example where the MPC algorithm is applied to a spacecraft stabilization problem. The numerical results confirm the theoretical error estimate and its sharpness.

2 Preliminaries

2.1 Strong sub-regularity

Here we introduce a notion which extends the property of *Strong (metric) sub-Regularity* (Ss-R) (see, e.g., [15, Chapter 3.9] and the recent paper [4]). Namely, we consider a general metric space $\mathscr{Y}$ in which two metrics are defined: d and d^*, and another metric space $\mathscr{Z}$ with a metric $d_{\mathscr{Z}}$. We denote by $\boldsymbol{B}(y;\alpha)$ the closed ball with radius $\alpha \geq 0$ in $(\mathscr{Y},d)$ centered at y, by $\boldsymbol{B}_*(y;\alpha)$ the ball with radius $\alpha \geq 0$ in $(\mathscr{Y},d^*)$, and similarly, $\boldsymbol{B}_{\mathscr{Z}}(z;\alpha)$ is the respective ball in $\mathscr{Z}$.

Definition 2.1. *A set-valued map $\Phi : \mathscr{Y} \rightrightarrows \mathscr{Z}$ is called Ss-R at $(\hat{y},\hat{z}) \in \mathscr{Y} \times \mathscr{Z}$ (with respect to the metrics d and d^*) if $\hat{z} \in \Phi(\hat{y})$ and there are positive constants α,β and κ (called further parameters of Ss-R) such that*

$$d^*(y,\hat{y}) \leq \kappa d_{\mathscr{Z}}(z,\hat{z}) \quad \text{for all } y \in \boldsymbol{B}(\hat{y};\alpha),\ z \in \Phi(y) \cap \boldsymbol{B}_{\mathscr{Z}}(\hat{z};\beta).$$

The Ss-R property plays a fundamental role in the error analysis of numerical methods. It was introduced under this name in [14], but has also been used under several other names (see also [19, Chapter 1] for the related but stronger property of strong upper regularity). A more detailed historical account can be found in [4, Section 1]. The extension with two metrics in $\mathscr{Y}$, presented above, is essential for the applications in the present paper.

The following simple claim is a modification of [4, Theorem 2.1] for the case of two metrics in Y.

Proposition 2.2. *Assume that $\mathscr{Z}$ is a linear space and $d_{\mathscr{Z}}$ is a shift-invariant metric in $\mathscr{Z}$. Assume, in addition, there exists a number $\gamma > 0$ such that $d(y_1,y_2) \leq \gamma d^*(y_1,y_2)$ for every $y_1,y_2 \in \mathscr{Y}$. Let $\Phi : \mathscr{Y} \rightrightarrows \mathscr{Z}$ be Ss-R at $(\hat{y},\hat{z})$ with parameters α',β' and κ'. Let the positive numbers $\varepsilon,\mu,\kappa,\alpha,\beta$ satisfy the relations*

$$\alpha \leq \alpha', \quad \beta + \mu\alpha + \varepsilon \leq \beta', \quad \mu\gamma\kappa' < 1, \quad \kappa = \frac{\kappa'}{1-\mu\gamma\kappa'}. \tag{3}$$

Then for every function $\varphi : \mathscr{Y} \to \mathscr{Z}$ that satisfies the conditions

$$d_{\mathscr{Z}}(\varphi(\hat{y}),0) \leq \varepsilon, \quad d_{\mathscr{Z}}(\varphi(y),\varphi(\hat{y})) \leq \mu\, d(y,\hat{y}) \quad \forall y \in \boldsymbol{B}(\hat{y};\alpha),$$

the map $\varphi + \Phi$ is strongly sub-regular at $(\hat{y},\hat{z}+\varphi(\hat{y}))$ with parameters α,β and κ.

Proof. Obviously, $(\hat{y},\hat{z}+\varphi(\hat{y})) \in \operatorname{graph}(\varphi+\Phi)$. Let us fix arbitrarily $z \in \boldsymbol{B}_{\mathscr{Z}}(\hat{z};\beta)$ and $y \in \boldsymbol{B}(\hat{y};\alpha) \subset \boldsymbol{B}(\hat{y};\alpha')$ such that $z \in \varphi(y)+\Phi(y)$. Then $z-\varphi(y) \in \Phi(y)$, and

$$\begin{aligned} d_{\mathscr{Z}}(z-\varphi(y),\hat{z}) &\leq d_{\mathscr{Z}}(z,\hat{z}) + d_{\mathscr{Z}}(\varphi(y),\varphi(\hat{y})) + d_{\mathscr{Z}}(\varphi(\hat{y}),0) \\ &\leq \beta + \mu\alpha + \varepsilon \leq \beta'. \end{aligned}$$

Due to the Ss-R property of Φ we estimate

$$
\begin{aligned}
d^*(y,\hat{y}) &\leq \kappa' d_{\mathscr{Z}}(z-\varphi(y),\hat{z}) \leq \kappa' d_{\mathscr{Z}}(z,\hat{z}+\varphi(\hat{y})) + \kappa' d_{\mathscr{Z}}(\varphi(y),\varphi(\hat{y})) \\
&\leq \kappa' d_{\mathscr{Z}}(z,\hat{z}+\varphi(\hat{y})) + \kappa' \mu\, d(y,\hat{y}) \leq \kappa' d_{\mathscr{Z}}(z,\hat{z}+\varphi(\hat{y})) + \kappa' \mu \gamma d^*(y,\hat{y}),
\end{aligned}
$$

which implies the claim of the theorem due to the definition of κ in (3). □

2.2 The optimality map

Problem $\mathscr{P}_p(0,x_0)$ given by (1)–(2) will be considered under the following assumptions.

Assumption (A1). The set U is convex and compact. The functions f and g are two times differentiable with respect to (x,u), these functions and their first and second derivatives in (x,u) are Lipschitz continuous (with respect to (p,x,u)) with a Lipschitz constant L.

For any $p \in \Pi$, along with problem $\mathscr{P}_p(0,x_0)$ we consider the family, denoted by $\mathscr{P}_p(\tau,x_\tau)$, consisting of problems which have the same form as (1)–(2) but with the initial time 0 replaced with any $\tau \in [0,T)$ and x_0 replaced with any $x_\tau \in \mathbb{R}^n$. Of course, then only the restriction of the parameter p to $[\tau,T]$ matters.

Assumption (A2). For every $u \in \mathscr{U}$, $x_0 \in \mathbb{R}^n$, and $p \in \mathit{\Pi}$ equation (2) has a solution x on $[0,T]$ (which is then unique due to Assumption (A1)). For every $\tau \in [0,T)$, $x_\tau \in \mathbb{R}^n$ and $p \in \mathit{\Pi}$ problem $\mathscr{P}_p(\tau,x_\tau)$ has an optimal solution.

Since the analysis in this paper is local, only local Lipschitz continuity of the functions mentioned in Assumption (A1) is needed; we assume global Lipschitz continuity to avoid routine technicalities. Assumption (A2) is also stronger than necessary, again for the sake of transparency. Local existence around a reference parameter and trajectory (to be introduced later) suffices.

Remark 2.3. Optimality in the last assumption means local optimality of the objective functional with respect to the L^1-norm of the controls. In fact, it is only needed that any solution (x,u) of $\mathscr{P}_p(\tau,x_\tau)$ satisfies, together with an absolutely continuous (*adjoint*) function $\lambda : [\tau,T] \to \mathbb{R}^n$, the *optimality (Pontryagin) system*

$$
\begin{aligned}
0 &= -\dot{x}(t) + f(p(t),x(t),u(t)), \quad x(\tau) - x_\tau = 0, &(4)\\
0 &= \dot{\lambda}(t) + \nabla_x H(p(t),x(t),\lambda(t),u(t)), \quad \lambda(T) = 0, &(5)\\
0 &\in \nabla_u H(p(t),x(t),\lambda(t),u(t)) + N_U(u(t)), &(6)
\end{aligned}
$$

where

$$
N_U(u) := \begin{cases} \{q \in \mathbb{R}^n \mid \langle q, v-u\rangle \leq 0 \text{ for all } v \in U\} & \text{if } u \in U, \\ \emptyset & \text{if } u \notin U \end{cases}
$$

is the normal cone to U at u, and the Hamiltonian H is defined as usual:

$$
H(p,x,\lambda,u) := g(p,x,u) + \langle \lambda, f(p,x,u)\rangle.
$$

Next, we reformulate the optimality system in functional spaces. The space $L^q(\tau,T)$, $q = 1,2,\ldots,\infty$, of vector functions on $[\tau,T]$ (with any fixed dimension) has the usual meaning, with the norm denoted by $\|\cdot\|_q$. The space of all absolutely continuous vector functions on $[\tau,T]$ is denoted by $W^{1,1}(\tau,T)$, with the norm

$\|x\|_{1,1} = \|x\|_1 + \|\dot{x}\|_1$. Moreover, $W_T^{1,1}(\tau,T)$ is the space of functions $\lambda \in W^{1,1}(\tau,T)$ for which $\lambda(T) = 0$. The notations of norms do not include the time horizon, but it will be clear from the context. For the same reason we often skip the time horizon from the notations of spaces.

