Report

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

Author(s): Muehlebach, Michael; D'Andrea, Raffaello

Publication Date: 2018

Permanent Link: https://doi.org/10.3929/ethz-b-000292793

Rights / License: In Copyright - Non-Commercial Use Permitted
On the Approximation of Constrained Linear Quadratic Regulator Problems and their Application to Model Predictive Control

Michael Muehlebach and Raffaello D’Andrea

Abstract

This article is concerned with the approximation of constrained continuous-time linear quadratic regulator problems, which are, for example, encountered in model predictive control. By representing input and state trajectories using basis functions, the underlying infinite-dimensional optimal control problems are reduced to convex finite-dimensional optimization problems that can be solved efficiently. The article quantifies the suboptimality and establishes convergence of the obtained approximations. The results are applied in the context of model predictive control. In particular, it will be shown that the truncation of the prediction horizon can be avoided, leading to recursive feasibility and closed-loop stability guarantees. On a quadruple integrator system, the resulting model predictive control algorithm is shown to outperform state-of-the-art algorithms in terms of average execution time needed to achieve a certain performance.

1 Introduction

Model predictive control (MPC) has become a well-known and widely used control strategy for solving challenging control problems. Unlike many other approaches, MPC addresses input and state constraints in a systematic way. It is based on repeatedly solving an optimal control problem, including the actual state as an initial condition and a prediction of the system’s evolution. This leads naturally to an implicit feedback law, providing robustness against modeling errors and disturbances, [1].

Due to the fact that an optimal control problem has to be solved at each sampling interval, online MPC, where the optimization is solved online, is computationally demanding. Thus, in order to render MPC computationally tractable, the underlying optimal control problem is simplified, typically by discretizing the dynamics and truncating the prediction horizon.

Herein, we propose an alternative approach that relies on a parametrization of input and state trajectories using basis functions. A Galerkin method is used to formulate the dynamics as an equality constraint relating the parameter vectors describing the input and the state. We will show on an example that this parametrized approach leads to different trade-offs between computational effort and achieved closed-loop cost. Moreover, the basis functions can be used to encode a priori knowledge of the system’s dynamics (e.g. different time scales), and even provide a means to retain an infinite prediction horizon. We will show that this leads to an MPC algorithm with inherent recursive feasibility and closed-loop stability guarantees. This contrasts the discrete-time approach where stability is often imposed indirectly using a combination of a terminal cost and a terminal set constraint. The proposed parametrized MPC approach is benchmarked against state-of-the-art discrete-time MPC strategies for underlining the numerical efficiency of the parametrized approach. Simulations of a quadruple integrator system are presented, where it is shown that the average

*Michael Muehlebach and Raffaello D’Andrea are with the Institute for Dynamic Systems and Control, ETH Zurich. The contact author is Michael Muehlebach, michaelmu@ethz.ch. This work was supported by ETH-Grant ETH-48 15-1.
execution time for achieving a given closed-loop cost can be reduced by up to one order of magnitude. This is particularly interesting for embedded systems with fast dynamics, where fast sampling is required but where the available computational power is limited.

**Outline of the paper:** The paper is divided into three parts. The first part is concerned with the approximation of constrained continuous-time linear quadratic regulator problems. The discussion is not restricted to infinite-horizon problems, which are often encountered in MPC, but also includes finite-horizon problems, with terminal costs and/or terminal state constraints. By approximating input and state trajectories using basis functions and exploiting duality, two different finite-dimensional optimization problems are derived, whose optimal costs yield upper and lower bounds on the cost of the underlying infinite-dimensional problem. By increasing the basis functions’ complexity (which might correspond to the polynomial order, for example), the resulting optimal costs are found to yield monotonic sequences approximating the underlying optimal control problem from above and below. This makes it possible to quantify the suboptimality of the obtained approximation, and, as we will show in the remainder, also bounds the $L^2$-distance from the optimal trajectories. Moreover, conditions guaranteeing convergence to the underlying optimal control problem will be established, thereby providing a theoretical justification of the proposed approach.

The second part deals with the application of the proposed approach to MPC. By choosing exponentially decaying basis functions, a truncation of the prediction horizon can be avoided, and as a result, closed-loop stability and recursive feasibility are shown to be inherent to the resulting parametrized MPC formulation.

The third part presents a simulation example that should highlight the different trade-offs between computation and achieved closed-loop cost of the proposed approach. A benchmark with state-of-the-art MPC algorithms is included.

Details regarding the numerical solution of the resulting finite-dimensional optimization problems, as well as technical results that do not really contribute to the basic understanding of the proposed approach are provided in the supplementary material, [2].

**Related work:** For our analysis of the approximations to constrained continuous-time linear quadratic regulator problems we adopt a similar point of view than presented in [3], where (weighted) Sobolev spaces are introduced as state space and (weighted) Lebesgue spaces are introduced as control spaces. By doing so, the author establishes a Pontryagin type of Maximum Principle for linear infinite-horizon optimal control problems. These problems have proven to be difficult to analyze as the standard transversality conditions cannot be extended directly to the infinite-horizon case, see for example [4] or [5, Ch. 3.7, Ch. 6.5].

In [6], polynomials are used for approximating continuous linear programs. Duality is exploited for constructing approximations yielding upper and lower bounds on the underlying continuous linear program. The resulting semi-infinite constraints are reformulated using sum-of-squares techniques yielding semidefinite programs. Our approach is similar in the sense that the lower bounds are also derived using duality. However, the optimal control problem that we consider cannot be cast as a continuous linear program, and as a result, our approach for constructing the lower bounds differs significantly. Moreover, we do not restrict ourselves to polynomial basis functions, and treat equality constraints in the form of linear ordinary differential equations by means of a Galerkin approach.

The optimal control problems that are discussed in the following can also be approximated by “standard” numerical optimization approaches such as shooting or collocation methods, see for example [7] or [8] and references therein. However, these approaches are typically tailored to nonlinear problems and as such, do not yield guarantees on the approximation quality in general. Moreover, these approaches tend to be computationally expensive.

Constrained linear quadratic infinite-horizon problems are often encountered in MPC, which is the main motivation for our work. In contrast to our formulation, the “standard” MPC approach relies mostly on a discrete-time finite-horizon formulation, [1]. In order to guarantee closed-loop stability, terminal cost and terminal state constraints are often needed, [9]. Moreover, as remarked in [10], truncating the prediction horizon leads to a discrepancy between the closed-loop performance objective and the finite-horizon open-

---

1Continuous linear programs are related to the constrained linear quadratic regulator problems considered herein by the fact that the constraints occurring in continuous linear programs could be used to encode linear dynamics.
loop performance objective that is minimized at every time step. An alternative approach is proposed in [11, Ch. 3, Ch. 6] and [12], where the finite differences (in the discrete-time setting), respectively the time-derivatives (in the continuous-time setting) of the control inputs are described with so-called Laguerre or Kauz basis functions. An analytical expression for the corresponding state trajectory as a function of the parametrized inputs is derived, eliminating thereby the state variables in the resulting optimization problem. Still, a finite prediction horizon is retained. Other input parametrizations in the context of MPC have been suggested in [13], [14], and [15], and depart from a dual-mode MPC-formulation. We use similar basis functions, but parametrize both, input and state trajectories and use the constrained linear quadratic regulator problem as a starting point. We show that the truncation of the prediction horizon can be avoided (without referring to a dual-mode MPC formulation) due to the decaying nature of the basis functions, resulting in inherent recursive feasibility and closed-loop stability guarantees. If the basis functions are well-chosen, only few basis functions are needed for obtaining a relatively good approximation. This leads to optimization problems with relatively few optimization variables that can be solved efficiently. Compared to state-of-the-art MPC solvers [16] and [17], the average execution time can be reduced by up to one order of magnitude for the quadruple integrator system presented in Sec. 4, without degrading closed-loop performance. Compared to the approach presented in [12], we parametrize the control inputs directly, which avoids lifting the system, and consequently reduces the number of variables. In our approach, the dynamics are represented by linear equality constraints, which may or may not be eliminated.

