TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

Sensitivity-based Warmstarting for Constrained Model Predictive Control

A.L. Dontchev, I.V. Kolmanovsky, D. Liao-McPherson, M.M. Nicotra, V.M. Veliov

Research Report 2019-08

May 2019

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

Sensitivity-based Warmstarting for Constrained Model Predictive Control*

A.L. Dontchev, I.V. Kolmanovsky, D. Liao-McPherson, M.M. Nicotra, V.M. Veliov

Abstract

Model predictive control (MPC) is of increasing interest in applications for constrained control of multivariable systems. However, one of the major obstacles to its broader use is the computation time and effort required to solve a possibly non-convex optimal control problem (OCP) online. This paper introduces a sensitivity-based warmstarting strategy for systems with nonlinear dynamics and polyhedral constraints with the goal of reducing the computational footprint of MPC controllers. It predicts changes in the solution of the parameterized OCP as the parameter varies, by calculating the semiderivative of the solution. We apply the theory of variational inequalities over polyhedral convex sets, thus avoiding restrictive conditions regarding the activity status of the constraints. A numerical study featuring MPC applied to unmanned aerial vehicles illustrates the proposed approach.

1 Introduction

In Model Predictive Control (MPC) [20, 23] control actions are computed by solving constrained optimal control problems (OCPs) in real-time. MPC can systematically handle nonlinearities and constraints, but it requires solving (possibly approximately) a potentially non-convex OCP at each sampling instance. This motivates research into advanced numerical methods to enable MPC implementation.

*A. Dontchev is with the American Mathematical Society and the University of Michigan, Email: {dontchev@umich.edu}. D. Liao-McPherson and I. Kolmanovsky are with the University of Michigan, Ann Arbor. Email:{dliaomcp, ilya}@umich.edu. M. Nicotra is with the University of Colorado Boulder, Email: {marco.nicotra@colorado.edu}. V. Veliov is with Technical University of Vienna, Austria, Email: {veliov@tuwien.ac.at}. This research is supported by the National Science Foundation Award Number CMMI 1562209, the Austrian Science Foundation (FWF) Grant P31400-N32, and the Australian Research Council (ARC) Project DP160100854.

At each sampling instance model predictive controllers measure or estimate the system state and then solve a discrete-time OCP, using the estimated state as an initial condition in the OCP, to determine the MPC-generated control action. As a result, the OCP is parameterized by the initial state. Often the states of the system at subsequent sampling instances are close so that, if the OCP satisfies appropriate regularity conditions, the solutions of the OCPs will be close as well. If we can determine an estimate of the change of the optimal solution, we can use this information to predict the optimal solution at the next time step and then start an optimization procedure from that prediction resulting in reduced computation time. The practice of exploiting sensitivity estimates to initialize an optimization algorithm is often referred to as sensitivity based warmstarting and is closely related to continuation/homotopy or solution tracking methods.

Many early sensitivity methods, e.g., CGMRES [5], are based on continuation methods for smooth nonlinear equations and cannot directly handle inequality constraints. In [27] Zavala and Biegler proposed an advanced step strategy which exploits the following result established in [21]: if the strong second order sufficient conditions (SSOSC), linear independence constraint qualification (LICQ) and strict complementarity slackness (SCS) condition hold, then the solution mapping is continuously differentiable in a vicinity of the solution. The derivative of the solution mapping, evaluated at the solution obtained at the previous sampling instance, can be used as a predictor for the optimal solution at the next sampling instance. However, the SCS condition is difficult to satisfy at all time instants. E.g., it cannot hold when an inequality constraint changes its activity status; typically, in such cases, the solution mapping is not differentiable with respect to the parameter. A similar method, IPA-SQP [11] computes a derivative based predictor using neighboring extremals and combines it with an sequential quadratic programming (SQP) based corrector. It handles constraint (de)activation using an active set strategy and requires that the SSOSC and LICQ hold.

In [9] Zavala and Anitescu developed a path-following strategy in the framework of parameterized generalized equations (GEs) using an augmented Lagrangian corrector-only approach. They assumed the SOSSC and LICQ but relaxed the SCS assumption, allowing the active set to vary. Similar approaches were proposed in [6] using a sequential convex programming based corrector, and in [7], where a predictor and corrector are derived using tools from nonsmooth analysis. A more elaborate analysis of a path-following method for tracking solution trajectories of parameterized variational inequalities is presented in [2].

If one replaces the LICQ with the weaker Mangasarian-Fromovitz constraint qualification (MFCQ) [3] but requires a strengthened form of the SSOSC, i.e., that the SSOSC holds for all Lagrange multipliers, then the optimal solution can be shown to be directionally differentiable [26]. This is exploited in a paper by Jäschke, Yang and Biegler [29] to enhance the advanced step warmstart by relaxing the SCS assumption used in [27]. However, this method requires solving additional linear and quadratic programming problems to handle the non-uniqueness of the Lagrange multipliers. A survey on sensitivity and solution tracking methods can be found in [30].

To the best of our knowledge, all existing sensitivity based methods require a constraint qualification of some sort, e.g., [6, 9, 11, 27] require the LICQ and [29] requires the MFCQ. Constraint qualifications are difficult verify, both a-priori and a-posteriori, because they require a very accurate estimate of the solution. Moreover, the absence of a constraint qualification can lead to numerical difficulties for most optimization algorithms; in extreme cases this can lead to failure of the optimization routine and the associated MPC controller.

In this paper, we present a novel warmstarting strategy based on a sensitivity analysis of the OCP's parameter to solution mapping; our predictor is based on the Bouligand or B-derivative [25], also known as the semiderivative; it reduces to the standard derivative when the SSOSC, LICQ, and SC all hold. This approach requires no constraint qualification, only a numerically verifiable second order sufficient condition. In exchange, we restrict ourselves to the case where the dynamics are nonlinear but the state and control inequality constraints are convex polyhedra.

