A soft constrained MPC formulation enabling learning from trajectories with constraint violations

Kim P. Wabersich, Raamadaas Krishnadas, and Melanie N. Zeilinger

***Abstract*— In practical model predictive control (MPC) implementations, constraints on the states are typically softened to ensure feasibility despite unmodeled disturbances. In this work, we propose a soft constrained MPC formulation supporting polytopic terminal sets in half-space and vertex representation, which significantly increases the feasible set while maintaining asymptotic stability in case of constraint violations. The proposed formulation allows for leveraging system trajectories that violate state constraints to iteratively improve the MPC controller's performance. To this end, we apply convex optimization techniques to obtain a data-driven terminal cost and set, which result in a quadratic MPC problem.**

***Index Terms*— Predictive control for linear systems, Constrained control, Iterative learning control**

I. Introduction

CONTROL problems can often be described as the minimization of a desired objective function subject to physical limitations on the control input and safety specifications on the states. As classical control approaches, e.g. Proportional-Integral-Derivative (PID) or Linear-Quadratic-Regulator (LQR) control, cannot explicitly consider such specifications, model predictive control gained popularity as a solution to constrained optimal control problems in both research and industry. In addition, more recent advances in learning-based MPC allow for an automated and data-driven design refinement of the MPC problem, see, e.g., [1], [2], to improve the closed-loop performance in an automated fashion.

Despite its benefit of offering theoretical guarantees in terms of constraint satisfaction and performance, an MPC controller implementation is often challenging, as it requires the solution of an optimization problem in real-time using incoming system measurements. While efficient numerical libraries for MPC implementation are available today, see e.g. [3], recursive feasibility guarantees commonly rely on the correctness of the model assumptions, as well as optimality when solving the MPC problem. Especially in the context of modern learning-based design mechanisms, it is, however, vital to deal with unknown disturbances in practice. As a result, system state constraints are often implemented in a soft constrained manner [4], penalizing constraint violations in the cost function. As soon as the softening is active, however, stability guarantees of the closed-loop system are generally lost and alternative formulations as presented, e.g., in [5] and [6] need to be employed. While the specific method in [5] offers desirable properties, i.e. constraint satisfaction if possible and input-to-state stability, even in the case of predicted or actual state constraint violations, the resulting MPC problem becomes a second-order cone program. This is due to constraining the last predicted state in an adaptively scaled ellipsoidal set that needs to be a subset of the softened state constraints. Despite the convexity of the resulting MPC problem, more advanced optimization algorithms are needed compared with more common linear or quadratic MPC problems, which can, e.g., be additionally pre-solved offline [7] to reduce the online computational load. Furthermore, modern learning MPC formulations [8], [9] based on polytopic terminal set enhancements are incompatible with these existing soft constrained MPC formulations.

Contributions: We propose an asymptotically stable soft constrained MPC scheme for linear systems with polytopic constraints using polytopic terminal sets in both half-space representation (H-representation) and vertex representation (V-representation). While maintaining the desirable closed-loop properties from [5], we reduce the MPC problem in both representations to a convex quadratic programming problem. Besides a less complex soft constrained MPC problem, the support of polytopic terminal sets in H-representation can significantly improve the control performance for small- to mid-sized systems by allowing for the use of a maximal forward invariant terminal set rather than an ellipsoidal inner approximation.

The support of terminal sets in V-representation facilitates rigorous softening of learning-based MPC formulations [1] to improve the performance of an iterative task using system trajectories, even in case of state constraint violations, by an enlargement of the terminal set and by approximating an optimal terminal value function. This allows for initializing the learning procedure with a poorly performing MPC, e.g., due to a short planning horizon, or a very basic unconstrained controller. We extend a learning-based formulation [1] based on an affine transformation of the system data, which preserves convexity and maintains a quadratic programming problem structure, thereby enabling a practical application to larger-

This work was supported by the Swiss National Science Foundation under grant no. PP00P2 157601/1.

Kim P. Wabersich (wkim@ethz.ch), Raamadaas Krishnadas (raamadaas.krishnadas@math.ethz.ch), and Melanie N. Zeilinger (mzeilinger@ethz.ch) are with the Institute for Dynamic Systems and Control, ETH Zurich, Switzerland.

scale systems. We illustrate the method using a small- and a large-scale numerical example.

Related Work: Related approaches include [6], which investigates recovery from infeasible states at the level of the optimization algorithm. The results are tailored to a limited class of algorithms, which prohibits the application of, e.g., conventional interior-point methods [10] and software tools like [11]. Compared to the learning-based MPC scheme presented in [8], which we use as basis for the proposed method, it should be noted that [9] provides a less conservative approach to leverage existing data. However, in the important special case of a quadratic stage cost function, a semi-definite programming problem is obtained that can be difficult to solve reliably in case of short sampling times.

II. Problem Statement and MPC Background

We consider linear discrete-time systems of the form

$$x(k+1) = Ax(k) + Bu(k), \quad k \in \mathbb{N}, \tag{1}$$

with states $x(k) \in \mathbb{R}^n$, inputs $u(k) \in \mathbb{R}^m$, and initial condition $x(0) = x_s \in \mathbb{R}^n$. The system (1) has physical input limitations $u(k) \in \mathbb{U}$ of the form $\mathbb{U} := \{u \in \mathbb{R}^m \mid A_u u \leq b_u\}$ with $A_u \in \mathbb{R}^{n_u \times m}$, $b_u \in \mathbb{R}^m$, and it should be operated within safety specifications of the form $x(k) \in \mathbb{X}$ with $\mathbb{X} := \{x \in \mathbb{R}^n \mid A_x x \leq b_x\}$, where $A_x \in \mathbb{R}^{n_x \times n}$, $b_x \in \mathbb{R}^{n_x}$.

