MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Safety Verification of Implicitly Defined MPC Feedback Laws

Holaza, J.; Takacs, B.; Kvasnica, M.; Di Cairano, S.

TR2015-086 July 2015

Abstract

For a closed-loop system composed of a linear controlled plant and an MPC feedback strategy we show how to verify that closed-loop state trajectories either enter or avoid a given set of unsafe states. The search for the safety certificate is formulated as a mixed-integer linear programming problem which yields non-conservative certificates. The optimal control commands generated by the MPC policy are represented by Karush-Kuhn-Tucker optimality conditions, which allow to perform the verification without the need to explicitly compute reachable sets.

2015 European Control Conference (ECC)

Safety Verification of Implicitly Defined MPC Feedback Laws

Juraj Holaza, Bálint Takács, Michal Kvasnica, and Stefano Di Cairano

***Abstract*— For a closed-loop system composed of a linear controlled plant and an MPC feedback strategy we show how to verify that closed-loop state trajectories either enter or avoid a given set of unsafe states. The search for the safety certificate is formulated as a mixed-integer linear programming problem which yields non-conservative certificates. The optimal control commands generated by the MPC policy are represented by Karush-Kuhn-Tucker optimality conditions, which allow to perform the verification without the need to explicitly compute reachable sets.**

I. Introduction

It is well known that Model Predictive Control (MPC) feedback strategies can provide an optimal operation of the plant while taking constraints into account [9]. However, certain safety specifications such as performance constraints (e.g., limits on overshoots and settling time) or obstacle avoidance constraints are difficult to impose in the standard context of convex MPC because they lead to non-convex formulations which are computationally expensive to implement in real time. If satisfaction of such safety bounds can be verified off-line, then the controller can be much simpler.

Given a model of the controlled plant $x(t+1) = f(x(t), u(t))$ and the MPC feedback strategy $u(t) = \kappa(x(t))$, the objective of this paper is to provide a rigorous certificate that the closed-loop system $f(x(t), \kappa(x(t)))$ is safe in the following sense: Given a set of initial conditions $\mathcal{I}$ and a set of unsafe states $\mathcal{U}$, determine whether there exists an initial condition $x(0) \in \mathcal{I}$ such that the MPC controller forces the closed-loop states to enter $\mathcal{U}$. If such an initial condition exists, the controller is not safe as it eventually forces the closed-loop system to violate design specifications. On the other hand, if no such $x(0) \in \mathcal{I}$ exists, the controller is deemed safe since the set of unsafe states will never be entered by the closed-loop system.

Such a safety verification task can be tackled mainly in two ways. The first set of approaches is based on so-called barrier certificates [12, 15], which are closely related to the concept of Lyapunov functions used for stability analysis. The downside of such approaches is that construction of the barrier certificates is usually achieved via convex relaxations to obtain a computationally tractable problem, and is thus conservative. Therefore such approaches might fail at finding the desired safety certificate even if one exists.

The second set of methods is based on reachability analysis where one investigates whether the set $\mathcal{U}$ is *reachable* by the closed-loop system from a given set of initial conditions $\mathcal{I}$ [4, 13, 1]. The reachability analysis is typically performed by computing forward reachable sets, followed by determining whether the intersection between such reachable sets and the set of unsafe states is empty or not. The reachability-based procedure has several limitations, though. First, and most importantly, it assumes that the analytic description of the closed-loop dynamics is known. MPC strategies, however, only describe the optimal control inputs *implicitly* as the optimal solution to an optimal control problem. Hence the analytic form of the closed-loop dynamics is not directly available. One way around this issue is to derive the *explicit* solution of MPC [3] by employing parametric optimization [5]. Such solutions, however, are often very complex and difficult to construct especially for problems of large dimensionality.

In this paper we take a different route which allows to investigate closed-loop systems without the need to compute the underlying explicit solution. Specifically, we show how to represent the closed-loop evolution of the feedback system by the Karush-Kuhn-Tucker (KKT) [6] conditions of the MPC optimization problem. However, even for MPC problems based on linear prediction models and with all constraints being linear, the KKT conditions are nonlinear. Therefore we show how to convert such nonlinearities, in a non-conservative manner, to linear inequalities which involve continuous and binary decision variables. This allows us to provide a non-conservative answer to the safety verification problem.

The second limitation of reachability-based approaches to safety verification is that they require computing either exact or approximate reachable sets [14]. Computation of exact reachable sets is expensive in large dimensions. Approximate sets [2, 13] are easier to construct, but may lead to conservative safety certificates. In this paper the construction of reachable sets is avoided altogether. Instead, the safety verification problem is posed as a series of mixed-integer linear program (MILP) of tractable size.

Finally, the common drawback of reachability-based approaches is that they only provide a certificate of safety for a finite number of time steps. In this paper we show that under mild assumptions on the terminal set included in the MPC problem, safety can be verified *ad infinitum*, i.e., for an infinite number of time steps.