1.1 Some useful mappings

This paper makes extensive use of several kinds of set valued mappings. A set-valued mapping $\mathcal{F}$ acting between $\mathbb{R}^k$ and $\mathbb{R}^l$ is denoted as $\mathcal{F} : \mathbb{R}^k \rightrightarrows \mathbb{R}^l$, to distinguish it from a function $f : \mathbb{R}^k \to \mathbb{R}^l$, while its inverse is defined as $y \mapsto \mathcal{F}^{-1}(y) = \{x \mid y \in \mathcal{F}(x)\}$. Given a closed convex set $C \subseteq \mathbb{R}^n$, the tangent cone to C at a point $x \in C$ is the set of all v such that $\frac{1}{\varepsilon_k}(x^k - x) \to v$ for some $x^k \to x, x^k \in C, \varepsilon_k \searrow 0$ the normal cone mapping of C is defined as

$$N_C(v) = \begin{cases} \{y \mid y^T(w - v) \leq 0 \quad \forall w \in C\} & \text{if} \quad v \in C, \\ \emptyset & \text{otherwise,} \end{cases}$$

The polar of a closed, convex cone K is

$$K^\circ = \{y \mid \langle y, x\rangle \leq 0, \forall x \in K\}, \tag{1}$$

and then the tangent cone $T_C(v) = N_C^\circ(v)$. The euclidean projection onto the set C is denoted by $\Pi_C(\cdot)$ and set addition/subtraction is defined as

$$K_1 \pm K_2 = \{z \mid z = z_1 \pm z_2, \quad z_1 \in K_1, z_2 \in K_2\}. \tag{2}$$

For any $x \in C$ and $v \in N_C(x)$ the critical cone to C at x for v is defined as

$$K_C(x, v) = \{y \mid y \in T_C(x), \; y^T v = 0\}. \tag{3}$$

Now suppose C is polyhedral; then, by definition, there exists a matrix Γ and a vector b of appropriate dimensions such that

$$C = \{x \mid \Gamma x \leq b\}. \tag{4}$$

To obtain a computationally tractable expression for inclusions of the type $y \in K_C(x,v)$, define the active constraint set,

(5) $$\mathcal{A}(x) = \{i \in [1,l] \mid \Gamma_i x = b_i\},$$

where l is the number of rows in Γ. Then the critical cone can be expressed as

(6) $$K_C(x,v) = \{y \mid \;\; \Gamma_i y \le 0,\; i \in \mathcal{A}(x),\; y^T v = 0\},$$

see e.g., [14, Theorem 2E.3]. This can be further simplified by noting that a constraint can be deactivated if it is locally redundant with respect to the other constraints. Given an active constraint i define the polyhedral set

(7) $$C_i = \{y \mid \Gamma_j y \le b_j,\; j \in [1,l] \setminus i\}$$

as the set that satisfies all other constraints. Due to the constraint $y^T v = 0$ in (6) if $v + N_{C_i}(x) \ni 0$, the constraint i is redundant, meaning that the critical cone will remain unchanged if one were to ignore the constraint $\Gamma_i y \le 0$. By defining the set of redundant constraints

(8) $$\bar{\mathcal{A}}(x,v) = \{i \in \mathcal{A}(x) \mid v + N_{C_i}(x) \ni 0\},$$

the critical cone (6) can be rewritten as

(9) $$K_C(x,v) = \{y \mid \;\; \Gamma_i y \le 0,\; i \in \bar{\mathcal{A}}(x,v),\; y^T v = 0 \Gamma_i y = 0,\; i \in \mathcal{A}(x) \setminus \bar{\mathcal{A}}(x,v)\},$$

where all the non-redundant constraints are now treated as equalities. In practice, the set $\bar{\mathcal{A}}(x,v)$ can be obtained by checking if $v = \Pi_{C_i}(v)$ for each $i \in \mathcal{A}(x)$.

2 Problem formulation

We consider the following discrete-time OCP:

(10a) $$\min_{x,u} \quad J(x,u) = \varphi(x_N) + \sum_{i=0}^{N-1} \ell(x_i, u_i)$$

subject to

(10b) $$x_{i+1} = f(x_i, u_i), \quad i = 0, \dots, N-1, \quad x_0 = p,$$

(10c) $$x_i \in X_i,\; i = 1, \dots, N,$$

(10d) $$u_i \in U_i,\; i = 0, \dots, N-1,$$

where N is a natural number denoting the discrete-time horizon, $x_i \in \mathbb{R}^n$, $u_i \in \mathbb{R}^m$, $x = (x_1, \dots, x_N)$ and $u = (u_0, \dots, u_{N-1})$ are the discrete-time state and control sequences, and the initial state is regarded as a parameter $p \in \mathbb{R}^n$. The functions $\ell : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $\varphi : \mathbb{R}^n \to \mathbb{R}$, and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are assumed twice continuously differentiable everywhere for simplicity. The families of closed and convex sets $X_i \subseteq \mathbb{R}^n$ and $U_i \subseteq \mathbb{R}^m$ describe state and control constraints that may vary in time. Throughout the paper we assume that each of the sets X_i, $i = 1, \dots, N$, and U_i, $i = 0, \dots, N-1$, is a *polyhedral set*; that is, it can be described by linear inequality constraints of the form

$$X_i = \{x \mid E_i x \le c_i\}, \quad U_i = \{u \mid M_i u \le d_i\} \tag{11}$$

for some matrices E_i, M_i and vectors c_i, d_i of compatible dimensions. We also assume that for a fixed reference value $\bar{p}$ of the parameter problem (10) has a solution $(\bar{x}, \bar{u})$.