The suboptimality of the “standard” MPC without terminal constraints and terminal cost with respect to the underlying infinite-horizon problem is discussed and quantified in [18, Ch. 6], [19], [20], and [21]. These approaches are based on approximate dynamic programming and typically involve finding control Lyapunov functions. The approach presented in the following is constructive in the sense that the suboptimality can be quantified by solving two finite-dimensional convex optimization problems.

Due to the fact that an infinite prediction horizon can be retained, the optimal open-loop cost of our approach always acts as an upper bound on the achieved closed-loop cost. Similarly, it is shown in [22] that the infinite-horizon closed-loop cost can be upper bounded by a corresponding discrete-time receding-horizon scheme including a terminal cost and a terminal state constraint. This leads naturally to stability guarantees and constraint satisfaction for all times (even between the sampling instants).

The authors from [23] and [24] propose to solve the discrete-time infinite-horizon linear quadratic regulator problem directly, by means of an operator splitting technique in [23] or by successively extending the prediction horizon of a finite-horizon approximation in [24]. These schemes require successive solutions of the discrete-time finite-horizon problem with varying prediction horizons. This contrasts the proposed approach, where the number of optimization variables is fixed and can be adjusted for trading off the computational complexity with the approximation quality.

Exploiting a parametrization of the input for reducing the computational complexity of MPC has already been explored by previous work, see for example [25], [26], [27], and [28]. In [25], the implications regarding closed-loop stability and recursive feasibility of an input parametrization are investigated in the context of nonlinear MPC. The input parametrization is required to be invariant to time shifts, which parallels the approach presented in the following. The formulation is based on a finite prediction horizon, and a terminal equality constraint (if the prediction horizon remains fixed) or a contraction property (if the prediction horizon enters the optimization) is required for guaranteeing closed-loop stability. In case the prediction horizon is fixed, the input parametrization is assumed to be translatable, see [25, Def. 1.5], which results either in a standard sample-and-hold parametrization, a nonlinear input parametrization, or requires additional assumptions compared to the parametrization presented in the following. In case the contraction property in combination with a varying prediction horizon is used for guaranteeing closed-loop stability, a nonlinear and in general non-convex optimization problem is obtained. We show that our parametrization evolves naturally from a time-shift requirement related to closed-loop stability and the fact that the open-loop trajectories should achieve a finite cost. In contrast to [25], the infinite-horizon formulation avoids the use of additional equality constraints for guaranteeing closed-loop stability and leads to a finite-dimensional
convex optimization problem with a quadratic cost and linear constraints.\(^2\)

In [29], multiresolution analysis is used for parametrizing the input trajectory. However, the approach is mainly applicable to open-loop stable systems, where the impulse response is assumed to be negligible after a certain time horizon. For dealing with unstable systems, the proposed approach would require additional terminal constraints on the unstable modes. Similarly, the authors from [30] apply the wavelet transformation for simplifying the control laws obtained with explicit model predictive control. They show that the resulting simplified control law is everywhere feasible and quantify the suboptimality.

Preliminary results to the ones presented herein appeared in the conference papers [31] and [32]. In [31], the application of our strategy to MPC and the implications regarding closed-loop stability and recursive feasibility are discussed in detail. In [32], the approximation quality with respect to the underlying optimal control problem is discussed. The results from [32] are extended herein by deriving a bound on the approximation error of the resulting optimal input and state trajectories in the \(L^2\)-norm. We state conditions guaranteeing that our approximations will actually converge to the solutions of the underlying optimal control problem, when increasing the basis functions complexity. Moreover, compared to [32], the results presented herein are derived using a different approach (for example not relying on conjugate functions), which we think is more accessible. We also present a benchmark against two state-of-the-art MPC solvers.

2 Part I: Theoretical foundation

2.1 Problem formulation

In a first step, we present and analyze approximations to the following optimal control problem

\[
J_\infty := \min \frac{1}{2} (||x||^2 + ||u||^2) + \psi(x_T) \\
\text{s.t. } \dot{x} = Ax + Bu, x(0) = x_0, x(T) = x_T, \\
C_xx + C_xu \leq b, x_T \in \mathcal{X}, \\
x \in L^2_n, u \in L^2_m,
\]

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

\[
x \rightarrow ||x||, \quad ||x||^2 := \int_I x^T x \, dt,
\]

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

\[
f = g, \quad f \leq g,
\]

when we mean \(f(t) = g(t)\), respectively \(f(t) \leq g(t)\) for all \(t \in I\) almost everywhere, with \(f, g \in L^2_q\), or equivalently,

\[
\int_I \delta p^T (f - g) dt = 0, \quad \int_I \delta \hat{p}^T (f - g) dt \leq 0,
\]

\(^2\)In the following, a continuous-time point of view is adopted, resulting in semi-infinite constraints due to the fact that the constraints are imposed over compact time intervals. These semi-infinite constraints can be avoided in a discrete-time setting. In [2, Sec. 4], a computationally efficient approach to deal with these semi-infinite constraints is presented.
for all smooth compactly supported test functions \( \delta p \) and \( \delta \hat{p} \), with \( \delta \hat{p}(t) \geq 0 \) for all \( t \in I \). The weak derivative of \( x \) is denoted by \( \dot{x} \). To simplify notation we abbreviate the domain of the objective function by

\[
X := L^2_n \times L^2_m \times \mathbb{R}^n. \tag{5}
\]

We assume throughout the article that the constraints in (1) are nonempty, i.e. there exist 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 MPC.

**Discussion of the assumptions:** The assumption of linear time-invariant dynamics will be important in the following, as it leads to approximate solutions of (1) that fulfill the equations of motion exactly. The assumption of a quadratic cost could be relaxed to include strongly convex running costs; we will comment on such extensions in due course. However, these extensions will generally increase the computational complexity needed for obtaining (approximate) numerical solutions. From a practical point of view, a quadratic running cost often represents a good compromise between generality and computational tractability.

The interval \( I \) is not restricted to have finite measure. The subsequent analysis remains valid even if \( T \to \infty \), with \( \psi = 0 \), \( X = \{0\} \), and \( x_T = \lim_{t \to \infty} x(t) = 0 \). Furthermore, the more general cost

\[
\frac{1}{2} \int_I x^T Q x + u^T R u \, dt \tag{6}
\]

can be cast into (1) by means of a linear coordinate transformation, provided that the matrices \( Q \in \mathbb{R}^{n \times n} \) and \( R \in \mathbb{R}^{m \times m} \) are positive definite. In addition, the \( L^2_q \) space can be replaced by the weighted Lebesgue space with squared norm

\[
\int_I x^T x e^{-\alpha t} \, dt, \tag{7}
\]

where \( \alpha > 0 \) is constant, without changing the subsequent analysis. The assumption that \( b \) is constant can be relaxed, for example by requiring \( b \) to be square integrable.

The strong convexity of the running cost and the terminal cost \( \psi \) is important for guaranteeing uniqueness of the corresponding minimizer. This will be established in Sec. 2.3, where we also argue that the minimum in (1) is indeed attained.

### 2.2 Finite-dimensional approximations of (1)

#### 2.2.1 Parametrization with basis functions

We will parametrize input and state trajectories using basis functions, that is,

\[
\hat{x}(t) = (I_n \otimes \tau^x(t))^T \eta_x, \quad \hat{u}(t) = (I_m \otimes \tau^u(t))^T \eta_u,
\]

where \( \eta_x \in \mathbb{R}^{ns} \) and \( \eta_u \in \mathbb{R}^{ms} \) are the parameter vectors, \( \tau^x(t) := (\tau_1(t), \tau_2(t), \ldots, \tau_s(t)) \in \mathbb{R}^s \) 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 it is clear from context, and we will indicate vectors as \( n \)-tuples, where the dimension and stacking can be inferred from context. The basis functions are required to satisfy the following assumptions:

1. \( \tau_i \in L^2_1, i = 1, 2, \ldots, s \) are linearly independent and orthonormal with respect to the standard \( L^2_1 \)-scalar product.

---

\( ^3 \)The equations of motion imply that \( \dot{x} \in L^2_2 \), 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 \( t \in [0, T] \).
A2) The basis functions fulfill $\dot{\tau}^s(t) = M_s^s \tau^s(t)$ for all $t \in I$, for some $M_s \in \mathbb{R}^{s \times s}$. 