To approximately minimize a control objective of the form $\sum_{k=0}^{\infty} \ell(x(k), u(k))$ with stage cost function $\ell(x,u) := x^\top Q x + u^\top R u$ and $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m \times m}$ positive semi-definite and positive definite, we consider a model predictive control scheme. At every time step k, an MPC problem is solved, which is given by

$$\min_{u_i} \quad J_{\mathrm{MPC}}(x, \{u_i\}) := \ell_f(x_N) + \sum_{i=0}^{N-1} \ell(x_i, u_i) \tag{2a}$$

$$\text{s.t.} \quad x_0 = x, \tag{2b}$$

$$x_{i+1} = Ax_i + Bu_i, \ \forall i = 0,..,N-1, \tag{2c}$$

$$A_u u_i \leq b_u, \ \forall i = 0,..,N-1, \tag{2d}$$

$$A_x x_i \leq b_x, \ \forall i = 0,..,N-1, \tag{2e}$$

$$\bar{A}_f x_N \leq \bar{b}_f, \tag{2f}$$

optimizing over an input and state sequence $\{u_i\}$ and $\{x_i\}$, with i denoting the prediction time step, but only the first input element is applied to the system. The resulting MPC control law is then given by $u(k) = \pi_{\mathrm{nom}}(x(k)) := u_0^*$, with u_i^* and x_i^* being solutions to (2), defined on the feasible set $\mathcal{X}_N^{\mathrm{nom}} := \{x \in \mathbb{R}^n | (2b) - (2f)\}$. In (2a) we use a finite horizon sum of stage costs up to a planning horizon N as the objective, with an additional terminal cost term ℓ_f. (2f) imposes a terminal set constraint of the form $\bar{\mathcal{X}}_f := \{x \in \mathbb{R}^n \mid \bar{A}_f x \leq \bar{b}_f\} \subseteq \mathbb{X}$, with $\bar{A}_f \in \mathbb{R}^{\bar{n}_f \times n}$ and $\bar{b}_f \in \mathbb{R}^{\bar{n}_f}$, satisfying the following standard assumption [12].

Assumption II.1. *Consider a terminal set $\bar{\mathcal{X}}_f \subseteq \mathbb{X}, 0 \in \mathrm{int}(\bar{\mathcal{X}}_f)$ and terminal cost function $\ell_f : \bar{\mathcal{X}}_f \to \mathbb{R}^+$. A linear state-feedback control law $u = Kx$, $K \in \mathbb{R}^{m \times n}$ exists, such that for all $x \in \bar{\mathcal{X}}_f$ it holds that $u = Kx \in \mathbb{U}$, $Ax + Bu \in \bar{\mathcal{X}}_f$, and $\ell_f(Ax + BKx) - \ell_f(x) \leq -\ell(x, Kx)$.*

Theorem II.2 (See, e.g., [12])**.** *If Assumption II.1 holds, then the origin is an asymptotically stable equilibrium point for the closed-loop system* (1) *under application of $u(k) = \pi_{\mathrm{nom}}(x(k))$ with a region of attraction given by $\mathcal{X}_N^{\mathrm{nom}}$.*

In addition to the often approximate linear model assumption in (1), we typically encounter unmodeled external disturbances resulting, e.g., from changing operating conditions, causing infeasibility of the MPC online problem (2). In the following, we therefore provide a soft constrained reformulation of (2) using polytopic terminal sets in half-space representation (H-representation) as in (2f), which is introduced in Section III. To also support terminal sets in vertex representation (V-representation) of the form

$$\bar{\mathcal{X}}_f := \mathrm{co}\left(\{\bar{x}_e^f\}_{e=0}^{N_e}\right) \text{ with } \bar{x}_0^f = 0, \tag{3}$$

we provide an alternative formulation in Section IV and thereby avoid the need for computationally expensive transformations between H- and V-representations. Most importantly, the formulation in V-representation allows us to derive a learning-based MPC in Section V, leveraging system trajectories even with constraint violations.

III. Soft constrained MPC with polytopic terminal sets in H-representation

As the constraints on the inputs are typically physical limitations that cannot be exceeded, a common approach to maintain feasibility of (2) is to soften the state constraints (2e) and to enlarge the terminal set from Assumption II.1 beyond the state constraints. Using the basic mechanisms presented in [5] to maintain stability, a soft constrained MPC problem with a polytopic terminal set reads

$$\min_{u_i, \xi_i, \alpha} \quad J_{\mathrm{MPC}}(x, \{u_i\}) + \ell_\xi(\xi_N) + \sum_{i=0}^{N-1} \ell_\xi(\xi_i + \xi_N) \tag{4a}$$

$$\text{s.t.} \quad (2b) - (2d), \ 1 \geq \alpha \geq 0, \ \xi_N \geq 0, \tag{4b}$$

$$\xi_i \geq 0, \ \forall i = 0,..,N-1, \tag{4c}$$

$$A_x x_i \leq b_x + \xi_i + \xi_N, \ \forall i = 0,..,N-1, \tag{4d}$$

$$A_f x_N \leq \alpha b_f, \tag{4e}$$

$$h_f([A_x]_j, \alpha) \leq [b_x + \xi_N]_j, \ \forall j = 1,..,n_x, \tag{4f}$$

with $[A_x]_j$ denoting the j-th row of A_x. In (4) we carry over the original objective (2a) as well as prediction and input constraints (4a) and (4b) from (2) and additionally introduce so-called slack-variables ξ_i and a terminal set scaling factor α with additional penalty terms in (4a) of the form $\ell_\xi(\xi) := c^\top \xi$ with $c > 0$ sufficiently large to represent an exact penalty function, see, e.g., [4]. While the slack variables ξ_i, $i = 0,..,N-1$ allow for constraint violations along the prediction horizon if necessary, (4f) represents the required terminal slack ξ_N w.r.t. a scaled terminal set $\alpha\mathcal{X}_f := \{x \in \mathbb{R}^n | A_f x \leq \alpha b_f\}$, $A_f \in \mathbb{R}^{n_f \times n}$, $b_f \in \mathbb{R}^{n_f}$. Importantly, the terminal set is not required to be a subset of the state constraints:

