



Certified Inductive Synthesis for Online Mixed-Integer Optimization

Marco Zamponi
IMT School for Advanced Studies
Lucca, Italy
marco.zamponi@imtlucca.it

Daniele Masti
GSSI Gran Sasso Science Institute
L'Aquila, Italy
daniele.masti@gssi.it

Emilio Incerto
IMT School for Advanced Studies
Lucca, Italy
emilio.incerto@imtlucca.it

Mirco Tribastone
IMT School for Advanced Studies
Lucca, Italy
mirco.tribastone@imtlucca.it

Abstract

In fields such as autonomous and safety-critical systems, online optimization plays a crucial role in control and decision-making processes, often requiring the integration of continuous and discrete variables. These tasks are frequently modeled as mixed-integer programming (MIP) problems, where feedback data are incorporated as parameters. However, solving MIPs within strict time constraints is challenging due to their \mathcal{NP} -complete nature. A promising solution to this challenge involves leveraging the largely invariant structure of these problems to perform most computations offline, thus enabling efficient online solving even on platforms with limited hardware capabilities. In this paper we present a novel implementation of this strategy that uses counterexample-guided inductive synthesis to split the MIP solution process into two stages. In the offline phase, we construct a mapping that provides feasible assignments for binary variables based on parameter values within a specified range. In the online phase, we solve the remaining continuous part of the problem by fixing the binary variables to the values predicted by this mapping. Our numerical evaluation demonstrates the efficiency and solution quality of this approach compared to standard mixed-integer solvers, highlighting its potential for real-time applications in resource-constrained environments.

CCS Concepts

• **Mathematics of computing** → **Mixed discrete-continuous optimization.**

Keywords

mixed-integer programming, counter-example guided inductive synthesis, cyber-physical systems, online optimization

ACM Reference Format:

Marco Zamponi, Emilio Incerto, Daniele Masti, and Mirco Tribastone. 2025. Certified Inductive Synthesis for Online Mixed-Integer Optimization. In *ACM/IEEE 16th International Conference on Cyber-Physical Systems (with CPS-IoT Week 2025) (ICCPs '25)*, May 6–9, 2025, Irvine, CA, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3716550.3722014>



This work is licensed under a Creative Commons Attribution 4.0 International License. *ICCPs '25, Irvine, CA, USA*

© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1498-6/2025/05
<https://doi.org/10.1145/3716550.3722014>

1 Introduction

Mathematical programming is a powerful tool to represent decision-making problems. Mixed-integer programming (MIP), in particular, has several applications in a wide spectrum of domains including job scheduling [23, 38], cyber-physical systems [9], and neural network verification [44]. Unfortunately, finding an optimal solution for MIP problems is known to be \mathcal{NP} -complete [47]. Exact techniques to solve this problem usually rely on branch-and-bound (B&B) algorithms, which explore the discrete variable state space efficiently by pruning nodes associated with infeasible or low-quality solutions [11]. Various techniques have been developed to improve the performance of B&B [32], including relaxation methods for lower bounds computation [18] and decomposition into smaller subproblems [21, 4]. Despite the significant improvements [2], however, solving MIP problems is still challenging.

The issue is especially problematic in those applications where a MIP problem is solved *online*, i.e., when its solution has to be delivered within a precise time window, as often required to ensure performance, reliability, and safety of a system [7]. In this context, much work has been devoted to developing schemes to accelerate finding (possibly sub-optimal) solutions, exploiting either some regularity of the problem [6] or domain-specific knowledge [36]. Indeed, in many real-time applications, one must repeatedly solve a problem with a fixed structure that depends on external (feedback) information, represented as a set of *parameters* for the problem, as seen in control applications for robotics [46] or autonomous driving [39], and also in state-space models of unknown dynamical systems extracted from data through system identification techniques [28, 17]. In particular, the approach in [17] constructs hybrid dynamical models, which are the ones considered in this paper. This structure can be exploited to partially precompute the solution in an *offline* phase, significantly reducing the operations to be carried out within real-time constraints.

Within the domain of cyber-physical system control, these approaches are commonly referred to as explicit Model Predictive Control [3] (MPC). In its original formulation, this technique exploited multi-parametric optimization arguments [37] to compute the explicit map between the external input information (the “parameters”) and the optimal decision associated with them [3]. Explicit MPC, however, suffers from the curse of dimensionality and is therefore impractical for most applications [5], although related

works have been devoted to reducing its computational and storage burden by approximating the exact map [8, 25].

A related line of research has focused on constructing a mapping between parameters and *feasible* (though not necessarily optimal) assignments of binary variables [26], allowing numerical solvers to compute the remaining real-valued decision variables during the online phase. These feasible, potentially sub-optimal solutions can be used either to "warm-start" the mixed-integer solvers or fixed directly in the optimization problem. This approach leverages the fact that many optimization problems, such as linear and quadratic programs, can be solved efficiently [13, 33]. Numerous studies have explored the use of machine learning to approximate solutions for these types of optimization problems. For example, [27] uses a neural network to learn approximate binary solutions for control tasks, while [10] predicts both binary decision variables and the set of active constraints, and [12] focuses on predicting the set of relaxed big-M constraints. However, these machine learning approaches often lack a formal soundness guarantee, necessitating fallback strategies to handle cases where the predictor may produce invalid outputs (or *hallucinate*). In most studies, when provided, fallback strategies rely heavily on domain-specific knowledge. For instance, [30] employs a rule-based approach drawn from established power-engineering principles to handle failures in their ML-based decision-making engine for microgrid applications.

In the context of functional synthesis, we cite Counter-Example Guided Inductive Synthesis (CEGIS), a family of approaches that has achieved significant success due to its ability to certify specification requirements, a critical aspect where machine learning approaches often fall short [15]. This iterative, data-driven approach generates functions that meet specified preconditions and postconditions using input-output pairs. CEGIS alternates between synthesis and verification steps: the synthesizer generates a candidate function from a set of template functions using a sample of input-output pairs, while the verifier checks if the function satisfies the required postconditions for all valid inputs. If the function is verified, the process concludes; otherwise, a counterexample that guides the next synthesis step is identified. CEGIS has been successfully applied across various domains of system verification and control, including the synthesis of controllers [40], Lyapunov functions [1], barrier certificates [35], and control Lyapunov functions [29].

In this paper, we present a novel application of CEGIS to synthesize an explicit map that links the input parameters of non-convex mixed-integer nonlinear programming problems with polynomial constraints to feasible binary variable assignments. The proposed approach provides a feasibility certificate for the output of the synthesized predictor, addressing the limitations of existing machine learning-based approximations that lack such guarantees. This map serves as the basis for a suboptimal decision-making engine, significantly reducing the computational load during the online phase. To our knowledge, this is the first application of CEGIS for generating a feasible binary solution map for MIP problems.

The proposed procedure is schematically depicted in Figure 1. The figure illustrates a two-phase approach for solving mixed-integer programming (MIP) problems, where the computational burden of real-time MIP solving is alleviated by pre-computing part of the problem offline. In the **offline phase** (left side of the figure), we initialize the function (or map) Δ^* that associates input

parameters belonging to a domain to feasible binary variable assignments and we refine it iteratively through cycles of synthesis and verification steps until it is verified. In the **online phase** (right side of the figure), the *certified* function Δ^* is evaluated for the given parameter values and the binary variables in the MIP problem are fixed to the function output, reducing it to an optimization problem over only continuous variables, for which the numerical solver quickly computes a *sub-optimal solution*.