Note that in case $I$ has infinite measure, $M_s$ is required to have strictly negative eigenvalues. This is a natural requirement, since a feasible state trajectory $x$ in (1) is guaranteed to decay due to the fact that it is required to be square integrable and to have a weak derivative in $L_2^n$, see [33, Cor. 8.9].

The assumption of linearly independent basis functions is necessary for the approximations to be unique. The assumption of orthonormal basis functions is without loss of generality, since orthonormal basis functions can be constructed from linearly independent ones via the Gram-Schmidt procedure. Assumption A2 is more restrictive. Well-known examples fulfilling Assumption A2 are sinusoids or polynomials. Assumption A2 implies, however, that the basis functions are able to capture an arbitrary time-shift, and can be used to conclude that the equations of motion are (depending on the formulation) fulfilled exactly by the parametrized input and state trajectories, [31]. If the basis functions are assumed to be continuously differentiable, the converse is also true, as we illustrate next. Thus, we set forth that the basis functions should be able to capture time-shifts, that is, for every vector $\eta \in \mathbb{R}^s$ and every time-shift $T_s \geq 0$ there exists a vector $\hat{\eta}(\eta, T_s)$ such that

$$
\tau(t - T_s)^T \eta = \tau(t)^T \hat{\eta}(\eta, T_s), \quad \forall t \in (T_s, T). \tag{8}
$$

We will now show that this implies that the basis functions must fulfill Assumption A2. In order to do so, we take the derivative with respect to $T_s$ and evaluate the resulting expression for $T_s \rightarrow 0$, leading to

$$
-\tau^T \eta = \tau^T \frac{\partial \hat{\eta}}{\partial T_s} \bigg|_{T_s \downarrow 0}, \quad \forall t \in I. \tag{9}
$$

We may choose the canonical unit vectors for $\eta$, which readily implies that $\dot{\tau}$ must be a linear combination of the vector $\tau$. This concludes that any set of continuously differentiable basis functions that can capture an arbitrary time-shift must fulfill Assumption A2. In the context of MPC, Assumption A2 is used to guarantee recursive feasibility and closed loop stability and will therefore be of paramount importance. In a discrete-time finite-horizon setting, the importance of the time-shift property regarding closed-loop stability has already been emphasized in [25], resulting in similar requirements on the basis functions, see for example [25, Def. 1.8]. A similar parametrization in the discrete-time context has also been suggested in [14], for example.

In the infinite-horizon case, that is for $T \rightarrow \infty$, examples fulfilling Assumptions A1 and A2 are given by exponentially decaying polynomials, or exponentially decaying sinusoids. In the case of polynomials, this leads to so-called Laguerre functions, which are given by

$$
\tau_i(t) = \sqrt{2\nu} L_i(2\nu t) e^{-\nu t}, \tag{10}
$$

where $L_i$ denotes the $i$th Laguerre polynomial, $i = 1, 2, \ldots, s$, and $\nu > 0$ is the rate of the exponential decay. The corresponding matrix $M_s$ has then the form

$$
M_s = \begin{pmatrix}
-\nu & 0 & 0 & \ldots \\
-2\nu & -\nu & 0 & \ldots \\
-2\nu & -2\nu & -\nu & \ldots \\
\vdots & \vdots & \ddots & \ddots
\end{pmatrix}. \tag{11}
$$

These basis functions will be used to approximate infinite-horizon problems arising in the context of MPC, see Sec. 3 and Sec. 4.

We will denote the finite-dimensional subspace spanned by the first $s$ basis functions as $X^s$,

$$
X^s := \{(x, u, x_T) \in X \mid \eta_x \in \mathbb{R}^{n^s}, \eta_u \in \mathbb{R}^{m^s}, x = (I_n \otimes \tau^s)^T \eta_x, u = (I_m \otimes \tau^s)^T \eta_u\}. \tag{12}
$$
The fact that $X^s$ is finite-dimensional can be used to conclude that $X^s$ is complete, i.e. that every Cauchy sequence in $X^s$ converges and has its limit in $X^s$. As a result, it follows that $X^s$ is a closed subspace of $X$.

This will become important for arguing that the minima of the resulting optimization problems are indeed attained. In addition, the following straightforward, but important relation

$$X^s \subset X^{s+1}$$

(13)

holds for all integers $s > 0$.

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 $\pi^{qs}: L^2_q \to \mathbb{R}^{qs}$, defined as

$$x \to \int_I (I_q \otimes \tau^s)x dt,$$

(14)

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

$$(x, u, x^T) \to (\pi^{ns}(x), \pi^{ms}(u), x^T).$$

(15)

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 $(\pi^{qs})^*: \mathbb{R}^{qs} \to L^2_q$ is given by

$$\eta \to (I_q \otimes \tau)^T \eta,$$

(16)

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 $(\pi^s)^*: \mathbb{R}^{ns} \times \mathbb{R}^{ms} \times \mathbb{R}^n \to X$ as

$$(\eta_x, \eta_u, x^T) \to ((\pi^{ns})^*(\eta_x), (\pi^{ms})^*(\eta_u), x^T).$$

(17)

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

### 2.2.2 Approximation of the constraints

In the following we seek to approximate the constraint

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

(18)

characterizing all square integrable functions $x$ and $u$, satisfying the inequality constraints. The first approximation is obtained by restricting the trajectories $x$ and $u$ to be spanned by the first $s$ basis functions, i.e.

$$\mathcal{C}^s_U := \{ (\tilde{x}, \tilde{u}, x^T) \in X^s \mid C_x \tilde{x} + C_u \tilde{u} \leq b, x^T \in \mathcal{X} \}.$$

In other words, $\mathcal{C}^s_U$ is defined as the intersection of $\mathcal{C}$ with $X^s$.

The set (18) can be reformulated using a variational formulation, leading to

$$\mathcal{C} = \{ (x, u, x^T) \in X \mid \int_I \delta p^T (-C_x x - C_u u + b) dt \geq 0$$

$$\forall \delta p \in L^2_{nc}: \delta p \geq 0, \left| \int_I \delta p^T b dt \right| < \infty; x^T \in \mathcal{X} \},$$

(19)

\[\text{In a metric space a set is closed if and only if it is sequentially closed.}\]
where in case $I$ has finite measure the Lebesgue integral of $\delta p^T b$ over $I$ is guaranteed to be finite. In the light of \eqref{eq:approximation2}, a second approximation is thus naturally obtained by restricting the test functions $\delta \tilde{p}$ to be spanned by the first $s$ basis functions, that is,

\[ C_s^L := \{ (x, u, x_T) \in X \mid \int_I \delta \tilde{p}^T (-C_x x + C_u u + b) dt \geq 0 \} \]

\begin{equation}
\forall \delta \tilde{p} = (I_{nc} \otimes \tau)^T \delta \eta_p : \delta \tilde{p} \geq 0, \delta \eta_p \in \mathbb{R}^{nc}; x_T \in \mathcal{X}. \quad (20)
\end{equation}

The sets $C_s^U$, $C_s^L$, and $C$ have the following properties:

B1) the sets $C_s^U$, $C_s^L$, and $C$ are closed and convex.

B2) the sets $\pi^s(C_s^U)$, $\pi^s(C_s^L)$ are closed and convex.

B3) $C_s^U \subset C_s^{s+1} \subset C$.

B4) $C_s^L \supset C_s^{s+1} \supset C$.

B5) $(\pi^s)^* \pi^s(C_s^U) \subset C_s^U$, $(\pi^s)^* \pi^s(C_s^L) \subset C_s^L$.

Property B1 ensures that the resulting optimizations over $C_s^U$, $C_s^L$, and $C$ will be convex, and that the corresponding minima will be attained. The sets $C_s^U$ and $C_s^L$ are represented as subsets of the Euclidean space through the map $\pi^s$. As a result, Property B2 ensures that the corresponding (finite-dimensional) optimization problems will be convex and that the corresponding minima will be attained. Property B3 is used to obtain a monotonically decreasing sequence bounding $J_\infty$ from above, whereas Property B4 is used to construct a monotonically increasing sequence bounding $J_\infty$ from below. Property B5 will guarantee consistency of the corresponding finite-dimensional optimization problems.