Assumption III.1. *The terminal set $\mathcal{X}_f$ satisfies all conditions in Assumption II.1 except for $\mathcal{X}_f \subseteq X$, i.e. neglecting state constraints.*

By allowing for $\mathcal{X}_f \nsubseteq \mathbb{X}$ we can enlarge the terminal set, i.e. $\bar{\mathcal{X}}_f \subseteq \mathcal{X}_f$, and can require the last predicted state to be contained in $\alpha\mathcal{X}_f$ through (4e) without being overly restrictive. The terminal slack ξ_N thereby ensures that $\alpha\mathcal{X}_f$ is fully contained inside the softened state constraints, i.e. $\alpha\mathcal{X}_f \subseteq \{x \in \mathbb{R}^n | A_x x \leq b_x + \xi_N\}$. The latter subset condition translates into (4f) using the support function

$$h([A_x]_j, \alpha) := \max_{x\in\mathbb{R}^n} \; [A_x]_j x \text{ s.t. } A_f x \leq \alpha b_f \tag{5}$$

for each state constraint half-space $j = 1,..,n_x$. Differently from a basic state constraint softening, e.g. [4], we employ both the stage-wise and terminal slacks at every prediction step in (4d) and in the stage-wise penalty (4a). As shown in [5] these components yield a strong incentive to first obtain terminal constraint satisfaction, which allows to establish asymptotic stability despite constraint violations.

We denote the soft constrained MPC control law resulting from (4) by $u(k) = \pi_{\text{hrep}}(x(k)) := u_0^*(x(k))$ with the enlarged feasible set $\mathcal{X}_N^{\text{hrep}} := \{x \in \mathbb{R} | (4b) - (4f)\} \supset \mathcal{X}_N^{\text{nom}}$. Note that a direct implementation of (4) would yield a non-linear optimization problem due to (4f), for which we provide a convex reformulation after the following theorem, showing recursive feasibility and asymptotic stability.

Theorem III.2. *If Assumption III.1 holds, then the origin is an asymptotically stable equilibrium point for the closed-loop system* (1) *under application of* $u(k) = \pi_{\text{hrep}}(x(k))$ *with a region of attraction* $\mathcal{X}_N^{\text{hrep}}$. *Furthermore, there exists a sufficiently large constant* $c \in \mathbb{R}^+$ *for the penalty* $\ell_\xi(\xi) = c^\top \xi$, *such that zero slack at some time step* k, $\xi_i^*(k) = 0$, *implies zero slack for all future time steps* $\bar{k} > k$, $\xi_i^*(\bar{k}) = 0$.

Proof. The proof is a direct extension of [5]. For every $x \in \mathcal{X}_N^{\text{hrep}}$, we construct a feasible candidate solution for the next state $x^+ = Ax + B\pi_{\text{hrep}}(x)$ as $u_i^+ = u_{i+1}^*$ for $i = 0,..,N-2$ with $u_{N-1}^+ = Kx_N^*$, $x_i^+ = x_{i+1}^*$, and $\xi_i^+ = \xi_{i+1}^*$ for $i = 0,..,N-1$ with $x_N^+ = (A+BK)x_N^*$ and $\xi_N^+ = \xi_N$, $\alpha^+ = \alpha$ since $\alpha\mathcal{X}_f$ is invariant for every $0 \leq \alpha \leq 1$ by convexity of $\mathbb{U}$, linearity of (1), and $0 \in \mathcal{X}_f$. Furthermore, the cost (4a) serves as a Lyapunov function and the cost difference at x^+ using $\{\xi_i^+, u_i^+, x_i^+, \alpha^+\}$ compared with the optimal cost at x can be bounded from above by $-\ell(x, \pi_{\text{hrep}}(x))$ using Assumption III.1, which implies the first part of the Theorem. The second part directly follows from the analysis above in combination with exact penalty arguments, see, e.g., [4]. □

Remark III.3. *While we show asymptotic stability for system* (1) *in this paper, the results can directly be extended to prove input-to-state stability of a system of the form* $x(k+1) = Ax(k)+Bu(k)+w(k)$ *with* $w(k) \in \mathcal{W}$ *by extending the results in [5, Thm. V.2].*

In the following, we introduce a convex reformulation of (4), which yields a quadratic program.

Lemma III.4. *The constraints* (4f) *are equivalent to*

$$\alpha\mu_j^* \leq [b_x + \xi_N]_j, \; \forall j = 1,..,n_x, \; \textit{with} \tag{6a}$$

$$\alpha\mu_j^* = \alpha \min_{\nu_j} \nu_j^\top b_f \; \textit{s.t.} \; \nu_j^\top A_f - [A_x]_j = 0, \; \nu_j \geq 0. \tag{6b}$$

Proof. The constraints (6) are obtained by formulating the dual problem of (5). The Lagrangian of (5) is given by $L(x,\nu) = -[A_x]_j x + \nu_j^\top(A_f x - \alpha b_f)$ with dual function

$$g(\nu_j) = \begin{cases} -\nu_j^\top b_f \alpha, & \text{if } \nu_j^\top A_f - [A_x]_j = 0, \\ -\infty, & \text{otherwise.} \end{cases}$$