The remainder of the paper is organized as follows: Section 2 introduces the mathematical concepts and notation. In Section 3, we describe our method. In particular, Section 3.1 describes the offline phase of our procedure, which is meant to construct our binary variables predictor; Section 3.2 details how such a function is exploited online. Section 4 evaluates the method using simple numerical examples to assess its practical advantages and limitations compared to techniques for mixed-integer optimization proposed in related works. Section 5 discusses the limitations of the proposed procedure, along with some possible mitigation strategies.

2 Preliminaries

We consider the following multi-parametric mixed-integer nonlinear programming problem (mp-MINLP):

$$\begin{aligned} (x_\theta^*, \delta_\theta^*) &:= \arg \min_{x, \delta} f(\theta, x, \delta) \\ \text{s.t.} \quad &g(\theta, x, \delta) \leq 0, \end{aligned} \quad (1)$$

where $\theta \in \Theta \subseteq \mathbb{R}^{N_\theta}$, is the parameter vector, $x \in \mathbb{R}^{N_x}$ is the continuous decision variable vector, and $\delta \in \{0, 1\}^{N_\delta}$ is the binary decision variable vector, with $N_\theta, N_x, N_\delta \in \mathbb{N}$ being the dimensions of corresponding vectors. The function f denotes a nonlinear objective function, while g is a vector-valued polynomial function representing problem constraints, with the inequality evaluated element-wise. Problem (1) is said to be feasible for a given parameter θ if there exists an assignment of decision variables (x, δ) that satisfies the constraints g . To simplify the exposition, we introduce a predicate, $feas(\theta)$, representing this condition, defined as follows

$$feas(\theta) := \exists x, \delta : g(\theta, x, \delta) \leq 0.$$

Additionally, we introduce the optimization problem in which binary decision variables are fixed to a constant value $\hat{\delta}$. This problem belongs to the class of multi-parametric nonlinear programming (mp-NLP) problems:

$$\begin{aligned} x_\theta^* &:= \arg \min_x f(\theta, x, \hat{\delta}) \\ \text{s.t.} \quad &g(\theta, x, \hat{\delta}) \leq 0. \end{aligned} \quad (2)$$

Solving this problem using numerical solvers is simpler than solving (1) as it avoids the need to traverse the discrete state space (e.g., via a B&B strategy [11]) to certify feasibility and optimality.

To distinguish between the feasibility conditions of the original mp-MINLP problem (1) and its counterpart (2), which has fixed binary variable values, we introduce the predicate $feas_{\hat{\delta}}(\theta)$

$$feas_{\hat{\delta}}(\theta) := \exists x : g(\theta, x, \hat{\delta}) \leq 0,$$

where $feas_{\hat{\delta}}(\theta)$ evaluates the feasibility of (2) based solely on the existence of at least one assignment of real decision variables x satisfying the constraints g .

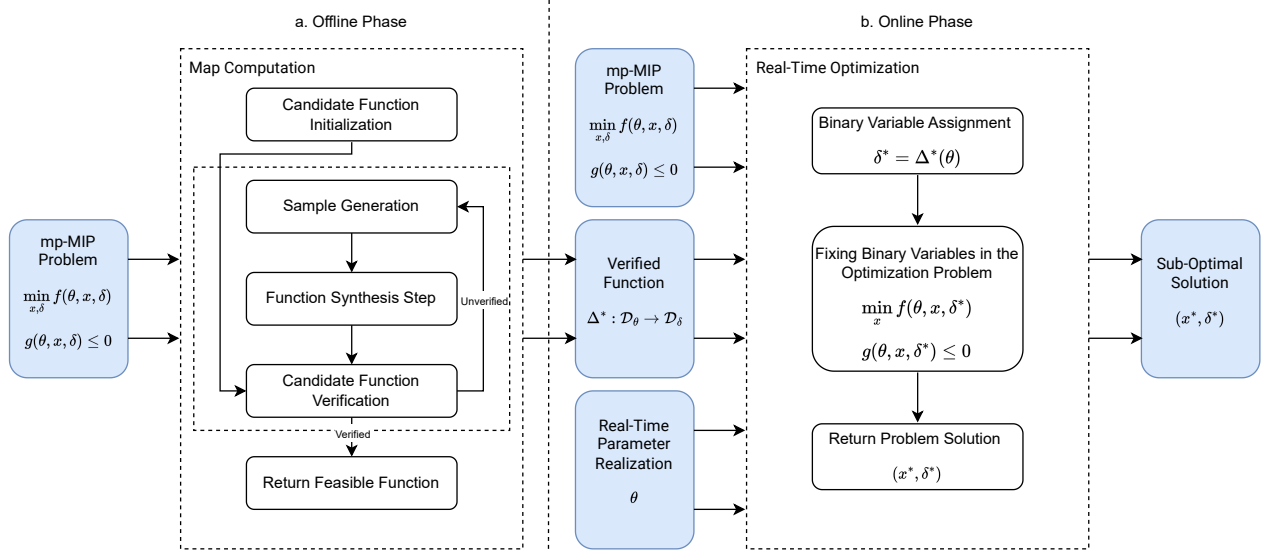


Figure 1: Overview of the proposed two-phase methodology for solving online optimization problems. (a) During the offline phase, a function between real-time parameter realizations and feasible binary variable assignments is synthesized through a CEGIS procedure. This phase is computationally expensive and is executed offline; thus, it does not impact operations during the system execution. (b) In the online phase, the pre-computed map is used to quickly compute feasible values for binary variables, while the remaining continuous decision variables are computed by numerical solvers in real time.

In what follows, we adopt the assumption [26, Assumption 1], where the overlap with respect to problem (1) and the set of feasible parameters Θ is positive, i.e., $\forall \theta \in \Theta$, there exists a neighborhood of the parameter such that its intersection with the feasible region associated with a binary value δ is non-degenerate.

2.1 Running example

The formalism in (1) is commonly used in the context of Model Predictive Control [19] (MPC) to formulate optimal control problems for hybrid dynamical systems. At each time step an optimization problem is solved to compute the optimal decision variables based on the current state of the system and a finite time horizon N . The optimal assignment of decision variables for the controls at the first time step is then applied as the action to the system.

To illustrate our approach, we introduce a running example involving a two-dimensional state-space model predictive control problem that tracks a fixed target point. This example features a circular input dead zone and polyhedral obstacle avoidance, both of which are represented using binary variables. We consider the following two-dimensional piece-wise affine (PWA) controlled dynamical system

$$\begin{bmatrix} p_x(k+1) \\ p_y(k+1) \end{bmatrix} = \begin{bmatrix} p_x(k) \\ p_y(k) \end{bmatrix} + \Delta t \cdot \begin{cases} \begin{bmatrix} 0, 0 \end{bmatrix}^T & \text{if } \|v(k)\|_2 < \underline{v} \\ \begin{bmatrix} v_x(k), v_y(k) \end{bmatrix}^T & \text{otherwise,} \end{cases} \quad (3a)$$

where $\Delta t \in \mathbb{R}$ is the discrete time step length, $p(k) = (p_x(k), p_y(k))$ is the system state representing the position in the plane, and $v(k) = (v_x(k), v_y(k))$ is the input velocity, which is known to have

a *deadzone*. We impose upper and lower bounds on the position

$$p \leq p(k) \leq \bar{p} \quad \forall k, \quad (3b)$$

with $p, \bar{p} \in \mathbb{R}^2$, while we represent the set of possible velocities as

$$\|v(k)\|_2 \in [\underline{v}, \bar{v}] \cup 0 \quad \forall k, \quad (3c)$$

where $\underline{v}, \bar{v} \in \mathbb{R}$. Additionally, we require this system to remain outside a specified polyhedral *exclusion zone*, thus we impose that

$$\begin{aligned} p(k) &\notin \mathbf{O} \quad \forall k, \\ \mathbf{O} &:= \{p \in \mathbb{R}^2 \mid Hp - w \leq 0\}, \end{aligned} \quad (3d)$$

where $H \in \mathbb{R}^{m \times 2}$ and $w \in \mathbb{R}^m$ represent the m half-spaces that define the polyhedron.

