



## Report

# On the Approximation of Constrained Linear Quadratic Regulator Problems and their Application to Model Predictive Control - Supplementary Notes

**Author(s):**

Muehlebach, Michael; D'Andrea, Raffaello

**Publication Date:**

2018

**Permanent Link:**

<https://doi.org/10.3929/ethz-b-000246553> →

**Rights / License:**

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

# On the Approximation of Constrained Linear Quadratic Regulator Problems and their Application to Model Predictive Control - Supplementary Notes

Michael Muehlebach and Raffaello D'Andrea\*

## Abstract

By parametrizing input and state trajectories with basis functions different approximations to the constrained linear quadratic regulator problem are obtained. These notes present and discuss technical results that are intended to supplement a corresponding journal article. The results can be applied in a model predictive control context.

## 1 Outline

Sec. 2 summarizes the notation. The resulting approximations are given by convex finite-dimensional optimization problems that have a quadratic cost, linear equality constraints, and linear semi-infinite inequality constraints. Sec. 3 suggests several strategies for dealing with these semi-infinite constraints. One of these strategies is discussed in more details in Sec. 4, as it is found to be computationally efficient in a model predictive control context. The remaining sections discuss several technical results. Sec. 5 and Sec. 6 are related to an upper, respectively lower bound approximation of the cost of the underlying constrained linear quadratic regulator problem. Sec. 7 presents a result that is related to the convergence of the different approximations. Sec. 8 shows recursive feasibility and closed-loop stability of the resulting model predictive control algorithm. Sec. 9 and Sec. 10 highlight important properties of the basis functions parametrizing input and state trajectories.

## 2 Notation

We are concerned with the approximation of the following optimal control problem

$$\begin{aligned} J_\infty &:= \min \frac{1}{2} (\|x\|^2 + \|u\|^2) + \psi(x_T) \\ \text{s.t. } \dot{x} &= Ax + Bu, x(0) = x_0, x(T) = x_T, \\ C_x x + C_u u &\leq b, x_T \in \mathcal{X}, \\ x &\in L_n^2, u \in L_m^2, \end{aligned} \tag{1}$$

where the space of square integrable functions mapping from the interval  $I := (0, T)$  to  $\mathbb{R}^q$  is denoted by  $L_q^2$ , where  $q$  is a positive integer; and the  $L_q^2$ -norm is defined as the map  $L_q^2 \rightarrow \mathbb{R}$ ,

$$x \rightarrow \|x\|, \quad \|x\|^2 := \int_I x^\top x \, dt, \tag{2}$$

---

\*Michael Muehlebach and Raffaello D'Andrea are with the Institute for Dynamic Systems and Control, ETH Zurich. The contact author is Michael Muehlebach, [michaemu@ethz.ch](mailto:michaemu@ethz.ch). This work was supported by ETH-Grant ETH-48 15-1.

with  $dt$  the Lebesgue measure. The function  $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$  is assumed to be positive definite and strongly convex,  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ ,  $C_x \in \mathbb{R}^{n_c \times n}$ ,  $C_u \in \mathbb{R}^{n_c \times m}$ , and  $b \in \mathbb{R}^{n_c}$  are constant, and the set  $\mathcal{X}$  is closed and convex. The dynamics as well as the stage constraints are assumed to be fulfilled almost everywhere. Thus we simply write

$$f = g, \quad f \leq g, \quad (3)$$

when we mean  $f(t) = g(t)$ , respectively  $f(t) \leq g(t)$  for all  $t \in I$  almost everywhere, with  $f, g \in L^2_q$ . The weak derivative of  $x$  is denoted by  $\dot{x}$ .<sup>1</sup> To simplify notation we abbreviate the domain of the objective function by

$$X := L^2_n \times L^2_m \times \mathbb{R}^n. \quad (4)$$

We assume throughout the article that the constraints in (1) are nonempty, i.e. there exists trajectories  $x$  and  $u$ , fulfilling the dynamics, the initial condition, the constraints, and thus achieve a finite cost.

The main motivation for studying problem (1) comes from the fact that (1) often serves as a starting point for model predictive control (MPC).

The problem (1) can be written in compact form as

$$\begin{aligned} J_\infty = \min & \|x\|^2 + \|u\|^2 + \psi(x_T) \\ \text{s.t. } & (x, u, x_T) \in \mathcal{C} \cap \mathcal{D}, \end{aligned} \quad (5)$$

where the set  $\mathcal{C}$  contains all trajectories  $x \in L^2_n, u \in L^2_m$  that fulfill the constraints and  $\mathcal{D}$  the trajectories that fulfill the dynamics and the initial condition, that is,

$$\mathcal{C} := \{(x, u, x_T) \in X \mid C_x x + C_u u \leq b, x_T \in \mathcal{X}\}, \quad (6)$$

$$\mathcal{D} := \{(x, u, x_T) \in X \mid \dot{x} = Ax + Bu, x(0) = x_0, x(T) = x_T\}. \quad (7)$$

It was shown that (5) can be approximated by two auxiliary optimization problems, where input and state trajectories are parameterized with basis functions,

$$\tilde{x}(t) = (I_n \otimes \tau^s(t))^\top \eta_x, \quad \tilde{u}(t) = (I_m \otimes \tau^s(t))^\top \eta_u.$$

The parameter vectors are denoted by  $\eta_x \in \mathbb{R}^{ns}$  and  $\eta_u \in \mathbb{R}^{ms}$ , whereas  $\tau^s(t) := (\tau_1(t), \tau_2(t), \dots, \tau_s(t)) \in \mathbb{R}^s$  contains the first  $s$  basis functions,  $\otimes$  denotes the Kronecker product, and  $I_q \in \mathbb{R}^{q \times q}$  refers to the identity matrix for any integer  $q > 0$ . The superscript  $s$  refers to the number of basis functions used for the approximation. For ease of notation the superscript  $s$  will be dropped whenever the number of basis functions is clear from context. The basis functions are required to satisfy the following assumptions:

A1) The basis functions  $\tau_i \in L^2_1, i = 1, 2, \dots, s$  are linearly independent and orthonormal with respect to the standard  $L^2_1$ -scalar product.

A2) The basis functions fulfill  $\dot{\tau}^s(t) = M_s \tau^s(t)$  for all  $t \in I$ , for some  $M_s \in \mathbb{R}^{s \times s}$ .

The finite-dimensional subspace spanned by the first  $s$  basis functions will be denoted by  $X^s$ ,

$$X^s := \{(x, u, x_T) \in X \mid \eta_x \in \mathbb{R}^{ns}, \eta_u \in \mathbb{R}^{ms}, x = (I_n \otimes \tau^s)^\top \eta_x, u = (I_m \otimes \tau^s)^\top \eta_u\}. \quad (8)$$

We can think of an element in  $X^s$  not only as an element in  $X$  (i.e. a tuple of a finite-dimensional vector and two square integrable functions), but also as a finite-dimensional vector given by the corresponding parameter vectors  $\eta_x$  and  $\eta_u$ . To make this distinction explicit, we introduce the map  $\pi^{qs} : L^2_q \rightarrow \mathbb{R}^{qs}$ , defined as

$$x \rightarrow \int_I (I_q \otimes \tau^s) x dt, \quad (9)$$

<sup>1</sup>The equations of motion imply that  $\dot{x} \in L^2_n$ , which can be used to conclude that  $x$  has a unique absolutely continuous representative defined on the closure of  $I$  (a classical solution of the equations of motion). With  $x(t)$  we refer to the value this unique absolutely continuous representative takes at time  $t \in [0, T]$ .

which maps an arbitrary element  $x \in L_n^2$  to its first  $s$  basis function coefficients. Similarly, we define  $\pi^s : X \rightarrow \mathbb{R}^{ns} \times \mathbb{R}^{ms} \times \mathbb{R}^n$  as

$$(x, u, x_T) \rightarrow (\pi^{ns}(x), \pi^{ms}(u), x_T). \quad (10)$$

As a consequence, we write  $\pi^s(x)$  for describing the finite dimensional representation of  $x \in X^s$ , that is, its representation in terms of the parameter vectors  $\eta_x$  and  $\eta_u$ . The adjoint map  $(\pi^{qs})^* : \mathbb{R}^{qs} \rightarrow L_q^2$  is given by

$$\eta \rightarrow (I_q \otimes \tau)^\top \eta, \quad (11)$$

and is used to obtain the trajectory corresponding to the vector  $\eta \in \mathbb{R}^{qs}$ , containing the first  $s$  basis function coefficients. Similarly, we define  $(\pi^s)^* : \mathbb{R}^{ns} \times \mathbb{R}^{ms} \times \mathbb{R}^n \rightarrow X$  as

$$(\eta_x, \eta_u, x_T) \rightarrow ((\pi^{ns})^*(\eta_x), (\pi^{ms})^*(\eta_u), x_T). \quad (12)$$

The composition  $(\pi^s)^* \pi^s : X \rightarrow X$  yields the projection of an element  $x \in X$  onto the subspace  $X^s \subset X$ .