Denote

$$Y_\tau := W^{1,1}(\tau,T) \times W_T^{1,1}(\tau,T) \times \mathscr{U}_\tau, \qquad Z_\tau := L^1(\tau,T) \times \mathbb{R}^n \times L^1(\tau,T) \times L^\infty(\tau,T),$$

where $\mathscr{U}_\tau = \{u \in L^\infty(\tau,T) \,:\, u(t) \in U \text{ for a.e. } t \in [\tau,T]\}$ is the set of admissible control functions on $[\tau,T]$, thus $\mathscr{U}_0 = \mathscr{U}$. We also set $Y := Y_0$ and $Z := Z_0$. The metrics in Y_τ and Z_τ are given in terms of norms as follows: for $y = (x,\lambda,u) \in Y_\tau$ and $z = (\xi,\nu,\eta,\rho) \in Z_\tau$

$$d(y,0) := \|y\| := \|x\|_{1,1} + \|\lambda\|_{1,1} + \|u\|_1, \qquad d_Z(z,0) := \|z\| := \|\xi\|_1 + |\nu| + \|\eta\|_1 + \|\rho\|_\infty.$$

In addition, we define in Y a second metric, d^*, as follows. Let $\Gamma \subset [0,T]$ be a fixed finite set. For $u_1,u_2 \in \mathscr{U}_\tau$, denote

$$d^*(u_1,u_2) := \inf\{\varepsilon > 0 \,:\, |u_1(t) - u_2(t)| \le \varepsilon \;\text{ for a.e. }\; t \in [0,T] \setminus (\Gamma + [-\varepsilon,\varepsilon])\}. \tag{7}$$

Somewhat overloading the notation, we define in Y the shift-invariant metric

$$d^*(y,0) := \|x\|_{1,1} + \|\lambda\|_{1,1} + d^*(u,0).$$

Lemma 2.4. *For every $u_1,u_2 \in \mathscr{U}$ it holds that*

$$\|u_1 - u_2\|_1 \le \gamma\, d^*(u_1,u_2),$$

where $\gamma := \max\{1, T + 2M \operatorname{diam}(U)\}$ and M is the number of points in Γ.

The proof is straightforward. Since $\gamma \ge 1$, we also have $\|y\| \le \gamma d^*_\tau(y,0)$ for any $y \in Y_\tau$.

Below we use the same notation for the Nemytskii operator and for its generating function: $f(p,x,u)(t) = f(p(t),x(t),u(t)$, $\nabla_x H(p,x,\lambda,u)(t) = \nabla_x H(p(t),x(t),\lambda(t),u(t))$, etc. For any $p \in \Pi$, $\tau \in [0,T)$, and $x_\tau \in \mathbb{R}^n$ define on Y_τ the set-valued map

$$\Phi_{(p,\tau,x_\tau)}(y) = F_{(p,\tau,x_\tau)}(y) + \begin{pmatrix} 0 \\ 0 \\ 0 \\ N_{\mathscr{U}_\tau}(u) \end{pmatrix}, \qquad F_{(p,\tau,x_\tau)}(y) = \begin{pmatrix} -\dot{x} + f(p,x,u) \\ x(\tau) - x_\tau \\ \dot{\lambda} + \nabla_x H(p,x,\lambda,u) \\ \nabla_u H(p,x,\lambda,u) \end{pmatrix},$$

where now $N_{\mathscr{U}_\tau}(u)$ is the normal cone to $\mathscr{U}_\tau$ at u in the space $L^1(\tau,T)$, that is,

$$\begin{aligned} N_{\mathscr{U}_\tau}(u) \;&:=\; \begin{cases} \{l \in L^\infty(\tau,T) \,:\, \int_\tau^T \langle l(t), v(t) - u(t)\rangle\, \mathrm{d}t \le 0 \text{ for all } v \in \mathscr{U}_\tau\} & \text{if } u \in \mathscr{U}_\tau, \\ \emptyset & \text{if } u \notin \mathscr{U}_\tau \end{cases} \\ &= \{l \in L^\infty(\tau,T) \,:\, l(t) \in N_U(u(t)) \text{ for a.e. } t \in [\tau,T]\}. \end{aligned}$$

With these notations one can recast the optimality system for problem $\mathscr{P}_p(\tau,x_\tau)$ as

$$0 \in \Phi_{(p,\tau,x_\tau)}(x,\lambda,u),$$

therefore the map $\Phi_{(p,\tau,x_\tau)}$ is called *optimality map*. Obviously, due to the compactness of the set U, $\Phi_{(p,\tau,x_\tau)}$ is a set-valued map from Y_τ to Z_τ.

Let us fix a reference parameter $\hat{p} \in \Pi$ and denote by $(\hat{x}, \hat{u})$ a solution of problem $\mathscr{P}_{\hat{p}}(0, x_0)$ (see Remark 2.3). Let $\hat{\lambda}$ be the corresponding adjoint function, so that the triplet $\hat{y} := (\hat{x}, \hat{\lambda}, \hat{u})$ satisfies the optimality system (4)–(6) with $\tau = 0$ and $x_\tau = x_0$, or equivalently, the inclusion $0 \in \Phi_{(\hat{p},0,x_0)}(\hat{y})$. The following assumption plays a key role in the error analysis of the MPC method presented in the next section.

Assumption (A3). The map $\Phi_{(\hat{p},0,x_0)} : Y \rightrightarrows Z$ is strongly sub-regular at $(\hat{y}, 0)$ (in the metrics $\|\cdot\|$ and d^* in Y) with parameters $\hat{\alpha}, \hat{\beta}$ and $\hat{\kappa}$.

The finite set Γ that appears in the definition of the metric d^* is arbitrary, but will be appropriately specified in Section 4 for several classes of problems, together with sufficient conditions for (A3).

3 The accuracy of the model predictive control method

This section presents the main result: the estimate of the accuracy of the model predictive control method, beginning with the description of the method in the context of finite horizon optimal control.

3.1 The model predictive control method

The MPC, applied to optimal control problems containing uncertain parameters, is a method for approximation of an optimal feedback control in real time by successively solving open-loop optimal control problems. Each of these open-loop problems involve measurements of the current system state and predictions for the uncertain parameters. In the next three paragraphs we present a version of the MPC method.

The optimal control problem into question is problem $\mathscr{P}_p(0, x_0)$, considered under Assumptions (A1)–(A3), where the parameter function $p \in \Pi$ is uncertain. It is assumed that for some parameter $\hat{p} \in \Pi$ equation (2) with $p = \hat{p}$ reproduces a "real" system, the states of which can be measured (with a measurement errors). As in the previous subsections we denote by $(\hat{x}, \hat{u})$ a reference optimal solution of $\mathscr{P}_{\hat{p}}(0, x_0)$.

Given a natural number N, we denote by $\{t_k\}_{k=0}^N$ the grid with step-size $h = T/N$, that is, $t_k = kh$, $k = 0, \ldots, N$. To describe the k-th stage of the MPC algorithm we assume that an admissible control function u^N is already determined on $[0, t_k]$ and applied to the "real" system. Denote by x^N the corresponding trajectory, that is the solution of (2) with $p = \hat{p}$ and $u = u^N$. Then the state of the "real" system is measured with a measurement error e_k, that is, the vector $x_k^0 = x^N(t_k) + e_k$ becomes available at time t_k. In addition, a prediction $p_k \in \Pi$ for the time horizon $[t_k, T]$ is made. Then an approximate solution $(\tilde{x}_k, \tilde{u}_k) \in W^{1,1} \times \mathscr{U}_{t_k}$ of the problem $\mathscr{P}_{p_k}(t_k, x_k^0)$ is found, and u^N is extended to $[0, t_{k+1}]$ as $u^N(t) = \tilde{u}_k(t)$ for $t \in (t_k, t_{k+1}]$.