The dual problem therefore results in (6b) by noting that the optimizer ν_j^* of (6b) is invariant under a linear scaling with $\alpha > 0$ and that the optimal value is 0 for $\alpha = 0$. Since the support function problem (5) is an LP and is always feasible, strong duality holds, i.e. $\alpha\mu_j^* = \max_{\bar{x} \text{ s.t. } A_f\bar{x} \leq \alpha b_f} [A_x]_j \bar{x}$ [13, Section 5.2.4], which completes the proof. □

Lemma III.4 allows us to pre-compute the support function (5) up to the scaling α through (6b) for all half-spaces $j = 1,..,n_x$, resulting in linear inequalities given by (6a), replacing (4f) and therefore we obtain a quadratic program as a soft constrained online MPC problem. Note that in relation to the overall number of constraints in an MPC problem (2), the additional n_x linear inequality constraints due to (6a) in (4) do not significantly increase the problem complexity.

Remark III.5. *Condition* (6a) *can equivalently be derived using the vertex representation* (3) *by reformulating* (4f) *as*

$$\max_{e_j = 1,..,N_e} [A_x]_j (\alpha x_{e_j}^f) \leq [b_x + \xi_N]_j, \; \forall j = 1,..,n_x$$

to obtain $\mu_j^* = [A_x]_j x_{e_j{}^*}^f$, *with optimal* $e_j{}^*$ *computed offline.*

IV. Soft constrained MPC with polytopic terminal sets in V-representation

Various learning-based procedures such as [8], [9] rely on a vertex representation of the terminal set in particular when allowing for enlargement of the terminal set based on available trajectories. Since the transformation between the H- and V-representations of $\mathcal{X}_f$ becomes computationally prohibitive for large-scale systems, we provide an alternative soft constrained MPC formulation in this section that directly supports terminal sets in the V-representation according to Assumption III.1 of the form (3). As a first step, we note that a simple integration of (3) into the soft constrained MPC problem as done in [8] would cast (4e)-(4f) with scaled terminal set $\alpha\mathcal{X}_f = \text{co}\left(\{\alpha x_e^f\}_{e=0}^{N_e}\right)$ into

$$A_x \alpha x_e^f \leq b_x + \xi_N, \lambda_e \geq 0, \Sigma_{e=0}^{N_e} \lambda_e = 1, x_N = \Sigma_{e=0}^{N_e} \lambda_e \alpha x_e^f$$

with multipliers λ_e, where x_N needs to be a convex combination of the vertices of $\alpha\mathcal{X}_f$. Due to the scaling through α in the soft constrained case, this introduces the bilinearity $\lambda_e\alpha$. We present a convex reformulation based on the following relation, which exploits that $x_0^f = 0$ according to (3):

Lemma IV.1. *Consider* (3). *For any* $0 \leq \alpha \leq 1$ *it holds that*

$$\alpha\mathcal{X}_f = \left\{ \Sigma_{e=0}^{N_e} \bar{\lambda}_e x_e^f | \Sigma_{e=0}^{N_e} \bar{\lambda}_e = 1, \bar{\lambda}_0 = 1 - \alpha \right\} := \bar{\mathcal{X}}_f^\alpha.$$

Using Lemma IV.1, we avoid the bilinearity by using the scaling $\alpha = 1 - \lambda_0$, which yields the convex MPC problem

$$\min_{\substack{u_i,\xi_i,\\ \lambda_e}} \quad J_{\mathrm{MPC}}(x, \{u_i\}) + \ell_\xi(\xi_N) + \sum_{i=0}^{N-1} \ell_\xi(\xi_i + \xi_N) \tag{7a}$$

$$\text{s.t.} \quad (2b) - (2d),\ \xi_N \geq 0, \tag{7b}$$

$$\xi_i \geq 0,\ \forall i = 0, .., N-1, \tag{7c}$$

$$A_x x_i \leq b_x + \xi_i + \xi_N,\ \forall i = 0, .., N-1, \tag{7d}$$

$$A_x(1-\lambda_0) x_e^f \leq b_x + \xi_N,\ \forall e = 0, .., N_e, \tag{7e}$$

$$\lambda_e \geq 0,\ \forall e = 0, .., N_e, \tag{7f}$$

$$x_N = \Sigma_{e=0}^{N_e} \lambda_e x_e^f, \text{ and } \Sigma_{e=0}^{N_e} \lambda_e = 1, \tag{7g}$$

preserving the theoretical properties from Theorem III.2.

V. Learning-based MPC from trajectories with constraint violations

In this section, we use the soft constrained MPC approach derived in Section IV to propose a learning-based MPC formulation for iterative tasks of horizon $\bar{N}$, starting at initial condition x_s. To this end, we provide a mechanism based on [8] to improve the closed-loop performance using available system trajectories by enlarging the terminal set and by improving the terminal cost estimate. Consider N_e different system trajectories of (1) satisfying

$$\mathcal{D}_{N_e} = \left\{ \{x_{k,e}, u_{k,e}, V_{k,e}\}_{k=0,e=0}^{k=\bar{N},e=N_e} \right\}, \tag{8a}$$

$$x_{\bar{N},e} = 0, x_{0,e} = x_s, u_{\bar{N},e} = 0, \forall e = 0, .., N_e \tag{8b}$$

$$u_{k,e} \in \mathbb{U},\ \forall k = 0, .., \bar{N}, \forall e = 0, .., N_e, \tag{8c}$$

$$V_{k,e} = \Sigma_{i=k}^{\bar{N}} \tilde{\ell}(x_{i,e}, u_{i,e}), \tag{8d}$$

where k denotes the time step of each state per trajectory sample e starting from x_s and reaching the origin within $\bar{N}$ time steps. Importantly, note that these trajectories can violate state constraints. This enables to learn from experiments, where temporary constraint violation has not caused the system to fail or where data has been gathered under more permissive state constraints, e.g., on a wider test track in case of autonomous driving. In (8d), we use $\tilde{\ell}(x_{i,e}, u_{i,e}) := \ell(x_{i,e}, u_{i,e}) + \ell_\xi(\max(0, A_x x_{i,e} - b_x))$ as the combined performance cost and constraint violation penalty.