The sets  $\mathcal{C}$  and  $\mathcal{D}$  are approximated in two different ways,

$$\mathcal{C}_U^s := \mathcal{C} \cap X^s, \quad (13)$$

$$\begin{aligned} \mathcal{C}_L^s := \{ (x, u, x_T) \in X \mid & \int_I \delta \tilde{p}^\top (-C_x x - C_u u + b) dt \geq 0 \\ & \forall \delta \tilde{p} = (I_{n_c} \otimes \tau)^\top \delta \eta_p : \delta \tilde{p} \geq 0, \delta \eta_p \in \mathbb{R}^{n_c s}; x_T \in \mathcal{X} \}, \end{aligned} \quad (14)$$

$$\mathcal{D}_U^s := \mathcal{D} \cap X^s, \quad (15)$$

$$\begin{aligned} \mathcal{D}_L^s := \{ (x, u, x_T) \in X \mid & \delta \tilde{p} = (I_n \otimes \tau)^\top \delta \eta_p, \\ & - \int_I \delta \dot{\tilde{p}}^\top x dt - \int_I \delta \tilde{p}^\top (Ax + Bu) dt - \delta \tilde{p}(0)^\top x_0 + \delta \tilde{p}(T)^\top x_T = 0, \forall \delta \eta_p \in \mathbb{R}^{ns} \}, \end{aligned} \quad (16)$$

resulting in the following two approximations to (5)

$$\begin{aligned} J_s := \inf & \|\tilde{x}\|^2 + \|\tilde{u}\|^2 + \psi(x_T) \\ \text{s.t. } & (\tilde{x}, \tilde{u}, x_T) \in \mathcal{C}_U^s \cap \mathcal{D}_U^s, \end{aligned} \quad (17)$$

and

$$\begin{aligned} \tilde{J}_s := \min & \|\tilde{x}\|^2 + \|\tilde{u}\|^2 + \psi(x_T) \\ \text{s.t. } & (\tilde{x}, \tilde{u}, x_T) \in \mathcal{C}_L^s \cap \mathcal{D}_L^s. \end{aligned} \quad (18)$$

By definition of the constraints  $\mathcal{C}_U^s$  and  $\mathcal{D}_U^s$ , the minimizer of (17) (for  $s \geq s_0$ ) is required to be an element of  $X^s$ . Consequently, the problem (17) is equivalent to

$$\begin{aligned} J_s = \inf & |\eta_x|^2 + |\eta_u|^2 + \psi(x_T) \\ \text{s.t. } & (\eta_x, \eta_u, x_T) \in \pi^s(\mathcal{C}_U^s) \cap \pi^s(\mathcal{D}_U^s), \end{aligned} \quad (19)$$

which corresponds to a convex finite-dimensional optimization problem. Likewise, it is found that the convex finite-dimensional problem

$$\begin{aligned} \tilde{J}_s = \min & |\eta_x|^2 + |\eta_u|^2 + \psi(x_T) \\ \text{s.t. } & (\eta_x, \eta_u, x_T) \in \pi^s(\mathcal{C}_L^s) \cap \pi^s(\mathcal{D}_L^s), \end{aligned} \quad (20)$$

is equivalent to (18) in the sense that its (unique) minimizer  $(\eta_x, \eta_u, \bar{x}_T)$  is related to the minimizer  $(x, u, x_T) \in X$  of (18) by  $x = (\pi^{ns})^*(\eta_x)$ ,  $u = (\pi^{ms})^*(\eta_u)$ ,  $\bar{x}_T = x_T$ , and achieves the same cost.

The proposed approximations can be applied in the context of MPC. We will show that by repeatedly solving the infinite-horizon optimal control problem (19) (with  $I = (0, \infty)$ ,  $\psi = 0$ ,  $x_T = \lim_{t \rightarrow \infty} x(t) = 0$ ),

recursive feasibility and closed-loop stability are inherent to the resulting MPC algorithm. For the sake of completeness, (19) is written out as

$$\begin{aligned}
& \min \eta_x^\top (Q \otimes I_s) \eta_x + \eta_u^\top (R \otimes I_s) \eta_u & (21) \\
& \text{s.t. } (I_n \otimes M_s^\top - A \otimes I_s) \eta_x - (B \otimes I_s) \eta_u = 0, \\
& \quad (I_n \otimes \tau(0))^\top \eta_x = x_0, \\
& \quad (C_x \otimes \tau(t))^\top \eta_x + (C_u \otimes \tau(t))^\top \eta_u \leq b, \forall t \in [0, \infty),
\end{aligned}$$

where the input and state costs are weighted with the matrices  $Q > 0$  and  $R > 0$ , which is common in MPC.

In the following an efficient numerical solution algorithm for (21) will be discussed.

### 3 Implementation of the semi-infinite constraint

The optimization (21) is a convex finite-dimensional optimization problem. However, it is not a quadratic program, as it includes the semi-infinite constraint

$$(C_x \otimes \tau(t))^\top \eta_x + (C_u \otimes \tau(t))^\top \eta_u \leq b, \forall t \in [0, \infty). \quad (22)$$

As a result, (21) cannot be solved by a standard quadratic programming solver. Three different approaches to deal with the semi-infinite constraint are immediate:

- 1) global polyhedral approximation
- 2) sum-of-squares approximation
- 3) local polyhedral approximation (active-set approach)

The first is based on a fixed polyhedral approximation, leading to a quadratic program. The second is based on exploiting polynomial basis functions for obtaining a characterization using linear matrix inequalities, whereas the third is based on an iterative constraint sampling scheme, resulting in a local polyhedral approximation. In the following subsections we will focus on approach 1 and 2. We will describe an efficient solution algorithm based on approach 3 in detail in Sec. 4.

It turns out that it is enough to check the constraint (22) over a compact time interval, instead of the unbounded interval  $t \in [0, \infty)$ . This is because the basis functions are assumed to be linearly independent and exponentially decaying according to Assumption A1 and A2. A formal proof of this claim can be found in Sec. 9. The compact time interval for which the constraint (22) has to be checked is denoted by  $[0, T_c]$  and depends on the choice of basis functions, on the order  $s$ , and in some cases also on the bound  $b$  (see Sec. 9).

#### 3.1 Global polyhedral approximation

In order to construct a global polyhedral approximation of the set  $C_U^s$  we assume that an upper bound on the achievable cost  $J_s$ , denoted by  $\bar{J}_s$ , is available. In order to simplify the discussion, we assume further that for now  $C_x$  and  $C_u$  are row vectors, and that  $b$  is a scalar. The proposed scheme can be readily extended to the case where  $C_x$  and  $C_u$  are matrices, and  $b$  is a vector. The approximation is based on constraint sampling, where we tighten the constraint slightly to

$$C_x \tilde{x}(t_i) + C_u \tilde{u}(t_i) \leq (1 - \epsilon)b, \quad (23)$$

with  $\epsilon > 0$ , and where  $t_i$ , denotes the constraint sampling instances,  $i = 1, 2, \dots, n_i$ , which are to be determined. The algorithm is based on the following two steps:

1) Compute

$$\begin{aligned}
h(t) &:= \max C_x \tilde{x}(t) + C_u \tilde{u}(t) - b \\
\text{s.t. } & C_x \tilde{x}(t_i) + C_u \tilde{u}(t_i) \leq (1 - \epsilon)b, \quad i = 1, \dots, n_i, \\
& (\tilde{x}, \tilde{u}, \lim_{t_e \rightarrow \infty} \tilde{x}(t_e)) \in \mathcal{D}_U^s, \\
& \|\tilde{x}\|^2 + \|\tilde{u}\|^2 \leq \bar{J}_s,
\end{aligned}$$

for all times  $t \in I_s$ , where  $I_s$  contains a finite number of sampling instances (to be made precise below).

2) Find the local peaks of  $h(t)$ , denoted by  $t_i^*$ . Add each  $t_i^*$  to the constraint sampling points if  $h(t_i^*) > -b\epsilon/2$ . Repeat the procedure until  $h(t) \leq -b\epsilon/2$  for all  $t \in I_s$ .

Note that the function  $h(t)$  is again only evaluated at the discrete time points  $t \in I_s$ . The index set  $I_s$  has to be chosen such that  $h(t) \leq -b\epsilon/2$  for all  $t \in I_s$  implies that  $h(t) \leq 0$  for all  $t \in [0, \infty)$ . As remarked earlier, due to the fact that the basis functions are exponentially decaying and linearly independent it is enough to check  $h(t) \leq 0$  for all  $t \in [0, T_c]$ , for a fixed time  $T_c$ , as  $h(t) \leq 0$  for all  $t \in (T_c, \infty)$  will be fulfilled automatically. Moreover, a Lipschitz constant of

$$C_x \tilde{x}(t) + C_u \tilde{u}(t) - b \quad (24)$$

can be found by using an upper bound on its time-derivative, that is, for example,

$$|\tau(t)^\top M^\top (C_x \eta_x + C_u \eta_u)| \leq |\tau(t)| |M^\top (C_x \eta_x + C_u \eta_u)|, \quad (25)$$

where the first term can be bounded for all  $t \in [0, \infty)$  due to the fact that the basis functions are exponentially decaying and the second term can be bounded using the fact that the cost  $J_s$  is below  $\bar{J}_s$ . We therefore choose the index set  $I_s$  as

$$I_s = \{t_k < T_c \mid t_k = k \frac{2L}{b\epsilon}, k = 0, 1, 2, \dots\}, \quad (26)$$

where  $L$  denotes a Lipschitz constant of (24).

It is important to note that the optimization in step 1 imposes the dynamics and the upper bound  $\bar{J}_s$  on the cost. Both constraints tend to reduce the number of constraint sampling points  $t_i$  greatly. The initial condition  $x_0$  enters the optimization as an optimization variable. The optimization in step 1 represents a quadratically constrained linear program for each time instant  $t \in I_s$ , and as such, it can be solved using standard software packages. The whole procedure for determining the constraint sampling points is done offline. Once these time instances are found, the optimization problem that is solved online reduces to a quadratic program. The number of constraint sampling points  $t_i$  is upper bounded by the cardinality of the index set  $I_s$ , and thus guaranteed to be finite. Due to the fact that the above procedure is greedy, it will not necessarily lead to the smallest number of constraint sampling points.