The process continues in the same way as long as $k < N$. The control u^N is called *MPC-generated control* and the corresponding trajectory x^N of the "real" system (2) with $u = u^N$ and $p = \hat{p}$ is called *MPC-generated trajectory*.

Two points are to be clarified. First, the quality of a prediction $p_k \in \Pi$ on $[t_k, T]$ will be measured by the norm $e_k^p := \|p_k - \hat{p}_{[t_k,T]}\|_\infty$. Second, the pair $(\tilde{x}_k, \tilde{u}_k)$ is an approximate solution of problem $\mathscr{P}_{p_k}(t_k, x_k^0)$ in the sense that for some absolutely continuous $\tilde{\lambda}_k$ the triplet $\tilde{y}_k := (\tilde{x}_k, \tilde{\lambda}_k, \tilde{u}_k)$ satisfies the inclusion (approximate optimality conditions)

$$\tilde{z}_k \in \Phi_{(p_k, t_k, x_k^0)}(\tilde{y}_k) \tag{8}$$

with some $\tilde{z}_k \in Z_\tau$. We mention that most of the numerical methods for optimal control give approximations with a small residual $\tilde{z}_k$. The norm $e_k^u := \|\tilde{z}_k\|$ of the residual will be used as a measure of the accuracy of the approximate solution $(\tilde{x}_k, \tilde{u}_k)$ of problem $\mathscr{P}_{p_k}(t_k, x_k^0)$.

3.2 The main theorem

The formulation of the main theorem uses the notations e_k, e_k^p, e_k^u, $\tilde{u}_k$, u^N and x^N introduced in the description of the MPC algorithm. In particular, (x^N, u^N) is the MPC-generated trajectory-control pair, which is compared in the next theorem with the reference optimal open-loop solution $(\hat{x}, \hat{u})$ of the "real" problem $\mathscr{P}_{\hat{p}}(0, x_0)$.

Theorem 3.1. *Let Assumptions (A1)–(A3) be fulfilled. Then there exists numbers* N_0, $\delta > 0$, C_1, C_2, *and* C_3 *such that for any natural number* $N \geq N_0$, *for any sequence of measurement errors* $\{e_k\}$, *for any sequence of predictions* $p_k \in \Pi$ *and approximation errors* $\{e_k^u\}$ *satisfying the conditions*

$$|e_k| + e_k^p + e_k^u \leq \delta, \qquad \|\tilde{u}_k - \hat{u}\|_1 \leq \delta, \quad k = 0, \ldots, N-1,$$

any MPC-generated trajectory-control pair (x^N, u^N) *satisfies the estimate*

$$\|u^N - \hat{u}\|_1 + \|x^N - \hat{x}\|_{1,1} \leq \begin{cases} C_1 \mathscr{E} & \text{if } \Gamma = \emptyset, \\ C_2 \sqrt{\mathscr{E}} + C_3 h & \text{if } \Gamma \neq \emptyset, \end{cases}$$

where

$$\mathscr{E} := \frac{1}{N} \sum_{k=0}^{N-1} (|e_k| + e_k^p + e_k^u)$$

is the averaged error appearing at the MPC steps.

The proof of the theorem is postponed to Section 5. Below in this subsection we discuss the obtained result and the assumptions.

Remark 3.2. *About Assumption (A3).* Assumptions (A1) and (A2) are standard and non-restrictive, although somewhat stronger than necessary, as noted after their formulation. Assumption (A3) has to be explained.

First of all, what is the finite set Γ in the definition of the metric d^* which is involved in (A3) through the definition of strong sub-regularity? This set may depend on the reference optimal control $\hat{u}$ of the unperturbed problem $\mathscr{P}_{\hat{p}}(0, x_0)$. Presumably, it consists of points of discontinuity of $\hat{u}$, but may be larger in order to include points at which the optimal control of a slightly disturbed problem may be discontinuous. Example 3.5 below illustrates this situation. The meaning of the metric d^* is that the distance between two control functions is small in this metric when their values are close to each other, possibly excepting points that are close to the set Γ. This property of the metric with which the Ss-R assumption (A3) is fulfilled is of key importance for that convergence and error analysis of the MPC method.

In the next section we shall provide sufficient conditions under which Assumption (A3) is fulfilled in particular classes of problems with empty or non-empty set Γ.

Remark 3.3. *Discussion on the theorem.* Theorem 3.1 estimates the error of the MPC-generated solution, compared with the optimal solution of the reference (unperturbed) problem, caused by prediction errors $\{e_k^p\}_k$, measurement errors $\{e_k\}_k$, approximation errors $\{e_k^u\}_k$, and the sampling size h. An important point is that the error estimate in the theorem depends on the average error, $\mathscr{E}$, which means that relatively large errors may occasionally appear at some MPC steps without a substantial influence on $\mathscr{E}$.

A similar result as in Theorem 3.1 is obtained in [10] in the case $\Gamma = \emptyset$ (see Subsection 4.1 of the present paper). Here we mention that in [10] a prediction made at the beginning is only used, that is, $p_k = p_0$ for all k. Moreover, the result in [10] only applies to the Euler method for approximate solving the auxiliary problems involved in the MPC algorithm.

Remark 3.4. *About the approximate solution of the auxiliary problems $\mathscr{P}_p(\tau, x_\tau)$.* Finding an approximate solution of the auxiliary problems is a separate issue that we do not address in detail in this paper. Numerical solutions usually involve time-discretization. Discretization methods with first and second order accuracy are known for coercive problems (see Subsection 4.1 for the last term), [7, 9], as well as for affine problems with purely bang-bang optimal controls, [1, 21]. The error in solving the resulting mathematical programming problems comes in addition. The above mentioned results are proved under assumptions that imply strong sub-regularity (for appropriate sets Γ) of the optimality maps associated with the considered problems.

Example 3.5. *Sharpness of the estimate.* Consider the problem

$$\min\{x^1(1) - x^2(1)\},$$

$$\begin{aligned} \dot{x}^1(t) &= p(t)x^2(t), \quad x^1(0) = 0, \\ \dot{x}^2(t) &= u(t), \qquad x^2(0) = 0, \qquad u(t)) \in [-1, 1]. \end{aligned}$$

The reference parameter is $\hat{p} \equiv 1$. The measurements are assumed exact, as well as the solutions of the auxiliary problems at the MPC steps. Thus $\mathscr{E} = h\sum_{k=0}^{N-1} \|p_k - \hat{p}\|_\infty$.

The solution of each problem $\mathscr{P}_p(t_k, \hat{x}(t_k))$ is straightforward: here

$$\tilde{\lambda}_k^2(t) = -1 + \int_t^1 p_k(s)\,\mathrm{d}s, \qquad \tilde{u}_k(t) = -\mathrm{sign}\,\tilde{\lambda}^2(t).$$

For $\hat{p}$ we have $\hat{\lambda}^2(t) < 0$ for all $t \in (0, 1]$, hence $\hat{u}(t) \equiv 1$. For $t = 1$ the control function $\hat{u}$ is not determined by the Pontryagin necessary optimality condition, because $\hat{\lambda}^2(0) = 0$. This does not matter from the control perspective, but suggests to define $\Gamma = \{0\}$ (see Remark 3.2). As it will become obvious in the next section, Assumption (A3) is fulfilled for this problem with this Γ.

Let us fix an arbitrary $\delta > 0$ and consider $h = 1/N$ with $N > 2$ and such that $2h \le \delta$. Define $p_0 = 1 + 2h$ and take all other predictions exact: $p_k = 1$, $k = 1, \ldots, N-1$. Then $\|p_0 - \hat{p}\|_\infty \le 2h \le \delta$. Moreover, $\mathscr{E} = 2h^2$.

On the other hand, we have

$$\tilde{\lambda}_0^2(t_1) = -1 + (1-h)(1+2h) = h(1-2h) > 0.$$

Since $\tilde{\lambda}_0^2$ is linear and $\tilde{\lambda}_0^2(1) = -1$, we obtain that $\tilde{\lambda}_0^2(t) > 0$ on $[0, t_1]$. Hence, $u^N(t) = \tilde{u}_0(t) = -1$ on $[0, t_1]$. Then

$$\|u^N - \hat{u}\|_1 \ge 2h.$$

Consequently,

$$\frac{\|u^N - \hat{u}\|_1}{\sqrt{\mathscr{E}}} \ge \frac{2h}{\sqrt{2}h} \ge \sqrt{2}.$$

Since $\mathscr{E}$ can be arbitrarily small (for small h), the estimation in the theorem is sharp.