The proof of Properties B1-B5 can be found in [2, Sec. 5]. A schematic illustration of the sets $C_s^U$, $C_s^L$, and $C_s^U$ is shown in Fig. 1.

It follows from the property of the set $C_s^U$ being contained in $C_s^{s+1}$, and $C_s^{s+1}$ being contained in $C_s^L$, that the limits

\[ \lim_{s \to \infty} C_s^U = \bigcup_{s=1}^{\infty} C_s^U, \quad \lim_{s \to \infty} C_s^L = \bigcap_{s=1}^{\infty} C_s^L \quad (21) \]

exist, [34, p. 18, p. 21]. Next, we will provide conditions asserting that the two limits agree. To that extent we introduce the following definition: The set $\mathcal{A}$ is an algebra, if it is closed under addition, scalar
multiplication, and multiplication, that is,
\[ f + g \in A, \quad fg \in A, \quad cf \in A, \]
for any \( f, g \in A, c \in \mathbb{R} \), [35, p. 161].

**Lemma 2.1** 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
\[ \lim_{s \to \infty} C^s_U = \lim_{s \to \infty} C^s_L. \]

**Proof** We consider the case where \( I \) has finite measure. The proof for the case where \( I \) has infinite measure is given in [2, Sec. 7]. We claim that \( \lim_{s \to \infty} C^s_U \supset \lim_{s \to \infty} C^s_L \). We prove the claim by contradiction. Let \((x, u, x_T) \in \lim_{s \to \infty} C^s_L \) be such that there exists an open set \( U \) with
\[ C_{kk}(t) + C_{uk}u(t) > b_k, \quad \forall t \in U \quad \text{a.e.}, \]
for some \( k \in \{1, 2, \ldots, n_c\} \), where \( C_{kk} \) and \( C_{uk} \) denotes the \( k \)th row of \( C_x \), respectively \( C_u \). Thus, it holds that
\[ \int_I \delta v(-C_{kk}x - C_{uk}u + b_k)dt < 0 \]
for all smooth test functions \( \delta v : I \to \mathbb{R} \), vanishing outside \( U \), with \( \delta v(t) > 0 \forall t \in U \). We pick one of these test functions and denote it by \( \delta \hat{p} \). Due to the fact that the basis functions are dense in the set of smooth functions with compact support, there exists a sequence \( \sqrt{\delta \hat{p}_i} \), which converges uniformly to \( \sqrt{\delta p} \), that is, for any given \( \epsilon > 0 \) there exists an integer \( N > 0 \) such that
\[ ||\sqrt{\delta \hat{p}_i} - \sqrt{\delta p}||_{\infty} < \epsilon, \quad \forall i > N, \]
where \( || \cdot ||_{\infty} \) denotes the supremum norm. From the assumption that the basis functions form an algebra that is closed under multiplication, we can infer that \( \delta \hat{p}_i \) lies likewise in the span of the basis functions. Moreover, we have that
\[ ||\delta \hat{p}_i - \delta p||_{\infty} = ||(\sqrt{\delta \hat{p}_i} - \sqrt{\delta p})(\sqrt{\delta \hat{p}_i} + \sqrt{\delta p})||_{\infty} \]
\[ \leq ||\sqrt{\delta \hat{p}_i} - \sqrt{\delta p}||_{\infty}||\sqrt{\delta \hat{p}_i} + \sqrt{\delta p}||_{\infty} \]
\[ < \epsilon ||\sqrt{\delta \hat{p}_i} - \sqrt{\delta p} + 2\sqrt{\delta p}||_{\infty} \]
\[ < \epsilon(\epsilon + 2)||\sqrt{\delta p}||_{\infty} \leq C_1 \epsilon, \]
for all integers \( i > N \), where \( C_1 > 0 \) is constant (for \( \epsilon \) sufficiently small). By assumption, \((x, u, x_T) \in C^s_L \), for all integers \( s > 0 \), and therefore
\[ \int_I \delta \hat{p}_i(-C_{kk}x - C_{uk}u + b)dt \geq 0, \]
for all integers \( i > 0 \). However, the above integral can be rewritten as
\[ 0 \leq \int_I \delta p(-C_{kk}x - C_{uk} + b)dt \]
\[ + \int_I (\delta \hat{p}_i - \delta p)^T(-C_{kk}x - C_{uk}u + b)dt, \]
where the last term can be bounded by (using Hölder’s inequality twice)
\[ \epsilon C_1 \int_I |C_{kk}x + C_{uk}u - b|dt \leq \epsilon C_1 ||C_{kk}x + C_{uk}u - b||_2 \sqrt{T}, \]
for all integers \( i > N \). The fact that the expression (32) converges to zero as \( i \to \infty \) leads to a contradiction with (24). It follows therefore that \( \lim_{s \to \infty} C^s_U \supset \lim_{s \to \infty} C^s_L \), which, combined with \( C^s_U \subset C^s_L \) for all integers \( s > 0 \), leads to the desired conclusion.
2.2.3 Approximation of the dynamics

We define the set
\[ \mathcal{D} := \{(x, u, x_T) \in X | \dot{x} = Ax + Bu, x(0) = x_0, x(T) = x_T\}, \]  
(34)
containing all trajectories \( x \) and \( u \) fulfilling the equations of motion in a weak sense. We obtain a first approximation to the set \( \mathcal{D} \) by restricting \( x \) and \( u \) to be spanned by the first \( s \) basis functions, that is,
\[ \mathcal{D}_U^s := \mathcal{D} \cap X^s. \]  
(35)
It was shown in [31] that the linearity of the dynamics implies that the set \( \mathcal{D}_U^s \) can be rewritten as the elements \( (\hat{x}, \hat{u}, x_T) \in X^s \), satisfying \( \hat{x}(0) = x_0, \hat{x}(T) = x_T \), and
\[ \int_I \delta \hat{p}^T (\dot{\hat{x}} - Ax - B\hat{u}) dt = 0, \]  
(36)
for all variations \( \delta \hat{p} \) that are spanned by the first \( s \) basis functions. In particular, (36) reduces to a linear equation in the coefficient vectors \( \eta_x \) and \( \eta_u \) defining the trajectories \( \hat{x} \) and \( \hat{u} \) compatible with the equations of motion.

In order to obtain a second approximation, we reformulate the dynamics in terms of the variational equality
\[ \int_I \delta p^T (\dot{x} - Ax - Bu) dt + \delta p(0)^T (x(0) - x_0) + \delta p(T)^T (x_T - x(T)) = 0, \]  
(37)
for all test functions \( \delta p \in H_n \), where \( H_n \) denotes the set of functions in \( L^2_n \) having a weak derivative in \( L^2_n \). As remarked earlier, \( \delta p \) has therefore a unique absolutely continuous representative to which we refer when writing \( \delta p(0), \delta p(T) \). In case \( I \) has infinite measure, the above equation reduces naturally to
\[ \int_I \delta p^T (\dot{x} - Ax - Bu) dt + \delta p(0)^T (x(0) - x_0) = 0, \]  
for all test functions \( \delta p \in H_n \). The formulation (37) is equivalent to the one in (34), which is implied by the fundamental lemma of the calculus of variations [36, p. 18]. Applying integration by parts results in
\[ -\int_I \delta p^T x dt - \int_I \delta p^T (Ax + Bu) dt - \delta p(0)^T x_0 + \delta p(T)^T x_T = 0, \quad \forall \delta p \in H_n, \]  
(38)
and is often referred to as weak formulation of the dynamics. The above equation is well-defined for all \( x \in L^2_n \) (and equivalent to (34)). Therefore, by restricting the variations in (38) to be spanned by the first \( s \) basis functions, we obtain
\[ \mathcal{D}_L^s := \{(x, u, x_T) \in X | \delta \hat{p} = (I_n \otimes \tau)^T \delta \eta_p, \]  
\[ -\int_I \delta \hat{p}^T x dt - \int_I \delta \hat{p}^T (Ax + Bu) dt - \delta \hat{p}(0)^T x_0 + \delta \hat{p}(T)^T x_T = 0, \forall \delta \eta_p \in \mathbb{R}^{ns}\}, \]  
(39)
as an approximation to \( \mathcal{D} \). Note that while the formulation (34) (and likewise (38)) implies \( x \in H_n \), the set \( \mathcal{D}_L^s \) contains also elements \( x \in L^2_n \) that do not necessarily have a weak derivative in \( L^2_n \). The use of the weak formulation for the definition of \( \mathcal{D}_L^s \) is motivated by the fact that the resulting set \( \mathcal{D}_L^s \) is closed (see below), which is important to ensure that the minimum is attained, when optimizing over \( \mathcal{D}_L^s \).