By restricting to a finite temporal horizon (i.e., considering only $k = 0, \dots, N$), we transform (3) into an equivalent Mixed Logical Dynamical (MLD) representation involving constraints over binary variables which is more suitable for optimization [9]. This modeling approach introduces auxiliary binary and continuous variables that numerical solvers subsequently optimize. We illustrate this modeling procedure by demonstrating how to convert the constraint (3c). Firstly, we impose an upper bound on the velocity norm:

$$\|v(k)\|_2 \leq \bar{v}. \quad (4)$$

We then introduce a binary variable $\delta_u(k)$ at each time step, which indicates whether $\|v(k)\|_2 < \underline{v}$, using the following constraints

$$M\delta_u(k) \geq \underline{v} - \|v(k)\|_2 \quad (5a)$$

$$M(1 - \delta_u(k)) > \|v(k)\|_2 - \bar{v}, \quad (5b)$$

where $M \in \mathbb{R}$ is a sufficiently large positive constant used to relax constraints. Notably, when $\delta_u(k)$ evaluates to true, the constraints ensure that $\|v(k)\|_2 < v$, while if $\delta_u(k)$ is false, they enforce $\|v(k)\|_2 \geq v$. Note that numerical solvers cannot handle strict inequalities, thus they are approximated as non-strict ones in practice. We can define a real auxiliary variable $z(k) = (z_x(k), z_y(k))$, which is set to $(0, 0)$ when $\delta_u(k)$ is false and to $(v_x(k), v_y(k))$ when $\delta_u(k)$ is true. These conditions are enforced using the following constraints:

$$-M\delta_u(k) \leq z_x(k) \leq M\delta_u(k) \quad (6a)$$

$$-M(1 - \delta_u(k)) \leq z_x(k) - v_x(k) \leq M(1 - \delta_u(k)) \quad (6b)$$

$$-M\delta_u(k) \leq z_y(k) \leq M\delta_u(k) \quad (6c)$$

$$-M(1 - \delta_u(k)) \leq z_y(k) - v_y(k) \leq M(1 - \delta_u(k)). \quad (6d)$$

We can implement the polyhedral exclusion zone constraints (3d) by adding a binary variable for each face of each obstacle at each time step. These auxiliary binary variables are evaluated as true if the system state is outside the corresponding plane of the obstacle. We introduce a binary variable δ_o^i for each of the $i = 1, \dots, m$ half-spaces defining an obstacle o and enforce their value through the following big-M constraints at each time step k :

$$M\delta_o^i(k) \geq H^i p(k) - w^i \quad (7a)$$

$$M(1 - \delta_o^i(k)) > -H^i p(k) + w^i, \quad (7b)$$

here H^i represents the row corresponding to the i -th half-space of the obstacle polyhedron. When the binary variable $\delta_o^i(k)$ is true, the condition $H^i p(k) - w^i > 0$ is enforced. Conversely, if it is false, the condition $H^i p(k) - w^i \leq 0$ holds. Additionally, we include a constraint ensuring that at least one of the binary variables associated with the faces of an obstacle must be true

$$\sum_{i=1}^m \delta_o^i(k) \geq 1. \quad (8)$$

Finally, we rewrite the system dynamics in terms of the newly introduced auxiliary variable:

$$\begin{bmatrix} p_x(k+1) \\ p_y(k+1) \end{bmatrix} = \begin{bmatrix} p_x(k) \\ p_y(k) \end{bmatrix} + \Delta t \cdot \left(\begin{bmatrix} v_x(k) \\ v_y(k) \end{bmatrix} - \begin{bmatrix} z_x(k) \\ z_y(k) \end{bmatrix} \right). \quad (9)$$

Optimal Control Problem. We can now encode our mission as an optimal control problem. We select a quadratic cost function that minimizes both the distance of the state from a target point and the state velocity. In particular, in our implementation, we are tracking the target point $p^T = (0, 0)$. The problem objective function is

$$\sum_{k=0}^{N-1} \begin{bmatrix} p(k) \\ v(k) \end{bmatrix}^T \mathbf{Q} \begin{bmatrix} p(k) \\ v(k) \end{bmatrix} + [p(N)]^T \mathbf{P} [p(N)]. \quad (10)$$

Here, matrices \mathbf{Q} and \mathbf{P} of appropriate dimensions represent weights for the objective function terms.

The resulting optimization problem belongs to the category of non-convex mixed-integer quadratically constrained problem (MIQCP), which is an instance of the optimization problem (1) and

Algorithm 1: CEGIS procedure

Data: P, Θ, Δ^*
Result: Δ^*

```

1  $cex \leftarrow null$ ;
2 do
3    $\Delta^* \leftarrow refine(\Delta^*, P, cex)$ ;
4    $cex \leftarrow verify(\Delta^*, P, \Theta)$ ;
5 while  $cex \neq null$ ;
6 return  $\Delta^*$ 
    
```

is stated as follows

$$(x_{\theta}^*, \delta_{\theta}^*) := \arg \min_{x, \delta} \quad (10) \quad (11a)$$

$$\text{s.t.} \quad (3b) (7) (8) \forall k = 0..N \quad (11b)$$

$$(4) (5) (6) (9) \forall k = 0..(N-1) \quad (11c)$$

$$p(0) = \theta, \quad (11d)$$

where x represents the vector of real decision variables, δ the vector of binary decision variables, and the vector of parameters θ represents the initial position of the system.

3 Methodology

In this section, we outline the methodology used in our approach, covering both the offline and online phases. We begin by describing the construction process for the certified predictor, using the optimization problem (11) as a representative example. Subsequently, we detail the steps of the online phase and discuss the complexity of the transformed optimization problem.

3.1 Offline Phase

The offline phase involves constructing the map, as illustrated in Figure 1(a). Our goal is to generate a function $\Delta^* : \Theta \rightarrow \{0, 1\}^{N\delta}$ that accepts parameters within the parameter set Θ and returns feasible assignments of binary variables for problem (1). A high-level outline of the algorithm is presented in Algorithm 1. The algorithm takes as input a problem instance P having the structure described as in (1), the parameter set Θ , and an initial candidate function Δ^* to be refined. The output is the certified function. In line 3, the *refine* function uses a learner to synthesize a candidate predictor based on pairs of parameter inputs and their corresponding binary optimizers, obtained by solving the optimization problem (1). In line 4, the function *verify* either certifies the correctness of the candidate predictor or identifies a counterexample that violates the desired properties. Figure 2 visually illustrates the steps of the iterative process.

3.1.1 Function Structure. In this work, we consider functions Δ^* that take the form of a piecewise constant map over subsets of the parameter set:

$$\Delta^*(\theta) = \begin{cases} \delta_1 & \theta \in \Omega_1 \\ \vdots & \\ \delta_n & \theta \in \Omega_n. \end{cases} \quad (12)$$

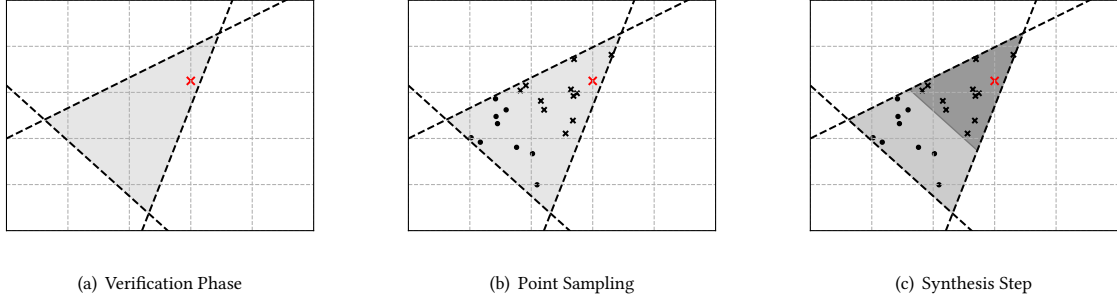


Figure 2: Example iterative verification and refinement process for synthesizing the piecewise function Δ^* . (a) The verification of the domain Ω (filled in gray) results in the identification of a counterexample (marked as a big cross). (b) Points from the domain Ω are randomly sampled for refinement (marked as dots and crosses depending on their optimal binary assignment). (c) The domain Ω is split into two subdomains $\Omega^{(1)}$ and $\Omega^{(2)}$, forming new branches of the piecewise function.