### 3.2 Sum-of-squares approximation

In case exponentially decaying polynomials are used as basis functions, sum-of-squares techniques can be applied. In particular, it is shown in [3] that the set

$$\{\eta \in \mathbb{R}^s \mid \eta^\top (1, t, \dots, t^{s-1}) \geq 0, \forall t \in [0, \infty)\} \quad (27)$$

can be expressed using matrix inequalities that are linear in the coefficients  $\eta$ . In the case of exponentially decaying polynomials it is therefore enough to approximate the exponential decay by a polynomial upper bound (for example by appropriately truncating a Taylor series expansion at 0), in order to approximate the constraint (22) in a slightly conservative manner. As a result, by applying the results from [3], the optimization problem (21) is approximated by a semidefinite program that can be solved using standard optimization routines.

## 4 A dedicated active-set approach

In the following section we present an efficient optimization routine for solving (19) (and likewise (21)). The method is an extension of traditional active set methods and generalizes to optimization problems with a linear quadratic cost function, linear equality constraints, and linear semi-infinite inequality constraints, i.e.

$$\hat{J}(I_c) := \min z^\top H z \quad (28)$$

$$\text{s.t. } A_{\text{eq}} z = b_{\text{eq}}, \quad (29)$$

$$l_b \leq (I_{n_c} \otimes \tau(t))^\top C_z z \leq l_u, \quad \forall t \in I_c, \quad (30)$$

where  $I_c$  is any subset of the non-negative real line. In case of (21), the interval  $I_c$  is taken to be  $[0, \infty)$ . Note that an optimization problem, whose objective function has a linear part, can be brought to the form (28) by completing the squares. It is assumed that the optimization problem (28) is feasible, that  $l_b < 0$  and  $l_u > 0$ , and that the Hessian  $H$  is positive definite, which guarantees existence and uniqueness of the corresponding minimizer<sup>2</sup>.

The method is based on the observation that if the set  $I_c$  consists merely of a collection of time instants (constraint sampling instances)  $t_i$ , (28) reduces to a quadratic program that can be solved efficiently. Moreover, due to the fact that the basis functions fulfill Assumptions A1 and A2, a trajectory parametrized with the basis functions has a finite number of maxima and minima, as is shown in Sec. 10. Consequently, (28) has only a finite number of active constraints. The collection of the time instants corresponding to these active constraints will be denoted by  $I_c^*$ . If this finite collection of constraint sampling instants is known ahead of time, one could simply solve (28) with respect to  $I_c^*$  instead of  $I_c$ , resulting in  $\hat{J}(I_c^*) = \hat{J}(I_c)$ . In addition, for any subset  $I_c^k$  of  $I_c$  it holds that  $\hat{J}(I_c) \geq \hat{J}(I_c^k)$ , and likewise, if  $I_c^k$  is a subset of  $I_c^{k+1}$  we have that  $\hat{J}(I_c^{k+1}) \geq \hat{J}(I_c^k)$ . Hence, a monotonically increasing sequence  $J(I_c^k)$ , bounded above by  $\hat{J}(I_c)$  can be constructed using any sequence of sets  $I_c^k$  that fulfill  $I_c^k \subset I_c^{k+1} \subset \dots \subset I_c$ . In particular, such sets are obtained by starting with an arbitrary initial guess  $I_c^0$  containing a finite number of constraint sampling points (or even the empty set), and by adding at least one constraint violation point at each iteration. Moreover, at each iteration the inactive constraints contained in the set  $I_c^k$  can be removed, as this will not alter the optimizer nor the optimal value  $\hat{J}(I_c^k)$ . In that way, the number of constraint sampling instances contained in  $I_c^k$  remains finite. This motivates Alg. 1, which solves (28) up to a given tolerance, by constructing an approximation to the set of active constraints  $I_c^*$ .

**Proposition 4.1** *Alg. 1 converges, that is,  $\lim_{k \rightarrow \infty} \hat{J}(I_c^k) = \hat{J}(I_c^*) = \hat{J}(I_c)$ . In order to achieve constraint violations smaller than  $\epsilon$  at most*

$$\frac{4c_\tau(\hat{J}(I_c) - \hat{J}(I_c^0))}{\sigma\epsilon^2} \quad (31)$$

*steps are required, where  $c_\tau$  is defined as*

$$c_\tau := \sup_{t \in I_c} |\tau(t)|^2, \quad (32)$$

*and  $\sigma$  denotes the smallest eigenvalue of the Hessian  $H$ .*

**Proof** From the above arguments it can be concluded that  $\hat{J}(I_c^k)$  is monotonically increasing whenever  $\hat{J}(I_c^k) < \hat{J}(I_c)$  and upper bounded by  $\hat{J}(I_c)$ . Therefore the sequence  $\hat{J}(I_c^k)$  converges. It remains to show that  $\hat{J}(I_c^k)$  converges to  $\hat{J}(I_c)$ . The strong convexity of the objective function can be used to establish

$$|z^{k+1} - z^k|^2 \leq 4\sigma^{-1}(\hat{J}(I_c^{k+1}) - \hat{J}(I_c^k)), \quad (33)$$

<sup>2</sup>This is due to the fact that the constraints describe a closed convex set and due to the strong convexity of the objective function.

---

**Algorithm 1** Iterative constraint sampling

---

**Initialize:** initial guess for the constraint sampling points:  $I_c^0 = \{t_1, t_2, \dots, t_N\}$ ; maximum number of iterations: MAXITER; constraint satisfaction tolerance:  $\epsilon$ ;

```
1:  $k = 0$ 
2: for  $k < \text{MAXITER}$  do
3:   solve (28) for  $I_c^k \rightarrow z^k, \hat{J}(I_c^k)$ 
4:   if infeasible then
5:     abort
6:   else if  $z^k$  fulfills (30) for all  $t \in I$  (with tol.  $\epsilon$ ) then
7:     algorithm converged
8:     return  $z^k$ 
9:   end if
10:  find at least one constraint violation instant  $\rightarrow t_c$ 
11:  remove inactive time instants in  $I_c^k$ 
12:   $I_c^{k+1} = I_c^k \cup \{t_c\}, k = k + 1$ 
13: end for
```

---

where  $z^{k+1}$  and  $z^k$  are the minimizer corresponding to  $\hat{J}(I_c^k)$  and  $\hat{J}(I_c^{k+1})$ . As a result we can conclude that  $z^k$  converges, and that  $\lim_{k \rightarrow \infty} z^k$  is well-defined and satisfies the constraint (30) (by a contradiction argument). It is therefore a feasible candidate for (28), implying that  $\lim_{k \rightarrow \infty} \hat{J}(I_c^k) \geq \hat{J}(I_c)$ , which, combined with  $\hat{J}(I_c^k) \leq \hat{J}(I_c)$  for all integers  $k$ , leads to  $\lim_{k \rightarrow \infty} \hat{J}(I_c^k) = \hat{J}(I_c)$ .

It remains to show that (31) is fulfilled. To that extent, let  $\epsilon > 0$  denote the smallest constraint violation that occurs within the first  $N$  steps. For all  $k \leq N - 1$ , there exists the constraint violation point  $t_j^k$ , which will be added to  $I_c^k$ , and therefore

$$\epsilon \leq |(e(t_j^k) \otimes \tau(t_j^k))^\top (z^k - z^{k+1})|, \quad (34)$$

where  $e(t_j^k)$  is a canonical unit vector. Combining the Cauchy-Schwarz inequality and the above bound on  $|z^{k+1} - z^k|$  results in

$$\epsilon^2 \leq 4c_\tau \sigma^{-1} (\hat{J}(I_c^{k+1}) - \hat{J}(I_c^k)). \quad (35)$$

By summing over the first  $N$  steps we arrive at

$$N\epsilon^2 \leq 4\sigma^{-1} c_\tau (\hat{J}(I_c^N) - \hat{J}(I_c^0)) \leq 4\sigma^{-1} c_\tau (\hat{J}(I_c^*) - \hat{J}(I_c^0)), \quad (36)$$

since the sequence  $\hat{J}(I_c^k)$  is strictly increasing and bounded above by  $\hat{J}(I_c^*)$ . Dividing by  $\epsilon^2$  on both sides concludes the proof.  $\blacksquare$

## 4.1 Implementation details

Alg. 1 can be naturally embedded in an active-set method. An introduction to active-set methods for solving quadratic programs can be found for example in [1, Ch. 10]. Starting from an initial guess of the active constraint sampling instants, which is denoted by  $I_c^0$ , the quadratic program with optimal cost  $\hat{J}(I_c^0)$  is solved: This is done by initially assuming that all constraints in the set  $I_c^0$  are active. The resulting optimization problem reduces to an equality constrained quadratic program, whose solution can be calculated by solving a linear system of equations. The Lagrange multipliers corresponding to (30), which are denoted by  $\mu(t)$ ,  $t \in I_c^0$ , are evaluated subsequently. If all the constraints are indeed active, then the optimizer to the quadratic program with cost  $\hat{J}(I_c^0)$  has been found. If, however, not all the constraints are found to be active, that is, if there are some Lagrange multipliers that are zero, the standard active-set procedure, see [1, Ch. 10] is used to find the subset of active constraints  $I_a \subset I_c^0$ . Provided that the active set  $I_a \subset I_c^0$  and the optimizer to the quadratic program with cost  $\hat{J}(I_c^0)$  has been found, the constraint (30) is then checked for all  $t \in I_c$ .



If no constraint violations occur, the solution to (28) has been found. If constraint violations occur, the time instant  $t_c$  for which a violation occurs, is added to the set of active constraints  $I_a$  resulting in  $I_c^1 = I_a \cup \{t_c\}$ . The above procedure is then repeated until convergence.