The sets \( \mathcal{D}_U^s, \mathcal{D}_L^s, \) and \( \mathcal{D} \) 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 \( \pi^*(\mathcal{D}_U^s) \) and \( \pi^*(\mathcal{D}_L^s) \) are closed and convex.

10
Moreover, the fact that in Sec. 2.2.3 we obtain the two auxiliary problems the infimum in (41) is attained and that the corresponding minimizer are unique for all integers $s \in \mathbb{D}$ is attained and that the corresponding minimizer is unique for $s$. We make the assumption that there exists an integer $\tilde{s}$.

The above properties are in complete analogy to the previous section, and will be used to draw analogous conclusions. In particular, Properties C1 and C2 ensure that the resulting optimization problems will have unique well-defined minimizers. Properties C3 and C4 will lead to a monotonically decreasing sequence bounding $J_\infty$ above, respectively to a monotonically increasing sequence bounding $\tilde{J}_\infty$ below. Property C5 will guarantee consistency. A proof of Properties C1-C5 can be found in [2, Sec. 6].

2.3 Resulting optimization problems

Using the definitions of the previous section we can rewrite (1) as

$$J_\infty = \min \|x\|^2 + \|u\|^2 + \psi(x_T)$$

s.t. $(x, u, x_T) \in \mathcal{C} \cap \mathcal{D}$.  \hspace{1cm} (40)

By assumption, there exists a feasible trajectory satisfying the dynamics and the constraints. Therefore the set $\mathcal{C} \cap \mathcal{D}$ is nonempty and the objective is bounded above. As a consequence, (40) reduces to an optimization over a closed convex and bounded set in the Banach space $X$ (the set $\mathcal{C} \cap \mathcal{D}$ is closed). Thus, the minimum in (40) is attained and due to the strong convexity of the objective function the corresponding minimizer $(x, u, x_T) \in X$ is unique, [37, Thm. 26, p. 93].

By combining the approximation of the constraints in Sec. 2.2.2 and the approximation of the dynamics in Sec. 2.2.3 we obtain the two auxiliary problems

$$J_s := \inf ||\dot{x}||^2 + ||\dot{u}||^2 + \psi(x_T)$$

s.t. $(\dot{x}, \dot{u}, x_T) \in \mathcal{C}_U^s \cap \mathcal{D}_U^s$, \hspace{1cm} (41)

and

$$\tilde{J}_s := \min ||\dot{x}||^2 + ||\dot{u}||^2 + \psi(x_T)$$

s.t. $(\dot{x}, \dot{u}, x_T) \in \mathcal{C}_L^s \cap \mathcal{D}_L^s$. \hspace{1cm} (42)

We make the assumption that there exists an integer $s_0 > 0$ large enough such that $\mathcal{C}_U^{s_0} \cap \mathcal{D}_U^{s_0}$ is nonempty. The closedness and convexity of the sets $\mathcal{C}_U^s$, $\mathcal{D}_U^s$, $\mathcal{C}_L^s$, and $\mathcal{D}_L^s$ leads to the conclusion that the infimum in (41) is attained and that the corresponding minimizer is unique for $s = s_0$. From the fact that $\mathcal{D}_U^s$ is contained in $\mathcal{D}_U^{s+1}$ and $\mathcal{C}_U^s$ is contained in $\mathcal{C}_U^{s+1}$ it follows that $J_{s+1} \leq J_s$ for all integers $s \geq s_0$. This implies that the infimum in (41) is attained and that the corresponding minimizer are unique for all integers $s \geq s_0$. Moreover, the fact that $\mathcal{D}_U^s \subset \mathcal{D}$ and $\mathcal{C}_U^s \subset \mathcal{C}$ implies that $J_s \geq J_\infty$ for all integers $s \geq s_0$. Similar arguments show that $\tilde{J}_s \leq \tilde{J}_{s+1} \leq \tilde{J}_\infty$ for all integers $s > 0$, that the minimum in (42) is indeed attained, and that the corresponding minimizers are unique for all integers $s > 0$. The results are summarized with the following lemma.

**Lemma 2.2** Let the sets $\mathcal{C} \cap \mathcal{D}$ and $\mathcal{C}_U^{s_0} \cap \mathcal{D}_U^{s_0}$ be nonempty for some integer $s_0 > 0$. Then the optimization problems (40), (41), and (42) are well defined and the corresponding minima are attained and are unique. Moreover, the costs $J_s$ form a monotonically decreasing sequence bounding $J_\infty$ above for all integers $s \geq s_0$, whereas the costs $\tilde{J}_s$ form a monotonically increasing sequence bounding $\tilde{J}_\infty$ below for all integers $s > 0$.
By definition of the constraints $C^s_t$ and $D^s_t$, the minimizer of (41) (for $s \geq s_0$) is required to be an element of $X^s$. Consequently, the problem (41) is equivalent to

$$J_s = \inf |\eta|_x^2 + |\eta|_u^2 + \psi(x_T)$$

s.t. $(\eta_x, \eta_u, x_T) \in \pi^s(C^s_t) \cap \pi^s(D^s_t),$ (43)

which corresponds to a convex finite-dimensional optimization problem. Note that the orthonormality of the basis functions can be used to conclude $|\eta|_x^2 = ||\tilde{x}||^2$ and likewise $|\eta|_u^2 = ||\tilde{u}||^2$. Similarly, the minimizer $(x, u, x_T) \in X$ of (42) is guaranteed to lie in $X^s$. This can be shown by contradiction: We assume therefore $(x, u, x_T) \in X \setminus X^s$. We construct $\tilde{x} := (\pi^{ns})^*(x)$, and $\tilde{u} := (\pi^{ns})^*(u)$. As a consequence of Properties B5 and C5 it follows that $(\tilde{x}, \tilde{u}, x_T) \in C^s_t \cap D^s_t$. Moreover, by orthonormality of the basis functions it holds that

$$\int T \tilde{x}^T(x - \tilde{x})dt = 0, \quad \int T \tilde{u}^T(u - \tilde{u})dt = 0,$$ (44)

which leads to

$$||x||^2 = ||x - \tilde{x}||^2 + ||\tilde{x}||^2, \quad ||u||^2 = ||u - \tilde{u}||^2 + ||\tilde{u}||^2.$$ (45)

This implies that $\tilde{x}$, $\tilde{u}$, and $x_T$ achieve a cost that is below $\tilde{J}_s$, contradicting the fact that $(x, u, x_T)$ are the minimizer of (42).

This shows that the convex finite-dimensional problem

$$\tilde{J}_s = \min |\eta|_x^2 + |\eta|_u^2 + \psi(x_T)$$

s.t. $(\eta_x, \eta_u, x_T) \in \pi^s(C^s_t) \cap \pi^s(D^s_t),$ (46)

is equivalent to (42) in the sense that its (unique) minimizer $(\eta_x, \eta_u, x_T)$ is related to the minimizer $(x, u, x_T) \in X$ of (42) by $x = (\pi^{ns})^*(\eta_x)$, $u = (\pi^{ns})^*(\eta_u)$, $x_T = x_T$, and achieves the same cost. By virtue of Lemma 2.2 we therefore conclude

**Theorem 2.3** Let the sets $C \cap D$ and $C^0_t \cap D^{s_0}_t$ be nonempty for some integer $s_0 > 0$. Then, the optimization problems (40), (43), and (46) are well-defined and the corresponding minima are attained and are unique. Moreover, the costs $J_s$ form a monotonically decreasing sequence bounding $J_{\infty}$ above for all integers $s \geq s_0$, whereas the costs $\tilde{J}_s$ form a monotonically increasing sequence bounding $J_{\infty}$ below for all integers $s > 0$.