Remark V.1. *Available trajectories that do not satisfy $x_{\bar{N},e} = 0$ can often be extended through a predicted open-loop trajectory via an MPC problem, starting from the last state of the trajectory and resulting in the origin.*

Different from the case discussed in Section IV, we do not assume access to a terminal set $\mathcal{X}_f$ and a terminal cost function ℓ_f, but construct a corresponding estimate using available data (8). Thereby, the main difference compared with similar learning-based approaches [8], [9], [14] is the data-driven terminal set and cost synthesis to ensure stability and constraint satisfaction if possible, based on trajectories that potentially violate constraints. In particular, we develop a convex mechanism to learn from cost samples $V_{k,e}$ that are available at potentially unsafe data locations $x_{k,e}$, which naturally would result in a non-convex bilinear MPC cost, similar as in [9]. In the following, we therefore derive a *convex scalable* barycentric terminal cost function from [8], [14] using data (8), which is given by

$$J_f^{*,\mathrm{BC}}(x) := \min_{\lambda_{k,e}} \quad \Sigma_{k,e} \lambda_{k,e} V_{k,e} \tag{9a}$$

$$\text{s.t.} \quad x = \Sigma_{k,e} \lambda_{k,e} x_{k,e}, \tag{9b}$$

$$\Sigma_{k,e} \lambda_{k,e} = 1, \tag{9c}$$

$$\lambda_{k,e} \geq 0,\ \forall k, e, \tag{9d}$$

where we use '$\Sigma_{k,e}$' as shorthand for '$\Sigma_{k=0,e=0}^{k=\bar{N},e=N_e}$' with a feasible terminal set $\mathcal{X}_f^{\mathrm{BC}} = \{x \in \mathbb{R}^n | (9b) - (9d)\}$, which can be scaled linearly by α through enforcing the convex constraint $1 - \Sigma_{e=0}^{N_e} \lambda_{\bar{N},e} = \alpha$, since $x_{\bar{N},e} = 0$ (8b), see Lemma IV.1.

Lemma V.2. *Consider the barycentric function $J_f^{*,\mathrm{BC}}(x)$ defined in* (9) *based on available system data according to* (8) *with the feasible set $\mathcal{X}_f^{\mathrm{BC}}$. For every $x \in \mathcal{X}_f^{\mathrm{BC}}$ with corresponding optimal solution $\lambda_{k,e}^*$ to* (9)*, there exists an input $u \in \mathbb{U}$ such that $Ax + Bu \in \mathcal{X}_f^{\mathrm{BC}}$ with feasible solution $\lambda_{k,e}^+$ satisfying $\Sigma_{e=0}^{N_e} \lambda_{\bar{N},e}^* \leq \Sigma_{e=0}^{N_e} \lambda_{\bar{N},e}^+$, and $J_f^{*,\mathrm{BC}}(Ax + Bu) - J_f^{*,\mathrm{BC}}(x) \leq -\ell(x,u) - \ell_\xi(\max(0, A_x x - b_x))$.*

Proof. Select $u = \Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e}^* u_{k,e}$ with $u_{k,e}$ available from (8) and note that $u \in \mathbb{U}$ due to convexity of $\mathbb{U}$. We have for $x^+ = Ax + Bu$ that

$$\begin{aligned} x^+ &= A\Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e}^* x_{k,e} + B\Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e}^* u_{k,e} \\ &= \Sigma_{k=0,e=0}^{k=\bar{N}-2,e=N_e} \lambda_{k,e}^* x_{k+1,e}, \end{aligned}$$

since $x_{\bar{N},e} = Ax_{\bar{N},e} + Bu_{\bar{N},e} = 0$ due to (8b). Define $\lambda_{k+1,e}^+ = \lambda_{k,e}^*$ for all $e = 0, .., N_e$, $k = 0, .., \bar{N}-2$, $\lambda_{0,e}^+ = 0$, and $\Sigma_{e=0}^{N_e} \lambda_{\bar{N},e}^+ := 1 - \Sigma_{k=0,e=0}^{k=\bar{N}-1,e=N_e} \lambda_{k,e}^+$. We obtain $\Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e}^+ x_{k,e} = x^+$, $\Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e}^+ = 1$ and $\lambda_{k,e}^+ \geq 0$ as required by (9b), (9c), and (9d). Further, note that it holds $\Sigma_{e=0}^{N_e} \lambda_{\bar{N},e}^* = 1 - \Sigma_{k=0,e=0}^{k=\bar{N}-1,e=N_e} \lambda_{k,e}^* \leq 1 - \Sigma_{k=0,e=0}^{k=\bar{N}-2,e=N_e} \lambda_{k,e}^* = \Sigma_{e=0}^{N_e} \lambda_{\bar{N},e}^+$, ensuring that $\lambda_{k,e}^+$ is a feasible solution for (9) at x^+ with desired properties. The cost decrease condition can be verified using $\lambda_{e,k}^+$ as follows:

$$\begin{aligned} &J_f^{*,\mathrm{BC}}(x^+) - J_f^{*,\mathrm{BC}}(x) \\ &= -\Sigma_{k=0,e=0}^{k=\bar{N}-1,e=N_e} \lambda_{k,e}^* \tilde{\ell}(x_{k,e}, u_{k,e}) \\ &= -\Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e}^* \tilde{\ell}(x_{k,e}, u_{k,e}) \\ &\leq -\tilde{\ell}(\Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e}^* x_{k,e}, \Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e}^* u_{k,e}), \\ &= -\tilde{\ell}(x, u), \end{aligned}$$

where we exploit the definition of $\lambda_{k,e}^+$, as well as $V_{\bar{N},e} = 0$, $x_{\bar{N},e} = 0$, $u_{\bar{N},e} = 0$, $\tilde{\ell}(x_{\bar{N},e}, u_{\bar{N},e}) = 0$ for all $e = 0, .., N_e$ by definition (8b), and convexity of $\tilde{\ell}$ in combination with Jensen's inequality, completing the proof. □

