

Online Constraint Tightening in Stochastic Model Predictive Control: A Regression Approach

Alexandre Capone, *Student Member, IEEE*, Tim Brüdigam, *Member, IEEE*,
and Sandra Hirche, *Fellow, IEEE*

Abstract—Solving chance-constrained stochastic optimal control problems is a significant challenge in control. This is because no analytical solutions exist for up to a handful of special cases. A common and computationally efficient approach for tackling chance-constrained stochastic optimal control problems consists of a deterministic reformulation, where hard constraints with an additional constraint-tightening parameter are imposed on a nominal prediction that ignores stochastic disturbances. However, in such approaches, the choice of constraint-tightening parameter remains challenging, and guarantees can mostly be obtained assuming that the process noise distribution is known a priori. Moreover, the chance constraints are often not tightly satisfied, leading to unnecessarily high costs. This work proposes a data-driven approach for learning the constraint-tightening parameters online during control. To this end, we reformulate the choice of constraint-tightening parameter for the closed-loop as a binary regression problem. We then leverage a highly expressive Gaussian process (GP) model for binary regression to approximate the smallest constraint-tightening parameters that satisfy the chance constraints. By tuning the algorithm parameters appropriately, we show that the resulting constraint-tightening parameters satisfy the chance constraints up to an arbitrarily small margin with high probability. Our approach yields constraint-tightening parameters that tightly satisfy the chance constraints in numerical experiments, resulting in a lower average cost than three other state-of-the-art approaches.

Index Terms—Autonomous systems, data-driven control, GPs, machine learning, online learning, optimal control, reinforcement learning, statistical learning, stochastic processes, uncertain systems

I. INTRODUCTION

SYSTEMS subject to constraint requirements are ubiquitous and pose considerable challenges for control design. A tool that has proven effective when dealing with such systems is Model Predictive Control (MPC), an optimization-based control approach that selects control inputs by mini-

mizing a finite horizon objective function subject to input and state constraints. However, although some control applications require specific constraints to hold at all times, in practice, it is seldom possible to guarantee constraint satisfaction if pronounced uncertainties are present, and we can only satisfy constraints with a specific probability. Furthermore, in non-safety-critical applications, it is often acceptable to allow a low probability of constraint violation, e.g., in building climate control [1], process control [2], [3], or power systems [4], [5]. These considerations have motivated the development of so-called Stochastic Model Predictive Control (SMPC) techniques, which allow constraint violations with a pre-specified probability [6]–[8]. This approach to system constraints, commonly referred to as *chance constraints*, significantly reduces conservatism compared to deterministic constraints, as possible worst-case uncertainty realizations do not necessarily need to be considered.

Although SMPC algorithms allow a less conservative approach to system constraints compared to conventional MPC techniques, their development is significantly more challenging. This is because the corresponding Stochastic Optimal Control Problem (SOCP) generally cannot be solved exactly, and approximations must be employed. A commonly encountered approximation consists of so-called scenario approaches, which solve an optimal control problem where the cost and rate of constraint satisfaction are averaged over a finite number of sampled scenarios [9]–[11]. While such methods can approximate the solution of the stochastic optimal control problem arbitrarily accurately, they often require a high number of sample scenarios, resulting in computationally expensive control algorithms. A comparatively inexpensive alternative consists of so-called analytic reformulations [8], which choose control inputs by solving a deterministic optimal control problem (OCP) based on the nominal model with tightened constraints. The tightened constraints are then chosen to guarantee closed-loop satisfaction of the chance constraints.

The most significant challenge in analytic reformulations of chance-constrained stochastic optimal control problems is the choice of constraint-tightening parameters. In particular, too loose values potentially lead to prohibitively frequent constraint violations. Conversely, excessively conservative parameter values will lead to an unnecessarily low number of constraint violations without fully exploiting the flexibility offered by chance constraints. For linear systems with zero-mean normally distributed uncertainties, we can reformulate

This work was supported in part by the European Research Council Consolidator Grant Safe data-driven control for human-centric systems (CO-MAN) under grant agreement number 864686.

Alexandre Capone is with the Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA. Sandra Hirche is with the School of Computation, Information and Technology, Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany. Tim Brüdigam is with the Chair of Automatic Control Engineering, Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany. (e-mail: acapone2@andrew.cmu.edu; tim.bruedigam@tum.de; hirche@tum.de.).

the chance constraints as deterministic constraints for the predicted mean, which allows us to obtain constraint-tightening parameters that satisfy the chance constraints for the closed-loop system [12], [13]. Similarly, for independent and identically distributed (iid) uncertainties with non-Gaussian distributions, Chebyshev inequalities can be applied to compute the worst-case magnitude of the stochastic uncertainty, which in turn allows us to compute constraint-tightening parameters that guarantee the satisfaction of the chance constraints [8], [14]. However, the resulting bound is potentially conservative, as the Chebyshev inequality is coarse for most distributions. Alternatively, sampling-based techniques can be brought to bear to compute the tightened constraints [15], [16]. A drawback of the abovementioned approaches is that an exact description of the system and underlying uncertainty are required. Moreover, unless the distribution of the underlying uncertainty is Gaussian, they rely on conservative estimates of the worst-case error to compute constraint-tightening parameters. This potentially does not fully exploit the flexibility offered by a chance-constrained approach. This issue can be addressed by employing simulation-based methods, where Monte Carlo simulations of the stochastic system model are performed to inform the choice of constraint-tightening parameter [17]–[20]. However, like in the non-simulation-based approaches, an exact system characterization is required to guarantee that the chance constraints are satisfied. Furthermore, full system simulations are computationally expensive and can only be performed for a finite horizon. As an alternative to simulation-based approaches, [21] updates the constraint-tightening parameters during control. However, the method assumes a linear system with strong controllability and reachability properties often unmet in practice. Moreover, it computes the constraint-tightening parameters based on the average number of constraint violations over time, potentially exhibiting slow convergence.

This article explores a purely data-driven approach for obtaining constraint-tightening parameters for analytic reformulations of chance-constrained stochastic optimal control problems. Our approach treats the choice of constraint-tightening parameters in a closed-loop setting as a binary regression problem, where the probability of constraint satisfaction is obtained via the stationary measure of the underlying system. The constraint-tightening parameter is then chosen by minimizing the corresponding entries subject to the chance constraints. Our approach does not require perfect knowledge of the nominal system plant nor the distribution of the underlying uncertainty, provided that the closed-loop system exhibits stable behavior. Furthermore, it is applied online, allowing potential improvements in performance during control.

A. Contribution and Structure

The main contributions of this article are as follows. (i) We provide sufficient conditions under which the choice of constraint-tightening parameter for the analytic reformulation of a stochastic optimal control problem can be treated as a binary regression problem. (ii) We present a GP binary regression-based approach for adapting optimization parameters in

an online fashion during control. To the best of our knowledge, it consists of the first binary regression-based attempt to learn chance constraints. It is considerably easier to implement than more common regression-based approaches, where full system knowledge and simulations are required. (iii) Our main theoretical result shows that our approach yields constraint-tightening parameters that satisfy the chance constraints up to an arbitrarily small value, provided that sufficient data is collected. We reach this result under a few assumptions that are not restrictive in practice. (iv) We compare our approach to different state-of-the-art approaches [8], [14], [15], which require more system knowledge than our approach and often yield conservative constraint-tightening parameters.