From the fact that $J_s$ and $\tilde{J}_s$ form monotonically increasing, respectively monotonically decreasing sequences it follows at once that

$$\lim_{s \to \infty} J_s, \quad \lim_{s \to \infty} \tilde{J}_s$$ (47)

exist. We will argue that not only the optimal cost, but also the optimal trajectories converge (strongly). In addition, this will provide a means to quantify the $L^2$-error of the input and state trajectories with respect to the trajectories corresponding to (1).

**Proposition 2.4** Let the assumptions of Thm. 2.3 be fulfilled. Let the optimal trajectories of (41), respectively (43) be denoted by $\tilde{x}^s$, $\tilde{u}^s$, $x_Ts$, and the optimal trajectories of (1) by $x$, $u$, $x_T$. It is further assumed that $\psi$ is $\mu$-strongly convex. Then, $\tilde{x}^s$, $\tilde{u}^s$, and $x_Ts$ converge strongly, and for all integers $s \geq s_0$ it holds that

$$||\tilde{x}^s - x||^2 + ||\tilde{u}^s - u||^2 + \frac{\mu}{2}|x_Ts - x_T|^2 \leq 2(J_s - J_{\infty}).$$ (48)

**Proof** We first note that the strong convexity of $\psi$ implies that for any real numbers $x_1$ and $x_2$ it holds that

$$\psi\left(\frac{1}{2}(x_1 + x_2)\right) \leq \frac{1}{2}(\psi(x_1) + \psi(x_2)) - \frac{\mu}{4}(x_1 - x_2)^2,$$ (49)
where $\mu$ is a strictly positive constant. We construct from the two optimizers $(\bar{x}^s, \bar{u}^s, x_{T_s})$ and $(\bar{x}^{s+1}, \bar{u}^{s+1}, x_{T_{s+1}})$ the feasible candidate
\begin{equation}
(\frac{1}{2}(\bar{x}^s + \bar{x}^{s+1}), \frac{1}{2}(\bar{u}^s + \bar{u}^{s+1}), \frac{1}{2}(x_{T_s} + x_{T_{s+1}})) \in C^{s+1} \cap D^{s+1}.
\end{equation}

Feasibility is guaranteed due to the convexity of both $C^{s+1}$ and $D^{s+1}$. This results in
\begin{equation}
J_{s+1} \leq \frac{1}{2} (\bar{x}^s - \bar{x}^{s+1})^2 + \frac{1}{2} (\bar{u}^s - \bar{u}^{s+1})^2 + \psi(x_{T_s} + x_{T_{s+1}}).
\end{equation}

Applying the relation (49), which holds with $\mu = 2$ in case of the $L^2$-norm, yields
\begin{equation}
J_{s+1} \leq \frac{1}{2} (J_s + J_{s+1}) - \frac{1}{4} (\bar{x}^s - \bar{x}^{s+1})^2 + \bar{u}^s - \bar{u}^{s+1})^2 - \frac{1}{2} \mu |x_{T_s} - x_{T_{s+1}}|^2.
\end{equation}

This implies
\begin{equation}
||\bar{x}^s - \bar{x}^{s+1}||^2 + ||\bar{u}^s - \bar{u}^{s+1}||^2 + \frac{\mu}{2} |x_{T_s} - x_{T_{s+1}}|^2 \leq 2(J_s - J_{s+1}).
\end{equation}

and from the convergence of $J_s$ it follows therefore that the elements $(\bar{x}^s, \bar{u}^s, x_{T_s}) \in X^s$ form a Cauchy sequence, which, due to the completeness of $X^s$, converges (strongly). The above argument can be repeated by replacing $\bar{x}^{s+1}$, $\bar{u}^{s+1}$ and $x_{T_{s+1}}$ by the optimizer of (1), resulting in the inequality (48).

The inequality (48) is particularly interesting, as it can be used to determine the quality of the approximation of (41) (or likewise (43)) with respect to (1). This is because the suboptimality in the cost can be bounded by
\begin{equation}
J_s - J_\infty \leq J_s - \tilde{J}_s,
\end{equation}
which stems from the fact that $\tilde{J}_s \leq J_\infty$ for all integers $s > 0$. As a result, by solving the two finite-dimensional problems (41) and (42) (or likewise (43) and (46)) not only the suboptimality of the cost $J_s$ compared to $J_\infty$ can be quantified, but also the $L^2$-distance of the corresponding optimal input and state trajectories.

The same reasoning can be applied to trajectories obtained by solving (42) or (46).

**Proposition 2.5** Let the assumptions of Thm. 2.3 be fulfilled. Let the optimal trajectories to (42), respectively (46) be denoted by $\tilde{x}^s, \tilde{u}^s, x_{T_s}$, and the optimal trajectories to (1) by $x, u, x_T$. It is further assumed that $\psi$ is $\mu$-strongly convex. Then, $\tilde{x}^s, \tilde{u}^s, x_{T_s}$ converge strongly, and for all integers $s \geq 0$ it holds that
\begin{equation}
||\tilde{x}^s - x||^2 + ||\tilde{u}^s - u||^2 + \frac{\mu}{2} |x_{T_s} - x_T|^2 \leq 2(J_\infty - \tilde{J}_s).
\end{equation}

In case the assumption of $\psi$ being strongly convex is dropped, strong convergence of $\tilde{x}^s$ and $\tilde{u}^s$ can still be established. Moreover, in the absence of the terminal cost $\psi$, for instance in the infinite-horizon case, the bounds (48) and (55) continue to hold, and it remains true that $\tilde{x}^s$ and $\tilde{u}^s$ converge strongly.

We now provide conditions under which $J_s$ and $\tilde{J}_s$ converge both to $J_\infty$.

**Theorem 2.6** Given that the basis functions form an algebra and are dense in the set of continuous functions with compact support in $I$, it holds that
\begin{equation}
\lim_{s \to \infty} J_s = \lim_{s \to \infty} \tilde{J}_s = J_\infty.
\end{equation}

**Proof** From Thm. 2.3 it can be inferred that $J_s$ and $\tilde{J}_s$ converge, and that the corresponding sequence of optimal input and state trajectories is bounded. Furthermore, by virtue of Lemma 2.1 we have that $\lim_{s \to \infty} C^0_{\ast} = \lim_{s \to \infty} C^1_{\ast}$, and as a consequence we can apply Prop. 3.6 in [32] (the proposition extends naturally to the finite-horizon case), which leads to the desired result.

By combining Thm. 2.6 with Prop. 2.4 or Prop. 2.5 it follows that not only the optimal value function but also the corresponding sequence of optimal trajectories converges strongly to the optimal trajectories of (1).
2.4 Remarks

In the previous section the basis functions were assumed to be continuous throughout the time interval $I$. It is, however, straightforward to extend the previous results in case the interval is split up, for example in $I_1, I_2, \ldots, I_N$, with $\bigcup_{i=1}^N I_i = I$, and piecewise continuous basis functions defined over the intervals $I_i$, $i = 1, 2, \ldots, N$ are used. The basis functions are then required to fulfill Assumptions A1 and A2 over the intervals $I_i$ separately, $i = 1, 2, \ldots, N$. This includes for example the case where polynomials are used as basis functions on $(0, 1)$ and exponentially decaying polynomials on $(1, \infty)$. Thereby, the basis functions complexity is in general increased, which potentially improves the approximation quality and leads to tighter upper and lower bounds on $J_{\infty}$. However, in the context of MPC, closed-loop stability and recursive feasibility is in general lost when splitting the interval $I$, due to the fact that the obtained solutions cannot be shifted in time (see Sec. 3).

Moreover, Thm. 2.3, Prop. 2.4, Prop. 2.5, and Thm. 2.6 can be generalized to a strongly convex running cost instead of a quadratic one. In practice, a quadratic running cost has the advantage of yielding a quadratic objective function, facilitating the numerical solution of the resulting optimization problem.