Remark 1. *In this paper we consider only the initial state as a parameter. The sensitivity analysis presented can be easily extended to OCPs in which the functions in the cost and state equations also depend on additional parameters e.g., a target state or a previewed time-varying input.*

3 Optimality and sensitivity

Problem (10) is a special case of the following parameterized optimization problem

$$\min_{v} \quad J(v) \tag{12a}$$

$$\text{subject to} \quad g(p, v) = 0, \quad v \in V, \tag{12b}$$

where $v = (u_0, x_1, u_1 \ \dots, \ u_{N-1}, x_N) \in \mathbb{R}^j$, $j = N(m+n)$, $J : \mathbb{R}^j \to \mathbb{R}$ is defined as in (10), the function $g : \mathbb{R}^{n+j} \to \mathbb{R}^d$, $d = nN$, is given by

$$g(p, v) = \begin{bmatrix} x_1 - f(p, u_0) \\ x_2 - f(x_1, u_1) \\ \vdots \\ x_N - f(x_{N-1}, u_{N-1}) \end{bmatrix}, \tag{13}$$

and where $V = U_0 \times X_1 \times U_1 \times \ldots \times U_{N-1} \times X_N$. The associated Lagrangian has the form

$$\mathcal{L}(p, v, q) = J(v) + q^T g(p, v), \tag{14}$$

where $q \in \mathbb{R}^{Nn}$ is the vector of the Lagrange multipliers associated with the equality constraints in (12); in the context of optimal control it is usually called the vector of costates.

Let $\bar{v} = (\bar{u}_0, \bar{x}_1, \bar{u}_1, \dots, \bar{u}_{N-1}, \bar{x}_N)$ be a local minimizer of (12) for a reference value $\bar{p}$ of the parameter. It is known, see e.g., [28, Theorem 6.14] and [15, Theorem 5.1.1], that under the constraint qualification condition

$$\text{the matrix } \nabla_v g(\bar{p}, \bar{v}) \in \mathbb{R}^{Nn \times N(m+n)} \text{ is surjective} \tag{15}$$

the first-order necessary conditions for optimality are

$$\nabla_v \mathcal{L}(p, v, q) + N_V(v) \ni 0, \tag{16}$$

$$g(p, v) = 0. \tag{17}$$

Noting that $\nabla_q \mathcal{L} = g$, defining $z = (v, q)$, $E = V \times \mathbb{R}^d$, and letting $F(p, z) = \nabla_z \mathcal{L}(p, v, q)$, we arrive at the following parameterized variational inequality (VI):

$$F(p, z) + N_E(z) \ni 0. \tag{18}$$

It is easy to show that (15) holds for problem (10). Indeed, denoting $B_0 = \nabla_{u_0} f(\bar{p}, \bar{u}_0)$, $A_i = \nabla_{x_i} f(\bar{x}_i, \bar{u}_i)$, and $B_i = \nabla_{u_i} f(\bar{x}_i, \bar{u}_i), i = 1, \ldots, N-1$, the surjectivity of the matrix $\nabla_v g(\bar{p}, \bar{v})$ becomes the condition that for every $\xi = (\xi_0, \ldots, \xi_{N-1}) \in \mathbb{R}^{Nn}$ the system

$$\begin{array}{rcl} x_1 - B_0 u_0 & = & \xi_0, \\ x_{i+1} - A_i x_i - B_i u_i & = & \xi_i, \quad i = 1, \ldots, N-1, \end{array}$$

has a solution. This condition clearly holds: simply choose an arbitrary sequence $(u_0, \ldots, u_{N-1})$ and determine $(x_1, \ldots, x_N)$ recursively.

Translated to the notation of the optimal control problem (10), the first-order necessary optimality conditions are represented by the following system involving the state equation, a difference variational inequality determined by a backward recursion, a variational inequality for the final state, and a variational inequality for the control:

$$\begin{array}{rcl} x_{i+1} & = & f(x_i, u_i), \quad i = 0, 1, \ldots, N-1, \ x_0 = p, \\ 0 & \in & q_{i-1} + \nabla_x H(x_i, u_i, q_i) + N_{X_i}(x_i), \quad i = 1, \ldots, N-1, \\ 0 & \in & q_{N-1} + \nabla \varphi(x_N) + N_{X_i}(x_N), \\ 0 & \in & \nabla_u H(x_i, u_i, q_i) + N_{U_i}(u_i), \quad i = 0, \ldots, N-1, \end{array} \tag{19}$$

where H is the Hamiltonian defined as

$$H(x, u, q) = \ell(x, u) - q^{\mathsf{T}} f(x, u).$$

In the sequel, we will use the short description (18) of the optimality system (19). The solution mapping of (18) is

$$p \mapsto S(p) = \{z \mid F(p, z) + N_E(z) \ni 0\}. \tag{20}$$

There is a well-developed theory for the properties of $S(p)$, most of which is collected in [14, Chapter 2]. We will use the following definitions:

Definition 1. *(Strong regularity) A set-valued mapping $\mathcal{F} : \mathbb{R}^k \rightrightarrows \mathbb{R}^l$ is said to be strongly regular at $\bar{x}$ for $\bar{y}$ if $\bar{y} \in \mathcal{F}(\bar{x})$ and the inverse $\mathcal{F}^{-1}$ has a Lipschitz localization around $\bar{y}$ for $\bar{y}$; that is, there exist neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that the truncated inverse mapping $V \ni \bar{y} \mapsto \mathcal{F}^{-1}(y) \cap U$ is a function which is Lipschitz continuous around $\bar{y}$.*

Definition 2. *(Semidifferentiability) A function* $\psi : \mathbb{R}^k \to \mathbb{R}^l$ *is said to be* semidifferentiable *at* $\bar{p}$ *if there exists a positively homogeneous function* $D\psi(\bar{p}) : \mathbb{R}^k \to \mathbb{R}^l$ *with the property that for every* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that for every* $\Delta p \in \mathbb{R}^k$ *with* $\|\Delta p\| \leq \delta$ *one has*

$$\|\psi(\bar{p} + \Delta p) - \psi(\bar{p}) - D\psi(\bar{p})(\Delta p)\| \leq \varepsilon \|\Delta p\|.$$

When $D\psi(\bar{p})$ happens to be linear, then it becomes the usual (Fréchet) derivative of ψ at $\bar{p}$. A semidifferentiable function is, in particular, directionally differentiable with the derivative in the direction of h satisfying $\psi'(\bar{p}; h) = D\psi(\bar{p})(h)$. The opposite implication holds if the function ψ is Lipschitz continuous. For more about semidifferentiable functions, including examples, see e.g. [14, Section 2D].