4 Sufficient conditions for strong sub-regularity of the optimality map

In this section we present some classes of problems for which Assumption (A3) has a more particular form with a specified set Γ, thus Theorem 3.1 is applicable. In addition, we further discuss the approximation issue mentioned in Remark 3.4.

4.1 The case of coercive problems

Following [8], in this subsection we consider the reference problem $\mathscr{P}_{\hat{p}}(0,x_0)$ under the so-called *coercivity condition*. To formulate it we use the following notational convention: we skip arguments of functions with "hat", shifting the "hat" over the notation of the function, e.g. $\hat{f}_x(t) := f_x(\hat{p}(t),\hat{x}(t),\hat{u}(t))$, $\hat{H}_{xx}(t) := H_{xx}(\hat{p}(t),\hat{x}(t),\hat{\lambda}(t),\hat{u}(t))$, etc. Here f_x and H_{xx} are the derivative of f and the Jacobian of H, respectively.

Assumption (B1). There is a constant $c_0 > 0$ such for any $v \in \mathscr{U} - \mathscr{U}$ the inequality

$$\int_0^T \left[\langle \hat{H}_{xx}(t)x(t),x(t)\rangle + 2\langle \hat{H}_{ux}(t)x(t),v(t)\rangle + \langle \hat{H}_{uu}(t)v(t),v(t)\rangle\right] \,\mathrm{d}t \geq c_0\|v\|_2^2$$

is fulfilled, where x is the (unique) solution of the equation $\dot{x}(t) = \hat{f}_x(t)x(t) + \hat{f}_u(t)v(t)$ with $x(0) = 0$.

It was proved in [8] that Assumption (B1), together with (A1) and (A2), implies (A3) with $\Gamma = \emptyset$, thus in this case the metric in $\mathscr{U}$ is $d^* = \|\cdot\|_\infty$, thus the metric in Y can be taken to be $d^*(y,0) = \|x\|_{1,\infty} + \|\lambda\|_{1,\infty} + \|u\|_\infty$.[1]

Even more, Assumptions (A1), (A2), (B1) imply the stronger property of *Strong metric Regularity* (SR), [13, Sect. 3.7] with respect to the norm $\|x\|_{1,\infty} + \|\lambda\|_{1,\infty} + \|u\|_\infty$ in the space Y. An important fact is, that the property SR is stable with respect to functional perturbations with a sufficiently small Lipschitz constant see, e.g., [13, Proposition 3G.2]). Then it is easy to see that for sufficiently small inaccuracies $|e_k|$, e_k^p and e_k^u all the maps $\Phi_{(p,t_k,x_k^0)}$ that appear in the MPC algorithm are SR, hence Ss-R, with constants that can be chosen uniformly with respect to k. In connection with Remark 3.4, we mention that thanks to the strong regularity of the optimality map (or the uniform strong sub-regularity) one can claim $O(h)$ uniform estimation of e_k^u if the Euler discretization with step size h is used in solving the problems $\mathscr{P}_p(t_k,x_k^0)$ (see [16]), and uniform convergence of the Newton method (see [2]). However, this issue is not at the focus of the present paper and we do not give precise formulations and details.

4.2 The case of affine problems with bang-bang optimal controls

In this subsection, we consider problem $\mathscr{P}_{\hat{p}}(0,x_0)$ to be affine, i.e., the objective integrand, g, and the right-hand side, f, in (2) are both affine with respect to u. The set U is assumed to be a convex compact polyhedron. Using geometric terminology, we denote by V the set of vertices of U, and by E the set of all unit vectors $e \in \mathbb{R}^m$ that are parallel to some edge of U. As usual, we define the so-called switching function $\hat{\sigma} : [0,T] \to \mathbb{R}^m$ by $\hat{\sigma}(t) := \hat{H}_u(t)$. Here and further we use the notational convention from the previous subsection: arguments of functions with "hat" are skipped and the "hat" is shifted over the notation of the functions.

Versions of the following assumption are standard in the literature on affine optimal control problems, see, e.g., [1, 5, 17, 20].

Assumption (C1). There exist numbers $\eta_0 > 0$ and $\mu_0 > 0$ such that if $s \in [0,T]$ is a zero of $\langle \hat{\sigma}, e\rangle$ for some $e \in E$, then

$$|\langle \hat{\sigma}(t), e\rangle| \geq \mu_0 |t - s|,$$

for all $t \in [s - \eta_0, s + \eta_0] \cap [0,T]$.

[1] The terminology of metric regularity was not used in [8] and the control system considered was stationary, but the result was easily extended to the non-stationary case in many subsequent contributions, see e.g. [12].

Assumption (C1) implies that $\hat{u}$ is bang bang and that, in particular, the set

$$\Gamma := \{s \in [0,T] : \langle \hat{\sigma}(s), e\rangle = 0 \text{ for some } e \in E\}$$

is finite. In what follows in this subsection, the metric d^* in Y is defined through this set Γ, see (7). As it will be seen in the proof of Theorem 3.1, the advantage of using this metric instead of the L^1-norm is that $d^*(u,\hat{u})$ being small not only implies that $\|u-\hat{u}\|_1$ is small, but also that $|u(t)-\hat{u}(t)|$ is small except on a small set around the zeros of the switching functions $\langle \hat{\sigma}(t), e\rangle$, $e \in E$. In this sense, u is structurally similar to $\hat{u}$.

Given $\varepsilon \geq 0$, we denote

$$\Sigma(\varepsilon) := [0,T] \setminus (\Gamma + [-\varepsilon, \varepsilon]).$$

We recall the following lemma proved in the recent paper [6].

Lemma 4.1. *[6, Lemma 2] Let Assumption (C2) be fulfilled. Then there exist positive numbers κ and ε such that for every functions $\sigma \in L^\infty$ with $\|\sigma - \hat{\sigma}\|_\infty \leq \varepsilon$ and for every $u \in \mathscr{U}$ satisfying $\sigma(t) + N_U(u(t)) \ni 0$ for a.e. $t \in [0,T]$ it holds that*

$$u(t) = \hat{u}(t) \text{ for a.e. } t \in \Sigma\big(\kappa \|\sigma - \hat{\sigma}\|_\infty\big).$$

With this lemma at hand, we are now ready to establish the following sufficient condition for the fulfillment of Assumption (A3).

Theorem 4.2. *Let problem $\mathscr{P}_{\hat{p}}(0,x_0)$ be affine and let Assumptions (A1), (A2), and (C1) be fulfilled. Then the following statements are equivalent:*

(i) The map $\Phi_{(\hat{p},0,x_0)} : Y \rightrightarrows Z$ is strongly sub-regular at $(\hat{y},0)$ in the single metric d in Y;

(ii) The map $\Phi_{(\hat{p},0,x_0)} : Y \rightrightarrows Z$ is strongly sub-regular at $(\hat{y},0)$ in the metrics d and d^* in Y.

Proof. The implication (ii) $\Rightarrow$ (i) is obvious. Let us prove the converse implication.

Let $\tilde{\alpha}, \tilde{\beta}$ and $\tilde{\kappa}$ be the parameters of Ss-R of $\Phi_{(\hat{p},0,x_0)}$ (in the single metric d in Y). Then

$$\|x-\hat{x}\|_{1,1} + \|\lambda - \hat{\lambda}\|_{1,1} + \|u-\hat{u}\|_1 \leq \tilde{\kappa} d_Z(z,0) \tag{9}$$

for all $y = (x,\lambda,u) \in B(\hat{y};\tilde{\alpha})$ and $z = (\xi,\nu,\eta,\rho) \in B_Z(0,\tilde{\beta})$ satisfying $z \in \Phi_{(\hat{p},0,x_0)}(y)$. Let us fix arbitrarily such a pair (y,z). Define $\sigma : [0,T] \to \mathbb{R}^m$ by $\sigma(t) := \nabla_u H(\hat{p}, y(t)) - \rho$. Clearly $\sigma(t) + N_U(u(t)) \ni 0$ for a.e. $t \in [0,T]$. Moreover, due to the affine structure of the problem, $\nabla_u H(p,y)$ is independent of u, thus we can estimate

$$\begin{aligned}
\|\sigma - \hat{\sigma}\|_\infty &\leq \|\nabla_u H(\hat{p},y) - \nabla_u H(\hat{p},\hat{y})\|_\infty + \|\rho\|_\infty \leq L_1(\|x-\hat{x}\|_\infty + \|\lambda - \hat{\lambda}\|_\infty) + \|\rho\|_\infty \\
&\leq L_1 \tilde{\kappa} d_Z(z,0) + d_Z(z,0) =: \tilde{c} d_Z(z,0),
\end{aligned}$$

where L_1 is the Lipschitz constant of $\nabla_u H(\hat{p},\cdot)$.