Each iteration requires solving equality constrained quadratic programs of the type

$$\min z^\top H z \quad (37)$$

$$\text{s.t. } A_{\text{eq}} z = b_{\text{eq}}, \quad (c(t) \otimes \tau(t))^\top C_z z = l_a(t), \quad \forall t \in I_a, \quad (38)$$

where  $c(t) \in \mathbb{R}^{n_c}$ ,  $l_a(t) \in \mathbb{R}$ ,  $t \in I_a$ , and  $I_a \subset I_c^k$  describe the active constraints corresponding to (30). Due to the fact that very few constraints are expected to be active, we use a range space approach, [1, p. 238]. To that extent, the equality constraint is eliminated and the optimizer  $z^*$  corresponding to (37) is rewritten as

$$z^* = \hat{b} + \hat{H} C_z^\top \sum_{t \in I_a} (c(t) \otimes \tau(t))^\top \mu(t), \quad (39)$$

where the dual variable  $\mu(t) \in \mathbb{R}$ , defined for  $t \in I_c^k$ , satisfies

$$(c(t_j) \otimes \tau(t_j))^\top C_z \hat{H} C_z^\top \sum_{t \in I_a} (c(t) \otimes \tau(t)) \mu(t) = l_a(t_j) - (c(t_j) \otimes \tau(t_j))^\top C_z \hat{b}, \quad (40)$$

for all  $t_j \in I_a \subset I_c^k$ , and  $\mu(t) = 0$  for all  $t \in I_c^k \setminus I_a$ , with

$$\hat{H} := H^{-1} A_{\text{eq}}^\top (A_{\text{eq}} H^{-1} A_{\text{eq}}^\top)^{-1} A_{\text{eq}} H^{-1} - H^{-1}, \quad (41)$$

$$\hat{b} := H^{-1} A_{\text{eq}}^\top (A_{\text{eq}} H^{-1} A_{\text{eq}}^\top)^{-1} b_{\text{eq}}. \quad (42)$$

The dual variable  $\mu(t)$  is therefore obtained by solving (40), and the optimizer  $z^*$  is then determined via (39). At each iteration, a single constraint is either added or removed. Therefore, in order to efficiently find a solution to (40), a  $LDL^\top$ -decomposition of the matrix

$$\{(c(t_j) \otimes \tau(t_j))^\top C_z \hat{H} C_z^\top (c(t_i) \otimes \tau(t_i))\}_{(t_j, t_i) \in I_a \times I_a} \quad (43)$$

is computed and adapted at each step by performing rank-1 updates. The matrix  $\hat{H}$  and the vector  $\hat{b}$  are precomputed. For additional details regarding the regularity of (43), and issues related to cycling and stalling we refer to [1, Ch. 10] and [4, p. 467].

## 4.2 Constraint check

It remains to explain how to efficiently check whether the constraint (30) is fulfilled for a given solution candidate  $z$ . We assume that the interval  $I_c$  has the form  $I_c = [0, T_c]$ . As it has been explained in Sec. 3, the constraint check over the interval  $[0, \infty)$  reduces to the check over a compact interval, provided that the basis functions fulfill Assumptions A1 and A2.

A straightforward approach would be to exploit the specific structure of the basis functions. For example, if the basis functions consist of exponentially decaying polynomials having a degree of at most 4, determining the stationary points of (30) amounts to solving a quartic equation, which can be done analytically. As a result, it would be enough to check the constraints at these stationary points in order to determine if the constraint is satisfied or not.

We propose a more general approach that is based on local Taylor approximations, and thus valid for arbitrary basis functions compatible with Assumptions A1 and A2. In order to simplify the discussion, we consider the special case of (30), where  $C_z$  is the identity and  $n_c = 1$ . The resulting algorithm extends naturally to the more general case. According to Taylor's theorem we obtain the following identity

$$\tau(t)^\top z = \tau(0)^\top z + \dot{\tau}(0)^\top z t + \ddot{\tau}(0)^\top z \frac{t^2}{2} + \tau^{(3)}(\bar{t})^\top z \frac{t^3}{6}, \quad (44)$$

where  $\bar{t} \in [0, t]$ . As will become clear in the following, a third-order Taylor expansion represents a good compromise between approximation quality and computational effort. The last term of the previous equation can be bounded by the Cauchy-Schwarz inequality leading to

$$|\tau^{(3)}(s)^\top z| \leq \sup_{\bar{t} \in I_c} |\tau(\bar{t})| |(M^\top)^3 z| =: R(z), \quad \forall s \in [0, t]. \quad (45)$$

As a consequence, the following upper and lower bounds are obtained

$$b_l(t) \leq \tau(t)^\top z \leq b_u(t), \quad (46)$$

for all  $t \in I_c$ , with

$$b_l(t) := \tau(0)^\top z + \dot{\tau}(0)^\top z t + \ddot{\tau}(0)^\top z \frac{t^2}{2} - R(z) \frac{t^3}{6}, \quad (47)$$

$$b_u(t) := \tau(0)^\top z + \dot{\tau}(0)^\top z t + \ddot{\tau}(0)^\top z \frac{t^2}{2} + R(z) \frac{t^3}{6}. \quad (48)$$

The situation is exemplarily depicted in Fig. 1 (right). Given that  $\dot{\tau}(0)^\top z \geq 0$  the lower bound attains its maximum at time

$$t_u := \frac{\ddot{\tau}(0)^\top z + \sqrt{(\ddot{\tau}(0)^\top z)^2 + 2R(z)\dot{\tau}(0)^\top z}}{R(z)} > 0, \quad (49)$$

whereas if  $\dot{\tau}(0)^\top z < 0$  the upper bound attains its minimum at time

$$t_l := \frac{-\ddot{\tau}(0)^\top z + \sqrt{(\ddot{\tau}(0)^\top z)^2 - 2R(z)\dot{\tau}(0)^\top z}}{R(z)} > 0. \quad (50)$$

Thus, if the lower bound exceeds  $l_u$  or the upper bound drops below  $l_l$ , that is, if

$$b_l(t_u) > l_u, \quad \text{or} \quad b_u(t_l) < l_l, \quad (51)$$

the constraint (30) is guaranteed to be violated at time  $t_u$ , respectively at time  $t_l$ . If this is not the case, we are guaranteed that the constraint (30) is satisfied in the interval  $[0, t_s]$ , with

$$t_s := \min\{t_1, t_2 \mid b_u(t_1) = l_u, b_l(t_2) = l_l\}. \quad (52)$$

Thus, finding the value  $t_s$  requires the solution of two cubic equations, which stems from the fact that a third order Taylor approximation was used as a starting point. By shifting the parameter vector in time by  $t_s$  and repeating the above procedure, the constraint is either found to be satisfied for all  $t \in I_c$  or a constraint violation is detected. Shifting the parameter vector by  $t_s$  amounts in multiplying  $z$  with the matrix exponential

$$e^{M^\top t_s} z \rightarrow z. \quad (53)$$

The procedure is illustrated by the flow chart depicted in Fig. 1 (left).

The efficiency of the proposed strategy can be improved via the following observation. A constraint violation occurring close to  $t = 0$ , is often found within the first few iterations. However, if no constraint violation occurs, the whole interval  $I_c$  needs to be traversed, which tends to increase computation. The computational effort may be reduced by including additional conservative constraint satisfaction checks. For example, upper and lower bounds can be tightened by a factor  $\gamma \in (0, 1)$  such that the satisfaction of

$$\gamma l_l \leq \tau(t_i)^\top z \leq \gamma l_u \quad (54)$$

for certain time instances  $t_i$  implies (30) (for all  $t \in I_c$ ). As a result, at each iteration, the above inequality is checked. If it is found to be fulfilled, then constraint satisfaction can be guaranteed.