Here, Ω_i , for $i = 1, \dots, n$, partition the parameter set Θ , that is, $\bigcup_{i=1}^n \Omega_i = \Theta$ and $\Omega_i \cap \Omega_j = \emptyset, \forall i \neq j \in \{1, \dots, n\}$. This template function can be refined indefinitely by further subdividing each subset, allowing for the approximation of functions with arbitrary precision. Various machine learning methods, such as decision trees, k-means clustering algorithms [24], or ReLU-activated neural networks [31], can be used to synthesize an initial candidate predictor in this form. Moreover, this template function is composed of disjoint regions, allowing the verification and refinement processes to be conducted independently and in parallel at each branch.

Building on the running example introduced earlier (Example 2.1), we illustrate a valid initial candidate function. In particular, a trivial way to initialize the predictor function could involve constructing a piecewise constant function with only one branch, defined over the entire domain Θ , and assigning a random binary variable δ such that $\Delta^*(\theta) = \delta, \forall \theta \in \Theta$.

3.1.2 Function Verification. The initial candidate function described in the previous section does not provide any formal certification of the feasibility of the returned binary solutions. The critical step in the procedure is the formal verification of the constructed predictor. We assume that every parameter within the domain Θ is feasible, i.e., $feas(\theta) \forall \theta \in \Theta$. It is important to note that this is not a strict assumption, as we can partition the domain of interest into multiple disjoint subdomains and synthesize a predictor for each.

We frame the requirements of our function in terms of pre-conditions and post-conditions, which are predicates that must be satisfied by the function inputs and outputs, respectively. Specifically, we require that the input parameters are contained within the subset of interest $\theta \in \Theta$. Additionally, we require that each output $\Delta^*(\theta)$ maintains the feasibility of the optimization problem. This condition can be expressed with the first-order formula $feas(\theta) \implies feas_{\Delta^*(\theta)}(\theta)$. Since we assumed each parameter θ in the domain is feasible, the implication simplifies to $feas_{\Delta^*(\theta)}(\theta)$, which is easier to verify. Given the desired pre-conditions and post-conditions, we define the first-order formula that must be verified to certify the predictor Δ^* :

$$\forall \theta \in \Theta : feas_{\Delta^*(\theta)}(\theta). \quad (13)$$

We are interested in identifying counterexamples where the formula (13) is not satisfied, as these counterexamples enable a more informed refinement process. Therefore, we query the negation of the formula (13):

$$\exists \theta \in \Theta : \neg feas_{\Delta^*(\theta)}(\theta). \quad (14)$$

If the verifier evaluates formula (14) as true, it will return the parameter θ^c that satisfies the formula, representing a counterexample. Conversely, if the formula is false, we obtain a certificate for the candidate function outputs, allowing us to use our predictor to generate feasible binary variable assignments confidently.

In this work, we focus on certifying piecewise constant functions such as (12), where each region can be verified independently by adjusting the pre-conditions of formula (14) accordingly. Specifically, for a particular branch i of the candidate function, defined in the region Ω_i , we modify the query (14) to search for a counterexample within that specific region:

$$\exists \theta \in \Omega_i : \neg feas_{\delta_i}(\theta). \quad (15)$$

If the formula (15) is evaluated as unsatisfiable, we can mark that branch as verified. Conversely, if a model satisfying the formula is found, it indicates that the branch domain contains at least one parameter associated with an unfeasible variable assignment and will need further refinement. Note that formula (15) includes polynomial terms involving variables θ and x , and consequently belongs to the theory of *Nonlinear Real Arithmetic* (NRA) with quantifiers. This class of satisfiability problem is decidable using procedures for the quantifier elimination [43], ensuring that every query will terminate in a finite number of steps.

Using our running example as a reference, we review the steps involved in verifying a specific branch of the predictor function. Consider the candidate function Δ^* introduced earlier, and suppose we want to verify its sole branch. The adapted formula from (15) for verifying the corresponding region is:

$$\exists \theta \in \Theta : \forall x : \neg((11b) \wedge (11c)),$$

where the constraints (11b) and (11c) are evaluated with fixed δ values given by $\Delta^*(\theta)$.

3.1.3 Function Synthesis. Suppose that the query of the formula (15) applied to a specific branch i of the piecewise function (12) results in a counterexample θ^c . This indicates that the candidate predictor needs to be refined. We conduct the refinement step using data-driven techniques, starting with the generation of a finite subset of parameters $\hat{\Omega}_i$ from the domain associated with the branch Ω_i . We derive a set of observations $C = \{(\theta, \delta_\theta^*) \mid \theta \in \hat{\Omega}_i\}$, where δ_θ^* represents the binary solution of the optimization problem (1). The synthesizer partitions the unverified region Ω_i into disjoint regions $\Omega_i^{(1)}, \dots, \Omega_i^{(n)}$. Each newly created region is assigned its corresponding value $\delta_i^{(j)}$ and is added as a new branch in the function (12), replacing the refined branch.

In the context of running example 2.1, assume that during the verification step of the candidate function Δ^* , a counterexample θ^c within the region Θ was found, necessitating the refinement of the function. We start by constructing a set of observations consisting of pairs $(\theta, \delta_\theta^*)$ for $\theta \in \Theta$, where δ_θ^* denotes the optimal assignment of binary variables obtained by solving problem (11) for an initial state θ . The synthesizer selects an appropriate hyperplane $(a, b) \in \mathbb{R}^2 \times \mathbb{R}$ to act as a splitter (using, for instance, an iteration of a linear classifier algorithm), thereby dividing the set Θ into two disjoint sets and refining the candidate function as follows:

$$\Delta^*(\theta) = \begin{cases} \delta_1 & \text{if } \theta \in \Theta \text{ and } a^T\theta - b \leq 0 \\ \delta_2 & \text{if } \theta \in \Theta \text{ and } a^T\theta - b > 0. \end{cases}$$

3.2 Online Phase

The online phase of the procedure, depicted in Figure 1(b), is responsible for solving the optimization problem during system execution and closely resembles methods used in related works [27, 26, 30]. Given the parameter values measured at a specific time frame and the map computed during the offline phase, we retrieve a suboptimal feasible binary solution, which is fixed within the original problem (11). Numerical solvers are then used to solve the resulting optimization problem for the remaining continuous variables.

3.2.1 Function Evaluation. The first step of the online phase involves evaluating the function $\Delta^*(\theta)$, which means finding the branch of the piecewise function (12) whose domain contains the point θ . An efficient representation for encoding the function Δ^* is a binary search tree [45], which allows us to query the function with $\mathcal{O}(h)$ comparisons, where h is the maximum height of the tree. Since each comparison involves evaluating a dot product in N_θ terms, the total complexity of function evaluation is $\mathcal{O}(N_\theta h)$.