The remainder of this article is structured in the following manner. We begin by formally describing the problem considered in this article in Section II. In Section III, we introduce the analytic reformulation-based MPC control law and discuss assumptions regarding the closed-loop system. Section IV introduces GPs for binary regression and describes how they are used to model the probability of constraint satisfaction. We show how to use GP-based binary regression to derive an online update rule for the constraint-tightening parameters in Section V. In Section VI, we employ a numerical simulation to illustrate the proposed method and compare it to other constraint-tightening approaches. Lastly, in Section VII, we provide some concluding remarks. The theoretical proofs of the main results can be found in Appendix - Proof of Theorem 1.

B. Notation

The symbols \mathbb{N} , \mathbb{R} , \mathbb{R}_+ , and \mathbb{R}_- denote the natural, real, non-negative real, and non-positive real numbers, respectively. Plain letters denote scalars. Lowercase/uppercase boldface letters denote vectors/matrices. For a vector $\mathbf{a} \in \mathbb{R}^d$, we use $\|\mathbf{a}\|_2$ to denote its Euclidean norm. For some integer d , the symbol \mathbf{I}_d denotes the d -dimensional identity matrix and $\mathbf{1}_{d_\gamma}$ denotes a d_γ -dimensional vector where every entry is equal to one. For a matrix \mathbf{A} , we employ $\text{eig}_{\max}(\mathbf{A})$ and $\text{eig}_{\min}(\mathbf{A})$ to denote the largest and smallest eigenvalue of \mathbf{A} , respectively. For an open set $\mathcal{O} \subseteq \mathbb{R}^d$, the symbol $\mathbb{I}_{\mathcal{O}}(\cdot)$ denotes the indicator function over the set \mathcal{O} , $\text{vol}(\mathcal{O})$ denotes the volume of \mathcal{O} . The symbol \emptyset denotes the empty set. The symbol $p(A|B)$ denotes the probability density function of event A conditioned on event B . For a set \mathcal{S} , we use $|\mathcal{S}|$ to denote its cardinality; for a scalar a , we use $|a|$ to denote its absolute value. For the states and inputs of a closed-loop system, the subscript $\tau|t$ refers to the prediction at time step $t + \tau$ given the state at time t . The symbol $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. For scalars a_1, \dots, a_N , the expression $\text{diag}(a_1, \dots, a_N)$ corresponds to the N -dimensional diagonal matrix with a_1, \dots, a_N as diagonal entries.

II. PROBLEM SETUP

Consider the stochastic discrete-time system

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \quad (1)$$

where $\mathbf{x}_t \in \mathcal{X} \subseteq \mathbb{R}^{d_x}$ and $\mathbf{u}_t \in \mathcal{U} \subseteq \mathbb{R}^{d_u}$ respectively denote the system's state and control input, and $\mathbf{w}_t \in \mathcal{W} \subseteq \mathbb{R}^{d_w}$ is iid process noise with (potentially unknown) probability density function $p_w(\cdot)$. For the system model for $f(\cdot, \cdot, \cdot)$, we assume to have a linear approximation of the form

$$f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \approx \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t. \quad (2)$$

We also assume that the closed-loop system is required to satisfy chance constraints. In this paper, we aim to satisfy the chance constraints in the long run, i.e., when the transient behavior has passed and the closed-loop system has reached some form of stationarity. This is expressed as

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \sum_{t=0}^T \mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x}_t)) \right] \geq 1 - \delta \quad (3)$$

where $\mathbf{h} : \mathcal{X} \rightarrow \mathbb{R}^{d_c}$ is a function representing nonlinear constraints and $\delta \in (0, 1)$ is a design parameter that specifies the desired risk. Unless specified otherwise, we employ $\mathbb{E}[\cdot]$ to denote the expected value with respect to the random variables \mathbf{w}_t , conditioned on the initial state \mathbf{x}_0 . In Section III, we show how (3) can be rewritten in the standard formulation for chance constraints using the concept of stationary distributions.

Our objective is then to design a control law that (approximately) minimizes the expected cost

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \sum_{t=0}^T l(\mathbf{x}_t, \mathbf{u}_t) \right], \quad (4)$$

with stage cost $l : \mathcal{X} \times \mathcal{U} \rightarrow \mathbb{R}_+$, while satisfying the long-term chance constraints (3).

III. CONTROL STRATEGY AND CLOSED-LOOP BEHAVIOR

In this section, we describe the control law used by our approach and discuss assumptions pertaining to the closed-loop behavior of the system.

A. Model predictive control and backup strategy

An MPC control strategy for stochastic systems typically involves solving a finite horizon stochastic optimal control problem. However, solving a stochastic finite-horizon reformulation of (3) is generally impossible since the linear model (2) seldom corresponds to the exact system dynamics \mathbf{f} and the probability density p_w of the stochastic disturbance \mathbf{w}_t is not known perfectly. Furthermore, computing the corresponding expected cost and chance constraints is often intractable even with perfect knowledge of \mathbf{f} and p_w . To address these issues, we employ a deterministic finite-time horizon reformulation of (3) and (4) with tightened constraints, which we use to find approximately optimal control inputs. This is commonly known as a finite-horizon analytic reformulation of the SOCP [8]. In addition, we employ a relaxed version of the MPC control strategy as a backup control law for points in the state space where the original MPC strategy is infeasible. The backup control law is computed by greedily including slack variables in the optimal control problem until feasibility is recovered. The standard MPC and backup strategies can then be subsumed into a single MPC-based control law. The

corresponding optimal control problem given time step t and state \mathbf{x}_t is

$$\min_{\mathbf{U}_{N-1|t}} \sum_{\tau=0}^{N-1} l(\mathbf{x}_{\tau|t}, \mathbf{u}_{\tau|t}) + l_N(\mathbf{x}_{N|t}) + c_{sl} \|\mathbf{s}_{\tau|t}^i\|_2^2 \quad (5a)$$

$$\text{s.t. } \mathbf{x}_{\tau+1|t} = \mathbf{A}\mathbf{x}_{\tau|t} + \mathbf{B}\mathbf{u}_{\tau|t}, \quad (5b)$$

$$\mathbf{x}_{0|t} = \mathbf{x}_t \quad (5c)$$

$$\mathbf{u}_{\tau|t} \in \mathcal{U}, \quad (5d)$$

$$\mathbf{h}(\mathbf{x}_{\tau|t}) \leq -\mathbf{g}_{i,\tau} + \mathbf{s}_{\tau|t}^i \quad (5e)$$

$$\mathbf{0} \leq \mathbf{s}_{\tau|t}^i, \quad \forall \tau \leq \rho_t \quad (5f)$$

$$\mathbf{0} = \mathbf{s}_{\tau|t}^i, \quad \forall \tau > \rho_t \quad (5g)$$

for all $\tau \in \{0, \dots, N-1\}$, where $c_{sl} > 0$, $\mathbf{s}_{\tau|t}^i$ are slack variables and ρ_t is an integer, selected by solving