II. Problem Statement

We consider the control of discrete-time linear time-invariant (LTI) systems of the form

$$x(t+1) = Ax(t) + Bu(t), \tag{1}$$

J. Holaza, B. Takács, and M. Kvasnica are with the Slovak University of Technology in Bratislava, Slovakia, {juraj.holaza, balint.takacs,michal.kvasnica}@stuba.sk. S. Di Cairano is with Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, dicairano@ieee.org.

with the state vector $x \in \mathbb{R}^{n_x}$ and the input vector $u \in \mathbb{R}^{n_u}$. The system is subject to polyhedral constraints

$$x \in \mathcal{X},\ u \in \mathcal{Z}, \tag{2}$$

where $\mathcal{X} \subset \mathbb{R}^{n_x}$ and $\mathcal{Z} \subset \mathbb{R}^{n_u}$ contain the origin in their respective interiors. The constrained finite-time optimal control problem for the prediction model in (1) is given by

$$U_{\text{ol}}^\star = \arg\min\ x_N^T Q_N x_N + \sum_{k=0}^{N-1} x_k^T Q_x x_k + u_k^T Q_u u_k \tag{3a}$$

$$\text{s.t. } x_{k+1} = Ax_k + Bu_k,\ k = 0, \dots, N-1, \tag{3b}$$

$$x_k \in \mathcal{X},\ k = 0, \dots, N-1, \tag{3c}$$

$$u_k \in \mathcal{Z},\ k = 0, \dots, N-1, \tag{3d}$$

$$x_N \in \mathcal{X}_f, \tag{3e}$$

where x_k and u_k are, respectively, predictions of states and inputs at the k-th step of the prediction horizon (denoted by N), initialized by x_0, the measurement (or estimate) of the current state. Moreover, $Q_N = Q_N^T \succeq 0$, $Q_x = Q_x^T \succeq 0$ and $Q_u = Q_u^T \succ 0$ denote weighting matrices, and $\mathcal{X}_f \subseteq \mathcal{X}$ is a polyhedral terminal set. Finally, $U_{\text{ol}}^\star$ denotes the open-loop sequence of optimal control moves, i.e., $U_{\text{ol}}^\star = [u_0^{\star T}, \dots, u_{N-1}^{\star T}]^T$, obtained by solving (3) for a particular initial condition x_0.

The receding-horizon implementation of the MPC feedback law is obtained by calculating the open-loop sequence $U_{\text{ol}}^\star$ for a particular initial condition $x_0 = x(t)$ at each sampling step, but only employing its first element, i.e., $u_0^\star$, as the closed-loop control action. Hence, the RHC feedback law $\kappa : \mathbb{R}^{n_x} \to \mathbb{R}^{n_u}$ is given by

$$\kappa(x(t)) = \underbrace{\begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix}}_{\Phi} U_{\text{ol}}^\star(x(t)). \tag{4}$$

Since $U_{\text{ol}}^\star(x(t))$ in (4) is implicitly defined as the solution of the numerical optimization problem (3), we refer to (4) as the *implicitly defined MPC feedback law*.

The closed-loop evolution of (1) subject to the MPC feedback in (4) is defined by

$$x_{\text{cl}}(t+1) = Ax_{\text{cl}}(t) + B\kappa(x_{\text{cl}}(t)). \tag{5}$$

When initialized from $x(0)$, the closed-loop state at any $t > 0$ is

$$x_{\text{cl}}(t) = A^t x(0) + \sum_{i=0}^{t-1} A^{t-i-1} B\kappa(x(i)). \tag{6}$$

The problem we aim at solving is to verify whether the parameters of (3) have been chosen such that the implicitly defined feedback law (4) forces the closed-loop state trajectory (6) *to avoid* a known set of unsafe states. If a trajectory entering the unsafe set exists, the controller is poorly designed as it does not exhibit required safety properties. Moreover, existence of such a unsafe closed-loop trajectory serves as a certificate of lack of safety. If no such trajectory entering the unsafe set exists, the controller is deemed safe. We distinguish between two versions of such a problem. One verifies whether the set of unsafe states can be reached from a given set of initial conditions in finite time:

Problem 2.1 (Finite-time safety verification): Let the LTI system (1), the implicitly defined MPC feedback law (4), the set of investigated initial conditions $\mathcal{I} \subseteq \mathbb{R}^{n_x}$ and the set of unsafe states $\mathcal{U} \subseteq \mathbb{R}^{n_x}$ be given. Moreover, let the integer $M < \infty$ be given. Provide a certificate that $x_{\text{cl}}(t) \notin \mathcal{U}$ for all $t \leq M$ with $x_{\text{cl}}(t)$ as in (6). □

Note that Problem 2.1 can only certify safety of the MPC controller up to $t = M$, but not for $t > M$. Therefore, the second more general and more useful version certifies safety *ad infinitum*, i.e., that the set of unsafe states cannot be reached in a possibly infinite number of time steps:

Problem 2.2 (Infinite-time safety verification): With the same inputs as in Problem 2.1, provide a certificate that $x_{\text{cl}}(t) \notin \mathcal{U}$ for all $t \leq \infty$. □

Two choices of the set of initial conditions $\mathcal{I}$ are typically considered. One option is to choose $\mathcal{I} = \text{dom}(\kappa)$ where $\text{dom}(\kappa)$ is the feasibility set of (3). In such a settings we verify the safety and properties for all feasible initial conditions. Alternatively, $\mathcal{I} \subseteq \text{dom}(\kappa)$, in which case only a subset of the feasible initial conditions is investigated (for instance the typical process operating conditions).