The *critical cone* for (18) at $(p, z) \in \operatorname{gph} S$ is

$$K(p, z) := K_E(z, F(p, z)) = \{w \in T_E(z) \mid w^T F(p, z) = 0\},$$

see Section 1.1 for more information on critical cones. Accordingly, the subspace defined as

$$K^+(p, z) = K(p, z) - K(p, z) \tag{21}$$

is the smallest subspace that includes the critical cone, while the subspace

$$K^-(p, z) = K(p, z) \cap [-K(p, z)] \tag{22}$$

is the largest subspace that is included in $K(p, z)$.

Using these concepts, we state the following theorem, which is a compilation of [14, Theorems 2E.6, 2E.8], and characterizes the strong regularity and the semidifferentiability properties of the solution mapping of (18):

Theorem 1. *Let* $\Lambda = \nabla_z F(\bar{p}, \bar{z})$ *and let* $K = K_E(\bar{z}, F(\bar{p}, \bar{z}))$ *be the corresponding critical cone. Then suppose that the mapping* $\Lambda + N_K$ *is strongly regular at* 0 *for* 0*, this being equivalent to the condition that the linear variational inequality*

$$\Lambda z + N_K(z) \ni r \tag{23}$$

has a unique solution $\bar{s}(r)$ *for each* $r \in \mathbb{R}^{j+d}$*. Then the solution mapping* S *of* (18) *with values* $S(p)$ *for* $p = \bar{p} + \Delta p$ *has a Lipschitz localization* s *at* $\bar{p}$ *for* $\bar{z}$ *which is semidifferentiable and its semiderivative* $Ds(\bar{p})(\Delta p)$ *is the solution of the following variational inequality:*

$$\Lambda z + N_K(z) + \nabla_p F(\bar{p}, \bar{z})\Delta p \ni 0. \tag{24}$$

Furthermore, in terms of the critical subspaces $K^+ = K^+(\bar{p}, \bar{z})$ *and* $K^- = K^-(\bar{p}, \bar{z})$*, defined in* (21) *and* (22)*, a sufficient condition for strong regularity of* $\Lambda + N_K$ *or, equivalently, single-valuedness of* $(\Lambda + N_K)^{-1}$*, is as follows:*

$$w \in K^+,\ \Lambda w \perp K^-,\ \langle w, \Lambda w \rangle \leq 0 \implies w = 0. \tag{25}$$

Thus, in order to determine the change of the solution for a given variation Δp of the parameter, one could compute the corresponding semiderivative, which in turn reduced to finding the critical cone K and solving a linear VI.

Let Δp be a change of the parameter (the initial state). From Theorem 1 it follows that the semiderivative

$$Ds(\bar{p})(\Delta p) = \begin{bmatrix} \Delta v^T & \Delta q^T \end{bmatrix}^T \tag{26}$$

is a solution of the linear VI

$$\begin{bmatrix} R & G^T \\ G & 0 \end{bmatrix} \begin{bmatrix} \Delta v \\ \Delta q \end{bmatrix} + N_W(\Delta v, \Delta q) \ni \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \tag{27}$$

where $R = \nabla_v^2 \mathcal{L}(\bar{z}, \bar{p})$, $G = \nabla_v g(\bar{v}, \bar{p})$, $W = K_V \times \mathbb{R}^n$, where the critical cone K_V is

$$K_V = \{v \in T_V(\bar{v}) \mid \nabla_v \mathcal{L}(\bar{p}, \bar{z})^T v = 0\}, \tag{28}$$

and

$$r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} -\nabla_{vp} \mathcal{L}(\bar{p}, \bar{z}) \\ -\nabla_p g(\bar{p}, \bar{v}) \end{bmatrix} \Delta p. \tag{29}$$

In order to apply Theorem 1, we need to adapt the sufficient condition for strong regularity (25) to our case. Denote

$$T = \begin{bmatrix} R & G^T \\ G & 0 \end{bmatrix}$$

and let $W^+ = W - W$ and $W^- = W \cap [-W]$ be the critical subspaces associated with the cone W. Then condition (25) becomes

$$w \in W^+, \quad Tw \perp W^-, \quad \langle w, Tw \rangle \leq 0 \quad \Longrightarrow \quad w = 0. \tag{30}$$

One should note that condition (30) is a special case of the more elaborate *critical face condition* obtained in [4], which characterizes the strong regularity of solution mappings associated with variational inequalities over polyhedral convex sets. Obtaining a sharp, numerically tractable, form of the critical face condition for state and control constrained discrete time optimal control problems is beyond the scope of the present paper and is left for future research. However, it is possible to derive a verifiable sufficient condition, observe that (30) holds provided that the matrix T is positive definite on W^+. This in turn holds if the matrix R is positive definite on the null-space of G; that is

$$v^T R v > 0 \;\text{ for all } v \neq 0 \text{ such that } \; Gv = 0. \tag{31}$$

This condition is a standard second order condition in optimization [3] and is always satisfied, for example, when the cost function is strongly convex and the system is linear.

The following statement summarizes the procedure for computing the semiderivatives of the solution mapping.