$$\rho_t = \min_{\tau} \tau, \quad \text{s.t. (5) is feasible.}$$

The symbol $\mathbf{U}_{N-1|t} := (\mathbf{u}_{0|t}, \dots, \mathbf{u}_{N-1|t})$ denotes the concatenated input sequence and $l_N : \mathcal{X} \rightarrow \mathbb{R}$ the terminal cost function. The constraint-tightening parameters

$$\mathbf{g}_{i,\tau} := \mathbf{g}_{t_i,\tau} \in \mathbb{R}^{d_c}, \quad (6)$$

which serve to compensate for deviations between the true stochastic system (1) and the deterministic model used in (5), are allowed to change at irregular time steps t_i and play a central role in our approach. In general, it is desirable to employ the lowest-possible values of $\mathbf{g}_{i,0}, \dots, \mathbf{g}_{i,N}$ that satisfy the chance constraints. This is because low values for $\mathbf{g}_{i,0}, \dots, \mathbf{g}_{i,N}$ imply a larger feasible region for (5) and, by extension, control inputs that potentially minimize the cost more effectively. In this article, we propose learning $\mathbf{g}_{i,0}, \dots, \mathbf{g}_{i,N}$ from data by iteratively estimating the resulting probability of constraint satisfaction and minimizing a weighted sum of $\mathbf{g}_{i,0}, \dots, \mathbf{g}_{i,N}$ accordingly. This is described detailedly in Section V. For simplicity of exposition, we henceforth subsume the constraint-tightening parameters after the update step t_i as

$$(\mathbf{g}_{i,0}^\top, \dots, \mathbf{g}_{i,N}^\top)^\top =: \boldsymbol{\gamma}_i \in \Gamma,$$

where $\Gamma \subset \mathbb{R}^{d_\gamma}$ is compact and $d_\gamma := Nd_c$.

We denote the minimizer of (5) given $\boldsymbol{\gamma}_i$ as

$$\mathbf{U}_{N-1|t}^*(\mathbf{x}_i, \boldsymbol{\gamma}_i) := (\mathbf{u}_{0|t}^*(\mathbf{x}_i, \boldsymbol{\gamma}_i), \dots, \mathbf{u}_{N-1|t}^*(\mathbf{x}_i, \boldsymbol{\gamma}_i)).$$

Following standard MPC, we recursively solve (5) given the current state \mathbf{x}_t and $\boldsymbol{\gamma}_i$, then apply the first entry $\mathbf{u}_{0|t}^*(\mathbf{x}_t, \boldsymbol{\gamma}_i)$ of $\mathbf{U}_{N-1|t}^*(\mathbf{x}_t, \boldsymbol{\gamma}_i)$ to the system. This implicitly specifies a control law that depends exclusively on \mathbf{x}_t and $\boldsymbol{\gamma}_i$, which we henceforth refer to as

$$\mathbf{u}^{\text{MPC}}(\mathbf{x}_t, \boldsymbol{\gamma}_i) = \mathbf{u}_{0|t}^*(\mathbf{x}_t, \boldsymbol{\gamma}_i). \quad (7)$$

B. Closed-loop dynamics for fixed constraints

In this paper, we do not require to know \mathbf{f} nor p_w . Instead, most of the assumptions leveraged by our approach pertain to the closed-loop dynamics of the system (1) under the control law (7) and a *fixed* vector of constraint-tightening

parameters γ . To introduce and discuss our assumptions, we require the closed-loop system for a fixed γ , given by

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}_t, \mathbf{w}_t) \\ &:= \mathbf{f}(\mathbf{x}_t, \mathbf{u}^{\text{MPC}}(\mathbf{x}_t, \gamma), \mathbf{w}_t). \end{aligned} \quad (8)$$

Furthermore, for any subset $\mathcal{O} \subseteq \mathcal{X}$ and state $\mathbf{x} \in \mathcal{X}$, we employ

$$\Pr(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w}) \in \mathcal{O} \mid \mathbf{x}) = \int_{\mathbb{R}^{d_x}} \mathbb{I}_{\mathcal{O}}(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w})) p_{\mathbf{w}}(\mathbf{w}) d\mathbf{w}$$

to denote the probability of the closed-loop system (8) reaching the set \mathcal{O} from \mathbf{x} in a single time step.

In rough terms, our main requirements are that the closed-loop system converges to a compact set in expectation and does not exhibit periodic behavior for all γ . This is expressed in the following assumption.

Assumption 1: There exists a compact subset $\mathcal{C} \subseteq \mathcal{X}$, a known measurable function $V : \mathcal{X} \rightarrow [1, \infty)$, a probability measure ν on \mathcal{X} , and positive constants $\beta, \tilde{\beta}, K < \infty$ and $\mu < 1$, such that the following conditions hold for all $\gamma \in \Gamma$:

i) $\Pr(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w}) \in \mathcal{O} \mid \mathbf{x}) \geq \tilde{\beta}\nu(\mathcal{O})$ holds for all $\mathbf{x} \in \mathcal{C}$ and all open subsets $\mathcal{O} \subseteq \mathcal{X}$.

ii) $\mathbb{E}(V(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w}))) \leq \begin{cases} \mu V(\mathbf{x}), & \text{if } \mathbf{x} \notin \mathcal{C}, \\ K, & \text{if } \mathbf{x} \in \mathcal{C}. \end{cases}$

iii) $\tilde{\beta}\nu(\mathcal{C}) \geq \beta$.

Condition i) in Assumption 1 requires that the probability distribution of \mathbf{x} be minorized by another probability distribution within the smaller level sets of V . It implies the existence of an open set of positive measure that can be reached from any point in \mathcal{C} with non-zero probability. Condition ii) is similar to drift conditions commonly found in stochastic control literature [22], where V plays a role akin to that of a Lyapunov function. Condition iii) states that a point starting in \mathcal{C} reaches a subset of \mathcal{C} of non-zero measure with probability larger than 0. It implies the closed-loop system does not exhibit periodic behavior with probability one.

Assumption 1 can be easily shown to hold for systems with additive Gaussian noise where the noiseless realization of the system is globally asymptotically stable. i) and iii) are not very restrictive, whereas ii) is similar to a stability condition. Note that $\nu, V, \beta, \tilde{\beta}, K$ and μ are assumed to be identical for all γ , and that the function V is assumed to be known. However, as will be shown later, these requirements are only necessary to derive a potentially conservative upper bound on the convergence speed of the closed-loop system to its stationary measure and are not strictly necessary in practice.

We now provide an example for a system that satisfies Assumption 1.

Example 1: Consider the setting where the closed-loop system equations for an arbitrary fixed $\gamma \in \Gamma$ are given by

$$\mathbf{x}_{t+1} = \mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}_t, \mathbf{w}_t) =: \tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x}_t) + \mathbf{w}_t, \quad (9)$$

where $\tilde{\mathbf{f}}_\gamma^{\text{MPC}} : \mathcal{X} \rightarrow \mathcal{X}$ and \mathbf{w} is normally distributed with zero mean and positive-definite covariance matrix $\Sigma_{\mathbf{w}}$.