3.2.2 Online Problem Solution. In the subsequent steps of the online phase, the feasible assignment $\delta^* = \Delta^*(\theta)$ is used to compute the complete solution to the problem. By fixing the binary variables in problem (1), we obtain the purely continuous optimization problem (2), which is guaranteed to be feasible. Finding the optimal solution for problem (2) in its general form remains an \mathcal{NP} -complete problem, as it is a non-convex NLP problem. However, since there is no longer a need to explore the discrete solution space for integer variables, we can focus solely on the continuous solution space using gradient-based methods, which significantly reduce the number of required operations. In specific cases, such as when problem (1) is a convex Mixed-Integer Quadratic Program

(MIQP), the continuous problem (2) simplifies to a convex QP. For this class of problems, efficient solvers suitable for embedded applications exist and can offer worst-case polynomial complexity certification [13], effectively reducing the complexity class of the problem during the *online* phase.

4 Experiments

In this section, we evaluate the effectiveness of the proposed online optimization procedure using numerical examples from hybrid system optimal control as benchmarks.

The primary goal of our proposed procedure is to transform the combinatorial optimization problem (1) into a continuous one (2). To verify the effectiveness of our approach, we measure the reduction in solving time between the original problem and its transformed version.

The synthesized function is designed to approximate the optimal binary solution function, similar to approaches in related works employing machine learning techniques [27, 30, 10]. However, those methods cannot guarantee the feasibility of the resulting solutions. To assess this aspect, we first empirically verify that our predictor consistently computes feasible solutions, whereas predictors constructed through machine learning techniques may fail to do so. Next, we compare the suboptimality of the objective function obtained from solving the continuous problem (2) with integer assignments generated by our predictor against the solution obtained by solving the original problem (1).

The experiments are conducted on two distinct problem benchmarks. The first benchmark is based on the previously introduced running example, chosen for its inclusion of nonlinear, non-convex constraints that existing methods cannot manage. The second benchmark extends the first, serving as a simple representative example of more commonly encountered problems.

4.1 Methodology Implementation

We now describe a practical implementation of the algorithm, coded in Python and tested on numerical benchmarks, along with an explanation of the related hyperparameters and their roles. Both the function synthesis procedure and result evaluations were conducted on a machine equipped with an Intel Xeon E7-4830 v4 CPU at 2.0 GHz and 512 GB RAM, running Python 3.12.

The Δ^* function is represented as a binary decision tree, initially constructed using an implementation of the *Classification and Regression Tree* (CART) algorithm with the Gini impurity metric, and trained iteratively until an initial specified height h is reached. To train the initial predictor, we sample N_i points from the parameter space Θ using the stochastic *Latin Hypercube Sampling* (LHS) technique [20] and solve the optimization problem (1) through the Gurobi mixed-integer solver [22], treating the resulting binary optimizers as labels for the supervised learning algorithm. Each leaf node of the constructed decision tree corresponds to a branch of the template function (12), where the predicted value is the most frequent value among the outputs in that node.

To certify the predictor, we use the Z3 SMT solver [16] to evaluate whether the formulas (15) have a satisfying model. The verification process takes a leaf of the binary decision tree as input and constructs the corresponding formulas for evaluation. We include

a small tolerance for formula satisfaction, that is selected to be within the bounds defined by the numerical solvers infeasibility thresholds, ensuring that feasibility guarantees are maintained. To speed up verification, we run multiple solvers in parallel, with each solver instance operating on the same leaf node. The first decisive result from the portfolio of solvers, whether a satisfying model or a certification of its absence, determines the verification outcome. The Z3 SMT solver supports quantified formulas in the theory of Nonlinear Real Arithmetic (NRA), enabling it to handle the class of problems discussed in this work.

If a counterexample is identified during the verification step, we solve the problem at that point and generate N_r additional samples: $N_r/2$ points within a small ε -neighborhood around the counterexample, where ε is adaptively determined as one-sixteenth of the unverified region bounding box length, and $N_r/2$ points from the whole unverified region, using stochastic LHS for both. We then split the current tree leaf by performing one iteration of the CART algorithm on the unverified branch, continuing until all leaf nodes are verified. Although the process can theoretically refine trees infinitely, for numerical stability, we stop the algorithm if the leaf domain becomes too small or if a configured maximum tree height h^m is reached during synthesis. This does not affect the usability of the technique, as sensors have limited bit resolutions and can only represent a finite set of distinct values.

The algorithm implementation is parallelized to improve efficiency. After the initial synthesis step, each leaf node can be processed independently, as the domains of the leaves are disjoint. We initiate multiple processes that take individual nodes as input, performing verification and synthesis steps as needed. These processes return updated nodes, which are then integrated into the binary tree. The refinement tasks continue until the termination conditions are met.

To further enhance the efficiency of constructing the predictor, as a pre-process step, we divide the parameter domain into multiple subdomains, and we synthesize certified predictors over each one of such domains through the process described above. The predictors are then merged into a single decision tree covering the entire domain of interest. Given the nondeterministic nature of the algorithm, we construct multiple trees for each subdomain and select the best one based on the maximum tree height and the number of nodes.

4.2 Nonlinear Controller Obstacle Avoidance

The first benchmark problem involves the optimal control of the dynamical system (3a) introduced earlier as the running example, with the associated MPC optimization problem (11). We consider a time horizon of $N = 3$ time steps, with a time step length of $\Delta t = 1$. The system position is constrained within the region bounded by the lower and upper limits $\underline{p} = (0, 0)$ and $\bar{p} = (1, 1)$ at each time step. Additionally, we impose an upper bound on the velocity norm $\bar{v} = 0.08$ and a lower bound $\underline{v} = 0.03$, below which the input dead zone is activated. Finally, we include a single rectangular obstacle represented by the lower and upper corner points $o^{(l)} = (0.1, 0.4)$ and $o^{(u)} = (0.7, 0.5)$.

4.2.1 Predictor Construction Details. To construct the candidate function, we divide the problem domain into four distinct

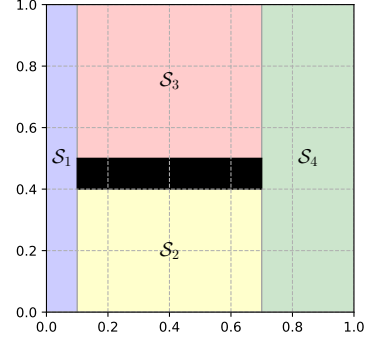


Figure 3: Domain partition for the nonlinear obstacle avoidance problem: the black area represents the infeasible region occupied by the rectangular obstacle, while the remaining states are divided into four feasible regions.

Minimum	Median	Maximum
18.8 %	92.4 %	97.7 %

Table 1: Nonlinear system relative decrease in solving times of problem (2) (fixed binary variables) relative to problem (1).

feasible regions, as shown in Figure 3, and synthesize a predictor for each subdomain. Initially, we sample $N_i = 60$ points to initialize each predictor; during each function refinement phase, we sample an additional $N_r = 20$ points, with half from the neighborhood of the generated counterexample. The decision tree, obtained by merging the best predictors for each subdomain, has a depth of 8 and was constructed using a total of 340 samples.

To establish a baseline for comparison, we compare the synthesized function with traditional machine learning approaches used to approximate the function yielding binary optimizers. Both models were implemented using the scikit-learn library [34]. Specifically, we trained an uncertified decision tree classifier with a maximum height of 8 (DT) and an uncertified random forest classifier consisting of 10 estimators, each with a maximum height of 8 (RF). Both models were trained using a total of 500 samples. These choices are representatives of predictors lacking formal feasibility guarantees with a comparable online evaluation computational complexity.