Theorem 2. *Let $\bar{z} = (\bar{v}, \bar{q})$ satisfy $\bar{z} \in S(\bar{p})$, let Δp be a variation of the parameter and assume the condition* (31) *holds at $(\bar{p}, \bar{z})$. Then the solution mapping $p \mapsto S(p)$ has a Lipschitz localization $s(p)$ at $\bar{p}$ for $\bar{z} = (\bar{v}, \bar{q})$ which is semidifferentiable at $\bar{p}$ and the corresponding semiderivative $Ds(\bar{p})(\Delta p) = (\Delta z, \Delta q)$, or, equivalently, the directional derivative $s'(\bar{p}; \Delta p)$, is the unique solution of the linear VI* (27)*.*

In Theorem 2 we assume that (30) is satisfied at the reference solution. This condition can be enforced by appropriately regularizing the cost function [22]. Moreover, it's possible to monitor if (31), which implies (30), holds by checking if

$$Z^T R Z \succ 0, \tag{32}$$

where the columns of Z form a basis for the nullspace of G and $(\cdot) \succ 0$ denotes positive definiteness. This is straightforward to check numerically by, e.g., forming Z using a QR decomposition of G^T and attempting to compute a Cholesky factorization of $Z^T R Z$, see [3, Section 16.1] for more details.

4 A predictor-corrector algorithm

Theorem 2 shows that the semiderivatives of the solution mapping can be computed by solving a linear VI. Note that (27) are the first-order necessary conditions (i.e., the optimality system) for the following quadratic program (QP),

$$\min_{\Delta v} \quad \frac{1}{2}\Delta v^T R \Delta v - r_1^T \Delta v, \tag{33a}$$

$$\text{subject to} \quad G\Delta v = r_2, \tag{33b}$$

$$\Delta v \in K_V, \tag{33c}$$

where R, G, K_V, r_1 and r_2 are defined in (27), (28) and (29). The critical cone constraint (33c) can be simplified by recalling that K_V can be expressed in terms of the index set of the active constraints, see Section 1.1. In addition, recall that we can express the set V as $V = \{v \mid Mv \leq h\}$ (see (11)) where

$$M = \begin{bmatrix} M_0 & & & \\ & E_0 & & \\ & & \ddots & \\ & & & E_N \end{bmatrix} \quad \text{and} \quad h = \begin{bmatrix} d_0 \\ c_1 \\ \vdots \\ c_N \end{bmatrix}. \tag{34}$$

The QP can thus be rewritten as

$$\min_{\Delta v} \quad \frac{1}{2}\Delta v^T R \Delta v + (P\Delta p)^T \Delta v, \tag{35a}$$

$$\text{subject to} \quad G\Delta v + Q\Delta p = 0, \tag{35b}$$

$$M_i \Delta v \leq 0, \quad i \in \bar{\mathcal{A}}, \tag{35c}$$

$$M_i \Delta v = 0, \quad i \in \mathcal{A}(\bar{v}) \setminus \bar{\mathcal{A}}, \tag{35d}$$

$$\nabla_v \mathcal{L}(\bar{p}, \bar{z})^T \Delta v = 0, \tag{35e}$$

where $\mathcal{A}(\bar{v})$ and $\bar{\mathcal{A}} = \bar{\mathcal{A}}(\bar{v}, F(\bar{p}, \bar{z}))$ are defined in Section 1.1, $P = \nabla_{pv} L(\bar{p}, \bar{z})$, $Q = \nabla_p g(\bar{p}, \bar{v})$ and $\bar{z} = (\bar{v}, \bar{q}) \in S(\bar{p})$. The solution of (27) can be obtained by solving (35) and extracting the primal solution and the multipliers associated with (35b).

This QP may be difficult to solve in practice since it is not necessarily convex. This can be addressed by modifying the Hessian as follows:

$$R \leftarrow R + \rho G^T G, \tag{36}$$

where $\rho \in (0, \infty)$. If (31) is satisfied then the modified R will be positive semidefinite provided ρ is chosen large enough [1, Proposition 4.8]. This modification of the Hessian changes the Lagrange multipliers associated with the co-state, but the original multipliers can be recovered as $\Delta q \leftarrow \Delta q^* - \rho Q \Delta p$ where Δq^* are obtained by solving (35) with $R \leftarrow R + \rho G^T G$, and extracting the multipliers associated with (35b), see e.g., [1, Section 4.2].

For the corrector we use the Josephy-Newton (JN) method. Given a fixed parameter value p, a VI of the form (18), and an initial guess z_0, the JN method constructs an iterative sequence by repeatedly solving the following linearized VI

$$\nabla_z F(p, z_k)(z_{k+1} - z_k) + N_E(z_{k+1}) \ni 0, \tag{37}$$

for z_{k+1}. It is well known that strong regularity of the VI implies local quadratic convergence of the JN method [18]. Denoting $\delta v = v^{k+1} - v^k$, and $\delta q = q^{k+1} - q^k$, (37) becomes

$$\nabla_v^2 \mathcal{L}(p, v_k, q^k)\delta v + \nabla_v \mathcal{L}(p, v_k, q^k) + \nabla_v g(p, v_k)^T \delta q + N_V(v_k + \delta v) \ni 0, \tag{38a}$$

$$\nabla_v g(p, v_k)\delta v + g(p, v_k) = 0. \tag{38b}$$

Note that this is the optimality system of the following QP:

$$\min_{\delta v} \quad \frac{1}{2}\delta v^T \nabla_v^2 \mathcal{L}(p, v_k, q_k)\delta v + \nabla_v \mathcal{L}(p, v_k, q_k)^T \delta v, \tag{39a}$$

$$\text{subject to} \quad \nabla_v g(p, v_k)\delta v + g(p, v_k) = 0, \tag{39b}$$

$$M\delta v \leq h - M v_k. \tag{39c}$$

When the Josephy-Newton iteration (37) is determined by solving the corresponding quadratic problem (39), this method is called Sequential Quadratic Programming (SQP). We can now summarize the predictor-corrector algorithm as Algorithm 1. At each sampling instance the system state is measured, and the sensitivity based predictor is used to estimate the updated iterate corresponding to the parameter change based on the solution of the OCP at the previous sampling instance. This estimate is then passed to an JN based corrector loop which stops when the norm of the residual

$$\pi(p, z) = ||z - \Pi_E[z - F(p, z)]||, \tag{40}$$

is within the specified tolerance. The control input is then extracted from the solution and applied to the system.