Lemma V.2 allows us to state a soft constrained MPC problem using the data-driven terminal cost $J_f^{*,\mathrm{BC}}$ and terminal set

$\mathcal{X}_f^{\mathrm{BC}}$ as follows:

$$\begin{aligned}
\min_{u_i,\xi_i,\lambda_{k,e}} \quad & \ell_\xi(\xi_N) + \Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e} V_{k,e} + \\
& \Sigma_{i=0}^{N-1} (\ell(x_i,u_i) + \ell_\xi(\xi_i + \xi_N)) && (10a)\\
\text{s.t.} \quad & (2b) - (2d), && (10b)\\
& \xi_N \geq 0,\ \xi_i \geq 0,\ \forall i = 0,..,N-1, && (10c)\\
& A_x(1 - \Sigma_{e=0}^{N_e} \lambda_{\bar{N},e}) x_{k,e} \leq b_x + \xi_N, && (10d)\\
& A_x x_i \leq b_x + \xi_i + \xi_N,\ \forall i = 0,..,N-1, && (10e)\\
& \lambda_{k,e} \geq 0, \forall k = 0,..,\bar{N}, e = 0,..N_e, && (10f)\\
& x_N = \Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e} x_{k,e}, && (10g)\\
& \Sigma_{k=0,e=0}^{k=\bar{N},e=N_e} \lambda_{k,e} = 1, && (10h)
\end{aligned}$$

with control law $\pi_{\mathcal{D}_{N_e}}(x(k)) := u_0^*$, feasible set $\mathcal{X}_N^{\mathcal{D}_{N_e}} := \{x \in \mathbb{R} | (10b) - (10h)\}$, and optimal cost $J^*_{\mathcal{D}_{N_e}}(x_s)$.

Corollary V.3. *If the data used to define the learning-based MPC problem* (10) *satisfies* (8)*, then the origin is an asymptotically stable equilibrium point for the closed-loop system* (1) *under application of* $u(k) = \pi_{\mathcal{D}_{N_e}}(x(k))$ *with region of attraction* $\mathcal{X}_N^{\mathcal{D}_{N_e}}$. *Furthermore, a sufficiently large constant* $c \in \mathbb{R}^+$ *exists for the penalty* $\ell_\xi(\xi) = c^\top \xi$ *such that zero slack at some time step* k, $\xi_i^*(k) = 0$, *implies zero slack for all future time steps* $\bar{k} > k$, $\xi_i^*(\bar{k}) = 0$.

Proof. The proof follows directly from the proof of Theorem III.2 together with Lemma V.2. □

Enlargement of the feasible set $\mathcal{X}_N^{\mathcal{D}_{N_e}}$ through data (8) comes at the cost of additional optimization variables $\lambda_{k,e}$ per trajectory e of length N_e and per time step k, which introduces N_e variables and $N_e + 1$ constraints in (10). In comparison, a soft constrained MPC with a standard terminal set within state constraints but with an enlarged horizon of $N + N_e$ would lead to additional mN_e input variables, $n_u N_e$ input constraints and $n_x N_e$ state constraints, see Section VI b) for a computational comparison.

Iterative performance improvement: Next, we investigate if the proposed soft constrained MPC controller according to (10) ensures iterative performance improvement despite leveraging system trajectories that potentially violate state constraints in combination with the online scaling mechanism of the terminal components. At every iteration, the data set $\mathcal{D}_{N_e}$ is extended using one closed-loop system trajectory satisfying (8) according to

$$\mathcal{D}_{N_e+1} = \mathcal{D}_{N_e} \cup \{x_{k,N_e+1}, u_{k,N_e+1}, V_{k,N_e+1}\}_{k=0}^{\bar{N}}, \quad (11)$$

with $u_{k,N_e+1} = \pi_{\mathcal{D}_{N_e}}(x_{k,N_e+1})$.

Theorem V.4. *Consider a dataset* (11) *satisfying* (8b)-(8d)*. It holds that*

$$V_{0,N_e+1} \leq J^*_{\mathcal{D}_{N_e}}(x_s) \text{ and } J^*_{\mathcal{D}_{N_e+1}}(x_s) \leq J^*_{\mathcal{D}_{N_e}}(x_s) \quad (12)$$

with optimal cost $J^*_{\mathcal{D}_{N_e}}(x_s)$ *according to* (10)*. If, in addition,* $\ell_\xi(\max(0, A_x x_{k,e} - b_x)) = 0$ *for all* $k \geq 0$ *and* $e \geq 0$ *it follows that*