Remark 2.3: The MPC problem (3) could be formulated to directly include the safety constraint $x_k \notin \mathcal{U}$ by imposing $x_k \in \mathcal{X} \backslash \mathcal{U}$ where "\" is the set difference operator. However, such constraints are, in general, non-convex and would make the MPC computationally impossible to solve, especially in real time. Moreover, unless additional conditions are also employed, the avoidance of the set of unsafe states would not be guaranteed *ad infinitum*. Finally, in this work we only aim at verifying whether $x_{\text{cl}}(t) \notin \mathcal{U}$ for a specific range of initial conditions $\mathcal{I}$, not for any $x(0) \in \mathcal{X}$. □

III. Safety Verification

In this section we propose a non-conservative procedure for solving Problems 2.1 and 2.2. The presented technical solution is based on the following assumptions:

Assumption 3.1: The set of initial conditions $\mathcal{I}$ contains at least one point $x_0 \in \mathcal{I}$ which is a feasible initial condition for (3).

Assumption 3.2: The set of unsafe states $\mathcal{U}$ is a convex polyhedron represented by $\mathcal{U} = \{x \in \mathbb{R}^{n_x} \,|\, Sx \leq s\}$. Moreover, $\mathcal{I}$ is also a polyhedron.

Assumption 3.3: The set of unsafe states $\mathcal{U}$ does not intersect the set of initial conditions $\mathcal{I}$, i.e., $\mathcal{U} \cap \mathcal{I} = \emptyset$.

Assumption 3.1 is non-restrictive and merely requires the user to choose the initial set which is consistent with constraints of the MPC problem (3). Assumption 3.2 is required to obtain a computationally tractable and non-conservative solution to the safety verification problems. Finally, Assumption 3.3 is quite natural and not restrictive as it merely excludes inconsistent scenarios where the MPC problem (3) is set up in such a way that it forces violation of safety bounds by starting from the unsafe set directly.

We start by converting the optimal control problem (3) into a quadratic program. With the substitution

$$x_{k+1} = A^k x_0 + \sum_{i=0}^{k-1} A^{k-i-1} B u_i, \tag{7}$$

the *open-loop* profile of predicted states in (3), i.e., $X_{\text{ol}} = [{x_0}^T, \dots, {x_N}^T]^T$ can be compactly written as

$$X_{\text{ol}} = \Gamma x_0 + \Psi U_{\text{ol}}, \tag{8}$$

with

$$\Gamma = \begin{bmatrix} I \\ A \\ A^2 \\ \vdots \\ A^N \end{bmatrix}, \Psi = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ B & 0 & 0 & \cdots & 0 \\ AB & B & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & \cdots & B \end{bmatrix}. \tag{9}$$

With the substitution (8), the open-loop optimal control problem (3) can be rewritten, after straightforward algebraic manipulations [5], into

$$U_{\text{ol}}^\star(x_0) = \arg\min_{U_{\text{ol}}} \frac{1}{2} U_{\text{ol}}^T H U_{\text{ol}} + x_0^T F U_{\text{ol}} \tag{10a}$$

$$\text{s.t. } G U_{\text{ol}} \leq w + E x_0, \tag{10b}$$

which is a strictly convex parametric quadratic program due to the assumption that $Q_u \succ 0$, $Q_N \succeq 0$, and $Q_x \succeq 0$. Moreover, we define

$$X_{\text{cl}}^M = [x_{\text{cl}}(0)^T, \dots, x_{\text{cl}}(M)^T]^T \tag{11}$$

as the *closed-loop* trajectory of (1) subject to the MPC feedback law (4) over M discrete time steps, with $x_{\text{cl}}(0) = x(0)$, and $x_{\text{cl}}(j)$ is given by (6) for each $j = 1, \dots, M$.

In what follows we show how to solve Problems 2.1 and 2.2 based on the assumption that the open-loop profile X_{ol} from (8) is equal to the closed-loop trajectory X_{cl}^N from (11), and $N = M$ in Problem 2.1.

Remark 3.4: Conditions under which $X_{\text{ol}} = X_{\text{cl}}^N$ are elaborated in [11]. One such a condition is that the terminal penalty Q_N, the terminal set $\mathcal{X}_f$, and the prediction horizon N are chosen such that the value function in (3a) is equal to the infinite-horizon value function. As shown, e.g., in [8], such a condition is satisfied if Q_N is the solution to the discret-time algebraic Riccati equation, $\mathcal{X}_f$ is the positively invariant set where the LQR controller satisfies the constraints, and the prediction horizon is sufficiently large. □