Note that both the predictor and corrector steps are realized by solving QPs. The predictor QP usually has significantly fewer constraints than the corrector QP, see Section 1.1, which can lead to reduced computation times. In addition, an initial feasible guess for the predictor QP is available, which can be helpful for both primal active-set and primal-barrier interior point methods.

Algorithm 1 Sensitivity based Predictor-Corrector MPC

Input: ε, p_k, p_{k-1}, v_{k-1}, q_{k-1}, κ
Output: $z_k = (v_k, q_k)$
Measure p_k
$\Delta p_k = p_k - p_{k-1}$
$z \leftarrow (q_{k-1}, v_{k-1})$
Solve (35) with $(\bar{v}, \bar{q}, \bar{p}) = (v_{k-1}, q_{k-1}, p_{k-1})$ to obtain $\Delta z = (\Delta v, \Delta q)$
$z \leftarrow z + \Delta z$
while $||\pi(p_k, v, q)|| > \varepsilon$ **do**
    Solve (38) with $(z_k, p_k) = (z, p_k)$ to obtain $(\delta v, \delta q)$
    $z \leftarrow z + \delta z$
end while
$z_k \leftarrow z$
Extract u_0 from z_k and apply it to the plant

5 Numerical simulations

To illustrate the theoretical results, we consider the following model of rotational and translational dynamics of an Unmanned Aerial Vehicle (UAV)

$$\begin{cases} \dot{p} = v, & m\dot{v} = T\ R(\theta)e_3 - mge_3, \\ \dot{\theta} = R(\theta)\omega, & J\dot{\omega} = -\omega \times (J\omega) + \tau, \end{cases} \tag{41}$$

where the state vector is given by the position $p \in \mathbb{R}^3$, velocity $v \in \mathbb{R}^3$, attitude $\theta \in (-\pi, \pi]^3$, and angular velocity $\omega \in \mathbb{R}^3$ of the UAV, whereas the input vector is composed of the total thrust $T \in \mathbb{R}^+$ and torques $\tau \in \mathbb{R}^3$ generated by the propellers. The mass and inertia matrix of the UAV are $m = 2\ kg$ and $J = \operatorname{diag}([0.82\ 0.82\ 1.62])10^{-2}\ kgm^2$, respectively, $e_3 = [0\ 0\ 1]^T$, and

$$R(\theta) = \begin{bmatrix} 1 & \sin(\theta_3)\tan(\theta_2) & \cos(\theta_3)\sin(\theta_2) \\ 0 & \cos(\theta_3) & -\sin(\theta_3) \\ 0 & \sin(\theta_3)/\cos(\theta_2) & \sin(\theta_3)/\cos(\theta_2) \end{bmatrix} \tag{42}$$

is the attitude kinematic matrix. The system dynamics are discretized using the forward Euler approximation with $T_s = 0.075\ s$. The MPC is designed using the quadratic cost function,

$$J(x, u) = \frac{1}{2} x_N^T P x_N + \sum_{i=0}^{N-1} \frac{1}{2} x_i^T Q x_i + \frac{1}{2} u_i^T R u_i,$$

where $Q = \operatorname{diag}([5\ 5\ 5\ 10\ 10\ 10\ 0.1\ 0.1\ 0.1\ 1\ 1\ 1])$, $R = \operatorname{diag}([0.1\ 0.01\ 0.01\ 0.01])$, and P was computed from these values using the LQR terminal cost for the system linearized about the origin. The system is subject to the input constraint $T \in [18, 22]$ and the state constraints $|v_j| \leq 0.4$, $|\theta_j| \leq 0.02$, $|\omega_j| \leq 0.2$, where $j \in \{1, 2, 3\}$ refers to each component of the vector.

Figure 1 displays the closed-loop response obtained using the MPC, which successfully enforces all the constraints. Figure 2 compares the computational cost obtained[1] with and without the prediction step presented in this paper. The natural residual was subject to the tolerance value $\pi(p, z) \leq 10^{-5}$.

The comparison in Figure 2 can be divided into three segments. In the first segment $[0, 11.5]$, the predictor+corrector strategy systematically achieves a lower computational time with respect to the case without the predictor. This is attributed to the fact that i) the predictor is successful at reducing the number of corrector iterations by one, and ii) the computational cost associated to the predictor QP (35) is an order of magnitude smaller than the computational cost of the corrector QP (39). In the second segment $[11.5, 16.5]$, the predictor step is unable to reduce the total number of corrector iterations. In this case, the predictor+corrector is slightly slower with respect to the case without the predictor, although the relatively small computational cost of the predictor step means that the overall loss is negligible with respect to what is gained when the predictor step is successful. In the third segment, $[16.5, 40]$, the prediction step is once again able to reduce the number of corrector iterations by one. In this particular example, this leads to an interesting behavior where the solution estimate is updated in "open loop" by performing a sequence of prediction steps. Once the accumulated error causes the residual to go above the desired threshold, the corrector activates and performs a single iteration to reset the solution.