4.2.2 Results. To evaluate the computational benefits provided by the procedure during online optimization, we sampled 1000 parameters from the feasible domain. We solved both the associated mixed-integer problem (1) (MIP) and the fixed delta problem (2) generated by our certified predictor (CP). Additionally, we considered the continuous problems created by fixing the binary variables to the values predicted by the uncertified decision tree classifier (DT) and the random forest classifier (RF).

In Figure 4, a scatter plot presents the solving times for both the mixed-integer problem and the continuous problem for the same set of parameters. Notably, continuous problems are solved more quickly than mixed-integer ones. This is further confirmed in Table 1, which provides descriptive statistics on the speedup.

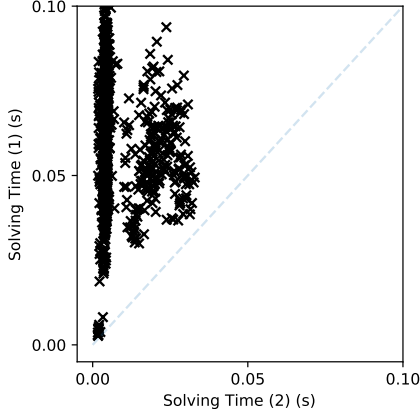


Figure 4: Solving times in the nonlinear system benchmark for optimization problems: (1), on y-axis and (2), with binary variables predicted by Δ^* on x-axis.

Method	# Feasible	# Unfeasible
MIP	1000	0
CP	1000	0
DT	978	22
RF	976	24

Table 2: Nonlinear system feasibility results for continuous problems, comparing the number of feasible instances when fixing binary values predicted by different methods.

We now empirically verify the soundness of the integer assignments returned by our certified predictor (CP) compared to those returned by models trained using unverified machine learning approaches. Table 2 shows the number of infeasible continuous problems constructed by fixing binary variables to the values predicted by the uncertified classifiers. Notably, the certified predictor (CP) maintains feasibility for all the considered parameters. Although the number of infeasible continuous problems from traditional machine learning model predictions may seem low, these occurrences can significantly impact online optimization, where the problem is solved multiple times per second.

Next, we examine the suboptimality of solutions obtained by solving the continuous problems with fixed binary variables compared to the optimal solutions of the mixed-integer problem (1). Relative suboptimality is measured using the formula $S^\theta = (V_{\text{sub}}^* - V_{\text{MIP}}^*)/V_{\text{MIP}}^*$, where V_{MIP}^* is the optimal objective value of problem (1), given that it is considered optimal if the optimality gap with the known lower bound is below 10^{-4} and V_{sub}^* is the objective value of the corresponding continuous problem (2). Table 3 summarizes the results for the various approaches studied, showing the count of suboptimality occurrences within different percentage value ranges for S^θ . The first range, labeled "Opt", corresponds to relative suboptimality values less than 10^{-4} , indicating near-optimal solutions. Notably, in most cases, the certified predictor Δ^* generates assignments of binary variables that result in negligible

Method	# Opt	# <1%	# <5%	# $\geq 5\%$
CP	820	53	40	87
DT	899	39	28	12
RF	905	33	27	11

Table 3: Nonlinear system suboptimality results for continuous problems, comparing the methods by the number of solutions within various suboptimality percentage ranges

suboptimality when constructing the continuous optimization problem. Although other tested machine learning approaches slightly outperform CP in terms of solution quality in the considered benchmark, it is important to note that the certified predictor can still undergo further refinement steps while maintaining feasibility.

4.3 Inertial System with Obstacle Avoidance

For the second benchmark, we extend the optimal control problem (11) previously presented. Specifically, we augment the system state to include the velocity of the object along each axis and control the force applied to the object. To simplify the system, we remove the circular input dead zone while retaining the polyhedral exclusion zone. The system is still tasked with tracking a fixed target point. Consequently, the resulting dynamical system is

$$\begin{bmatrix} p_x(k+1) \\ p_y(k+1) \\ v_x(k+1) \\ v_y(k+1) \end{bmatrix} = \begin{bmatrix} p_x(k) \\ p_y(k) \\ v_x(k) \\ v_y(k) \end{bmatrix} + \Delta t \cdot \begin{bmatrix} v_x(k) \\ v_y(k) \\ a_x(k) \\ a_y(k) \end{bmatrix}, \quad (16)$$

where $\Delta t \in \mathbb{R}$ represents the discrete time step length, $p(k) = (p_x(k), p_y(k))$ denotes the current position in the plane, $v(k) = (v_x(k), v_y(k))$ represents the current velocity, and the newly added term $a(k) = (a_x(k), a_y(k))$ represents the input acceleration. The state position and velocity are bounded:

$$p \leq p(k) \leq \bar{p} \quad v \leq v(k) \leq \bar{v}, \quad (17)$$

with $\bar{p}, \bar{v} \in \mathbb{R}^2$. Additionally, we impose upper and lower bounds on the acceleration:

$$a \leq a(k) \leq \bar{a}, \quad (18)$$

with $\bar{a}, \bar{a} \in \mathbb{R}^2$. The obstacle constraints are encoded using constraints (7) and (8), as in the previously introduced non-inertial dynamical system.

The objective function aims to track the reference point $p^r = (0, 0)$ while minimizing the norm of the velocity and acceleration and is given by

$$\sum_{k=0}^{N-1} \begin{bmatrix} p(k) \\ v(k) \\ a(k) \end{bmatrix}^T \mathbf{Q} \begin{bmatrix} p(k) \\ v(k) \\ a(k) \end{bmatrix} + \begin{bmatrix} p(N) \\ v(N) \end{bmatrix}^T \mathbf{P} \begin{bmatrix} p(N) \\ v(N) \end{bmatrix}. \quad (18)$$

Here, matrices \mathbf{Q} and \mathbf{P} of appropriate dimensions represent weights for the objective function terms.

The resulting optimization problem, which falls into the category of Mixed-Integer Quadratic Programming (MIQP), is formulated as

follows:

$$\begin{aligned}
 (x_{\theta}^*, \delta_{\theta}^*) &:= \arg \min_{x, \delta} \quad (19) \\
 \text{s.t.} \quad &(17) \quad (7) \quad (8) \quad \forall k = 0..N \\
 &(18) \quad (16) \quad \forall k = 0..(N-1) \\
 &p(0), v(0) = \theta,
 \end{aligned}$$

where x denotes the vector of real decision variables, δ represents the vector of binary decision variables, and θ is the vector of parameters, which corresponds to the initial state of the system, including both the initial position and velocity.

In this model, we consider a time horizon of $N = 5$ time steps, with a time step length $\Delta t = 0.2$. The system position is bounded by lower and upper bounds $\underline{p} = (0, 0)$, $\bar{p} = (1, 1)$ at each time step, while velocity within $\underline{v} = (-0.05, -0.05)$, $\bar{v} = (0.05, 0.05)$, and the input acceleration within $\underline{a} = (-0.05, -0.05)$, $\bar{a} = (0.05, 0.05)$. We consider a single rectangular obstacle, represented by the lower and upper points $o^{(l)} = (0.1, 0.4)$, $o^{(u)} = (0.7, 0.5)$.

4.3.1 Predictor Construction Details. We consider a subset of the problem domain Θ consisting of the rectangle enclosed by the points $\underline{\theta} = (0.01, 0.01, -0.01, -0.01)$ and $\bar{\theta} = (0.99, 0.39, 0.01, 0.01)$. We divide the problem domain into the four quadrants of the velocity plane, and synthesize a predictor for each subdomain. Initially, we sample $N_i = 60$ points to initialize each predictor, and during each function refinement phase, we sample an additional $N_r = 20$ points, with half sampled from the neighborhood of the generated counterexample. The decision tree obtained by merging the best predictors for each subdomain has a depth of 10 and was constructed using a total of 900 samples.