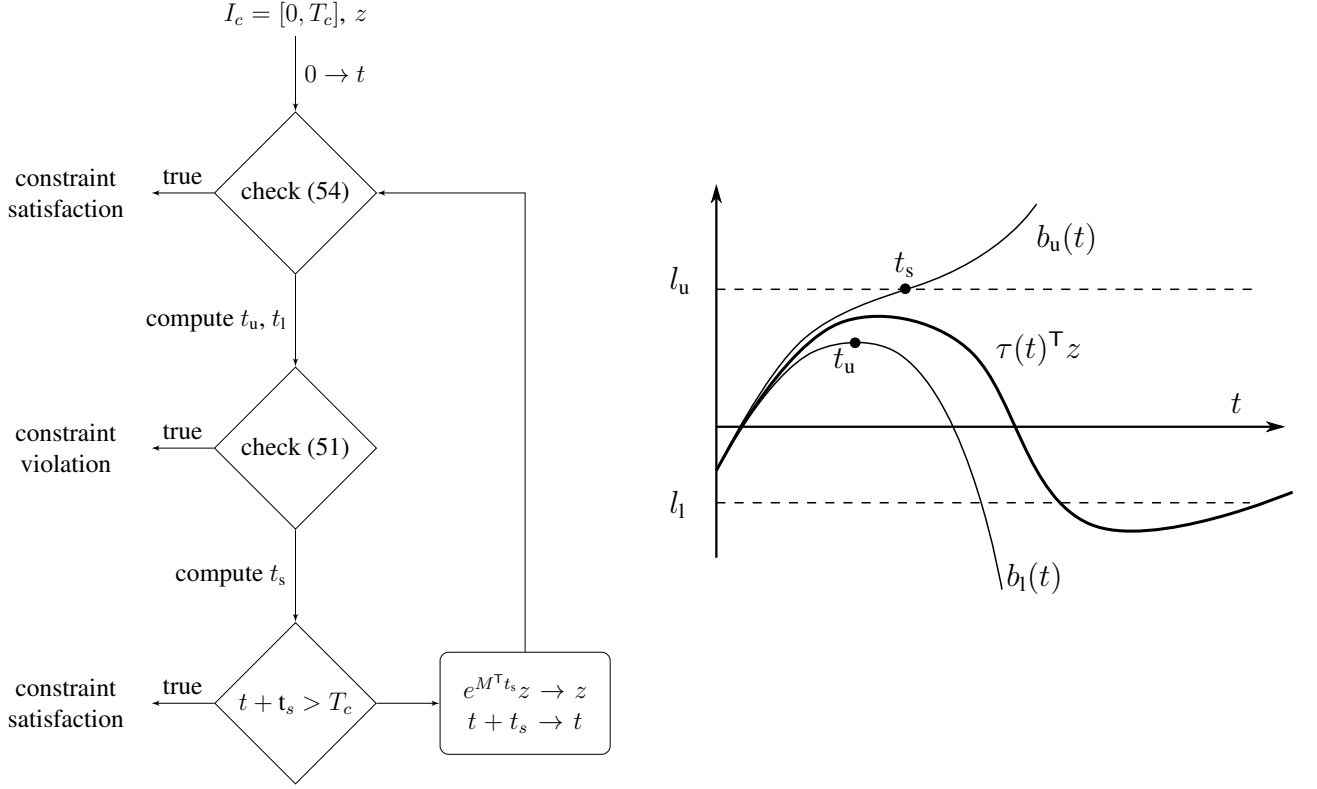


Figure 1: Left: Flow chart that illustrates the proposed constraint satisfaction check. Right: Illustration of the upper and lower bounds  $b_u(t)$  and  $b_l(t)$  obtained from the Taylor expansion of  $\tau(t)^T z$ . In that case the test (51) is indecisive, and constraint satisfaction over the interval  $[0, t_s]$  can be guaranteed.

## 5 Properties B1-B5

The sets  $\mathcal{C}$ ,  $\mathcal{C}_U^s$ , and  $\mathcal{C}_L^s$  have the following properties

- B1) the sets  $\mathcal{C}_U^s$ ,  $\mathcal{C}_L^s$ , and  $\mathcal{C}$  are closed and convex.
- B2) the sets  $\pi^s(\mathcal{C}_U^s)$ ,  $\pi^s(\mathcal{C}_L^s)$  are closed and convex.
- B3)  $\mathcal{C}_U^s \subset \mathcal{C}_U^{s+1} \subset \mathcal{C}$ .
- B4)  $\mathcal{C}_L^s \supset \mathcal{C}_L^{s+1} \supset \mathcal{C}$ .
- B5)  $(\pi^s)^* \pi^s(\mathcal{C}_U^s) \subset \mathcal{C}_U^s$ ,  $(\pi^s)^* \pi^s(\mathcal{C}_L^s) \subset \mathcal{C}_L^s$ .

We will sketch the proofs of Properties B1-B5 in the following. It follows from the linearity of the stage constraint, and the fact that the terminal constraint  $\mathcal{X}$  is convex, that the sets  $\mathcal{C}$ ,  $\mathcal{C}_U^s$ , and  $\mathcal{C}_L^s$  are likewise convex. We will sketch the proof that the set  $\mathcal{C}_L^s$  is closed. The argument can be translated to the set  $\mathcal{C}$  by using the following variational formulation

$$\mathcal{C} = \{(x, u, x_T) \in X \mid \int_I \delta p^T (-C_x x - C_u u + b) dt \geq 0, \forall \delta p \in L_{n_c}^2 : \delta p \geq 0, \left| \int_I \delta p^T b dt \right| < \infty; x_T \in \mathcal{X}\}. \quad (55)$$

We will argue indirectly, i.e. that the complement of  $\mathcal{C}_L^s$  is open. To that extent we choose  $(x, u, x_T) \in X \setminus \mathcal{C}_L^s$ . As a result, there exists a test function  $\delta \tilde{p}$ , with  $\delta \tilde{p} \geq 0$ , which is spanned by the first  $s$  basis functions and is such that

$$\int_I \delta \tilde{p}^T (-C_x x - C_u u + b) dt < 0. \quad (56)$$

For any  $\hat{x} \in L_n^2$ ,  $\hat{u} \in L_m^2$ , and  $\hat{x}_T \in \mathbb{R}^n$ , with  $\|\hat{x} - x\| < \epsilon$ ,  $\|\hat{u} - u\| < \epsilon$ , and  $|\hat{x}_T - x_T| < \epsilon$ , it follows that

$$\int_I \delta\tilde{p}^\top (-C_x \hat{x} - C_u \hat{u} + b) dt = \int_I \delta\tilde{p}^\top (-C_x x - C_u u + b) dt + \int_I \delta\tilde{p}^\top (C_x(x - \hat{x}) + C_u(u - \hat{u})) dt, \quad (57)$$

where the last integral can be bounded by (using the Cauchy-Schwarz inequality)

$$\|\delta\tilde{p}\|(|C_x|\|x - \hat{x}\| + |C_u|\|u - \hat{u}\|) < \|\delta\tilde{p}\|(|C_x| + |C_u|)\epsilon. \quad (58)$$

We can infer from  $\mathcal{X}$  being closed, that there exists an open ball centered at  $x_T$ , which does not intersect  $\mathcal{X}$ . As a result, by choosing  $\epsilon$  small enough, it can be concluded that

$$\int_I \delta\tilde{p}^\top (-C_x \hat{x} - C_u \hat{u} + b) dt < 0, \quad \hat{x}_T \notin \mathcal{X}, \quad (59)$$

and therefore  $(\hat{x}, \hat{u}, x_T) \in X \setminus \mathcal{C}_L^s$ , for all  $\hat{x} \in L_n^2$ ,  $\hat{u} \in L_m^2$ ,  $\hat{x}_T \in \mathbb{R}^n$  with  $\|x - \hat{x}\| < \epsilon$ ,  $\|u - \hat{u}\| < \epsilon$ , and  $|\hat{x}_T - x_T| < \epsilon$ . Hence, the complement of  $\mathcal{C}_L^s$  is open, and therefore  $\mathcal{C}_L^s$  is closed.

Note that the set  $\mathcal{C}_U^s$  is given by the intersection of the set  $\mathcal{C}$  with the linear subspace  $X^s$  spanned by the first  $s$  basis functions. Both of these sets are closed<sup>3</sup> implying that  $\mathcal{C}_U^s$  is closed as well.

The projection  $\pi^s$  is linear, which asserts the convexity of the sets  $\pi^s(\mathcal{C}_U^s)$  and  $\pi^s(\mathcal{C}_L^s)$ . Moreover, it is surjective, and hence, by the open mapping theorem, it follows directly from  $\mathcal{C}_L^s$  and  $\mathcal{C}_U^s$  being closed that  $\pi^s(\mathcal{C}_L^s)$  and  $\pi^s(\mathcal{C}_U^s)$  are closed as well.

The inclusion  $\mathcal{C}_U^s \subset \mathcal{C}_U^{s+1}$  follows directly from  $\mathcal{C}_U^s = \mathcal{C} \cap X^s$  and the inclusion  $X^s \subset X^{s+1}$ . In other words, given  $(x, u, x_T) \in \mathcal{C}_U^s$ , the parameter vectors  $\eta_x$  and  $\eta_u$ , corresponding to the state and input trajectories  $x$  and  $u$ , can be extended with zeros resulting in trajectories  $\hat{x}$ ,  $\hat{u}$  spanned by  $s+1$  basis functions. But  $\hat{x} = x$  and  $\hat{u} = u$  and therefore  $(x, u, x_T) \in \mathcal{C}_U^{s+1}$ . The inclusion  $\mathcal{C}_L^s \supset \mathcal{C}_L^{s+1}$  follows from the fact that the dimension of the subspace to which the test functions  $\delta\tilde{p}$  are constrained increases with  $s$ .

The claim  $(\pi^s)^* \pi^s(\mathcal{C}_U^s) \subset \mathcal{C}_U^s$  follows from the fact that  $(\pi^s)^* \pi^s$  is a projection from  $X$  onto  $X^s \subset X$  and that  $\mathcal{C}_U^s \subset X^s$ . The claim that  $(\pi^s)^* \pi^s(\mathcal{C}_L^s) \subset \mathcal{C}_L^s$  follows from the linearity of the stage constraints. More precisely, it follows by noting that for any  $\delta\tilde{p} = (I_{n_c} \otimes \tau)^\top \delta\eta_p$  and any  $x \in L_n^2$ ,

$$\begin{aligned} \int_I \delta\tilde{p}^\top C_x x dt &= \delta\eta_p^\top C_x \pi^{ns}(x) \\ &= \int_I \delta\tilde{p}^\top C_x (\pi^{ns})^* \pi^{ns}(x) dt \end{aligned} \quad (60)$$

holds.

## 6 Properties C1-C5

The sets  $\mathcal{D}$ ,  $\mathcal{D}_U^s$ , and  $\mathcal{D}_L^s$  have the following properties

- C1) the sets  $\mathcal{D}_U^s$ ,  $\mathcal{D}_L^s$ , and  $\mathcal{D}$  are closed and convex.
- C2) the sets  $\pi^s(\mathcal{D}_U^s)$  and  $\pi^s(\mathcal{D}_L^s)$  are closed and convex.
- C3)  $\mathcal{D}_U^s \subset \mathcal{D}_U^{s+1} \subset \mathcal{D}$ .
- C4)  $\mathcal{D}_L^s \supset \mathcal{D}_L^{s+1} \supset \mathcal{D}$ .
- C5)  $(\pi^s)^* \pi^s(\mathcal{D}_U^s) \subset \mathcal{D}_U^s$ ,  $(\pi^s)^* \pi^s(\mathcal{D}_L^s) \subset \mathcal{D}_L^s$ .

<sup>3</sup>  $X^s$  is finite-dimensional, thus complete, and hence also closed.