Furthermore, let

$$V(\mathbf{x}) = 1 + \frac{1}{2} \mathbf{x}^\top \mathbf{P} \mathbf{x},$$

where \mathbf{P} is a positive-definite matrix, and assume there exists a quadratic matrix $\tilde{\mathbf{A}}$ and positive definite matrix \mathbf{M} , such that $\tilde{\mathbf{A}}^\top \mathbf{P} \tilde{\mathbf{A}} - \mathbf{P} + \mathbf{M}$ is negative-definite, $\mathbf{P} - \mathbf{M}$ is positive-definite, and

$$(\tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x}))^\top \mathbf{P} (\tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x})) \leq \mathbf{x}^\top \tilde{\mathbf{A}}^\top \mathbf{P} \tilde{\mathbf{A}} \mathbf{x} \quad (10)$$

holds for all $\mathbf{x} \in \mathcal{X}$ and all $\gamma \in \Gamma$. Then, the closed-loop system (9) can be shown to satisfy Assumption 1 as follows.

We begin with Condition ii) of Assumption 1. We have

$$\begin{aligned} \mathbb{E}[V(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w}))] &= \mathbb{E}[V(\tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x}) + \mathbf{w})] \\ &\leq \mathbb{E}\left[1 + \frac{1}{2}(\mathbf{x}^\top \tilde{\mathbf{A}}^\top \mathbf{P} \tilde{\mathbf{A}} \mathbf{x} + (\tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x}))^\top \mathbf{P} \mathbf{w} \right. \\ &\quad \left. + \mathbf{w}^\top \mathbf{P} \tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x}) + \mathbf{w}^\top \mathbf{P} \mathbf{w})\right] \\ &= \mathbb{E}\left[1 + \frac{1}{2}(\mathbf{x}^\top \tilde{\mathbf{A}}^\top \mathbf{P} \tilde{\mathbf{A}} \mathbf{x} + \mathbf{w}^\top \mathbf{P} \mathbf{w})\right] \\ &= 1 + \frac{1}{2}(\mathbf{x}^\top \tilde{\mathbf{A}}^\top \mathbf{P} \tilde{\mathbf{A}} \mathbf{x} + \text{tr}(\Sigma_{\mathbf{w}} \mathbf{P})) \\ &\leq 1 + \frac{1}{2}(\mathbf{x}^\top (\mathbf{P} - \mathbf{M}) \mathbf{x} + \text{tr}(\Sigma_{\mathbf{w}} \mathbf{P})) \\ &\leq 1 + \frac{1}{2}(\text{eig}_{\max}(\mathbf{P}^{-\frac{1}{2}}(\mathbf{P} - \mathbf{M})\mathbf{P}^{-\frac{1}{2}}) \mathbf{x}^\top \mathbf{P} \mathbf{x} \\ &\quad + \text{tr}(\Sigma_{\mathbf{w}} \mathbf{P})), \end{aligned} \quad (11)$$

where the last inequality is due to

$$\frac{\mathbf{x}^\top (\mathbf{P} - \mathbf{M}) \mathbf{x}}{\mathbf{x}^\top \mathbf{P} \mathbf{x}} \leq \text{eig}_{\max}(\mathbf{P}^{-\frac{1}{2}}(\mathbf{P} - \mathbf{M})\mathbf{P}^{-\frac{1}{2}}).$$

Define

$$\mu := \frac{1}{2} \left(1 + \text{eig}_{\max}(\mathbf{P}^{-\frac{1}{2}}(\mathbf{P} - \mathbf{M})\mathbf{P}^{-\frac{1}{2}})\right) \quad (12)$$

and note that $\mu < 1$. Moreover, let

$$\mathcal{C} = \left\{ \mathbf{x} \mid \mathbf{x}^\top \mathbf{P} \mathbf{x} \leq \frac{1}{1 - \mu} \text{tr}(\Sigma_{\mathbf{w}} \mathbf{P}) + 2 \right\}. \quad (13)$$

For any $\mathbf{x} \notin \mathcal{C}$, from (11) we then obtain

$$\begin{aligned} \mathbb{E}[V(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w}))] &< 1 + \frac{1}{2}((2\mu - 1)\mathbf{x}^\top \mathbf{P} \mathbf{x} + (1 - \mu)\mathbf{x}^\top \mathbf{P} \mathbf{x} - 2 + 2\mu) \\ &= \mu + \frac{1}{2}(\mu \mathbf{x}^\top \mathbf{P} \mathbf{x}) = \mu V(\mathbf{x}). \end{aligned}$$

Furthermore, for $\mathbf{x} \in \mathcal{C}$, we have

$$\begin{aligned} \mathbb{E}[V(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w}))] &\leq 1 + \frac{1}{2}((2\mu - 1)\mathbf{x}^\top \mathbf{P} \mathbf{x} + \text{tr}(\Sigma_{\mathbf{w}} \mathbf{P})) \\ &\leq 1 + \frac{1}{2} \left(\frac{1}{1 - \mu} \text{tr}(\Sigma_{\mathbf{w}} \mathbf{P}) + 2 \right) =: K. \end{aligned} \quad (14)$$

Hence, Condition ii) of Assumption 1 holds with \mathcal{C} as in (13), μ as in (12), and K as in (14).

We now show that Condition i) holds for ν equal to the uniform distribution on \mathcal{C} , i.e.,

$$\nu(\mathcal{O}) = \frac{1}{\text{vol}(\mathcal{C})} \int_{\mathcal{O} \cap \mathcal{C}} d\mathbf{x}.$$

Note that, for any \mathcal{O} with $\mathcal{O} \cap \mathcal{C} = \emptyset$,

$$\Pr(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w}) \in \mathcal{O} \mid \mathbf{x}) \geq \nu(\mathcal{O}) = 0$$

holds trivially. Hence, we assume $\mathcal{O} \subseteq \mathcal{C}$ without loss of generality. Note that, due to (10) and the definition of \mathcal{C} ,

$$\begin{aligned} \|\tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x})\|_2^2 &\leq \frac{\text{eig}_{\max}(\tilde{\mathbf{A}}^\top \tilde{\mathbf{A}})}{\text{eig}_{\min}(\mathbf{P})} \left(\frac{1}{1-\mu} \text{tr}(\boldsymbol{\Sigma}_w \mathbf{P}) + 2 \right) \\ \|\mathbf{x}\|_2 &\leq \frac{1}{\text{eig}_{\min}(\mathbf{P})} \left(\frac{1}{1-\mu} \text{tr}(\boldsymbol{\Sigma}_w \mathbf{P}) + 2 \right) \end{aligned}$$

holds for all $\mathbf{x} \in \mathcal{C}$. Since p_w is a zero-mean multivariate Gaussian distribution, this implies that there exists a constant $c > 0$, such that

$$c \leq p_w(\tilde{\mathbf{w}} - \tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x})) \quad (15)$$

holds for all $\tilde{\mathbf{w}} \in \mathcal{C}$ and all $\mathbf{x} \in \mathcal{C}$. For any $\mathbf{x} \in \mathcal{C}$, we then have