Define

$$\hat{\alpha} = \tilde{\alpha}, \quad \hat{\beta} = \min\left\{\tilde{\beta}, \varepsilon/\tilde{c}\right\}, \quad \hat{\kappa} = \tilde{\kappa} + \tilde{c}\kappa,$$

where ε and κ are the numbers in Lemma 4.1. For the pair (y,z) we additionally assume that $z \in B_Z(0,\hat{\beta})$. Then by Lemma 4.1,

$$u(t) = \hat{u}(t) \text{ for a.e. } t \in \Sigma\big(\kappa \tilde{c} d_Z(z,0)\big),$$

which directly implies $d^*(u,\hat{u}) \le \kappa\tilde{c}d_Z(z,0)$. Together with (9), we obtain that

$$d^*(y,\hat{y}) \le \tilde{c}\kappa d_Z(z,0) + \tilde{\kappa} d_Z(z,0) = \hat{\kappa} d_Z(z,0),$$

which proves (ii) with Ss-R constants $\hat{\alpha}$, $\hat{\beta}$, $\hat{\kappa}$. □

A general sufficient condition for strong sub-regularity of the map $\Phi_{(\hat{p},0,x_0)} : Y \rightrightarrows Z$ in the single metric d in Y is given in [20, Theorem 3.1]. It involves the following assumption.

Assumption (C2). There is a constant $c_0 > 0$ such for any $v \in \mathscr{U} - \hat{u}$ the inequality

$$\int_0^T \langle \hat{H}_u(t), v(t)\rangle \, \mathrm{d}t + \int_0^T \left[\langle \hat{H}_{xx}(t)x(t), x(t)\rangle + 2\langle \hat{H}_{ux}(t)x(t), v(t)\rangle\right] \, \mathrm{d}t \ge c_0\|v\|_1^2 \tag{10}$$

is fulfilled, where x is the (unique) solution of the equation $\dot{x}(t) = \hat{f}_x(t)x(t) + \hat{f}_u(t)v(t)$ with $x(0) = 0$.

In [20, Theorem 3.1] it is proved that Assumption (C2), together with (A1), (A2), and the affine structure of the problem, implies metric sub-regularity of the optimality map in the single metric d in Y. In contrast to the L^2 coercivity condition in the previous subsection, Assumption (C2) requires "coercivity" with respect to the L^1-norm. It is well known that Assumption (B1) does not hold for affine problems, see [12, Lemma 3]. Another difference between (B1) and (C2) is that the inequality in (C2) involves not only a quadratic, but also a linear form of v. It is remarkable that alone this linear term can ensure fulfillment of Assumption (C1). Indeed, in [20, Proposition 4.1] it is proved that Assumption (C1) implies the inequality

$$\int_0^T \langle \hat{H}_u(t), v(t)\rangle \, \mathrm{d}t \ge c_1\|v\|_1^2$$

for a constant $c_1 > 0$ and all $v \in \mathscr{U} - \hat{u}$. In particular, if the quadratic form in (10) is nonnegative for $v \in \mathscr{U} - \hat{u}$, then (A1), (A2), (C1) imply (C2), hence also Assumption (A3).

4.3 A numerical example

In this section we illustrate the result obtained in Theorem 3.1 by considering a problem of axisymmetric spacecraft spin stabilization from [24, p. 353]. The transversal angular velocity components ω_1 and ω_2 of the spacecraft satisfy

$$\dot{\omega}_1 = \lambda\,\omega_2 + \frac{M_d}{J_t},$$
$$\dot{\omega}_2 = -\lambda\,\omega_1 + \frac{M_c}{J_t},$$

where $\lambda = \frac{J_t - J_3}{J_t} n$, J_t is the spacecraft transversal moment of inertia, J_3 is the spacecraft moment of inertia about the spin axis, n is the spin rate, M_d is the disturbance torque, which can be caused by thruster misalignment, and M_c is the control moment. Rescaling the time $(t \to \lambda t)$, denoting $x_1 = \omega_1$, $x_2 = \omega_2$, $p = \frac{M_d}{J_t}$, adding initial conditions, and considering $u := \frac{M_c}{J_t}$ as a control variable, we reformulate the model as

$$\begin{cases} \dot{x}_1 = x_2 + p, & x_1(0) = 1, \\ \dot{x}_2 = -x_1 + u, & x_2(0) = 1, \\ -a \le u \le a, \end{cases} \tag{11}$$

where $p(\cdot)$ is a time-dependent parameter and a is a positive constant. The MPC algorithm is applied in [11] to the following optimal control problem with the dynamic (11):

$$\min\left\{|x(T)|^2 + \alpha \int_0^T (u(t))^2 \,\mathrm{d}t\right\},$$

where α is a positive weighting parameter. This problem is coercive in the sense of Assumption (B1), which makes the analysis in [11] possible. Here we consider the alternative objective functional

$$\min\left\{|x(T)|^2 + \alpha \int_0^T |u(t)| \,\mathrm{d}t\right\}, \tag{12}$$

which may be more realistic in case of direct transformation of fuel into force, as in reaction engines. The optimal control problem (11)-(12) is not coercive, nor does it fit in the framework of affine problems. However, following [23, Remark 3.3], we transform it to an affine problem by substituting

$$u = u_1 - u_2, \quad |u| = u_1 + u_2, \quad \text{where} \quad u_1, u_2 \in [0,1].$$

Thus, the affine optimal control problem we will consider is

$$\min\left\{|x(T)|^2 + \alpha \int_0^T [u_1(t) + u_2(t)] \,\mathrm{d}t\right\},$$

subject to

$$\begin{cases} \dot{x}_1 = x_2 + p, & x_1(0) = 1, \\ \dot{x}_2 = -x_1 + u_1 - u_2, & x_2(0) = 1, \\ 0 \le u_1, u_2 \le a. \end{cases}$$

We consider the last problem with the specifications $T = 4\pi$, $\alpha = 0.25$, $a = 0.2$, and reference parameter $\hat{p}(t) \equiv 0$. In the MPC simulation, the measurement error e_k is sampled randomly from a uniform distribution, with $|e_k| \le 0.1$. The parameter $p(\cdot)$ is piecewise constant on the uniform mesh of 3200 points in $[0,T]$; its values in every subinterval are chosen randomly in the interval $[-0.05, 0.05]$ with uniform distribution. For solving the auxiliary problems $\mathscr{P}_p(t_k, x_k^0)$ we use the Euler discretization scheme, which provides an error e_k^u of order $O(h)$ see, e.g., [1, 20]); recall that $e_k^u := \|\tilde{z}_k\|$ and that $\tilde{z}_k$ is the residual in (8).

We run the MPC algorithm with different mesh sizes N. Using the notations in Theorem 3.1, we consider the quantity

$$RE = \frac{||u^N - \hat{u}||_1 + ||x^N - \hat{x}||_{1,1}}{\sqrt{\frac{1}{N}\sum_{k=0}^{N-1}(|e_k| + e_k^p + h) + h}}, \tag{13}$$

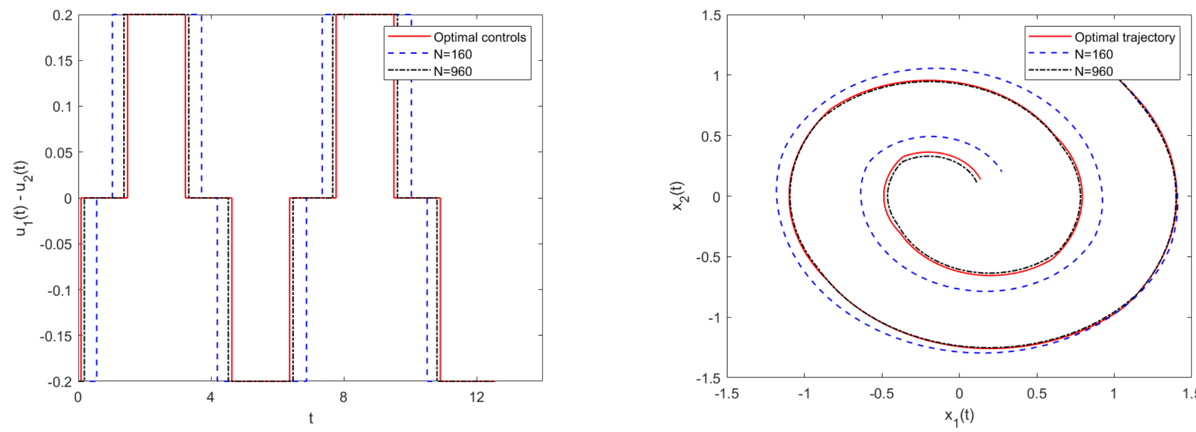
which represents the relative error in the MPC-generated solution (x^N, u^N). According to the error estimate in Theorem 3.1, the quantity RE should be bounded. The numerical experiment confirms this, as can be seen in Table 1. Moreover, the result suggest that the value RE stays away from zero when N increases, which indicates that the estimate in Theorem 3.1 is sharp for this example.