[1]DELL Latitude 7390 2-in-1, Intel Core i7-8650U, 2.11 GHz, 16 GB

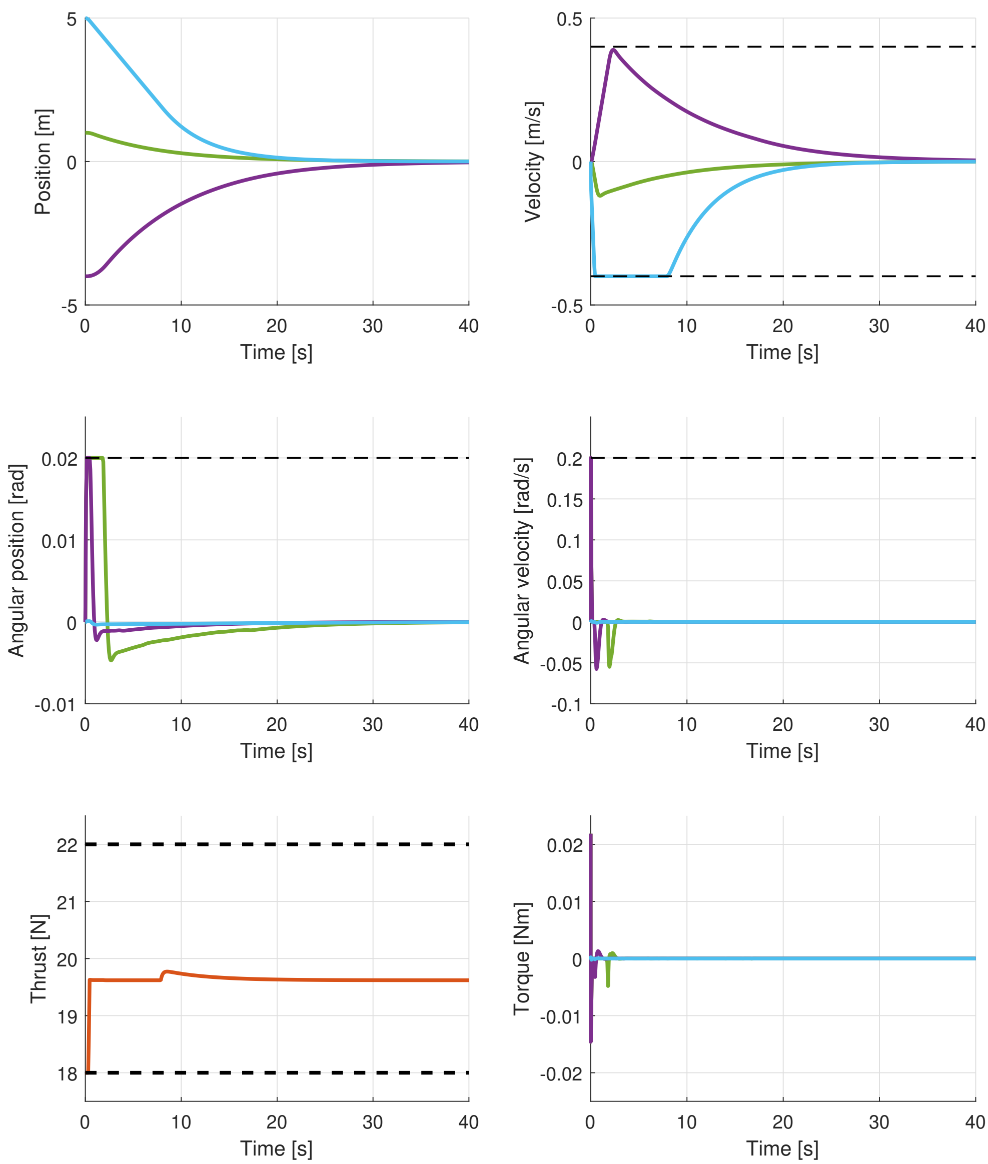

Figure 1: The closed-loop response of the UAV. All state and input constraints are respected.

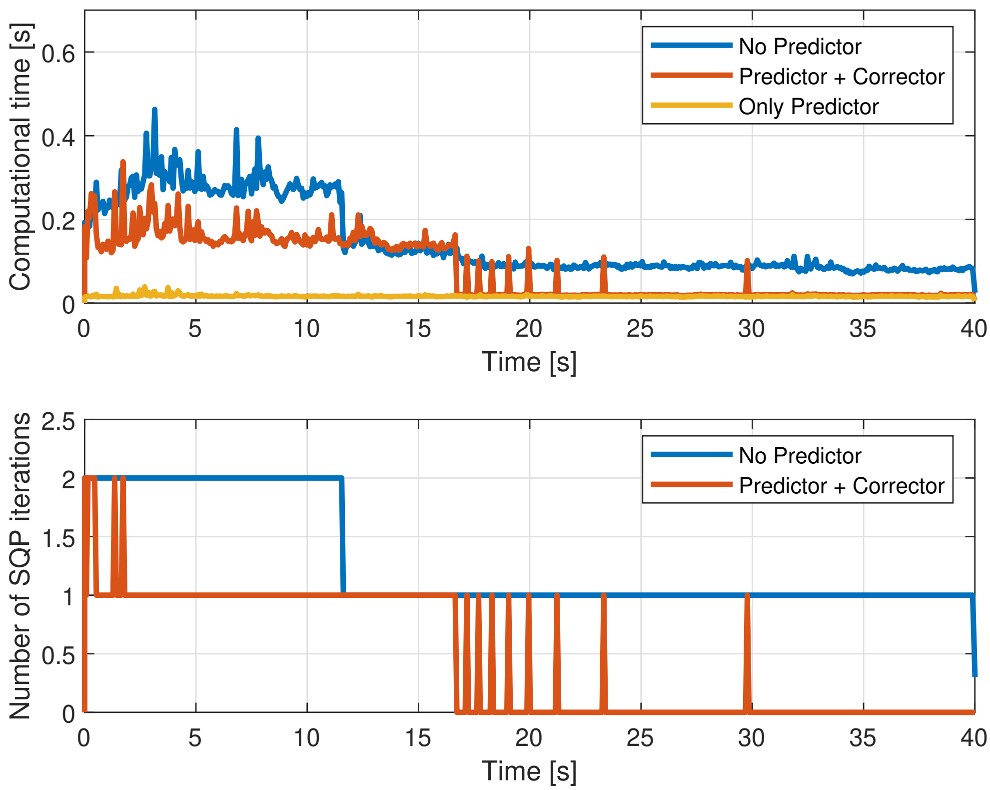

Figure 2: A comparison between nonlinear MPC implemented using SQP, i.e., using corrector iterations only, and using the proposed predictor-corrector scheme.