$$\begin{aligned} \Pr(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w}) \in \mathcal{O} \mid \mathbf{x}) &= \Pr(\tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x}) + \mathbf{w} \in \mathcal{O} \mid \mathbf{x}) \\ &= \int_{\mathcal{X}} \mathbb{I}_{\mathcal{O}}(\tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x}) + \mathbf{w}) p_w(\mathbf{w}) d\mathbf{w} \\ &= \int_{\mathcal{X}} \mathbb{I}_{\mathcal{O}}(\tilde{\mathbf{w}}) p_w(\tilde{\mathbf{w}} - \tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x})) d\tilde{\mathbf{w}} \\ &\geq \text{vol}(\mathcal{O}) \inf_{\mathbf{x}, \tilde{\mathbf{w}} \in \mathcal{C}} p_w(\tilde{\mathbf{w}} - \tilde{\mathbf{f}}_\gamma^{\text{MPC}}(\mathbf{x})) \geq c \text{vol}(\mathcal{O}) \\ &= \frac{c}{\text{vol}(\mathcal{C})} \nu(\mathcal{O}), \end{aligned}$$

hence Condition i) holds for $\tilde{\beta} = \frac{c}{\text{vol}(\mathcal{C})}$. Condition iii) trivially follows with $\beta = \tilde{\beta}$.

Assumption 1 allows us to establish convergence of the distribution of the closed-loop system to a stationary probability measure. Before we state this formally, we introduce the definition of a stationary probability measure for the closed-loop system (8).

Definition 1: Consider the closed-loop dynamical system (8) for a fixed $\gamma \in \Gamma$. A probability measure π_γ on \mathcal{X} is said to be stationary under $\mathbf{f}_\gamma^{\text{MPC}}$ if, for every open subset $\mathcal{O} \subseteq \mathcal{X}$,

$$\int_{\mathcal{X}} \Pr(\mathbf{f}_\gamma^{\text{MPC}}(\mathbf{x}, \mathbf{w}) \in \mathcal{O} \mid \mathbf{x}) \pi_\gamma(d\mathbf{x}) = \pi_\gamma(\mathcal{O}). \quad (16)$$

Intuitively, a stationary measure is a probability measure that does not change under the system dynamics. This means that if the initial condition is distributed according to the stationary measure, then the probability distribution of the state remains constant for all future time steps. Given Assumption 1, it can be shown that such a unique stationary measure exists for the closed-loop system (8) and that the distribution of the state converges to that of the stationary distribution. This is stated in the following.

Lemma 1 (23, Theorem 1.1): Let Assumption 1 hold. Then, for every $\gamma \in \Gamma$, the closed-loop system (8) has a unique stationary probability measure π_γ . Furthermore, there exist positive constants $\varphi \in (0, 1)$ and $\vartheta \in \mathbb{R}_+$ that are independent of γ , such that

$$\left| \mathbb{E} \left[\mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x}_t)) \right] - \mathbb{E}_{\pi_\gamma} \left[\mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x})) \right] \right| \leq \vartheta V(\mathbf{x}_0) \varphi^t \quad (17)$$

holds for all $t \in \mathbb{N}$ and all $\mathbf{x}_0 \in \mathcal{X}$. Here \mathbf{x}_{t+1} obeys (8) with initial condition \mathbf{x}_0 and \mathbb{E}_{π_γ} corresponds to the expected value subject to \mathbf{x} being distributed according to the stationary distribution π_γ .

Lemma 1 states that the closed-loop system (8) converges exponentially to a steady state in the sense that it obeys the time-independent probability distribution π_γ . When deriving a data-driven model of the probability of constraint satisfaction, we leverage Lemma 1 by waiting long enough until the distribution of the state has approximately reached π_γ . This allows us to estimate the long-term behavior of the system from samples over a finite horizon.

Given Lemma 1, we can then rewrite the left-hand side of the long-term chance constraints (3) as

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \sum_{t=0}^{T-1} \mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x}_t)) \right] = \mathbb{E}_{\pi_\gamma} \left[\mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x})) \right], \quad (18)$$

where \mathbf{x}_{t+1} obeys (8). This allows us to rewrite the chance constraints (3) using the more common formulation (3)

$$H(\gamma) := \Pr_{\pi_\gamma}(\mathbf{h}(\mathbf{x}) \leq \mathbf{0}) \geq 1 - \delta. \quad (19)$$

Remark 1: It is possible to employ a result similar to Lemma 1 to reformulate the closed-loop long-term cost (4) as a function of γ , similarly to (19). However, the resulting function is generally difficult to learn since the available measurements correspond to a stochastic process in a continuous space, indexed over γ . A way of potentially tackling this would be to treat the measurements as noisy data, then model the expected value using, e.g., a parametric model.

Our approach, discussed in Section V, iteratively learns an approximation of H , then updates the constraint-tightening parameters by minimizing a weighted sum subject to the estimated chance constraints. Doing so while providing rigorous guarantees for the chance constraints is impossible without further assumptions. Hence, we make assumptions regarding H and the space of constraint-tightening parameters Γ , which we explain in the following.

We assume that the space of constraint-tightening parameter Γ is compact and contains parameters that strictly satisfy the chance constraints.

Assumption 2: The space of constraint-tightening parameters Γ is compact. Furthermore, there exists a constant $\varepsilon_{\text{feas}} > 0$, such that the chance constraint are strictly satisfied with margin $\varepsilon_{\text{feas}}$, for some γ in Γ i.e.,

$$\exists \gamma \in \Gamma : H(\gamma) > 1 - \delta + \varepsilon_{\text{feas}}.$$

Assumption 2 is not very restrictive, since Γ can still be very large, and $\varepsilon_{\text{feas}}$ very small. It implies the problem is well posed, as the chance constraints can be satisfied strictly. In

practice, choosing Γ such that Assumption 2 is satisfied can be achieved, e.g., by choosing Γ such that very high values for γ are permissible, as high values typically lead to more conservative control inputs and fewer constraint violations. Alternatively, if Γ does not satisfy Assumption 2 initially, we can use our approach to estimate the maximal probability of constraint satisfaction and use this information to increase Γ .

We also make the following assumption regarding H .

Assumption 3: There exists a known Lipschitz continuous sigmoid function $s : \mathbb{R} \rightarrow [0, 1]$ and an unknown function $q^* : \Gamma \rightarrow \mathbb{R}$, such that the following conditions hold for all $\gamma \in \Gamma$.

- The composition of s and q^* yields the long-term probability of constraint satisfaction, i.e.,

$$H(\gamma) = s(q^*(\gamma)).$$

- The function $q^*(\cdot)$ belongs to the reproducing kernel Hilbert space (RKHS) with reproducing kernel

$$k(\gamma, \gamma') = \psi^{-1} k_0(\lambda\gamma, \lambda\gamma'), \quad (20)$$

where $k_0 : \mathbb{R}^{d_\gamma} \times \mathbb{R}^{d_\gamma} \rightarrow \mathbb{R}_+$ is a nonsingular kernel, the reciprocal signal variance $\psi \in \mathbb{R}_+$ scales the kernel, and the reciprocal lengthscale $\lambda \in \mathbb{R}_+$ determines the bandwidth of $k(\cdot, \cdot)$. Furthermore, $k(\gamma, \cdot)$ has continuous partial derivatives up to order $2\alpha + 2$ for some $\alpha \in \mathbb{N}$.