Assume that Q_N, $\mathcal{X}_f$ and N have been chosen such that the equivalence between X_{ol} and X_{cl}^N is established per Remark 3.4. Since $X_{\text{ol}} = X_{\text{cl}}^N$ is assumed, the closed-loop state profile X_{cl}^N in (11) is equal to X_{ol} from (8) where the optimal open-loop sequence of control inputs, i.e., $U_{\text{ol}}^\star$, is employed. Introduce $\mathcal{M}_j \in \mathbb{R}^{n_x \times N n_x}$ as

$$\mathcal{M}_j = \begin{bmatrix} 0_{n_x \times (j-1)n_x} & I_{n_x \times n_x} & 0_{n_x \times (N-j)n_x} \end{bmatrix}. \tag{12}$$

Then $\mathcal{M}_j X_{\text{cl}}^N = x(j)$. In other words, the matrix $\mathcal{M}_j$ extracts from the closed-loop state trajectory in (11) its j-th element. Reachability/unreachability of the set of unsafe states $\mathcal{U}$ in exactly j steps with $0 \leq j \leq N$ can then be stated as

$$\text{find } x(0) \tag{13a}$$

$$\text{s.t. } x(0) \in \mathcal{I}, \tag{13b}$$

$$\mathcal{M}_j(\Gamma x(0) + \Psi U_{\text{ol}}^\star) \in \mathcal{U}, \tag{13c}$$

$$U_{\text{ol}}^\star = \arg\min_{U_{\text{ol}}} 1/2 U_{\text{ol}}^T P U_{\text{ol}} + x(0)^T Q U_{\text{ol}} \tag{13d}$$

$$G U_{\text{ol}} \leq w + E x(0), \tag{13e}$$

$$x(0) \in \mathcal{I}, \tag{13f}$$

where (13c) translates to $x(j) \in \mathcal{U}$ via (8) and (12).

Problem (13) is a *bilevel* optimization problem where $U_{\text{ol}}^\star$ in (13c) is the optimal solution of the lower-level problem (13d)−(13f). This lower-level optimization problem implicitly defines the open-loop sequence of control inputs which are optimal for a particular value of the initial condition $x(0)$, which is investigated in the higher-level problem. Note that the higher-level problem, represented by (13a)−(13c), is related to the lower-level problem via $x(0)$. Therefore as $x(0)$ changes in the higher-level problem, a different $U_{\text{ol}}^\star$ will be generated by the lower-level problem and vice versa.

Our first two results provide conditions under which a positive or a negative answer to Problem 2.1 exists.

Theorem 3.5: Let the sets $\mathcal{I}$ and $\mathcal{U}$ satisfy Assumptions 3.1 and 3.3, respectively. If the bilevel problem (13) is feasible for some $j = 1, \dots, N$ then there exists $x(0) \in \mathcal{I}$ such that $x_{\text{cl}}(j) \in \mathcal{U}$ for some $0 < j \leq N$ (i.e., the set $\mathcal{U}$ is reachable from $\mathcal{I}$ in, at most, N steps).

Proof: First note that constraints (13b) and (13d)−(13f) can always be satisfied by a suitable choice of $x(0)$ due to Assumption 3.1. Therefore feasibility of (13) depends only on feasibility of (13c). Since $X_{\text{ol}} = X_{\text{cl}}^N$ is assumed, the j-th open-loop predicted state x_j is equal to the actual closed-loop state $x_{\text{cl}}(j)$. Thus (13c) translates to $x_{\text{cl}}(j) \in \mathcal{U}$ due to (12). Therefore if (13) is feasible for some value of j, we have that $x(0) \in \mathcal{I}$ and $x_{\text{cl}}(j) \in \mathcal{U}$, which shows that feasibility of (13) implies reachability of $\mathcal{U}$ from $\mathcal{I}$. ■

Theorem 3.5 provides a way for finding the counter-example for the safety properties investigated in Problem 2.1. Such a counter-example is represented by the existence of the initial condition $x(0)$, along with the number of time steps j the closed-loop system takes to reach the set of unsafe states.

The following result is a direct corollary of Theorem 3.5 and establishes conditions under which $\mathcal{U}$ cannot be reached from $\mathcal{I}$ in, at least, N steps, and thus provides a positive certificate of controller's safety according to Problem 2.1:

Corollary 3.6: If the bilevel problems (13) is infeasible for $j = 1, \dots, N$, then there *does not* exist any $x(0) \in \mathcal{I}$ for which $x_{\text{cl}}(t) \in \mathcal{U}$ for some $t \leq N$, i.e., the set $\mathcal{U}$ is *not* reachable from $\mathcal{I}$ in, at least, N steps, thus $x_{\text{cl}}(t) \notin \mathcal{U}$ for $t \leq N$. ■

Next we show that the infinite-time safety verification task of Problem 2.2 can be answered in finite time providing the following assumption hold:

Assumption 3.7: The terminal set $\mathcal{X}_f$ in (3e) is a positively invariant set with $\mathcal{U} \cap \mathcal{X}_f = \emptyset$. □

Existence of a positive invariant terminal set is a standard assumption in MPC to obtain closed-loop stability guarantees. Here, in addition we required that the terminal set is chosen not to intersect the unsafe set, i.e., the terminal set is guaranteed to be safe.