We will sketch the proof of Properties C1-C5 below.

The convexity of the sets  $\mathcal{D}_U^s$ ,  $\mathcal{D}_L^s$ , and  $\mathcal{D}$  follows directly from the linearity of the dynamics.

The fact that the set  $\mathcal{D}_L^s$  is closed can be seen by a similar argument used for showing closedness of  $\tilde{\mathcal{C}}^s$  in Sec. 5, that is, showing that  $X \setminus \mathcal{D}_L^s$  is open. The set  $\mathcal{D}$  can be characterized using the variational equality,

$$-\int_I \delta \dot{p}^\top x dt - \int_I \delta p^\top (Ax + Bu) dt - \delta p(0)^\top x_0 + \delta p(T)^\top x_T = 0, \quad \forall \delta p \in H_n, \quad (61)$$

where  $H_n$  denotes the set of functions in  $L_n^2$  having a weak derivative in  $L_n^2$ . Thus, again a similar argument can be applied to show that  $\mathcal{D}$  is closed. It follows that  $\mathcal{D}_U^s$  is closed, since  $\mathcal{D}_U^s$  is defined as the intersection of  $\mathcal{D}$  with the closed set  $X^s$ .

The linearity and surjectivity of the map  $\pi^s$  implies that  $\pi^s(\mathcal{D}_U^s)$  and  $\pi^s(\mathcal{D}_L^s)$  are indeed closed (by the open mapping theorem) and convex.

The inclusion  $\mathcal{D}_U^s \subset \mathcal{D}_U^{s+1} \subset \mathcal{D}$  for all  $s$  follows directly from the fact that  $X^s \subset X^{s+1} \subset X$  and  $\mathcal{D}_U^s = \mathcal{D} \cap X^s$ . The inclusion  $\mathcal{D}_L^{s+1} \subset \mathcal{D}_L^s$  for all  $s$  can be seen by noting that the variational equality in (16) has to hold for variations spanned by more and more basis functions as  $s$  increases. The claim that  $\mathcal{D}$  is contained in  $\mathcal{D}_L^s$  follows from the equivalence of (61) with the formulation in (7). The properties  $(\pi^s)^* \pi^s(\mathcal{D}_U^s) \subset \mathcal{D}_U^s$  and  $(\pi^s)^* \pi^s(\mathcal{D}_L^s) \subset \mathcal{D}_L^s$  can be shown using the same arguments as in 5, where the latter relies on the linearity of the dynamics.

## 7 Convergence results

We prove the following result for the case where  $I$  has infinite measure.

**Lemma 7.1** *Given that the basis functions form an algebra and that the basis functions are dense in the set of smooth functions with compact support in  $I$ , it holds that*

$$\lim_{s \rightarrow \infty} \mathcal{C}_U^s = \lim_{s \rightarrow \infty} \mathcal{C}_L^s.$$

**Proof** The idea of the proof is the following: We claim that  $\lim_{s \rightarrow \infty} \mathcal{C}_U^s \supset \lim_{s \rightarrow \infty} \mathcal{C}_L^s$ . We assume that the claim is incorrect and show that this leads to a contradiction. Thus, we choose  $(x, u, x_T) \in \lim_{s \rightarrow \infty} \mathcal{C}_L^s$ , such that there exists an open set  $U$  (bounded) and a  $k \in \{1, 2, \dots, n_c\}$ , for which

$$\int_I \delta v (-C_{xk} x - C_{uk} u + b_k) dt < 0 \quad (62)$$

holds for all smooth test functions  $\delta v : I \rightarrow \mathbb{R}$ ,  $\delta v \geq 0$ , with support in  $U$ , and  $\delta v(t_0) > 0$  for some  $t_0 \in U$ . Due to the smoothness of the test functions,  $\delta v(t_0) > 0$  readily implies that there is an open neighborhood of  $t_0$ , denoted by  $\mathcal{N}(t_0)$ , such that  $\delta v(t) > 0, \forall t \in \mathcal{N}(t_0)$ . The above integral exists, since  $\delta v$  is bounded, has compact support and  $x \in L_n^2, u \in L_m^2$ . We fix  $t_0 \in U$  and pick one of these variations that is positive, strictly positive at time  $t_0$ , and has support in  $U$ , which we name  $\delta p$ . Due to the fact that the basis functions are dense in the set of smooth functions with compact support, there exists a sequence  $\sqrt{\delta \tilde{p}_i}$  that converges uniformly to  $\sqrt{\delta p}$ . Due to the fact that the basis functions form an algebra,  $\delta \tilde{p}_i$  lies likewise in the span of the basis functions. Moreover, for a given  $\epsilon > 0$  (small enough) there exists an integer  $N > 0$  such that

$$\|\delta \tilde{p}_i - \delta p\|_\infty < C_1 \epsilon$$

holds for all integers  $i > N$ , where  $C_1 > 0$  is constant.

We claim that there is an integer  $p > 0$  such that the basis function  $\tau_p : I \rightarrow \mathbb{R}$  is nonzero for  $t_0$ . This can be shown by a contradiction argument: If the claim was not true, then all basis functions  $\tau^p = (\tau_1, \tau_2, \dots, \tau_p)$  would be zero at time  $t_0$ . The basis functions fulfill Assumption A2, and hence the first order differential equation  $\dot{\tau}^p = M_p \tau^p$ , which is guaranteed to have unique solutions. From  $\tau^p(t_0) = 0, \dot{\tau}^p(t_0) =$

0 we can infer that  $\tau^p(t) = 0$  for all  $t \in (0, \infty)$  is the (unique) solution to  $\dot{\tau}^p = M_p \tau^p$ , contradicting Assumption A1.

Thus, we can choose the basis function  $\tau_p$  that is nonzero for  $t_0$  and hence also nonzero in a neighborhood around  $t_0$ , due to the smoothness of the basis functions. The same applies to the function  $\tau_p^2$ , which is likewise contained in the set of basis functions (the basis functions form an algebra that is closed under multiplication). Moreover, due to the fact that the basis functions are orthonormal, it follows that

$$\int_I \tau_p^2 dt = 1. \quad (63)$$

In the same way, the function  $\delta \tilde{p}_i(t) \tau_p^2(t)$  is also contained in the set of basis functions, is non-negative for all  $t \in I$ , and is bounded and integrable for all integers  $i > 0$  ( $|\tau^p|$  is bounded, see 10). Thus, it follows that the integral

$$\int_I \tau_p^2 \delta \tilde{p}_i (-C_{xk}x - C_{uk}u + b_k) dt \quad (64)$$

exists. By assumption  $(x, u, x_T) \in \lim_{s \rightarrow \infty} \mathcal{C}_L^s$ , and therefore

$$0 \leq \int_I \tau_p^2 \delta \tilde{p}_i (-C_{xk}x - C_{uk}u + b_k) dt \quad (65)$$

$$= \int_I \tau_p^2 \delta p (-C_{xk}x - C_{uk}u + b_k) dt \quad (66)$$

$$+ \int_I \tau_p^2 (\delta \tilde{p}_i - \delta p) (-C_{xk}x - C_{uk}u + b_k) dt, \quad (67)$$

where the last term can be bounded by

$$\epsilon C_1 \int_I \tau_p^2 |C_{xk}x + C_{uk}u - b_k| dt. \quad (68)$$

The above integral is bounded due to the fact that  $\tau_p^2$  is bounded and integrable,  $x \in L_n^2$ ,  $u \in L_m^2$ , and that  $b_k$  is bounded, and therefore (67) can be made arbitrarily small by sufficiently increasing  $i$ . However, this leads to a contradiction, since (66) is strictly negative, according to (62). This proves that  $\lim_{s \rightarrow \infty} \mathcal{C}_U^s \supset \lim_{s \rightarrow \infty} \mathcal{C}_L^s$ . The desired result is then established due to the fact that  $\mathcal{C}_U^s \subset \mathcal{C}_L^s$  holds for all integers  $s > 0$ . ■

## 8 Recursive feasibility and closed-loop stability

**Proposition 8.1** *Provided that the optimization (21) is feasible at time  $t = 0$ , it remains feasible for all times  $t > 0$ , and the resulting closed-loop system is guaranteed to be asymptotically stable.*

**Proof** The proof is taken from [2] and included for completeness. The following notation is introduced: The closed-loop state and input trajectories are denoted by  $x(t)$  and  $u(t)$ . The predicted trajectories are referred to as  $\tilde{x}(t_p|t)$ ,  $\tilde{u}(t_p|t)$ , where  $t_p > 0$  denotes the prediction horizon. For  $t_p = 0$ , the prediction matches the true trajectory, that is  $\tilde{x}(0|t) = x(t)$ ,  $\tilde{u}(0|t) = u(t)$  for all  $t \in [0, \infty)$ . The predictions  $\tilde{x}(t_p|t)$ ,  $\tilde{u}(t_p|t)$  are obtained by solving (21) subject to the initial condition  $x_0 = x(t)$ , which yields the parameters  $\eta_x$  and  $\eta_u$  defining  $\tilde{x}(t_p|t)$  and  $\tilde{u}(t_p|t)$  by

$$\tilde{x}(t_p|t) = (I_n \otimes \tau(t_p))^T \eta_x, \quad \tilde{u}(t_p|t) = (I_m \otimes \tau(t_p))^T \eta_u. \quad (69)$$

In order to highlight the dependence on the initial condition, the resulting optimal cost of (21) is denoted by  $J^{\text{MPC}}(x(t))$ .