For most functions H , we can easily find functions q^* and s that satisfy the first requirement of Assumption 3. For example, whenever H assumes values strictly between 0 and 1, any invertible sigmoid function s can be employed, e.g., the logistic function, yielding $q^*(\gamma) = s^{-1}(H(\gamma))$. The second requirement of Assumption 3 restricts q^* to a RKHS, specified by a kernel of the form (20). This assumption is not very restrictive, since the RKHSs with reproducing kernels that can be expressed as in (20) and have derivatives up to order $2\alpha + 2$ is very rich. Examples of kernels that satisfy this requirement include the squared-exponential or Matérn kernels, which can approximate continuous functions uniformly and arbitrarily accurately in compact spaces [24]. Note that Assumption 3 is considerably less restrictive than assuming a parametric structure for $q^*(\cdot)$, which would yield a significantly less expressive function space.

IV. LEARNING PROBABILITY OF CONSTRAINT SATISFACTION

In order to choose constraint-tightening parameters γ that satisfy the chance constraints (19), we aim to learn an approximation \hat{H}_i of H using binary regression and n measurements collected during control up to time t_i , where the number of collected data is potentially smaller than the number of time steps, i.e., $n \leq t_i$. We achieve this using GPs, which we introduce in this section. The idea behind learning H using binary regression is shown in Figure 1.

A GP is an infinite collection of random variables, of which any finite subset is jointly normally distributed [25]. It is fully specified by a prior mean, which we set to zero without loss of generality, and positive-definite kernel function $k : \Gamma \times \Gamma \mapsto \mathbb{R}$, which we assume to have the same form

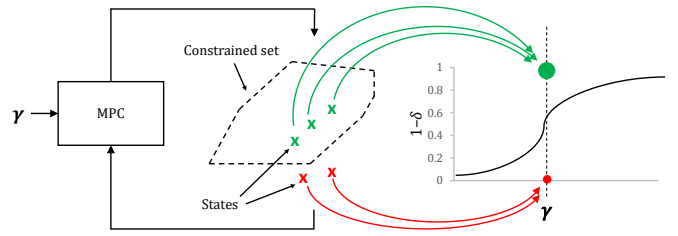


Fig. 1: Sketch of binary regression-based learning. The left-hand side of the figure shows the closed-loop system given γ , which results in states that are either within the constrained set (green) and outside of it (red). The goal is to learn a mapping from the constraint-tightening vector γ to the corresponding probability of constraint satisfaction. To this end, the number of constraint satisfactions (green) and violations (red) given γ are used as measurement data. In other words, γ is used as a training input, whereas instances where constraints are satisfied (green) and violated (red) are assigned zeroes and ones, respectively, and used as training targets. The training data is then used to fit a GP binary regression model, shown on the right-hand side of the figure.

as (20). A GP q_{GP} is then a random variable, such that any finite number of evaluations $q_{\text{GP}}(\gamma_1), \dots, q_{\text{GP}}(\gamma_n)$ corresponds to a multivariate Gaussian random variable

$$(q_{\text{GP}}(\gamma_1), \dots, q_{\text{GP}}(\gamma_n))^{\top} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}) \quad (21)$$

with covariance matrix entries $[\mathbf{K}]_{ij} = k(\gamma_i, \gamma_j)$. In order to obtain an expressive function space, we additionally assign prior probability density functions p_ψ and p_λ to ψ and λ , respectively, as opposed to assuming fixed values for ψ and λ . This way, we can express our belief about the smoothness and amplitude of $q_{\text{GP}}(\cdot)$ without completely ruling out other possibilities. This paper does not assume a specific form for p_ψ and p_λ . However, we require that the tails satisfy the following conditions.

Assumption 4: There exist sequences of positive scalars ψ_j and λ_j , $j \in \mathbb{N}$, such that the priors on ψ and λ satisfy

$$\int_0^{\psi_j} p_\psi(\psi) d\psi \leq \exp(-cj), \quad \int_{\lambda_j}^{\infty} p_\lambda(\lambda) d\lambda \leq \exp(-cj),$$

for some $c > 0$. Furthermore, there exist $b_1, b_2 \in \mathbb{R}_+$, and a sequence of positive scalars M_j , $j \in \mathbb{N}$, such that

$$M_j^2 \psi_j \lambda_j^{-2} \geq b_1 j \quad \text{and} \quad M_j^{d_\gamma \alpha - 1} \leq b_2 j,$$

where α is chosen as in Assumption 3.

Assumption 4 is not restrictive, as the priors p_ψ and p_λ are purely design choices. It dictates how fast p_ψ and p_λ have to decay relative to each other. Assumption 4 is satisfied, e.g., by priors with exponentially decaying tails [26]. In rough terms, Assumption 4 limits how fast the posterior can change as new measurements are collected, which is useful for establishing convergence of the learned model.

The idea behind GP-based binary regression is to squash the GP through a sigmoid function $s : \mathbb{R} \rightarrow [0, 1]$, obtaining the

prior probability associated with the value of a binary random variable

$$p(y = 1|\gamma) = s(q_{\text{GP}}(\gamma)). \quad (22)$$

This way, GP evaluations with a high positive value are assigned a high probability and vice versa. Given training data

$$\mathcal{D}_i = \{\gamma_j, y_j\}_{j=1, \dots, n} =: \{\mathbf{G}, \mathbf{y}\}$$

with binary training labels $y_j \in \{0, 1\}$ and fixed hyperparameters ψ, λ , the GP binary regression model is then conditioned on \mathcal{D}_i , yielding the predictive model

$$p(y = 1|\mathcal{D}_i, \gamma) = \int_{\mathbb{R}} s(q_{\text{GP}})p(q_{\text{GP}}|\mathcal{D}_i, \gamma)dq_{\text{GP}}, \quad (23)$$

where γ denotes the test input. The posterior distribution of the latent variable $q_{\text{GP}}(\gamma)$ given γ is computed as [25]

$$p(q_{\text{GP}}|\mathcal{D}_i, \gamma) = \int_{\mathbb{R}^n} p(q_{\text{GP}}|\gamma)p(q_{\text{GP}}|\mathcal{D}_i)dq_{\text{GP}}, \quad (24)$$

where $\mathbf{q}_{\text{GP}} := (q_{\text{GP}}(\gamma_1), \dots, q_{\text{GP}}(\gamma_t))^\top$, the posterior of the latent variable values is given by

$$p(\mathbf{q}_{\text{GP}}|\mathcal{D}_i) = \frac{p(\mathbf{y}|\mathbf{q}_{\text{GP}})p(\mathbf{q}_{\text{GP}}|\mathbf{G})}{p(\mathbf{y}|\mathbf{G})},$$

and the terms $p(\mathbf{y}|\mathbf{q})$, $p(\mathbf{q}|\mathbf{G})$, $p(\mathbf{y}|\mathbf{G})$ are in turn computed using (21) and (22). To choose the hyperparameters ψ and λ , we can sample from the corresponding posterior given the data

$$p(\psi, \lambda|\mathcal{D}_i) = \frac{p(\mathbf{y}|\mathbf{G}, \psi, \lambda)p(\psi)p(\lambda)}{p(\mathbf{y}|\mathbf{G})}. \quad (25)$$

Alternatively, the hyperparameters ψ and λ can be chosen by maximizing the posterior (25).