2.5 A remark on the discrete-time formulation

The main results presented earlier can be translated to the discrete-time case. Due to the high similarity, we will restrict the discussion to the following few remarks.

In the discrete-time case the trajectories $x, u$ corresponding to the discrete-time counterpart of (1) are approximated via

$$
\tilde{x}(k) = (I_n \otimes \tau(k))^T \eta_x, \quad \tilde{u}(k) = (I_m \otimes \tau(k))^T \eta_u,
$$

for all $k \in I_d$, where $I_d$ is a subset of all non-negative integers. We will slightly abuse notation and denote both discrete-time and continuous-time trajectories with the same variables, that is, $\tilde{x}, \tau, \text{etc.}$ The discrete-time analogue of Assumption A2 is given by

A2D) The basis functions fulfill $\tau^*(k+1) = M_{ds} \tau^*(k)$ for all $k \in I_d$.

The subscript 'd' highlights that the matrix $M_{ds}$ and $M_s$ are a priori unrelated. However, we may choose $M_{ds} = e^{M_s T_d}$, for a fixed time $T_d > 0$, in which case the discrete-time basis function $\tau(k)$ matches the corresponding continuous-time basis function $\tau(t)$ at time $t = kT_d$ for all $k \in I_d$. In complete analogy to the continuous-time case, Assumption A2D leads to an invariance of the basis functions with respect to shifts in the index $k$. Hence, in the context of MPC, closed-loop stability guarantees can be shown by the same arguments as in the continuous-time setting (see Sec. 3).

The major difference compared to the continuous-time formulation is that the inequality constraints in the discrete-time version of (1) are at most enforced at a countable number of time indices. No matter whether $I$ has finite or infinite cardinality, this always leads to a finite number of inequality constraints that need to be enforced in the corresponding approximations.\(^6\) Hence, in the discrete-time setting the resulting approximations can always be reduced to standard quadratic programs.

3 Part II: Model predictive control

The proposed approximations can be applied in the context of MPC. We will show that by repeatedly solving the infinite-horizon optimal control problem (43) (with $I = (0, \infty)$, $\psi = 0$, $x_T = \lim_{t \to \infty} x(t) = 0$), recursive feasibility and closed-loop stability are inherent to the resulting MPC algorithm. Due to the fact

\(^6\)In case $I$ has not a finite cardinality, the basis functions are required to be exponentially decaying, due to Assumption A2D and the fact that they are square summable. Hence, as will be shown in the following (the results translate to the discrete setting), it is enough to check the inequality constraints at a finite number of points.
that the basis functions are decaying the constraint \( \lim_{t \to \infty} \ddot{x}(t) = 0 \) is satisfied by construction and does not lead to an additional terminal constraint. For the sake of completeness, (43) is written out as

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

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

The MPC algorithm consists of two steps: In a first step, input and state trajectories \( \dot{x} \) and \( \dot{u} \) are obtained by solving (58) subject to the current state as initial condition \( x_0 \). In a second step, the first portion of the input \( \dot{u} \) is applied to the system, and the procedure is repeated in the next sampling interval. As a consequence, feedback control is achieved.

We recall that Assumption A2 implies that the basis functions can capture arbitrary time-shifts in the sense that for every time-shift \( T_d \) and any given trajectory \( f(t) = \tau_x(t)^T\eta \), where \( \eta \in \mathbb{R}^s \) is the parameter vector, there exists a different set of parameters \( \tilde{\eta}_x \) such that \( f(t - T_d) = \tau_x(t - T_d)^T\tilde{\eta}_x \) for all times \( t \). This result can be easily established by rewriting the basis functions in terms of a matrix exponential as done in [31].

We now discuss the stability properties of the proposed control strategy. Without loss of generality we set \( t = 0 \). According to the first step of the MPC algorithm we solve (58) to obtain the optimal trajectories \( \dot{x} \) and \( \dot{u} \). The input \( \dot{u} \) is then applied to the system over the time span \([0, T_d]\), where \( T_d \) is the sampling time. In the absence of modeling errors, the system will evolve along the predicted trajectory \( \dot{x} \), which is due to the fact that the predictions \( \dot{x} \) and \( \dot{u} \) are exact (as shown in [31]). Due to the time-shift property of the basis functions, there exist parameters \( \tilde{\eta}_x \) and \( \tilde{\eta}_u \) for expressing the shifted trajectories \( \dot{x}(t + T_d) \), \( \dot{u}(t + T_d) \) as a linear combination of the same basis functions, that is

\[
\dot{x}(t + T_d) = (I_n \otimes \tau(t))^T\tilde{\eta}_x, \quad \dot{u}(t + T_d) = (I_m \otimes \tau(t))^T\tilde{\eta}_u,
\]

for all times \( t \in [0, \infty) \). The trajectories \( \dot{x} \) and \( \dot{u} \) are guaranteed to satisfy the equations of motion and the stage constraints for all times and therefore the parameters \( \tilde{\eta}_x \) and \( \tilde{\eta}_u \) are feasible candidates for the optimization (58) with \( x_0 = \dot{x}(T_d) \) at the time instant \( t = T_d \). Recursive feasibility follows then by induction. Moreover, the resulting optimal cost (obtained by solving (58) at time \( t = T_d \)) is certainly lower than the cost achieved by the feasible candidates \( \tilde{\eta}_x \) and \( \tilde{\eta}_u \) corresponding to the trajectories \( \dot{x}(t + T_d) \) and \( \dot{u}(t + T_d) \). The optimal cost at time \( t = T_d \) is therefore bounded by the difference of the optimal cost at time \( t = 0 \) with the integral of the running cost over the interval \([0, T_d]\). As a consequence, the cost is guaranteed to decrease over time, acts therefore as a Lyapunov function, and can be used to conclude closed-loop stability. The previous argument is summarized with the following proposition.

**Proposition 3.1** Provided that the optimization (58) 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** A formal proof can be found in [31] and is included in [2, Sec. 8].

The previous result continues to hold even if (58) is not solved to full optimality (which is often not practicable), provided that the numerical solution algorithm is monotonic in the cost. Given that the numerical solution algorithm is initialized with feasible trajectories at time \( t = 0 \), a single iteration of the solver at each time-step is enough for the stability guarantee to hold, as follows from the above argument.

### 4 Simulation example

The proposed approach is illustrated on a quadruple integrator system, that is,

\[
x_q^{(4)} = u, \tag{60}
\]
where \( x_{qi} \) corresponds to the quadruple integrator state and \( u \) to the input. The example is used to highlight the potential of the proposed MPC approach. However, we do not claim that our approach is superior in general, but believe that it leads to a different trade-off (basis function complexity versus computation) compared to “standard” MPC, which might be beneficial for some applications. We define the state vector as \( x := (x_{qi}, \dot{x}_{qi}, \ddot{x}_{qi}, \dddot{x}_{qi}) \) and consider the task of driving the system from \( x(0) = x_0 \) back to the origin. We penalize input and state deviations with the following cost

\[
J = \int_0^\infty \frac{1}{2} \tau^T x + \frac{1}{2} u^2 \, dt,
\]

and constrain the input \( u \) to lie in \([-0.5, 0.5]\) and the state \( x_{qi} \) to be non-negative, that is, \( x_{qi} \geq 0 \). The basis functions \( \tau \) are designed to be orthonormal and spanned by

\[
\tau \in \exp(-\nu t) \, \text{span}(1, t, t^2, \ldots, t^{s-1}),
\]

where \( \nu \) is set to \( 0.7 \, \text{s}^{-1} \) (this corresponds approximately to the closed-loop poles of an LQR design). This leads to so-called Laguerre functions, c.f. (10) that fulfill Assumptions A1 and A2. Note that the theorem of Stone-Weierstrass, [38, p. 147], states that the basis functions given by (62) are dense in the set of continuous functions vanishing at infinity. The set of smooth compactly supported functions is contained in the set of continuous functions vanishing at infinity, [39, p. 70] and as a result, the assumptions of Lemma 2.1 are fulfilled.