We also observe in Table 1 that the objective values for the MPC-generated solutions decrease when N increases, which is to be expected because of the more frequent measurements.

In Figure 1, we compare the obtained MPC-generated controls and the corresponding trajectories with the open-loop solution $\hat{u} = \hat{u}_1 - \hat{u}_2$. The auxiliary controls $\hat{u}_1$ and $\hat{u}_2$ are of bang-bang type, while the resulting optimal control u in problem (11)-(12) also takes value zero. The MPC-generated control u^N differs from the optimal open-loop one in small intervals around the switching points of the latter, which is consistent with the choice of the metric in d^* in case of affine problems.

N	160	320	480	640	800	960
Obj. val.	0.3253	0.2423	0.2233	0.2220	0.2151	0.2119
RE	0.4343	0.2086	0.1535	0.2481	0.1923	0.1765

Table 1: Objective values and relative errors *RE* of the MPC-generated solutions with different mesh sizes.



(a) Control solutions

(b) Trajectories

Figure 1: Red: optimal open-loop control and trajectory; Blue and Black: MPC generated solutions with N=160 and N=960, correspondingly.

5 Proof of the main theorem

In this section we use the notations introduced in Sections 2 and 3. In addition, $y \in Y_\tau$ we define

$$d^*_\tau(y,\hat{y}) := d^*(\tilde{y},\hat{y}), \qquad \text{where } \tilde{y}(t) = \begin{cases} \hat{y}(t) & \text{for } t \in [0,\tau), \\ y(t) & \text{for } t \in [\tau,T]. \end{cases}$$

Proposition 5.1. *Let assumptions (A1)–(A3) be fulfilled. Then there exist numbers* $\delta_0 > 0$*,* $\alpha_0 > 0$*,* $\beta_0 > 0$*, and* c_0 *such that for every* $\tau \in [0,T)$*,* $x_\tau \in \mathbb{R}^n$ *and* $p \in \Pi$ *satisfying*

$$|x_\tau - \hat{x}(\tau)| \le \delta_0, \qquad \|p - \hat{p}\|_\infty \le \delta_0, \tag{14}$$

and for every $y = (x,\lambda,u) \in Y_\tau$ *with* $\|u - \hat{u}\|_1 \le \alpha_0$ *and* $z_\tau \in \Phi_{(p,\tau,x_\tau)}(y) \cap \boldsymbol{B}_{Z_\tau}(0;\beta_0)$ *it holds that*

$$d^*_\tau(y,\hat{y}) \le c_0 \left(\|z_\tau\| + \|p - \hat{p}\|_\infty + |x_\tau - \hat{x}(\tau)|\right).$$

Proof. We shall fix the numbers $\delta_0 > 0$, $\alpha_0 > 0$, $\beta_0 > 0$ and c_0 below, in a way that they depend only on $\hat{\alpha}$, $\hat{\beta}$, $\hat{\kappa}$, L and T and may be viewed as constants. The numbers c_1, c_2, ..., that will appear later will also be appropriate constants in the same sense.

Let us fix arbitrarily $\tau \in [0,T)$, $x_\tau \in \mathbb{R}^n$, and $p \in \Pi$ satisfying (14), along with $y_\tau = (x(\cdot),\lambda(\cdot),u(\cdot)) \in \boldsymbol{B}(\hat{y};\alpha_0)$ and $z_\tau = (\xi,\nu,\eta,\rho) \in \boldsymbol{B}_{Z_\tau}(0;\beta_0)$ such that $z_\tau \in \Phi_{(p,\tau,x_\tau)}(y_\tau)$. We define

$$\tilde{p}(t) = \begin{cases} \hat{p}(t) & \text{for } t \in [0,\tau), \\ p(t) & \text{for } t \in [\tau,T], \end{cases} \qquad \tilde{u}(t) = \begin{cases} \hat{u}(t) & \text{for } t \in [0,\tau), \\ u(t) & \text{for } t \in [\tau,T]. \end{cases}$$

Obviously $\|\tilde{p}-\hat{p}\|_\infty = \|p-\hat{p}\|_\infty$, $\|\tilde{u}-\hat{u}\|_1 = \|u-\hat{u}\|_1$, and $d_0^*(\tilde{u},\hat{u}) = d_\tau^*(u,\hat{u})$. Similarly, we extend ξ and η as zero on $[0,\tau)$, denoting the resulting elements of Z_0 by $\tilde{\xi}$ and $\tilde{\eta}$. Moreover, we define $\tilde{x}$ as the (unique) solution on $[0,T]$ of the equations

$$\dot{\tilde{x}} = f(\tilde{p},\tilde{x},\tilde{u}) - \xi, \quad \tilde{x}(\tau) = x_\tau + \nu,$$

and the function $\tilde{\lambda}$ as the solution of

$$-\dot{\tilde{\lambda}} = \nabla_x H(\tilde{p},\tilde{x},\tilde{\lambda},\tilde{u}) - \eta, \quad \tilde{\lambda}(T) = 0.$$

Notice that $\tilde{x}(t) = x(t)$ and $\tilde{\lambda}(t) = \lambda(t)$ for $t \in [\tau,T]$.

Let us estimate $\|\tilde{y}-\hat{y}\|$. Using Assumption (A1) and the Grönwall inequality we obtain that

$$\|\tilde{x}-\hat{x}\|_\infty \le c_1\left(\|\tilde{p}-\hat{p}\|_\infty + \|\tilde{\xi}\|_1 + |x_\tau - \hat{x}(\tau)| + |\nu|\right) \le c_2\left(\delta_0 + \|\xi\|_1 + \delta_0 + |\nu|\right) \le c_3(\delta_0 + \|z_\tau\|).$$

From here one can also estimate $\|\dot{\tilde{x}}-\dot{\hat{x}}\|_1$, which gives

$$\|\tilde{x}-\hat{x}\|_{1,1} \le c_4(\delta_0 + \|z_\tau\|).$$

Similarly we estimate

$$\|\tilde{\lambda}-\hat{\lambda}\|_{1,1} \le c_5(\delta_0 + \|z_\tau\|).$$

Moreover,

$$\|\tilde{u}-\hat{u}\|_1 = \|u-\hat{u}\|_1 \le \alpha_0.$$

The last three estimates imply that

$$\|\tilde{y}-\hat{y}\| \le c_6(\delta_0 + \alpha_0 + \beta_0) \le \hat{\alpha},$$

provided that the positive numbers δ_0, α_0 and β_0 are chosen sufficiently small.