Theorem 3.8: Let the sets $\mathcal{I}$ and $\mathcal{U}$ satisfy Assumptions 3.1 and 3.3, and let $\mathcal{X}_f$ in (3e) fullfil the conditions of Assumption 3.7. If the bilevel problems (13) are infeasible for each $j = 1, \dots, N$, then the set $\mathcal{U}$ is *not* reachable from $\mathcal{I}$ in any number (including infinity) of steps, i.e., $x_{\text{cl}}(t) \notin \mathcal{U}$ for all $t > 0$.

Proof: If (13) is infeasible for all $j = 1, \dots, N$, then either $\mathcal{U}$ is unreachable in at least N steps, or it could be reached with more than N steps. The former case is already covered by Corollary 3.6 and thus $x_{\text{cl}}(t) \notin \mathcal{U}$ for $t = 1, \dots, N$. The latter case is impossible under Assumption 3.7. To see this, note that positive invariance of $\mathcal{X}_f$ means that $x_{\text{cl}}(t+k) \in \mathcal{X}_f$ for any $k > 0$ once $x_{\text{cl}}(t) \in \mathcal{X}_f$. Since $X_{\text{ol}} = X_{\text{cl}}^N$ is assumed, x_N (which is equal to $x_{\text{cl}}(N)$) will be contained in the terminal set $\mathcal{X}_f$ via (3e). Thus from positive invariance of the terminal set we have $x_{\text{cl}}(N+k) \in \mathcal{X}_f$ for any $k > 0$. Finally, since $\mathcal{U} \cap \mathcal{X}_f = \emptyset$ by Assumption 3.7, we have that $x_{\text{cl}}(N+k) \notin \mathcal{U}$ for all $k > 0$. Therefore $x_{\text{cl}}(t) \notin \mathcal{U}$ for all $t > 0$. ■

The safety verification taks of Problems 2.1 and 2.2 for the scenario discussed here can thus be solved by determining the feasibility of the bilevel optimization problem (13) for $j = 1, \dots, N$, where N is the prediction horizon in (3). If the problem is feasible for a particular j, the answer to Problems 2.1 and 2.2 is that the set $\mathcal{U}$ of unsafe states is reachable from some $x(0) \in \mathcal{I}$, and the controller is thus not safe. In such a case further values of j need not be considered. Moreover, a feasible solution to the bilevel optimization problem also provides the initial condition which serves as a counter-example to safety verification.

On the other hand, if the bilevel problem (13) are infeasible for all $j = 1, \dots, N$, then the answer to Problem 2.1 is that $\mathcal{U}$ is not reachable, and the controller is thus safe for at least N steps. Answering Problem 2.2 requires that the terminal set satisfies Assumption 3.7. In such a case the infinite-dimensional problem reduces to a finite-dimensional one.

Next we determine feasibility of (13) in a non-conservative fashion by converting it to a mixed-integer linear program. To do so, we first formulate the Karush-Kuhn-Tucker (KKT) conditions [6] of the lower-level problem in (13d)−(13f):

$$HU_{\text{ol}}^\star + F^T x(0) + G^T \lambda = 0, \tag{14a}$$
$$GU_{\text{ol}}^\star \leq w + Ex(0), \tag{14b}$$
$$\lambda \geq 0, \tag{14c}$$
$$\lambda_k (G_k U_{\text{ol}}^\star - w_k - E_k x(0)) = 0, \tag{14d}$$

where (14a) is the stationarity condition, (14b) represents primal feasibility, (14c) is the dual feasibility, and (14d) stands for the complementary slackness condition, which is imposed for $k = 1, \dots, n_c$, where n_c is the number of rows of G. Moreover, G_k, w_k, E_k denote the k-th row of the corresponding matrix. Since the lower-level problem is a strictly convex parametric QP, the KKT conditions (14) are necessary and sufficient [6]. However, they are nonlinear due to product between the Lagrange multipliers λ and the decision variables $U_{\text{ol}}^\star$ in (14d).

Such a nonlinearity can be worked around by realizing that for (14d) to hold, either $\lambda_k = 0$ or $G_k U_{\text{ol}}^\star - w_k - E_k x(0) = 0$ for all $k = 1, \dots, n_c$. One can introduce binary indicators $\delta_k \in \{0, 1\}$ and $\gamma_k \in \{0, 1\}$ such that

$$(\delta_k = 1) \;\Leftrightarrow\; (\lambda_k = 0) \tag{15a}$$
$$(\gamma_k = 1) \;\Leftrightarrow\; (G_k U_{\text{ol}}^\star - w_k - E_k x(0) = 0). \tag{15b}$$