For comparison, we also trained an uncertified decision tree classifier (DT) with a maximum height of 10 and an uncertified random forest classifier (RF) composed of 10 estimators, each with a maximum height of 10. Both models were trained using a total of 1000 samples.

4.3.2 Results. We conduct these tests similarly to the nonlinear dynamical system case. Specifically, we sample 1000 parameters from the domain used to construct the verified tree and solve both problem (1) (MIP) and the problem (2). In the latter, integer assignments are generated by our certified predictor (CP), the decision tree classifier (DT), or the random forest classifier (RF). For this test, we use the OSQP solver [41] and its extension for solving MIQP problems [42], which are well-suited for embedded settings where optimization problems need to be solved with limited computational resources, as is common in control systems.

In Figure 5, a scatter plot illustrating the solving times for both the mixed-integer problem 1 and the continuous problem 2 is presented, while Table 4 provides the corresponding descriptive statistics. Notably, the reduction in solving times between the two cases is significant, as the exploration of the binary variable state space is eliminated. It is essential to note that, in this case, the problem 2 is a convex QP instance, thus finding an optimal solution has polynomial complexity [13].

We verify the feasibility of integer assignments from our certified predictor (CP) against uncertified machine learning models. Table 5 shows the count of infeasible instances for binary variables predicted by the considered classifiers. The certified predictor (CP)

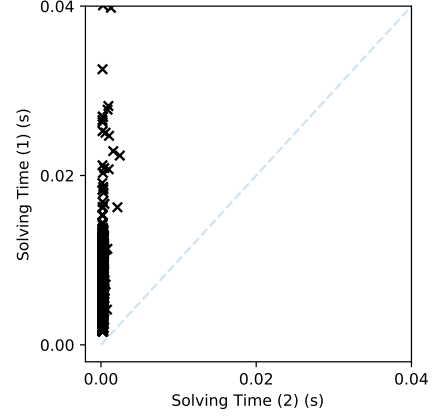


Figure 5: Solving times in the inertial system benchmark for optimization problems: ((1)), on y-axis and ((2)), with binary variables predicted by Δ^* on x-axis.

Minimum	Median	Maximum
81.0 %	96.8 %	99.7 %

Table 4: Inertial system relative decrease in solving times of problem (2) (fixed binary variables) relative to problem (1).

Method	# Feasible	# Unfeasible
MIP	1000	0
CP	1000	0
DT	988	12
RF	994	6

Table 5: Inertial system feasibility results for continuous problems, comparing the number of feasible instances when fixing binary values predicted by different methods.

Method	# Opt	# <1%	# <5%	# $\geq 5\%$
CP	979	10	10	1
DT	981	5	2	0
RF	975	9	6	4

Table 6: Inertial system suboptimality results for continuous problems, comparing the methods by the number of solutions within various suboptimality percentage ranges

maintained feasibility for all parameters, while infeasible predictions occurred with uncertified predictors.

We also assess suboptimality by comparing solutions from the continuous problem with fixed binary variables to those from the optimal mixed-integer problem. Table 6 reports suboptimality counts within different percentage value ranges for S^θ . Notably, the solution quality from the certified predictor is comparable to that of other uncertified machine learning approaches in this case.

5 Limitations

During the verification phase, we evaluate the formula (15), which belongs to the theory of *Nonlinear Real Arithmetic* with quantifiers. This theory is decidable but has doubly exponential time complexity [14]. The Z3 SMT solver used in our implementation employs incomplete heuristic decision procedures for the NRA theory with quantifiers, which may not always converge. To prevent infinite computations, we impose a timeout on the verification process. In such cases, we mitigate the issue by sampling points from the unverified region and applying the refinement cycle again, although this approach lacks termination guarantees.

Regarding algorithm termination, in [26, Definition 4], the concept of overlap for an optimization problem (1) and the parameter space Θ is defined. This condition certifies algorithm divergence if the overlap is zero, indicating that the predictor would require infinite precision. Our procedure cannot synthesize a certified predictor for such problems. One potential mitigation involves detecting these cases and explicitly pre-computing a parameter space partition where each subset has an overlap greater than zero.

Due to the computational complexity of the verification procedure and the curse of dimensionality in constructing the approximator function, our approach is currently suitable only for applications with a limited number of parameters and decision variables. Moreover, at present, constructing a satisfactory predictor requires running the algorithm multiple times and selecting the best one from the results, resulting in a computationally intensive process. Future work could address these limitations by exploring alternative implementations for verifiers and synthesizers, potentially leveraging specific structures inherent in the considered optimization problems. Such improvements would enhance the quality of the predictors, reduce the number of CEGIS iterations required, and ultimately decrease the number of replicas needed to compute a satisfactory predictor.

6 Conclusion

We propose an approach based on the CEGIS method to accelerate mixed-integer nonlinear programming problems by computing a map that generates certified feasible integer solutions. The importance of the synthesized function is demonstrated by the ability to construct purely continuous optimization problems, which are easier to solve, albeit possibly leading to suboptimal solutions. We validate the effectiveness of our method through two numerical examples, highlighting the computational time gains in solving optimization problems and empirically verifying both the soundness of the predictor and the solution quality of the returned binary assignments. Future work will focus on applying post-processing techniques to reduce the number of nodes in the decision tree while maintaining feasibility, thereby improving evaluation times for the function Δ^* , and refining the predictor to enhance the suboptimality of the returned binary assignments, as well as exploring alternative implementations of verification and synthesis strategies.

Acknowledgments

This work has been partially funded by the European Union – NextGenerationEU under the Italian Ministry of University and Research (MUR) through the National Innovation Ecosystem grant

ECS00000041 - VITALITY (CUP: D13C21000430001), the SERICS project (SEcurity and RIghts in the CyberSpace, PE0000014, PNRR M4C2 I.1.3), and the 'Resilienza Economica e Digitale' project (CUP: D67G23000060001), awarded as part of the 'Department of Excellence' initiative (Dipartimenti di Eccellenza 2023-2027, Ministerial Decree no. 230/2022). Additional support was provided by project RDS - PTR22-24 P2.1 Cybersecurity. Daniele Masti is also part of INdAM/GNAMPA.