Overall, these numerical simulations show that the proposed predictor has the potential to significantly reduce the computational cost at most time instances and without incurring in significant drawbacks in the other instances. This in turn can translate to reduced power consumption and e.g., extended flight range for the UAV.

6 Conclusions

In this paper we propose a new sensitivity-based warmstarting strategy for model predictive control of systems with nonlinear dynamics and linear state and control inequality constraints. The strategy involves a predictor which utilizes the semiderivative of the solution of the optimality system. The method exploits the polyhedrality of the constraint set and requires fewer assumptions than comparable methods in the literature. Specifically, it doesn't require a difficult to verify constraint qualification. Numerical simulations demonstrate the potential of the strategy. Future work includes moving from polyhedral to more general constraints, relaxing the conditions for strong regularity, and developing tailored quadratic programming solvers for the predictor.

References

[1] A. F. Izmailov, M.V. Solodov. *Newton-type methods for optimization and variational problems*. Springer, 2014.

[2] A.L. Dontchev, M.I. Krastanov, R.T. Rockafellar and V.M. Veliov. An Euler–Newton continuation method for tracking solution trajectories of parametric variational inequalities, *SIAM Journal on Control and Optimization*, **51**(3):1823–1840, 2013.

[3] J. Nocedal and S. Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[4] A.L. Dontchev and R.T. Rockafellar. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM Journal on Optimization*, **6**(4):1087–1105, 1996.

[5] T. Ohtsuka. A continuation/GMRES method for fast computation of nonlinear receding horizon control. *Automatica*, **40**(4):563–574, 2004.

[6] Q.T. Dihn, C. Savorgnan and M: Diehl. Adjoint-based predictor-corrector sequential convex programming for parametric nonlinear optimization, *SIAM Journal on Optimization*, **22**(4):1258–1284, 2012.

[7] D. Liao-McPherson, M. Nicotra and I. Kolmanovsky. A semismooth predictor corrector method for real-time constrained parametric optimization with applications in

model predictive control, *Proceedings of the 2018 Conference on Decision and Control (CDC)*, 2018.

[8] F.H. Clarke. *Optimization and nonsmooth analysis.* SIAM, 1990.

[9] V.M. Zavala and M. Anitescu. Real-time nonlinear optimization as a generalized equation. *SIAM Journal on Control and Optimization*, **48**(8):5444–5467, 2010.

[10] E.L. Allgower and K. Georg. *Numerical continuation methods: an introduction.* Springer Science & Business Media, Vol. 13, 2012.

[11] R. Ghaemi, J. Sun, and I.V. Kolmanovsky. An integrated perturbation analysis and sequential quadratic programming approach for model predictive control, *Automatica*, **45**(10):2412–2418, 2009.

[12] J.-H. Hours and C.N. Jones. A parametric nonconvex decomposition algorithm for real-time and distributed NMPC, *IEEE Transactions on Automatic Control*, **61**(2):287–302, 2016.

[13] I. Das and F.A. Potra. Subsequent convergence of iterative methods with applications to real-time model-predictive control, *Journal of optimization theory and applications*, **119**(1):37–47, 2003.

[14] A.L. Dontchev and T.R. Rockafellar, *Implicit functions and solution mappings*, Second edition. Springer; New York, Heidelberg, Dordrecht, London, 2014,.

[15] A.D. Ioffe and V.M. Tihomirov. *Theory of extremal problems.* Studies in Mathematics and its Applications, Vol. 6, North-Holland Publishing Co.; Amsterdam; New York, 1979.

[16] S.M. Robinson. Strongly regular generalized equations, *Mathematics of Operations Research*, **5**(1):43–62, 1980.

[17] M. Diehl, H.G. Bock, and J.P. Schlöder. A real-time iteration scheme for nonlinear optimization in optimal feedback control, *SIAM Journal on control and optimization*, **43**(5):1714–1736, 2005.

[18] N.H. Josephy. *Newton's method for generalized equations.* University of Wisconsin-Madison Mathematics Research Center, 1979.

[19] G.C. Goodwin, M.M. Seron, and J.A. De Doná. *Constrained control and estimation: an optimisation approach.* Springer Science & Business Media, 2006.

[20] J.B. Rawlings, D.Q. Mayne, and M. Diehl. *Model predictive control: theory, computation and design.* Nob Hill Publ., 2018.

[21] A.V. Fiacco. Sensitivity analysis for nonlinear programming using penalty methods. *Mathematical programming*, **10**(1):287–311, 1976.

[22] L.T. Bigler and D.M. Thierry. Large-scale optimization formulations and strategies for nonlinear model predictive control. In *6th IFAC Conference on Nonlinear Model Predictive Control*, 2018.

[23] L. Grüne and J Pannek. *Nonlinear Model Predictive Control*. Springer, 2017.

[24] V. Kungurtsev and M. Diehl. SQP Methods for Parametric Nonlinear Optimization. *Optimization Online*, pp. 1–35, 2014.

[25] S.M. Robinson. Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity, *Nonlinear Analysis and Optimization*, pp. 45–66, 1987.

[26] D. Ralph and S. Dempe. Directional derivatives of the solution of a parametric nonlinear program. *Mathematical programming*, **70**(1-3):159–172, 1995.

[27] V.M. Zavala and L.T. Biegler. The advanced-step NMPC controller: Optimality, stability and robustness, *Automatica*, **45**(1):86–93, 2009.

[28] R.T. Rockafellar and R.J-B. Wets. *Variational analysis*. Springer Science & Business Media, Vol. 317, 2009.

[29] J. Jäschke, X. Yang, and L.T. Biegler. Fast economic model predictive control based on NLP-sensitivities. *Journal of Process Control*, **24**(8):1260–1272, 2014.

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

[31] J.F. Bonnans. Local analysis of Newton-type methods for variational inequalities and nonlinear programming. *Applied Mathematics and Optimization*, **29**(2):161–186, 1994.