By assumption, (21) is feasible at time  $t = 0$ . The resulting trajectories  $\tilde{x}(t_p|0)$ ,  $\tilde{u}(t_p|0)$  fulfill the equations of motion, the initial condition  $\tilde{x}(0|0) = x(0)$ , and the constraints and hence, the system evolves according to  $x(t) = \tilde{x}(t|0)$ ,  $u(t) = \tilde{u}(t|0)$ ,  $\forall t \in [0, T_d]$ . Due to the time-shift property of the basis functions implied by Assumption A2, the feasible candidates

$$\begin{aligned}\tilde{x}(t_p + T_d|0) &= (I_n \otimes \tau(t_p)^\top)(I_n \otimes \exp(M_s T_d)^\top)\eta_x, \\ \tilde{u}(t_p + T_d|0) &= (I_m \otimes \tau(t_p)^\top)(I_m \otimes \exp(M_s T_d)^\top)\eta_u\end{aligned}\tag{70}$$

for the optimization at time  $T_d$  can be constructed from the optimizer  $\eta_x$  and  $\eta_u$  at time 0. As a result, recursive feasibility of (21) follows by induction.

We will show that the optimal cost  $J^{\text{MPC}}$  acts as a Lyapunov function. The function  $J^{\text{MPC}}$  is a valid Lyapunov candidate since  $J^{\text{MPC}}(x) > 0$  for all  $x \neq 0$  and  $J^{\text{MPC}}(x) = 0$  if and only if  $x = 0$ . Due to the fact that the shifted trajectories  $\tilde{x}(t_p + T_d|0)$  and  $\tilde{u}(t_p + T_d|0)$  (as defined in (70)) are feasible for the optimization at time  $T_d$ , the following upper bound on  $J^{\text{MPC}}(x(T_d))$  can be established

$$J^{\text{MPC}}(x(T_d)) \leq \int_{T_d}^{\infty} \tilde{x}(t_p|0)^\top Q \tilde{x}(t_p|0) + \tilde{u}(t_p|0)^\top R \tilde{u}(t_p|0) dt_p.\tag{71}$$

The right-hand side can be rewritten as

$$J^{\text{MPC}}(x(0)) - \int_0^{T_d} \tilde{x}(t_p|0)^\top Q \tilde{x}(t_p|0) + \tilde{u}(t_p|0)^\top R \tilde{u}(t_p|0) dt_p,\tag{72}$$

resulting in

$$\begin{aligned}J^{\text{MPC}}(x(T_d)) - J^{\text{MPC}}(x(0)) &\leq \\ &- \int_0^{T_d} \tilde{x}(t_p|0)^\top Q \tilde{x}(t_p|0) + \tilde{u}(t_p|0)^\top R \tilde{u}(t_p|0) dt_p.\end{aligned}\tag{73}$$

The right-hand side of (73) is guaranteed to be strictly negative, except for  $x(0) = 0$ , and thus, by induction,  $J^{\text{MPC}}(x(kT_d))$  is strictly decreasing, which concludes the proof.  $\blacksquare$

## 9 Reduction of the semi-infinite constraint

The following section discusses the reduction of the semi-infinite constraint

$$a_l \leq \tau(t)^\top \eta \leq a_u\tag{74}$$

over the interval  $t \in [0, \infty)$ , with  $a_l, a_u \in \mathbb{R}$ ,  $a_l < 0$ ,  $a_u > 0$  to a compact interval. Thereby we consider the symmetric case, where  $|a_l| = |a_u|$  first, before discussing the asymmetric case  $|a_l| \neq |a_u|$ . It is shown that the length of this compact time interval depends only on the properties of the basis functions and on the ratio between  $|a_u|$  and  $|a_l|$ . Both  $a_l$  and  $a_u$  are assumed to be finite.

### 9.1 The symmetric case

**Proposition 9.1** *Provided that the basis functions fulfill Assumptions A1 and A2 for all  $t \in [0, \infty)$  there exists a positive real number  $T_c$  such that*

$$\sup_{t \in [0, \infty)} |\tau(t)^\top \eta| = \max_{t \in [0, T_c]} |\tau(t)^\top \eta|\tag{75}$$

holds for all parameter vectors  $\eta \in \mathbb{R}^s$ .

**Proof** Without loss of generality we restrict the parameter vectors to have unit magnitude, that is,  $|\eta| = 1$ .<sup>4</sup>

We prove the claim in 4 steps. We first derive an exponentially decaying upper bound on the Euclidean norm of the basis function vector  $\tau$ . We then use this bound to argue that the basis functions are linearly independent over the interval  $[0, T_i]$  (the scalar  $T_i$  will be determined). The third step consists of constructing a lower bound on

$$\max_{t \in [0, T_i]} |\tau(t)^\top \eta| \quad (76)$$

that holds for all parameter vectors  $\eta$  with  $|\eta| = 1$ . Linear independence of the basis functions on  $[0, T_i]$  will be used to argue that this lower bound is strictly positive. In the last step, we show that if  $t$  is sufficiently large,  $|\tau(t)^\top \eta|$  will be below this lower bound, which concludes the proof.

Step 1): The fact that  $M_s$  is asymptotically stable implies that there exists a quadratic Lyapunov function that decays exponentially. This provides a means to establish the following bound

$$|\tau(t)|^2 \leq C_2 e^{-c_2 t}, \quad \forall t \in [0, \infty), \quad (77)$$

where  $C_2 > 0, c_2 > 0$  are constant.

Step 2): We use the Gram matrix to argue that the basis functions are linearly independent over the interval  $[0, T_i]$ . According to [5, p. 2, Thm. 3] it holds that the basis functions are linearly independent in the set  $[0, T_i]$  if and only if the matrix

$$\int_0^{T_i} \tau \tau^\top dt \quad (78)$$

has full rank. This is the case if the bilinear form

$$v^\top \int_0^{T_i} \tau \tau^\top dt v \quad (79)$$

is strictly positive for all  $v \in \mathbb{R}^s$  with  $|v| = 1$ . Combining the fact that the basis functions are orthonormal with the Cauchy-Schwarz inequality, leads to the following lower bound of the above bilinear form,

$$1 - \int_{T_i}^{\infty} (v^\top \tau)^2 dt \geq 1 - \int_{T_i}^{\infty} |\tau|^2 dt. \quad (80)$$

Using the upper bound (77) we therefore obtain

$$v^\top \int_0^{T_i} \tau \tau^\top dt v \geq 1 - \int_{T_i}^{\infty} C_2 e^{-c_2 t} dt = 1 - \frac{C_2}{c_2} e^{-c_2 T_i}, \quad (81)$$

for all  $v \in \mathbb{R}^s$  with  $|v| = 1$ . Thus, we fix  $T_i > 0$ , such that

$$1 > \frac{C_2}{c_2} e^{-c_2 T_i}, \quad (82)$$

implying that the matrix (78) is positive definite and has therefore full rank. Consequently, the basis functions are guaranteed to be linearly independent on the interval  $[0, T_i]$ .

Step 3): We claim that

$$c^* := \inf_{|\eta|=1} \max_{t \in [0, T_i]} |\tau(t)^\top \eta| \quad (83)$$

is well-defined and strictly positive. To that extent we first prove that the function  $g : \mathbb{R}^s \rightarrow [0, \infty)$ ,

$$g(\eta) := \max_{t \in [0, T_i]} |\tau(t)^\top \eta| \quad (84)$$

---

<sup>4</sup>The claim holds trivially for  $\eta = 0$ ; in case  $\eta \neq 0$  we can always normalize  $\eta$ .



is continuous (in fact Lipschitz-continuous). Therefore we consider two parameter vectors  $\eta_1$  and  $\eta_2$  with  $g(\eta_1) \leq g(\eta_2)$  (without loss of generality). From the fact that  $g(\eta_2)$  can be rewritten as  $g(\eta_1 + (\eta_2 - \eta_1))$  and by invoking the triangle inequality it can be inferred that

$$g(\eta_1 + (\eta_2 - \eta_1)) \leq \max_{t \in [0, T_i]} |\tau(t)^\top \eta_1| + |\tau(t)^\top (\eta_2 - \eta_1)|. \quad (85)$$

By noting that the maximum of the above sum is smaller than the sum of the summand's maxima it can be concluded that

$$g(\eta_2) \leq g(\eta_1) + \max_{t \in [0, T_i]} |\tau(t)^\top (\eta_2 - \eta_1)|. \quad (86)$$

Moreover, by combining the Cauchy-Schwarz inequality and the bound (77) we obtain

$$|g(\eta_2) - g(\eta_1)| \leq \sqrt{C_2} |\eta_2 - \eta_1|, \quad (87)$$

showing that the function  $g$  is indeed (Lipschitz) continuous. As a result, the Bolzano-Weierstrass theorem asserts that the infimum in (83), is attained and well-defined. It remains to argue that  $c^* > 0$ . For the sake of contradiction, we assume  $c^* = 0$ . This implies the existence of the minimizer  $\eta^*$ , with  $|\eta^*| = 1$ , which fulfills

$$\max_{t \in [0, T_i]} |\tau(t)^\top \eta^*| = 0. \quad (88)$$

As a result, it follows that  $\tau(t)^\top \eta^*$  is zero for all  $t \in [0, T_i]$ , which contradicts the fact that the basis functions  $\tau$  are linearly independent on  $[0, T_i]$ .

Step 4): From the upper bound (77) and the Cauchy-Schwarz inequality it follows that

$$|\tau(t)^\top \eta| \leq \sqrt{C_2} e^{-\frac{c_2 t}{2}}, \quad \forall t \in [0, \infty), \quad (89)$$

and for all  $\eta \in \mathbb{R}^s$  with  $|\eta| = 1$ . Clearly,  $c^* \leq \sqrt{C_2}$  and therefore we can choose the time  $T_c$  such that

$$c^* = \sqrt{C_2} e^{-\frac{c_2 T_c}{2}}, \quad (90)$$

implying

$$\sup_{t \in (T_c, \infty)} |\tau(t)^\top \eta| < c^* \leq \max_{t \in [0, T_i]} |\tau(t)^\top \eta|, \quad (91)$$

for all  $\eta \in \mathbb{R}^s$  with  $|\eta| = 1$ . This proves the claim.