$$V_{0,N_e+1} \leq V_{0,N_e}. \quad (13)$$

Proof. We first consider (12): Since (10) using $\mathcal{D}_{N_e+1}$ is feasible for any solution of (10) using $\mathcal{D}_{N_e}$, it holds that $J^*_{\mathcal{D}_{N_e+1}}(x_s) \leq J^*_{\mathcal{D}_{N_e}}(x_s)$. Using the candidate sequence from the proof of Theorem III.2 we can conclude that $J^*_{\mathcal{D}_{N_e}}(x_{k+1,N_e+1}) - J^*_{\mathcal{D}_{N_e}}(x_{k,N_e+1}) \leq -\ell(x_{k,N_e+1}, u_{k,N_e+1}) - \ell_\xi(\max(0, A_x x_{k,N_e+1} - b_x)$ since ℓ_ξ is affine. This allows us to derive the upper bound

$$\begin{aligned}
&\Sigma_{k=0}^{\bar{N}-1} J^*_{\mathcal{D}_{N_e}}(x_{k+1,N_e+1}) - J^*_{\mathcal{D}_{N_e}}(x_{k,N_e+1}) \leq \\
&\Sigma_{k=0}^{\bar{N}-1} - \ell(x_{k,N_e+1}, u_{k,N_e+1}) - \ell_\xi(\max(0, A_x x_{k,N_e+1} - b_x)
\end{aligned}$$

implying $J^*_{\mathcal{D}_{N_e}}(x_{\bar{N},N_e+1}) - J^*_{\mathcal{D}_{N_e}}(x_s) = -J^*_{\mathcal{D}_{N_e}}(x_s) \leq -V_{0,N_e+1}$ by exploiting the telescoping sum and $x_{\bar{N},N_e+1} = u_{\bar{N},N_e+1} = 0$. To prove (13), we show $J^*_{\mathcal{D}_{N_e}}(x_s) \leq V_{0,N_e}$ implying with (12) that $V_{0,N_e+1} \leq V_{0,N_e}$: $V_{0,N_e} = \Sigma_{k=0}^{N} \tilde{\ell}(x_{k,N_e}, u_{k,N_e}) + \Sigma_{k=N+1}^{\bar{N}} \tilde{\ell}(x_{k,N_e}, u_{k,N_e}) \geq J^*_{\mathcal{D}_{N_e}}(x_s)$, since $\ell_\xi(\max(0, A_x x_{k,N_e+1}) - b_x)) = 0$ for all $k \geq 0$ allowing to select $x_i = x_{i,N_e}$, $u_i = u_{i,N_e}$, $\xi_i = 0$, $\xi_N = 0$, and $\lambda_{\bar{N},N_e} = 1$, $\lambda_{k \neq N,e} = 0$ for all k, e as feasible candidate solution to (10) completing the proof. □

Remark V.5. *As long as we observe a performance improvement* (13)*, we can exchange old trajectories with new trajectories, maintaining the complexity of the MPC problem, while refining the data used for design. A corresponding practical implementation strategy would be to keep adding trajectories until* $\ell_\xi(\max(0, A_x x_{k,N_e+1}) - b_x)) = 0$ *for all* $k \geq 0$ *holds, i.e. until there are no closed-loop constraint violations, followed by only updating old trajectories.*

Remark V.6. *While establishing convergence to the global optimal solution as shown in [8, Theorem 3] is beyond the scope of this paper, note that the analysis of the proposed method can be reduced to the special case considered in [8] if an iterative refinment of the terminal set according to Remark V.5 allows to eliminate constraint violating trajectories.*

VI. Numerical Examples

In this section we present small- and large-scale simulation examples to demonstrate the proposed method. A matlab script with further details and plots can be found online [15].

a) Illustrative 2D example: We consider a Euler-discretized 2D mass-spring-damper system with 0.05 seconds sampling time, spring constant 1, mass 1, damping factor 0.1, and cost $Q = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0.1 \end{smallmatrix}\right)$, $R = 1$. The position and velocity are constrained to the interval $[-1, 1]$ and an external control force, acting on the mass, to $[-2, 2]$. We construct a terminal set and cost function according to Section IV using closed-loop trajectories resulting from a saturated LQR controller. As depicted in Figure 1, we initialize the system outside the state constraints such that the corresponding scaling of the terminal set is required to be larger than the system's state constraints for obtaining a feasible solution. As depicted by the shaded blue rectangles, the state constraints need to be softened besides the required terminal slack values. Nevertheless, the closed-loop system under the proposed soft-constrained MPC controller asymptotically converges to the origin, and the terminal set becomes smaller at every time step until it is fully contained in the state constraints. In a second example, we initialize the

e	$V_{0,e}$,VI a)	$J^*_{\mathcal{D}_e}(x_s)$,VI a)	$V_{0,e}$,VI b)	$J^*_{\mathcal{D}_e}(x_s)$,VI b)
1	2885.16	7762.08	3144.87	5781.10
2	8.56	1287.43	2912.45	3083.58
3	7.47	9.04	2905.22	2908.93
4	7.36	7.37	2903.65	2904.90
10	7.36	7.36	2902.15	2902.19

TABLE I

CLOSED-LOOP COST $\boldsymbol{V_{0,e}}$ AND MPC COST $\boldsymbol{J^*_{\mathcal{D}_e}(x_s)}$ STARTING FROM INITIAL STATE $\boldsymbol{x_s}$ DURING DIFFERENT EPISODES $\boldsymbol{e}$ USING THE ITERATIVE TERMINAL SET ENLARGEMENT ACCORDING TO (11).