Now we estimate the residual $r := (r_\xi, r_\nu, r_\eta, r_\rho) \in Z_0$ which $\tilde{y} := (\tilde{x},\tilde{\lambda},\tilde{u})$ gives in $\Phi_{(\hat{p},0,x_0)}$. We have

$$\begin{aligned}
\|r_\xi\|_1 &= \left\|f(\hat{p},\tilde{x},\tilde{u}) - f(\tilde{p},\tilde{x},\tilde{u}) + \tilde{\xi}\right\|_1 \le TL\|p-\hat{p}\|_\infty + \|\xi\|_1, \\
|r_\nu| &= |\tilde{x}(0) - x_0| \le c_1(\|p-\hat{p}\|_\infty + |x_\tau - \hat{x}(\tau)| + |\nu|), \\
\|r_\eta\|_1 &= \left\|\nabla_x H(\hat{p},\tilde{y}) - \nabla_x H(\tilde{p},\tilde{y}) + \tilde{\eta}\right\|_1 \le TL\|p-\hat{p}\|_\infty + \|\eta\|_1, \\
\|r_\rho\|_\infty &= \|\nabla_u H(\hat{p},\tilde{y}) - \nabla_u H(\tilde{p},\tilde{y}) + \tilde{\rho}\|_\infty \le L\|p-\hat{p}\|_\infty + \|\rho\|_\infty.
\end{aligned}$$

Summarizing, we obtain that

$$\|r\| \le c_7\left(\|z_\tau\| + \|p-\hat{p}\|_\infty + |x_\tau - \hat{x}(\tau)|\right).$$

We can choose δ_0 and β_0 smaller if needed so that $\|r\| \le \hat{\beta}$. Due to Assumption (A3) we have that

$$d_0^*(\tilde{y},\hat{y}) \le \hat{\kappa}c_7\left(\|z_\tau\| + \|p-\hat{p}\|_\infty + |x_\tau - \hat{x}(\tau)|\right).$$

Since $d_\tau^*(\tilde{y},\hat{y}) \le d_0^*(\tilde{y},\hat{y})$ and $\tilde{y} = y$ on $[\tau,T]$, we obtain the desired result with $c_0 = \hat{\kappa}c_7$. □

An alternative way to prove the last proposition is first to show that Assumption (A3) holds for all maps $\Phi_{(\hat{p},\tau,\hat{x}(\tau))}$ with $\tau \in [0,T)$, and then to apply Proposition 2.2.

Now we continue with the proof of Theorem 3.1.

Let δ_0, α_0, β_0 and c_0 be the constants from Proposition 5.1. Define the following constants: M is number of elements of Γ (equals zero if $\Gamma = \emptyset$) and

$$D := \text{ diam}(U), \qquad \bar{C}_1 := c_0 LTe^{LT}, \qquad \bar{C}_2 := 2DLe^{TL}\sqrt{c_0 TM}, \qquad \bar{C}_3 := 6MDLe^{LT}. \tag{15}$$

Let the numbers N_0 and $\delta > 0$ be defined in such a way that

$$\hat{\delta} + \delta \le \delta_0, \qquad \delta \le \alpha_0, \qquad \delta \le \beta_0,$$

where $\hat{\delta} := \bar{C}_1\delta + \bar{C}_2\sqrt{\delta} + \bar{C}_3 h$, which is obviously possible. Moreover, denote

$$\mathscr{E}_i := |e_i| + e_i^p + e_i^u, \;\; i = 0, \ldots, N-1.$$

Since for any $i \in \{0, \ldots, N-1\}$ the triplet $y = \tilde{y}_i$ satisfies (8), we shall apply Proposition 5.1 with $y = \tilde{y}_i$, $\tau = t_i$ $x_\tau = x_i^0 = x^N(t_i) + e_i$, $p = p_i$, $z_\tau = \tilde{z}_i$. We have

$$\|p_i - \hat{p}\|_\infty \le e_i^p \le \delta \le \delta_0,$$

$$\|\tilde{u}_i - \hat{u}\|_1 \le \delta \le \alpha_0,$$

$$\|\tilde{z}_i\| = e_i^u \le \delta \le \beta_0.$$

Proposition 5.1 gives

$$d^*_{t_i}(\tilde{u}_i, \hat{u}) \le c_0(|e_i| + e_i^p + e_i^u) = c_0\mathscr{E}_i, \tag{16}$$

provided that

$$|x^N(t_i) + e_i - \hat{x}(t_i)| \le \delta_0. \tag{17}$$

Let us fix an arbitrary $k \in \{1, \ldots, N-1\}$ and denote

$$d_i := \|\tilde{u}_i - \hat{u}\|_{L^\infty(t_i, t_{i+1})}, \qquad \Delta(t) := |x^N(t) - \hat{x}(t)|, \qquad \Delta_i = \Delta(t_i), \;\; i = 0, \ldots, k.$$

Assume inductively that

$$\Delta_k \le \hat{\delta} \quad \text{ and } \; d^*_{t_i}(\tilde{u}_i, \hat{u}) \le c_0\mathscr{E}_i, \;\; i = 0, \ldots, k. \tag{18}$$

For $k = 0$ we have $\Delta_0 = |x^N(0) - \hat{x}(0)| = 0$. Thus (17) is fulfilled because $|e_0| \le \delta \le \delta_0$. The second inequality in (18) is fulfilled due to (16), thus the inductive assumption is fulfilled for $k = 0$.

Due to Assumption (A1) and the construction of u^N in the MPC method, we have for $t \in [t_k, t_{k+1}]$

$$\Delta(t) \le \Delta_k + \int_{t_k}^t |f(\hat{p}(s), x^N(s), \tilde{u}_k(s)) - f(\hat{p}(s), \hat{x}(s), \hat{u}(s))| \,\mathrm{d}s \le \Delta_k + \int_{t_k}^t (L\Delta(s) + L|\tilde{u}_k(s) - \hat{u}(s)|) \,\mathrm{d}s.$$

Using the Grönwall inequality we obtain that

$$\Delta(t) \le e^{Lh}(\Delta_k + hLd_k).$$

Applied to $\Delta_{k+1} = \Delta(t_{k+1})$, this recursive inequality implies in a standard way that

$$\Delta_{k+1} \le hL\Big(e^{(k+1)hL}d_0 + e^{khL}d_1 + \ldots + e^{hL}d_k\Big) \;\le\; e^{TL}Lh\sum_{i=0}^{k} d_i. \tag{19}$$

The key part of the proof is to estimate $\sum_{i=0}^{k} d_i$. Let us denote

$$\begin{aligned} \bar{K} &:= \{i \in \{0,\dots,k\} \,:\, |\tilde{u}_i(t) - \hat{u}(t)| \le d^*_{t_i}(\tilde{u}_i,\hat{u}) \text{ for a.e. } t \in [t_i,t_{i+1}]\}, \\ K &:= \{0,\dots,k\} \setminus \bar{K}. \end{aligned}$$

Then

$$d_i \le \begin{cases} D & \text{for } i \in K, \\ d^*_{t_i}(\tilde{u}_i,\hat{u}) & \text{for } i \in \bar{K}. \end{cases} \tag{20}$$

Denoting by s the number of elements of K, we have due to the inductive assumption (18), that

$$\sum_{i=0}^{k} d_i = \sum_{i\in K} d_i + \sum_{i\in\bar{K}} d_i \le sD + \sum_{i\in\bar{K}} d^*_{t_i}(\tilde{u}_i,\hat{u}) \le sD + \sum_{i\in\bar{K}} c_0 \mathscr{E}_k \le sD + \frac{c_0 T}{h}\mathscr{E}. \tag{21}$$

Let us assume that $\Gamma \ne \emptyset$, that is, $M > 0$. The definition of the metric d^* in (7) implies that for each $i \in K$ there exists $t \in (t_i,t_{i+1})$ such that

$$\mathrm{dist}(t,\Gamma) \le d^*_{t_i}(\tilde{u}_i,\hat{u}). \tag{22}$$

Let $m(i)$ be the minimal natural number (also including 0) such that

$$((t_i,t_{i+1}) + h[-m(i),m(i)]) \cap \Gamma \ne \emptyset.$$

Then in the case $m(i) > 0$ we have that

$$((t_i,t_{i+1}) + h[-m(i)+1,m(i)-1]) \cap \Gamma = \emptyset,$$

hence

$$(t + h[-m(i)+1,m(i)-1]) \cap \Gamma = \emptyset.$$

Due to (22) we obtain that

$$d^*_{t_i}(\tilde{u}_i,\hat{u}) \ge h(m(i)-1), \quad i \in K. \tag{23}$$

Denote by l_j, $j = 0,1,\dots,N$ the number of those indexes $i \in K$ for which $m(i) = j$. The following relations are apparently satisfied:

$$\begin{aligned} &0 \le l_0 \le M, \\ &0 \le l_j \le 2M, \quad j = 1,\dots,N, \\ &\sum_{j=0}^{N} l_j = s. \end{aligned} \tag{24}$$

Then, having in mind (23),

$$\sum_{i\in K} d^*_{t_i}(\tilde{u}_i,\hat{u}) \;\ge\; h\sum_{i\in K} \max\{0,m(i)-1)\} \ge h\Big(0.l_0 + 0.l_1 + \sum_{j=2}^{N} (j-1)l_j\Big). \tag{25}$$