Remark 2: In practice, (24) typically does not have an analytical solution, and we have to resort to numerical approximations, e.g., Laplace's approximation or Markov Chain Monte Carlo. For the latter approach, convergence guarantees can be obtained; see, e.g., [27].

V. ONLINE UPDATE OF CONSTRAINT PARAMETERS

We now present our approach, followed by corresponding theoretical guarantees. The goal of our approach is to find a solution to the optimization problem

$$\min_{\gamma \in \Gamma} \mathbf{a}^\top \gamma, \quad \text{s.t. } H(\gamma) \geq 1 - \delta, \quad (26)$$

where $\mathbf{a} \in \mathbb{R}_+$ is a vector of positive weights. This choice of cost function is motivated by the fact that if the entries of γ are small, the feasible region of (5) is large, leading to potentially more effective control inputs. Our method applies the MPC control law (7) using γ_i , which in turn is updated at time steps t_i , specified in the following. To this end, our approach employs a GP-based approximation \hat{H}_i of H , trained using collected measurements. This is repeated until I data points have been collected, after which a final optimization step is performed to determine the approximately optimal constraint-tightening parameters.

Algorithm 1 SMPC with Online Constraint Parameter Update

Input: Initial parameter γ_0 , GP model \hat{H}_0 , desired risk δ , search space Γ

Output: Constraint-tightening parameters γ_I

- 1: Initialize data set $\mathcal{D}_0 = \{\}$
 - 2: Set $t_i = 0$, set $\mathbf{x}_{t_i} = \mathbf{x}_0$
 - 3: **while** $i < I$ **do**
 - 4: Apply $\mathbf{u}^{\text{MPC}}(\mathbf{x}_t, \gamma_i)$ to system using (7), measure \mathbf{x}_{t+1}
 - 5: Update time step $t = t + 1$
 - 6: **if** $t_i + T_{\text{wait}}(\mathbf{x}_{t_i}) \leq t < t_i + T_{\text{wait}}(\mathbf{x}_{t_i}) + T_{\text{col}}$ **then**
 - 7: Collect data and update data set

$$\mathcal{D}_i = \mathcal{D}_i \cup \{\gamma_i, \mathbb{I}_{\mathbb{R}^-}(\mathbf{h}(\mathbf{x}_t))\}$$
 - 8: **else if** $t = t_i + T_{\text{wait}}(\mathbf{x}_i) + T_{\text{col}}$ **then**
 - 9: Update $i = i + 1$, $t_i = t$, $\mathbf{x}_{t_i} = \mathbf{x}_t$
 - 10: Use \mathcal{D}_i to train GP model \hat{H}_i
 - 11: **if** $i = n_{\text{c}_{\text{rand}}}$, $n \in \mathbb{N}$ or (27) is infeasible **then**
 - 12: Sample γ_i from uniform distribution on Γ
 - 13: **else**
 - 14: Update γ_i by solving (27)
 - 15: **end if**
 - 16: **end if**
 - 17: **end while**
 - 18: Choose γ_I by solving (28)
-

Given the state \mathbf{x}_t and a vector γ_i of constraint-tightening parameters, our method applies the MPC control law (7) using γ_i . To guarantee that the collected data is representative of the long-term closed-loop behavior, γ_i is kept constant for $T_{\text{wait}}(\mathbf{x}_{t_i})$ time steps, where \mathbf{x}_{t_i} corresponds to the state when γ_i was last updated, and $T_{\text{wait}}: \mathcal{X} \rightarrow \mathbb{R}_+$ is a function to be specified in the following. The reason for waiting $T_{\text{wait}}(\mathbf{x}_{t_i})$ time steps is that it ensures that the distribution of the state is sufficiently close to that of the stationary distribution π_{γ_i} when we start collecting data. After waiting $T_{\text{wait}}(\mathbf{x}_{t_i})$ steps, our approach then collects T_{col} measurements before updating γ_i . The role of T_{col} is threefold. First, it ensures that we obtain sufficient information about the distribution of $\mathbb{I}_{\mathbb{R}^{\text{dc}}}(\mathbf{h}(\mathbf{x}_t))$ for each training input. Secondly, it limits the frequency with which γ_i is updated, which can be leveraged to reduce the overall computational load. Lastly, it allows us to keep collecting data without restarting the procedure, which would require us to wait until the new stationary distribution given the updated vector γ_i has been approximately reached. After T_{col} measurements have been collected, the data collected up until the current time step is used to learn a GP-based approximation \hat{H}_i of H , as described in Section IV. The constraint-tightening parameters γ_{i+1} are then updated by approximately solving the optimization problem

$$\gamma_{i+1} = \arg \min_{\gamma \in \Gamma} \mathbf{a}^\top \gamma, \quad \text{s.t. } \hat{H}_i(\gamma) \geq 1 - \delta. \quad (27)$$

In addition to the optimization-based update step (27), we also include a random update step that takes place every c_{rand} updates, where $c_{\text{rand}} \in \mathbb{N} \setminus \{0\}$ is arbitrary but fixed, i.e.,