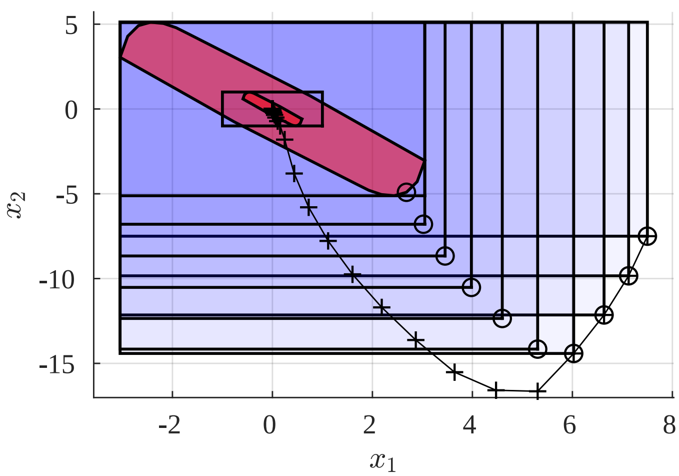

Fig. 1. Visualization of numerical example VI a): The predicted state trajectory is shown, starting at $\boldsymbol{x} = [\mathbf{7.5}, \ -\mathbf{7.5}]$ (circles) with predicted state constraint softening (shaded blue rectangles) and scaled terminal set (shaded red polytope). The black line with crosses shows the closed-loop trajectory with final state constraint (small rectangle) and corresponding terminal set (small red polytope).

terminal set with $\{0\}$ and perform iterative learning with initial state $x_s = [0.75, 0.75]^\top$ according to Section V, Remark V.5. The corresponding MPC problem and closed-loop cost values are given in Table I, which confirm the bounds presented in Theorem V.4 and converge to the global optimal solution. Finally, we compare the volume of $\mathcal{X}_N^{\mathcal{D}_{N_e}}$ with the feasible set of its nominal counterpart, i.e. $\xi_i = 0 \ \forall i = 1, ..N$, recovering the approach in [8], and observe that the volume of the region of attraction increases by a factor of ≈ 64.4.

b) Large-scale thermal application: Based on the model used in [16], we apply the iterative learning-based approach to a larger-scale constrained server-cooling example with 64 dynamically coupled states and 64 inputs. Thereby, we initialize the terminal set with a single trajectory, where we start from a random initial condition and simulate the system under application of an input saturated LQR controller. We select an MPC planning horizon $N = 3$ and data horizon $\bar{N} = 50$ according to Section V, resulting in performance improvements as given in Table I. Compared to a nearly optimal controller, represented by an MPC using an overall planning horizon of $N = 53$, the learning-based formulation yields 99.995% of the closed-loop performance after 3 learning episodes, with a ≈ 12 times faster solve time of the MPC problem.

VII. Conclusion

In this paper, we introduced a soft constrained MPC formulation in the form of a quadratic programming problem, supporting polytopic terminal sets in H- and V-representation while preserving desirable asymptotic stability guarantees. The approach allows to use a maximal terminal set and can be combined with recent learning-based MPC techniques to improve the performance using available system trajectories that potentially violate the state constraints.

References

[1] U. Rosolia, X. Zhang, and F. Borrelli, "Data-Driven Predictive Control for Autonomous Systems," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, no. 1, pp. 259–286, 2018.

[2] L. Hewing, K. P. Wabersich, M. Menner, and M. N. Zeilinger, "Learning-Based Model Predictive Control: Toward Safe Learning in Control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, no. 1, pp. 269–296, 2020.

[3] R. Verschueren, G. Frison, D. Kouzoupis, N. van Duijkeren, A. Zanelli, B. Novoselnik, J. Frey, T. Albin, R. Quirynen, and M. Diehl, "Acados: A modular open-source framework for fast embedded optimal control," *arXiv:1910.13753*, 2019.

[4] E. C. Kerrigan and J. M. Maciejowski, "Soft Constraints And Exact Penalty Functions In Model Predictive Control," in *Proc. UKACC Int. Conf. (Control)*. Citeseer, 2000.

[5] M. N. Zeilinger, M. Morari, and C. N. Jones, "Soft constrained model predictive control with robust stability guarantees," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1190–1202, 2014.

[6] C. Feller and C. Ebenbauer, "A stabilizing iteration scheme for model predictive control based on relaxed barrier functions," *Automatica*, vol. 80, pp. 328–339, 2017.

[7] A. Alessio and A. Bemporad, "A survey on explicit model predictive control," in *Nonlinear model predictive control*. Springer, 2009, pp. 345–369.

[8] U. Rosolia and F. Borrelli, "Learning Model Predictive Control for Iterative Tasks: A Computationally Efficient Approach for Linear System," in *IFAC-PapersOnLine*, vol. 50, no. 1, 2017, pp. 3142–3147.

[9] F. D. Brunner, M. Lazar, and F. Allgöwer, "Stabilizing model predictive control: On the enlargement of the terminal set," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 15, pp. 2646–2670, 2015.

[10] A. Domahidi, A. U. Zgraggen, M. N. Zeilinger, M. Morari, and C. N. Jones, "Efficient interior point methods for multistage problems arising in receding horizon control," in *Proceedings of the IEEE Conference on Decision and Control*, 2012, pp. 668–674.

[11] H. J. Ferreau, C. Kirches, A. Potschka, H. G. Bock, and M. Diehl, "qpoases: A parametric active-set algorithm for quadratic programming," *Mathematical Programming Computation*, vol. 6, no. 4, pp. 327–363, 2014.

[12] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, jun 2000.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

[14] C. N. Jones and M. Morari, "Polytopic approximation of explicit model predictive controllers," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2542–2553, 2010.

[15] K. P. Wabersich, R. Krishnadas, and M. Zeilinger, "A soft constrained MPC formulation enabling learning from trajectories with constraint violations [Source Code]," https://doi.org/10.24433/CO.7967683.v1, 2021.

[16] K. P. Wabersich and M. Zeilinger, "Bayesian model predictive control: Efficient model exploration and regret bounds using posterior sampling," in *Learning for Dynamics and Control*. PMLR, 2020, pp. 455–464.