The minimum of the sum in the right-hand side with respect to $\{l_j\}$ subject to the relations around (24) is attained at

$$l_0 = M,\ l_j = 2M,\ j = 1,\dots,r, \quad l_{r+1} = s - (M + 2Mr),$$

where $r = \left[\frac{s-M}{2M}\right]$ and $[a]$ means the integer part of a. Substituting l_j in (25) we obtain that

$$\sum_{i\in K} d^*_{t_i}(\tilde{u}_i, \hat{u}) \ge 2hM \sum_{j=2}^{r} (j-1) = hMr(r-1) \ge hM\left(\left[\frac{s}{2M}\right]-2\right)^2.$$

From here and the second inequality in (18) we obtain that

$$\left([\frac{s}{2M}]-2\right)^2 \le \frac{c_0}{Mh}\sum_{i=0}^{k}\mathscr{E}_i = \frac{c_0 T}{Mh^2}\mathscr{E},$$

hence

$$s \le 6M + \frac{2}{h}\sqrt{c_0 T M \mathscr{E}}.$$

From (21) we obtain that

$$h\sum_{i=0}^{k} d_i \le D\Big(6Mh + 2\sqrt{c_0 T M}\sqrt{\mathscr{E}}\Big) + c_0 T \mathscr{E}, \tag{26}$$

which combined with (19) and (15) gives

$$\Delta_{k+1} \le \bar{C}_1\mathscr{E} + \bar{C}_2\sqrt{\mathscr{E}} + \bar{C}_3 h. \tag{27}$$

This inequality was obtained in the case $M > 0$. However, in the case $M = 0$ the first term in the final inequality in (21) is missing and the analysis simplifies, resulting in the same estimation for Δ_{k+1} but with $M = 0$, which implies $\bar{C}_2 = \bar{C}_3 = 0$.

Now, in order to verify the inductive assumption (18) we observe that $\mathscr{E} \le \delta$, hence from (27)

$$\Delta_{k+1} \le \bar{C}_1\delta + \bar{C}_2\sqrt{\delta} + \bar{C}_3 h = \hat{\delta},$$

thus the first inequality in (18) is satisfied for $k+1$. The second inequality in (18) for $k+1$ follows from (16), which is fulfilled for $i = k+1$ because (17) holds:

$$|x^N(t_{k+1}) + e_{k+1} - \hat{x}(t_{k+1})| \le \Delta_{k+1} + |e_{k+1}| \le \hat{\delta} + \delta \le \delta_0.$$

This completes the induction step. As a result we have obtained that for any $t \in [0,T]$ if $t \in [t_k, t_{k+1}]$ then

$$\Delta(t) \le \Delta_{k+1} \le \bar{C}_1\mathscr{E} + \bar{C}_2\sqrt{\mathscr{E}} + \bar{C}_3 h. \tag{28}$$

Now we estimate

$$\begin{aligned}\|u^N - \hat{u}\|_1 &= \sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} |u^N(t) - \hat{u}(t)|\,\mathrm{d}t \le \sum_{i=0}^{N-1} h\|u^N - \hat{u}\|_{L^\infty(t_i,t_{i+1})} = \sum_{i=0}^{N-1} h\|\tilde{u}_i - \hat{u}\|_{L^\infty(t_i,t_{i+1})} = h\sum_{i=0}^{N-1} d_i \\ &\le D\Big(6Mh + 2\sqrt{c_0 T M}\sqrt{\mathscr{E}}\Big) + c_0 T\mathscr{E}, \end{aligned} \tag{29}$$

where in the last inequality we use (26) for $k = N-1$. Finally, we have

$$\|\dot{x}^N - \dot{\hat{x}}\|_1 \le \int_0^T |f(\hat{p}(t), x^N(t), u^N(t)) - f(\hat{p}(t), \hat{x}(t), \hat{u}(t))|\,\mathrm{d}t \le LT\|\Delta\|_C + L\|u^N - \hat{u}\|_1.$$

Combining this inequality with (28) and (29), and considering separately the case $M = 0$, we obtain existence of constants C_1, C_2 and C_3 for which the claim of the theorem holds.

References

[1] W. Alt, U. Felgenhauer, M. Seydenschwanz. Euler discretization for a class of nonlinear optimal control problems with control appearing linearly. *Compututaional Optimization and Applications*, **69**:825–856, 2018.

[2] F.J. Aragón, A.L. Dontchev, M. Gaydu, M.H. Geoffroy, V.M. Veliov. *SIAM Journal on Control and Optimization*, **49**(2):339–362], 2011.

[3] J.F. Bonnans, A. Shapiro. *Perturbation analysis of optimization problems*. Springer, 2000.

[4] R. Cibulka, A.L. Dontchev, A.Y. Kruger. Strong metric subregularity of mappings in variational analysis and optimization. *Journal of Mathematical Analysis and Application*, **457**:1247–1282, 2018.

[5] A. Domínguez Corella., M. Quincampoix., V.M. Veliov.: Strong bi-metric regularity in affine optimal control problems. To appear in *Pure and Applied Functional Analysis*. Available at https://orcos.tuwien.ac.at/research/

[6] A. Domínguez Corella., V.M. Veliov. Hölder Regularity in Bang-Bang Type Affine Optimal Control Problems. To appear in *Large-scale scientific computing*. Available at https://orcos.tuwien.ac.at/research/

[7] A. L. Dontchev. An a priori estimate for discrete approximations in nonlinear optimal control. *SIAM J. Control Optim.*, **34**:1315–1328, 1996.

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

[9] A.L. Dontchev, W.W. Hager, V.M. Veliov. Second-order Runge-Kutta approximations in control constrained optimal control. *SIAM J. Numerical Anal.*, **38**(1):202–226, 2000.

[10] A.L. Dontchev, I.V. Kolmanovsky, M. I. Krastanov, V.M. Veliov. Approximating open-loop and closed-loop optimal control by model predictive control. *European Control Conference (ECC)*, 2020, DOI: 10.23919/ECC51009.2020.9143615.

[11] A.L. Dontchev, I.V. Kolmanovsky, M.I. Krastanov, V.M. Veliov, P.T. Vuong. Approximating optimal finite horizon feedback by model predictive control. *Systems & Control Letters*, **139**:104666, May 2020.

[12] A.L. Dontchev, M.I. Krastanov, V.M. Veliov. On the existence of Lipschitz continuous optimal feedback control. *Vietnam Journal of Mathematic*, **47**(3):579–597, 2019.

[13] A.L. Dontchev, R.T. Rockafellar. Characterizations of Lipschitz stability in nonlinear programming *Mathematical programming with data perturbations*, 65–82, Lecture Notes in Pure and Appl. Math., 195, Dekker, New York, 1998.

[14] A.L. Dontchev, R.T. Rockafellar. Regularity and conditioning of solution mappings in variational analysis. *Set-Valued Analysis*, **12**:79–109, 2004.

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

[16] A.L Dontchev, V.M. Veliov. Metric regularity under approximations. *Control and Cybernetics*, **38**(4):1283–1303 , 2009.

[17] U. Felgenhauer.: On stability of bang-bang type controls. SIAM J. of Control and Optimization, **41**(6), 1843-1867 (2003)

[18] L. Grüne, J. Panek. *Nonlinear model predictive control*. Springer, 2011.

[19] D. Klatte, B. Kummer. *Nonsmooth equations in optimization*. Kluwer Academic Publisher, 2002.

[20] N.P. Osmolovskii, V.M. Veliov. Metric sub-regularity in optimal control of affine problems with free end state. ESAIM Control Optim. Calc. Var., **26**, Paper No. 47 (2020)

[21] A. Pietrus, T. Scarinci, V.M. Veliov. High order discrete approximations to Mayer's problems for linear systems. *SIAM J. Control Optim.*, **56**(1):102–119, 2018.

[22] S. V. Raković, W. S. Levine, Eds. *Handbook of Model Predictive Control.* Birkhäuser 2018.

[23] S.P Sethi, G. L. Thompson. Optimal control theory. Applications to management science and economics *Kluwer Academic Publishers, Boston, MA*, 2000.

[24] B. Wie. *Space vehicle dynamics and control*, AIAA , 1998.

[25] I.J. Wolf, W. Marquardt. Fast NMPC schemes for regulatory and economic NMPC–A review. *Journal of Process Control*, **44**:162–183, 2016.