By applying standard rules of propositional logic [16], also known as the *big-M* technique, the equivalences in (15) can be furthermore rewritten into a set of inequalities that are linear in the decision variables λ_k, $U_{\text{ol}}^\star$, δ_k, and γ_k,

$$-Z(1-\delta_k) \leq \lambda_k \leq Z(1-\delta_k), \tag{16a}$$
$$-Z(1-\gamma_k) \leq G_k U_{\text{ol}}^\star - w_k - E_k x(0) \leq Z(1-\gamma_k), \tag{16b}$$

where Z is a sufficiently large constant. It is trivial to verify that if $\delta_k = 1$ in (16a), then $\lambda_k = 0$ is the only feasible value. If $\delta_k = 0$, then (16a) is inactive. Similar reasoning holds for (16b). Then the complementarity slackness condition (14d) can be equivalently written as the propositional logic statement of the form $\delta_k \vee \gamma_k$ (i.e., either the k-th Lagrange multiplier is zero, or the k-th constraint is active), or, equivalently, be written as $\delta_k + \gamma_k \geq 1$. Therefore the KKT conditions (14) can be equivalently written as

$$HU_{\text{ol}}^\star + F^T x(0) + G^T \lambda = 0, \tag{17a}$$
$$GU_{\text{ol}}^\star \leq w + Ex(0), \tag{17b}$$
$$\lambda \geq 0, \tag{17c}$$
$$-Z(1-\delta_k) \leq \lambda_k \leq Z(1-\delta_k), \tag{17d}$$
$$-Z(1-\gamma_k) \leq G_k U_{\text{ol}}^\star - w_k - E_k x(0) \leq Z(1-\gamma_k), \tag{17e}$$
$$\delta_k + \gamma_k \geq 1, \tag{17f}$$

where (17d)−(17f) are imposed for $k = 1, \dots, n_c$.

Abbreviate (17) by $\text{KKT}(x(0), U_{\text{ol}}^\star, \lambda, \delta, \gamma) \leq 0$. Then the bilevel optimization problem (13) can be equivalently written as

$$\text{find } x(0) \tag{18a}$$
$$\text{s.t. } x(0) \in \mathcal{I}, \tag{18b}$$
$$S(\mathcal{M}_j(\Gamma x(0) + \Psi U_{\text{ol}}^\star)) \leq s, \tag{18c}$$
$$\text{KKT}(x(0), U_{\text{ol}}^\star, \lambda, \delta, \gamma) \leq 0, \tag{18d}$$

where (18c) is equivalent to (13c) since $\mathcal{U}$ is assumed to be a polyhedron, cf. Assumption 3.2. Since all constraints are linear (cf., (17)), problem (18) is a mixed-integer feasibility problem with continuous decision variables $x(0) \in \mathbb{R}^{n_x}$, $U_{\text{ol}}^\star \in \mathbb{R}^{N n_u}$, $\lambda \in \mathbb{R}^{n_c}$, and binary decision variables $\delta \in \{0,1\}^{n_c}$ and $\gamma \in \{0,1\}^{n_c}$, where n_c is the number

of constraints of the pQP formulation of the MPC problem in (10).

Remark 3.9: Note that showing safety of the MPC feedback per Corollary 3.6 and Theorem 3.8 relies on infeasibility of (18) for each $j = 1, \dots, N$. To prevent numerical problems which might lead to false indication of infeasibility, we propose to soften the hard constraints (18c) by

$$S(\mathcal{M}_j(\Gamma x(0) + \Psi U_{\text{ol}}^\star)) \leq s + \omega, \tag{19}$$

where $\omega \in \mathbb{R}^{n_S}$ are the slack variables (here, n_S is the number of rows of S in Assumption 3.2), satisfying $\omega \geq 0$. Moreover, the objective (18a) should be replaced by $\min \|\omega\|_1$. Such a modified problem is always feasible. If $\omega = 0$ in the modified problem, (18) is feasible by Assumption 3.1. If $\omega_i > 0$ for at least one component of ω in the modified problem, then (18) is infeasible. □

Finally, we remark that even though the MILP problem (18) is non-convex due to presence of binary decision variables, its feasibility can always be determined in finite time, and the optimal solution of the modified problem per Remark 3.9 can always be found in finite time, e.g. by branch-and-bound methods.

IV. Case Study

A. Four Tanks System

We apply the procedure of Section III to verify safety properties of an MPC controller which governs inflows into a four tank system depicted in Fig. 1. The linearized dynamics of such a system is represented by [7]

$$\dot{x} = \begin{bmatrix} -\frac{1}{T_1} & 0 & \frac{F_3}{(F_1 T_3)} & 0 \\ 0 & -\frac{1}{T_2} & 0 & \frac{F_4}{(F_2 T_4)} \\ 0 & 0 & -\frac{1}{T_3} & 0 \\ 0 & 0 & 0 & -\frac{1}{T_4} \end{bmatrix} x + \begin{bmatrix} \frac{\gamma_1}{F_1} & 0 \\ 0 & \frac{\gamma_2}{F_2} \\ 0 & \frac{(1-\gamma_2)}{F_3} \\ \frac{(1-\gamma_1)}{F_4} & 0 \end{bmatrix} u, \tag{20}$$

with

$$T_i = \frac{F_i}{k_i} \sqrt{\frac{2h_i^{\text{s}}}{g}}, \tag{21}$$

where $x = (h - h^{\text{s}})$, $u = (q - q^{\text{s}})$ are the deviations of states and inputs from respective steady-state values. Here, h_i are the liquid levels of the corresponding tanks, and q_a and q_b are the manipulated inflows. Finally, γ_1 and γ_2 are constants which govern the split of inflows into the lower and upper level tanks In this case study we have used $h_i^s = 0.2\,\text{m}$, $i = 1, \dots, 4$, $q_a^s = q_b^{\text{s}} = 1 \cdot 10^{-4}\,\text{m}^3\text{s}^{-1}$, $F_i = 0.06\,\text{m}^2$, $i = 1, \dots, 4$, $k_1 = 8.7932 \cdot 10^{-4}$, $k_2 = 7.3772 \cdot 10^{-4}$, $k_3 = 6.3495 \cdot 10^{-4}$, $k_4 = 4.3567 \cdot 10^{-4}$, $g = 9.81\,\text{ms}^{-2}$, $\gamma_1 = 0.2$, $\gamma_2 = 0.4$, and discretization of (20) with sampling time of 5 seconds.