References

- [1] Alessandro Abate, Daniele Ahmed, Mirco Giacobbe, and Andrea Peruffo. 2020. Formal synthesis of Lyapunov neural networks. *IEEE Control Systems Letters*, 5, 3, 773–778.
- [2] Tobias Achterberg and Roland Wunderling. 2013. Mixed integer programming: Analyzing 12 years of progress. In *Facets of Combinatorial Optimization: Festschrift for Martin Grötschel*. Springer, 449–481.
- [3] Alessandro Alessio and Alberto Bemporad. 2009. A survey on explicit model predictive control. *Nonlinear Model Predictive Control: Towards New Challenging Applications*, 345–369.
- [4] Cynthia Barnhart, Ellis L Johnson, George L Nemhauser, Martin WP Savelsbergh, and Pamela H Vance. 1998. Branch-and-price: Column generation for solving huge integer programs. *Operations research*, 46, 3, 316–329.
- [5] Farhad Bayat, Tor Arne Johansen, and Ali Akbar Jalali. 2011. Flexible piecewise function evaluation methods based on truncated binary search trees and lattice representation in explicit MPC. *IEEE Transactions on Control Systems Technology*, 20, 3, 632–640.
- [6] A Giovanni Beccuti, Tobias Geyer, and Manfred Morari. 2004. Temporal lagrangian decomposition of model predictive control for hybrid systems. In *2004 43rd IEEE Conference on Decision and Control*. Vol. 3. IEEE, 2509–2514.
- [7] Calin Belta and Sadra Sadraddini. 2019. Formal methods for control synthesis: An optimization perspective. *Annual Review of Control, Robotics, and Autonomous Systems*, 2, 1, 115–140.
- [8] Alberto Bemporad and Carlo Filippi. 2006. An algorithm for approximate multi-parametric convex programming. *Computational optimization and applications*, 35, 87–108.
- [9] Alberto Bemporad and Manfred Morari. 1999. Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35, 3, 407–427.
- [10] Dimitris Bertsimas and Bartolomeo Stellato. 2022. Online mixed-integer optimization in milliseconds. *INFORMS Journal on Computing*, 34, 4, 2229–2248.
- [11] Dimitris Bertsimas and Robert Weismantel. 2005. *Optimization Over Integers*. Dynamic Ideas.
- [12] Abhishek Cauligi, Preston Culbertson, Edward Schmerling, Mac Schwager, Bartolomeo Stellato, and Marco Pavone. 2021. Coco: Online mixed-integer control via supervised learning. *IEEE Robotics and Automation Letters*, 7, 2, 1447–1454.
- [13] Gionata Cimini and Alberto Bemporad. 2017. Exact complexity certification of active-set methods for quadratic programming. *IEEE Transactions on Automatic Control*, 62, 12, 6094–6109.
- [14] George E Collins. 1976. Quantifier elimination for real closed fields by cylindrical algebraic decomposition: a synopsis. *ACM SIGSAM Bulletin*, 10, 1, 10–12.
- [15] Cristina David and Daniel Kroening. 2017. Program synthesis: challenges and opportunities. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375, 2104, 20150403.
- [16] Leonardo De Moura and Nikolaj Björner. 2008. Z3: An efficient SMT solver. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer, 337–340.
- [17] Filippo Fabiani, Bartolomeo Stellato, Daniele Masti, and Paul J Goulart. 2025. A neural network-based approach to hybrid systems identification for control. *Automatica*, 174, 112130.
- [18] Marshall L Fisher. 1981. The Lagrangian relaxation method for solving integer programming problems. *Management science*, 27, 1, 1–18.
- [19] Carlos E Garcia, David M Prett, and Manfred Morari. 1989. Model predictive control: Theory and practice—A survey. *Automatica*, 25, 3, 335–348.
- [20] Sushant S Garud, Iftekhar A Karimi, and Markus Kraft. 2017. Design of computer experiments: A review. *Computers & Chemical Engineering*, 106, 71–95.
- [21] Arthur M Geoffrion. 1972. Generalized benders decomposition. *Journal of optimization theory and applications*, 10, 237–260.
- [22] Gurobi Optimization, LLC. 2024. Gurobi optimizer reference manual. (2024).
- [23] Wen-Yang Ku and J Christopher Beck. 2016. Mixed integer programming models for job shop scheduling: A computational analysis. *Computers & Operations Research*, 73, 165–173.

- [24] J MacQueen. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of 5-Th Berkeley Symposium on Mathematical Statistics and Probability/University of California Press*.
- [25] Danylo Malyuta and Behçet Açıkmeşe. 2019. Approximate multiparametric mixed-integer convex programming. *IEEE Control Systems Letters*, 4, 1, 157–162.
- [26] Danylo Malyuta, Behçet Açıkmeşe, Martin Cacan, and David S Bayard. 2019. Partition-based feasible integer solution pre-computation for hybrid model predictive control. In *18th European Control Conference (ECC)*. IEEE, 2776–2782.
- [27] Daniele Masti and Alberto Bemporad. 2019. Learning binary warm starts for multiparametric mixed-integer quadratic programming. In *18th European Control Conference (ECC)*. IEEE, 1494–1499.
- [28] Daniele Masti and Alberto Bemporad. 2021. Learning nonlinear state–space models using autoencoders. *Automatica*, 129, 109666.
- [29] Daniele Masti, Filippo Fabiani, Giorgio Gnecco, and Alberto Bemporad. 2023. Counter-example guided inductive synthesis of control Lyapunov functions for uncertain systems. *IEEE Control Systems Letters*, 7, 2047–2052.
- [30] Daniele Masti, Tomas Pippia, Alberto Bemporad, and Bart De Schutter. 2020. Learning approximate semi-explicit hybrid MPC with an application to microgrids. *IFAC-PapersOnLine*, 53, 2, 5207–5212.
- [31] Guido F Montufar, Razvan Pascanu, Kyunghyun Cho, and Yoshua Bengio. 2014. On the number of linear regions of deep neural networks. *Advances in neural information processing systems*, 27.
- [32] David R Morrison, Sheldon H Jacobson, Jason J Sauppe, and Edward C Sewell. 2016. Branch-and-bound algorithms: A survey of recent advances in searching, branching, and pruning. *Discrete Optimization*, 19, 79–102.
- [33] Yuri Nesterov and Arkadii Nemirovskii. 1994. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM.
- [34] Fabian Pedregosa et al. 2011. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12, 2825–2830.
- [35] Andrea Peruffo, Daniele Ahmed, and Alessandro Abate. 2021. Automated and formal synthesis of neural barrier certificates for dynamical models. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer, 370–388.
- [36] Tomás Pippia, Joris Sijs, and Bart De Schutter. 2019. A single-level rule-based model predictive control approach for energy management of grid-connected microgrids. *IEEE Transactions on Control Systems Technology*, 28, 6, 2364–2376.
- [37] Efstratios N Pistikopoulos, Nikolaos A Diangelakis, and Richard Oberdieck. 2020. *Multi-Parametric Optimization and Control*. John Wiley & Sons.
- [38] Paul Pop, Michael Lander Raagaard, Silviu S Craciunas, and Wilfried Steiner. 2016. Design optimisation of cyber-physical distributed systems using IEEE time-sensitive networks. *IET Cyber-Physical Systems: Theory & Applications*, 1, 1, 86–94.
- [39] Rien Quirynen, Sleiman Safaoui, and Stefano Di Cairano. 2024. Real-time mixed-integer quadratic programming for vehicle decision-making and motion planning. *IEEE Transactions on Control Systems Technology*.
- [40] Hadi Ravanbakhsh and Sriram Sankaranarayanan. 2016. Robust controller synthesis of switched systems using counterexample guided framework. In *Proceedings of the 13th International Conference on Embedded Software*, 1–10.
- [41] Bartolomeo Stellato, Goran Banjac, Paul Goulart, Alberto Bemporad, and Stephen Boyd. 2020. OSQP: An operator splitting solver for quadratic programs. *Mathematical Programming Computation*, 12, 4, 637–672.
- [42] Bartolomeo Stellato, Vihangkumar V Naik, Alberto Bemporad, Paul Goulart, and Stephen Boyd. 2018. Embedded mixed-integer quadratic optimization using the OSQP solver. In *18th European Control Conference (ECC)*. IEEE, 1536–1541.
- [43] Alfred Tarski. 1998. A decision method for elementary algebra and geometry. In *Quantifier Elimination and Cylindrical Algebraic Decomposition*. Springer, 24–84.
- [44] Vincent Tjeng, Kai Xiao, and Russ Tedrake. 2017. Evaluating robustness of neural networks with mixed integer programming. *arXiv preprint arXiv:1711.07356*.
- [45] Petter Tøndel, Tor Arne Johansen, and Alberto Bemporad. 2003. Evaluation of piecewise affine control via binary search tree. *Automatica*, 39, 5, 945–950.
- [46] Shangming Wei, Milos Zefran, and Raymond A DeCarlo. 2008. Optimal control of robotic systems with logical constraints: Application to UAV path planning. In *IEEE International Conference on Robotics and Automation*. IEEE, 176–181.
- [47] Laurence A Wolsey and George L Nemhauser. 2014. *Integer and Combinatorial Optimization*. John Wiley & Sons.