## 9.2 The asymmetric case

**Proposition 9.2** *Provided that the basis functions fulfill Assumptions A1 and A2 for all  $t \in [0, \infty)$  there exists a positive real number  $T_c$  such that*

$$a_l \leq \tau(t)^\top \eta \leq a_u, \quad \forall t \in [0, T_c] \quad (92)$$

implies

$$a_l \leq \tau(t)^\top \eta \leq a_u, \quad \forall t \in [0, \infty), \quad (93)$$

for any  $\eta \in \mathbb{R}^s$ , where  $a_l, a_u \in \mathbb{R}$ ,  $a_l < 0, a_u > 0$ .

**Proof** We define  $\tilde{f} := \tau^\top \eta$  and establish upper and lower bounds on  $\tilde{f}$ . Without loss of generality we assume  $\eta \neq 0$ . Combining Assumption A2 with the Caley-Hamilton theorem leads to

$$\tilde{f}^{(s)}(t) + a_1 \tilde{f}^{(s-1)}(t) + \dots + a_s \tilde{f}(t) = \eta^\top \tau^{(s)}(t) + a_1 \eta^\top \tau^{(s-1)}(t) + \dots + a_s \eta^\top \tau(t) \quad (94)$$

$$= \eta^\top (M^s + a_1 M^{s-1} + \dots + a_s M) \tau(t) = 0, \quad (95)$$

where  $a_1, a_2, \dots, a_s$  are the coefficients of the characteristic polynomial of the matrix  $M$ . Thus, the trajectory  $\tilde{f}$  and its time derivatives fulfill the following set of differential equations

$$\dot{f} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ -a_s & -a_{s-1} & -a_{s-2} & \dots & -a_1 \end{pmatrix}}_{:=\hat{M}} f, \quad f(0) = \underbrace{\begin{pmatrix} \tau(0)^\top \\ \tau(0)^\top M^\top \\ \vdots \\ \tau(0)^\top (M^{(s-1)})^\top \end{pmatrix}}_{:=\hat{H}} \eta, \quad (96)$$

where  $f := (\tilde{f}, \tilde{f}^{(1)}, \dots, \tilde{f}^{(s-1)})$ . The matrix  $\hat{M}$  is Hurwitz and therefore, due to the Lyapunov theorem, there exists a symmetric matrix  $P > 0$ ,  $P \in \mathbb{R}^{s \times s}$  that satisfies

$$P\hat{M} + \hat{M}^\top P + Q = 0 \quad (97)$$

for any symmetric matrix  $Q > 0$ ,  $Q \in \mathbb{R}^{s \times s}$ . We fix the positive definite matrix  $Q$  and consider the quadratic Lyapunov function  $V(t) = f(t)^\top P f(t)$ , where  $P$  satisfies (97). The time derivative of  $V$  can be upper bounded by

$$\dot{V} = -f^\top Q f \leq -\lambda_Q |f|^2 \leq -\frac{\lambda_Q}{\lambda^P} f^\top P f \quad (98)$$

$$\leq -\frac{\lambda_Q}{\lambda^P} V, \quad (99)$$

where the minimum eigenvalue of  $Q$  is denoted by  $\lambda_Q$  and the maximum eigenvalue of  $P$  is denoted by  $\lambda^P$ . As a result, this yields the upper bound

$$V(t) \leq V(0) e^{-\frac{\lambda_Q}{\lambda^P} t} \leq \lambda^P \lambda^{\hat{H}^\top \hat{H}} |\eta|^2 e^{-\frac{\lambda_Q}{\lambda^P} t}, \quad (100)$$

where  $\lambda^{\hat{H}^\top \hat{H}}$  denotes the maximum eigenvalue of the matrix  $\hat{H}^\top \hat{H}$ . According to the proof of Prop. 9.1, we may choose  $T_i$  such that

$$1 > \frac{C_2}{c_2} e^{-c_2 T_i}, \quad (101)$$

implying that the basis functions are linearly independent on the interval  $[0, T_i]$  (see (82) and Prop. 9.1 for the definition of the constants  $c_2$  and  $C_2$ ). Linear independence can be used to establish the following lower bound, c.f. Prop. 9.1:

$$\int_0^{T_i} |\tilde{f}|^2 dt = \eta^\top \int_0^{T_i} \tau \tau^\top dt \eta \geq c_3 |\eta|^2, \quad (102)$$

where the constant  $c_3 > 0$  denotes the minimum eigenvalue of the matrix

$$\int_0^{T_i} \tau \tau^\top dt. \quad (103)$$

Without loss of generality it is assumed that  $|a_1| \leq |a_u|$ . Choosing the real number  $T_c > T_i$  implies that the constraint is imposed over the interval  $[0, T_i]$  and therefore

$$\int_0^{T_i} |\tilde{f}|^2 dt \leq T_i |a_u|^2. \quad (104)$$

Combined with the above upper bound on  $|\eta|^2$  it follows that

$$|\eta|^2 \leq \frac{T_i |a_u|^2}{c_3}. \quad (105)$$

This can be used to upper bound the squared magnitude of  $\tilde{f}(t)$ , that is,

$$|\tilde{f}(t)|^2 \leq \frac{1}{\lambda_P} V(t) \leq \frac{\lambda^P \lambda^{\hat{H}^\top \hat{H}} |a_u|^2 T_i}{\lambda_P c_3} e^{-\frac{\lambda_Q}{\lambda^P} t}, \quad (106)$$

where  $\lambda_P$  denotes the minimum eigenvalue of the matrix  $P$ . As a result, we may choose  $T_c$  such that the above upper bound is below  $|a_1|^2$ . This may be achieved by choosing  $T_c > \max\{\hat{t}, T_i\}$ , where

$$\hat{t} := -\frac{\lambda^P}{\lambda_Q} \left( 2 \ln \left( \frac{|a_u|}{|a_1|} \right) + \ln \left( \frac{c_3 \lambda_P}{T_i \lambda^P \lambda^{\hat{H}^\top \hat{H}}} \right) \right) \quad (107)$$

and  $\ln$  refers to the natural logarithm.

## 10 Additional properties

In the following we will discuss some of the properties of the basis functions that fulfill Assumptions A1 and A2.

### 10.1 Conditions imposed on $M_s$

The fact that the basis functions are orthonormal on the interval  $I = (0, T)$  implies

$$\tau(T)\tau(T)^\top - \tau(0)\tau(0)^\top = \int_I \frac{d}{dt}(\tau\tau^\top) dt \quad (108)$$

$$= M_s + M_s^\top. \quad (109)$$

In case the interval  $I$  has infinite measure, that is,  $I = (0, \infty)$ , the above formula reduces naturally to

$$-\tau(0)\tau(0)^\top = M_s + M_s^\top. \quad (110)$$

### 10.2 Bounds on the Euclidean norm

In case the interval  $I$  has infinite measure, that is,  $I = (0, \infty)$ , it holds that

$$|\tau(t)| \leq |\tau(0)|, \quad \forall t \in [0, \infty). \quad (111)$$

This results from

$$\frac{d}{dt}(\tau(t)^\top \tau(t)) = \tau(t)^\top (M_s + M_s^\top) \tau(t) \quad (112)$$

$$= -|\tau(0)^\top \tau(t)|^2 \leq 0, \quad (113)$$

for all  $t \in [0, \infty)$ , which follows from (110).

### 10.3 Finite number of minima and maxima

In the following we will argue that a function that is not everywhere zero and spanned by basis functions fulfilling Assumptions A1 and A2, has a finite number of minima and maxima. The function will be denoted by  $\tilde{f} := \tau^\top \eta$ ,  $\eta \in \mathbb{R}^s$ ,  $\eta \neq 0$ , where the function and the basis functions are defined over the interval  $t \in I = (0, T)$ . If  $I$  has infinite measure, it follows from Sec. 9 that  $\tilde{f}$  takes its maxima and minima within a compact interval, and hence, without loss of generality, we assume in the following that  $I$  has finite measure. Moreover, due to the Caley-Hamilton theorem it holds that

$$\tilde{f}^{(s)}(t) + a_1 \tilde{f}^{(s-1)}(t) + \dots + a_s \tilde{f}(t) = 0, \quad (114)$$

for all  $t \in I$ , where the  $a_k$ ,  $k = 1, 2, \dots, s$  are the coefficients of the characteristic polynomial of the matrix  $M_s$ . According to [6], the time derivative of  $\tilde{f}$  has at most  $s - 1$  zeros on any subinterval of length  $l$ , where

$$\sum_{k=1}^s \frac{a_k l^k}{k!} < 1. \quad (115)$$

This proves readily that  $\tilde{f}$  has a finite number of minima and maxima in the interval  $I$ .

## References

- [1] R. Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, second edition, 1987.
- [2] M. Muehlebach and R. D'Andrea. Parametrized infinite-horizon model predictive control for linear time-invariant systems with input and state constraints. *Proceedings of the American Control Conference*, pages 2669–2674, 2016.
- [3] Y. Nesterov. Squared functional systems and optimization problems. In H. Frenk, K. Roos, T. Terlaky, and S. Zhang, editors, *High Performance Optimization*, chapter 17, pages 405–440. Springer, 2000.
- [4] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, second edition, 2006.
- [5] G. Sansone. *Orthogonal Functions*. Interscience Publishers, 1959.
- [6] S. Yakovenko. On functions and curves defined by ordinary differential equations. *Fields Institute Communications*, 24:497–525, 1999.