B. Control Objective

The control objective is to manipulate the liquid levels in all four tanks to their respective steady-state values (i.e., for the deviation states x_i to reach zero levels), while satisfying state constraints $-0.2 \leq x_i \leq 0.2$, $i = 1, \dots, 4$ (which corresponds to $0\,\text{m} \leq h_i \leq 0.4\,\text{m}$) and input constraints $-1 \cdot 10^{-4} \leq u_j \leq 1 \cdot 10^{-4}$ (which translate to $0\,\text{m}^3\text{s}^{-1} \leq q_j \leq 2 \cdot 10^{-4}\,\text{m}^3\text{s}^{-1}$). This is achieved by devising an MPC feedback strategy which solves (3) with $Q_x = \text{diag}(1, 1, 1, 1)$, $Q_u = \text{diag}(1, 1)$, Q_N equal to the solution of the algebraic Riccati equation, $\mathcal{X}_f$ being the constraint admissible set of the plant in closed-loop with the LQR controller, obtained for the same cost function of the MPC. Finally, $N = 8$. The closed-loop trajectory of the system in (20) subject to the MPC policy (4) is provided in Fig. 2. As can be seen, the response of x_1 and x_2 (which represent levels in the bottom tanks) exhibits a non-minimum phase behavior, which is a consequence of $\gamma_1 < 0.5$ and $\gamma_2 < 0.5$.

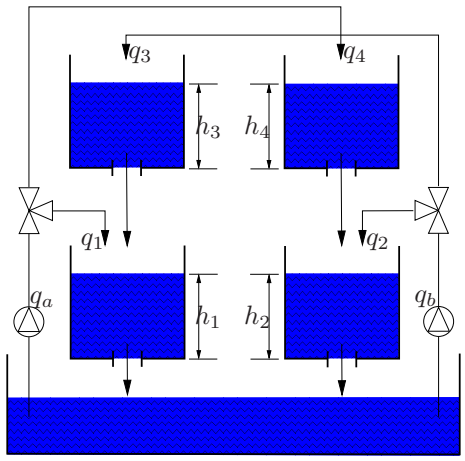
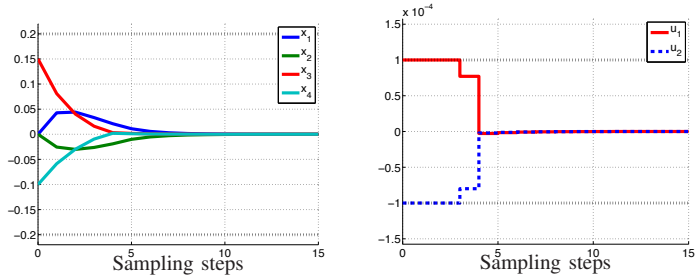


Fig. 1. The four tanks system.



(a) Closed-loop profile of tanks' states. (b) Closed-loop control actions.

Fig. 2. Closed-loop regulation of the four tanks system from the initial condition $x_0 = [0,\ 0,\ 0.15,\ -0.1]^T$. Black dashed lines represent corresponding constraints.

C. Safety Verification

We wish to verify that the MPC policy provides that the overshoots and undershoots in lower tanks due to the non-minimum phase behavior do not exceed prescribed bounds. Specifically, for the set of initial conditions $\mathcal{I} = \{x \mid -0.2 \leq x_{3,4} \leq 0.2,\ x_{1,2} = 0\}$ (i.e., bottom tanks at their steady-state levels with upper tanks being filled up to arbitrary levels within constraints) we aim at verifying that the closed-loop system avoids the sets $\mathcal{U}_1 = \{x \mid x_1 \geq 0.05\}$, $\mathcal{U}_2 = \{x \mid x_1 \leq -0.05\}$, $\mathcal{U}_3 = \{x \mid x_2 \geq 0.05\}$, and $\mathcal{U}_4 = \{x \mid x_2 \leq -0.05\}$. The reasoning behind such a choice is verifying whether the MPC controller rejects disturbances in the upper tanks without large changes of the levels in the bottom tanks.

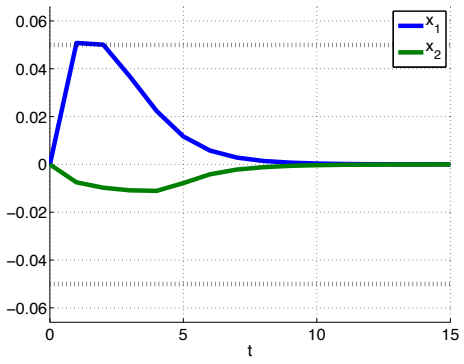


Fig. 3. Example of an unsafe trajectory of lower tanks which starts from $x_0 = [0,\ 0,\ -0.13,\ 0.2]^T$ and violates the limits of maximal under/over-shoots, represented by the dotted black lines.