The open-loop state and input trajectories resulting from solving (43) with \( x_0 = (0.3, 0.3, 0.3, 0.3) \) and \( s = 12 \) are shown in Fig. 2. The resulting cost amounts to \( J_{12} = 13.95 \). By solving the problem (46), we obtain the lower bound \( J_{12} = 11.00 \). Thus we can conclude that the cost corresponding to (1) is at most 20% below \( J_{12} \).

Fig. 2 compares open-loop and closed-loop trajectories. The closed-loop trajectories are obtained when resolving the optimization problem (43) every \( T_d = 20 \, \text{ms} \), and applying the obtained input trajectory \( \hat{u}(t) \) in between. In practice this could be realized with two different processes running at different frequencies, one solving (43) at a slower rate and one applying the input \( \hat{u}(t) \) at a higher rate.\(^8\) Closed-loop and open-loop trajectories are significantly different, which is due to the fact that a high polynomial order is required to approximate the bang-bang behavior accurately. The achieved closed-loop cost is \( J_{12cl} = 13.06 \), lying between \( J_{12} \) and \( J_{12} \).

Next, the proposed parametrized MPC approach is benchmarked against the discrete-time MPC approach used by FORCES Pro, [16], and qpOASES (version 3.2.1), [17]. The MPC solver FORCES implements an interior point method that exploits the so-called multistage formulation obtained from the discrete-time MPC formulation. The quadratic programming solver qpOASES implements an active set method tailored to MPC. A terminal cost that matches an LQR design is included in the discrete-time formulation. No terminal set constraint is added, hence closed-loop stability is not guaranteed in the discrete-time approach. The evolution of the system, starting from \( x_0 \) is simulated over 20s and is used to compute the closed-loop cost according to

\[
\sum_{k=1}^{1000} \left( \frac{1}{2} x(kT_d)^T x(kT_d) + \frac{1}{2} u(kT_d)^2 \right) T_d.
\]

The time horizon (in the discrete-time approach) is increased from 93 to 130 samples with a step length of two samples. A time horizon of 92 samples was found to yield an unstable closed-loop trajectory for the given initial condition. The optimization routine FORCES was run with the standard settings, including an

\(^7\)The exponential decay can be expressed using polynomials, which shows that the basis functions form an algebra.

\(^8\)We applied the proposed parametrized approach in practice, see [40], [41], by using zero-order hold. Although the stability guarantees are lost when applying zero-order hold, we did not experience any issues due to the robustness of MPC. By relying on a discrete-time formulation of our parametrized approach the zero-order hold can be accounted for and closed-loop stability can be guaranteed with zero-order hold.
absolute tolerance of $10^{-6}$ for the duality gap and a constraint satisfaction tolerance of $10^{-6}$. The standardized MPC-settings were used for qpOASES. For the parametrized approach the number of basis functions is increased from 8 to 12. The semi-infinite constraint is handled via the active-set method proposed in [2, Sec. 4], where an absolute tolerance of $10^{-6}$ is used for constraint satisfaction. The results are displayed in Fig. 3, where the average execution time (averaged over the $20s$ simulation) is plotted as a function of the achieved closed-loop cost. The parametrized MPC approaches is shown to outperform FORCES and qpOASES by up to one order of magnitude in terms of the average execution time. The achieved closed-loop cost with parametrized approach increases for $s = 11, 12$ compared to $s = 10$. This can most probably be attributed to the relative large discrepancy between closed-loop and open-loop trajectories.

The benchmark is repeated for 100 random initial conditions, that are uniformly distributed in $[0, 0.2] \times [0, 0.2] \times [0, 0.2]$. The corresponding results are shown in Fig. 4 (average execution time) and Fig. 5 (maximum execution time). Note that the initial conditions are guaranteed to be stabilizable, and were indeed stabilized by the parametrized MPC approach. In the discrete-time case, a time horizon below 36 was found to yield unstable closed-loop trajectories. In addition, Fig. 4 displays the sensitivity of the execution time of the parametrized approach with respect to the chosen constraint satisfaction tolerance. The constraint satisfaction tolerance was also varied for FORCES, but found to influence the execution time only insignificantly. In the result shown the constraint satisfaction tolerance for FORCES was set to $10^{-6}$.

In order to demonstrate the scalability of the parametrized approach with the number of states, a chain of $n$ integrators is considered,

$$x_{ni}^{(n)} = u,$$

where $x_{ni}$ corresponds to the first integrator state and $u$ to the input. We define the state vector to be $x := (x_{ni}, \dot{x}_{ni}, \ldots, x_{ni}^{(n-1)})$, and penalize input and state deviations with the cost function (61). The basis function are chosen to be orthonormal and spanned according to (62) with $\nu = 0.7s^{-1}$ and $s = 12$. The input constraint $u \in [-0.5, 0.5]$ and the state constraint $x_{ni} \geq 0$ is included. The execution time required to compute a single solution of (43) subject to the initial condition $x_0 = (0.1, 0.1, \ldots, 0.1)$ is shown in
Figure 3: Shown is the average execution time as a function of the closed-loop cost obtained from a 20s simulation starting at $x_0$. The dashed line indicates the 68% confidence interval of the execution time (plus one standard deviation) for the parametrized approach. The execution time was found to vary only insignificantly for FORCES and qpOASES, and hence, the standard deviation is not shown. To trade-off the execution time with the closed-loop cost, the prediction horizon is changed from 93 to 130 (first in steps of two, then in steps of five) in the discrete-time formulation (used by FORCES and qpOASES), whereas the number of basis functions is increased from 8 to 12 in the parametrized formulation.

Fig. 6. No feasible solution was found with the given basis functions for values of $n$ larger than 7. For $n = 7$ a time horizon of approximately 450 samples is required to achieve closed-loop stability with the discrete-time formulation, which amounts to 3607 optimization variables. In contrast, for $s = 12$ and $n = 7$ the optimization problem resulting from the parametrized approach includes 91 optimization variables.

Summarizing, we can conclude that the parametrized approach might be promising, in particular for systems with marginally stable or unstable dynamics that require high sampling frequencies. On the example of the quadruple integrator system, the parametrized approach outperformed the standard discrete-time approach in terms of (averaged) execution time, without necessarily degrading performance. We believe that the computational advantages stem from a reduction in the number of optimization variables and the fact that only very few constraints are typically active (consider for example the open-loop trajectory shown in Fig. 2), which is exploited by our active-set approach. Hence, it is conjectured that the computational benefits are even higher for systems with higher state and/or input dimensions.

5 Conclusion

The article discussed approximations to the constrained linear quadratic regulator problem, which are based on representing input and state trajectories with basis functions. In particular, a sequence of lower and upper bounds on the cost of the underlying optimal control problem is derived. The approximations are shown to converge. The proposed framework is applied to MPC, where it is shown that an infinite prediction horizon can be retained, leading to recursive feasibility and closed-loop stability. The results are illustrated on a quadruple integrator system. The proposed approach leads to different computational trade-offs compared to “standard” MPC, which might be beneficial for some applications. In case of the quadruple integrator system, it is shown to outperform the state-of-the-art discrete-time solvers in terms of average execution time, without degrading performance.
Figure 4: Simulation of 100 random initial conditions. Depicted is the average execution time as a function of the average closed-loop cost. The dashed line indicates the 68% confidence interval of the average execution time for the parametrized approach. The execution time was found to vary only insignificantly for FORCES and qpOASES, and therefore the corresponding standard deviation is not shown. In order to trade-off the execution time with the closed-loop cost, the prediction horizon is changed from 36 to 60 (36,38,40,42,44,45,50,55,60) in the discrete-time formulation (used by FORCES and qpOASES), whereas the number of basis functions is increased from 8 to 12 in the parametrized formulation.

Figure 5: Simulation of 100 random initial conditions. Depicted is the maximum execution time as a function of the average closed-loop cost. In order to trade-off the execution time with the closed-loop cost, the prediction horizon is changed from 36 to 60 (36,38,40,42,44,45,50,55,60) in the discrete-time formulation (used by FORCES and qpOASES), whereas the number of basis functions is increased from 8 to 12 in the parametrized formulation.
Figure 6: Execution time required for solving (43) subject to the initial condition $x_0 = (0.1, 0.1, \ldots, 0.1)$ for the $n$th order integrator system.

References