Remark 4.1: It may not be desirable to include $-0.05 \leq x_{1,2} \leq 0.05$ as hard constraints in (3) since it would render the MPC problem infeasible for several initial conditions. The objective here is to verify whether the MPC policy is tuned in such a way that it "voluntarily" maintains these limits for a *specific* range of initial conditions. □

Since the conditions of Remark 3.4 are satisfied, the open-loop predicted sequence is equal to the actual closed-loop response, thus we can use the procedures of Section III to solve the infinite-time verification task of Problem 2.2 by employing Theorems 3.5 and 3.8. Specifically, we have formulated (18) using YALMIP [10] and solved the resulting MILPs by CPLEX. After 1.0 seconds[1] a feasible solution to (18) was found which, according to Theorem 3.5 means that the safety specifications are violated. The associated counter-example is represented by the initial condition $x_0 = [0,\ 0,\ -0.13,\ 0.2]^T$, for which the MPC feedback forces x_2 to exceed 0.05, as can be seen in Fig. 3

V. Conclusion

We have proposed non-conservative methods which allow to verify whether a closed-loop system, composed of a linear controlled plant and an MPC controller, avoids a certain set of unsafe states. The procedure was based on exploiting the Karush-Kuhn-Tucker optimality conditions, which characterize the optimal control inputs. Subsequently, an optimization problem was set up to determine whether there exists an initial condition for which the closed-loop response enters the unsafe set. The safety certificate was then based on infeasibility of such a problem. Under mild conditions on the terminal set employed in the MPC setup, safety properties can be verified to hold *ad infinitum*, i.e., for an infinite number of time steps. Moreover, we have proposed an alternative formulation in which infeasibility is replaced by feasibility of softened constraints. The resulting safety verification problem to be solved is a mixed-integer linear problem for which efficient solvers exist.

[1]On a 1.8 GHz CPU running Matlab R2013a.

Acknowledgments

J. Holaza, B. Takács and M. Kvasnica gratefully acknowledge the contribution of the Scientific Grant Agency of the Slovak Republic under grant 1/0403/15 and the contribution of the Slovak Research and Development Agency under the project APVV 0551-11. This research was supported by Mitsubishi Electric Research Laboratories, under a Collaborative Research Agreement.

References

[1] E. Asarin, T. Dang, and O. Maler. The d/dt tool for verification of hybrid systems. In *Computer Aided Verification*, pages 365–370. Springer, 2002.

[2] A. Bemporad, C. Filippi, and F. D. Torrisi. Inner and outer approximation of polytopes using boxes. *Computational Geometry: Theory and Applications*, 27(2):151–178, 2004.

[3] A. Bemporad, M. Morari, V. Dua, and E.N. Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 38(1):3–20, January 2002.

[4] A. Bemporad, F. D. Torrisi, and M. Morari. Optimization-Based Verification and Stability Characterization of Piecewise Affine and Hybrid Systems. In B. H. Krogh and N. Lynch, editors, *Proc. of the Intern. Workshop on Hybrid Systems: Computation and Control*, volume 1790, pages 45–58, Pittsburgh, USA, March 2000. Springer-Verlag.

[5] F. Borrelli. *Constrained Optimal Control of Linear and Hybrid Systems*, volume 290. Springer-Verlag, 2003.

[6] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[7] I. Drca. Nonlinear model predictive control of the four tank process, 2007.

[8] P. Grieder, F. Borrelli, F. D. Torrisi, and M. Morari. Computation of the constrained infnite time linear quadratic regulator. *Automatica*, 40:701–708, 2004.

[9] Maciejowski J.M. *Predictive Control with Constraints*. Prentice-Hall, 2001.

[10] J. Löfberg. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. In *Proc. of the CACSD Conference*, Taipei, Taiwan, 2004. Available from http://users.isy.liu.se/johanl/yalmip/.

[11] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, June 2000.

[12] S. Prajna and A. Jadbabaie. Safety verification of hybrid systems using barrier certificates. In *Hybrid Systems: Computation and Control*, pages 477–492. Springer, 2004.

[13] O. Stursberg and B.H. Krogh. Efficient representation and computation of reachable sets for hybrid systems. In *Hybrid Systems: Computation and Control*, pages 482–497. Springer, 2003.

[14] F. D. Torrisi. *Modeling and Reach-Set Computation for Analysis and Optimal Control of Discrete Hybrid Automata*. Dr. sc. thesis, ETH Zurich, Zürich, Switzerland, March 2003.

[15] P. Wieland and F. Allgöwer. Constructive safety using control barrier functions. In *Proceedings of the 7th IFAC Symposium on Nonlinear Control Systems*, pages 462–467, 2007.

[16] H.P. Williams. *Model Building in Mathematical Programming*. John Wiley & Sons, Third Edition, 1993.