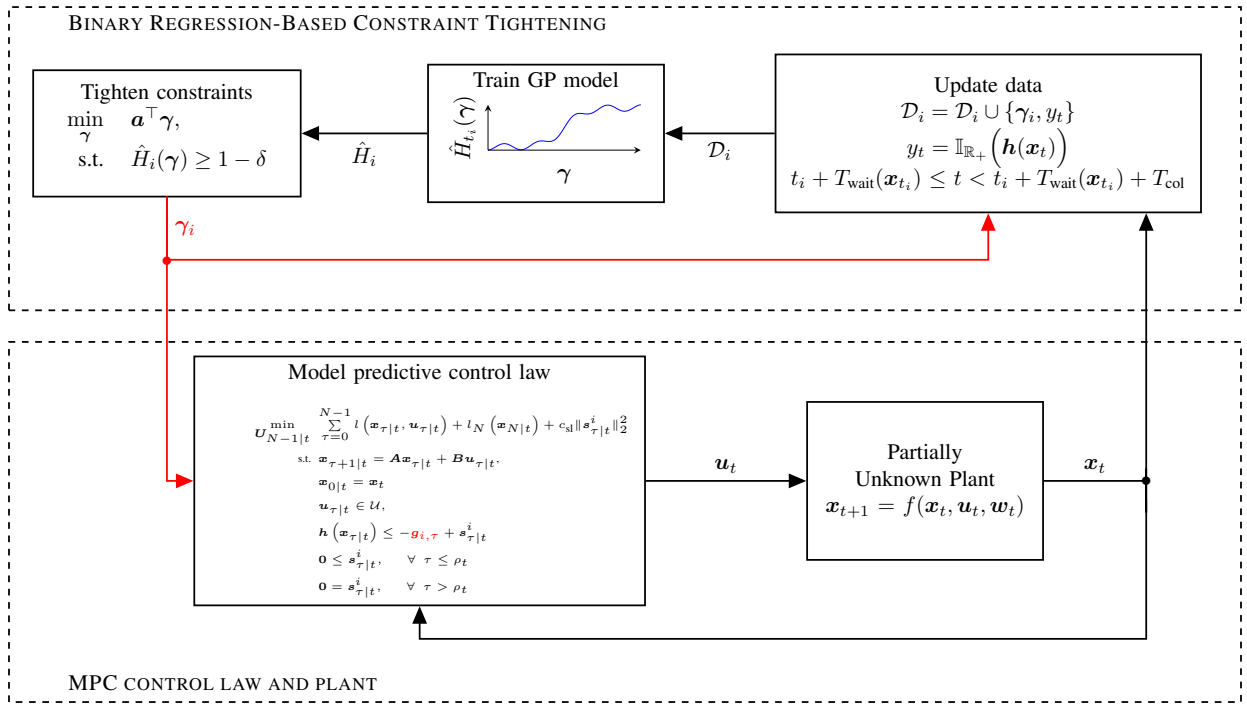


Fig. 2: Closed-loop control with constraint-tightening.

whenever

$$t = t_i, \quad i = nc_{\text{rand}}, \quad n \in \mathbb{N},$$

and whenever (27) is infeasible. The role of random updates is to ensure sufficient coverage of the constraint-tightening parameter space Γ . This procedure is repeated until I data points have been collected, after which we choose the final vector of constraint-tightening parameters from all previously collected inputs

$$\gamma_I = \arg \min_{\gamma \in \{\gamma_1, \dots, \gamma_I\}} \mathbf{a}^T \gamma, \quad \text{s.t. } \hat{H}_I(\gamma) \geq 1 - \delta. \quad (28)$$

We require this final step because the class of functions that satisfies Assumption 3 is very general: our model \hat{H}_i converges to the true function H over all collected data points $\gamma_1, \dots, \gamma_i$, but we cannot generally exclude pathological cases where this does not imply uniform convergence, e.g., if the Lipschitz constant of \hat{H}_i grows unbounded. Although this step poses no practical issues, it is not required, e.g., if the Lipschitz constant of \hat{H}_i is bounded for all i or if Γ is finite. These steps are summarized in Algorithm 1. An illustration without the random update step and final optimization step (28) is given in Figure 2.

Remark 3: The method proposed in this paper can be directly extended to the episodic setting, where a finite-horizon OCP problem is addressed, provided that the initial state distribution is iid. In this case, we can employ Algorithm 1 by setting $T_{\text{wait}}(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathcal{X}$ and T_{col} equal to the episode horizon. The only difference is that the state \mathbf{x}_t is reset every T_{col} steps according to the initial state distribution. The equivalent of the stationary distribution π_{γ} is then the joint distribution of the states within the entire horizon T_{col} .

We now show that if we choose T_{col} , I and the function T_{wait} high enough, the approximately optimal parameter γ_I

satisfies the chance constraints (3) up to an arbitrarily small margin. This corresponds to our main theoretical result and is stated in the following.

Theorem 1: Let Assumptions 1 to 4 hold. For any I , choose T_{wait} and T_{col} such that

$$T_{\text{wait}}(\mathbf{x}) \geq \frac{1}{-\log(\varphi)} (\log(\vartheta V(\mathbf{x})) + I \log(2)), \quad (29)$$

$$T_{\text{col}} \geq c_{\text{col}} I, \quad (30)$$

where φ and ϑ are as in Lemma 1, and $c_{\text{col}} > 0$ is an arbitrary but fixed parameter. Then, there exists a lower bound $I_{\text{min}} \in \mathbb{N}$, such that, for any $\rho_1 > 0$, any $\rho_2 \in (0, 1)$, and any $I \in \mathbb{N}$ with $I \geq I_{\text{min}}$, with probability at least $1 - \rho_2$, (28) is feasible and the corresponding solution satisfies

$$H(\gamma_I) \geq 1 - \delta - \rho_1. \quad (31)$$

Proof: See Appendix - Proof of Theorem 1. ■

Theorem 1 states that if the number of collected data points I is high enough, then, with high probability $1 - \rho_2$, the final optimization step in Algorithm 1 is feasible, and we obtain a vector of constraint-tightening parameters γ_I that satisfies the chance constraints (3) up to an arbitrarily small margin ρ_1 . However, to achieve this, we must choose the function that specifies the waiting time T_{wait} and T_{col} and such that (29) and (30) are satisfied, respectively. The bounds (29) and (30) grow linearly with I . In the case of T_{col} , this is because our model \hat{H}_I converges to the true function H on average, hence ensuring that T_{col} corresponds to a fixed fraction c_{col} of I allows us to recover pointwise bounds. The bound on T_{wait} (29) is directly tied to Lemma 1, which states that the data becomes more representative of the stationary measure π_{γ_i} of the closed-loop system as time increases. It implies that we can extract more information from a higher amount of data

T_{wait} if the corruption in the data is sufficiently low, which is an intuitive result.

Remark 4: In practice, determining the parameters φ and ϑ required to compute the bound (29) for T_{wait} can be difficult. Furthermore, the resulting bound can be conservative, stipulating a long waiting time before we can start collecting data. However, the bound (29) corresponds only to a sufficient condition, not a necessary one, and in practice, a lower value can be potentially picked. In some cases, we can determine convergence to the stationary measure π_{γ_i} empirically, e.g., by checking whether the state has converged to a neighborhood of a specific point.

Remark 5: The bound (30), the random update step, and the final optimization step in Algorithm 1 are all related to the nonparametric nature of our the GP-based model. In the case of strictly parametric models, stronger convergence results can be obtained [28], simplifying the requirements of Theorem 1. However, such a model is generally more difficult to justify than the nonparametric model used in this paper, as the corresponding function space is typically considerably less expressive.

Remark 6: Although we do not provide a result that establishes convergence of γ_I to the optimum of (26), the same techniques employed in this paper can be employed to derive such a result, e.g., using Berger's maximum theorem [29]. However, to this end, we require the additional assumption that the subset of Γ that satisfies $H(\gamma) < 1 - \delta$ is convex in δ . In our experiments, we observed that H is locally monotonically increasing, which is potentially sufficient. Furthermore, this property caused γ_I to converge to the minimizer of (26) in all experiments.

A. Discussion

A clear strength of the proposed approach is that we can obtain guarantees while taking the backup strategy into account, which is of high practical relevance. It is also worth noting that other forms of backup controller are also allowed, provided that the closed-loop system satisfies Assumption 1. Our approach could potentially also be used to inform such a choice in practice, e.g., by applying and comparing backup control strategies other than the slack variable-based approach presented here.

Our approach adapts the constraint-tightening parameter γ by exploring the parameter space Γ , where the observed constraint violations inform the exploration. Although our approach is guaranteed to converge a γ that satisfies the chance constraints due to Theorem 1, the corresponding algorithm might choose values of γ that will momentarily lead to a high rate of constraint violations. In cases where this is undesirable, it can be mitigated by adopting a safe strategy, e.g., by adaptively increasing the space of constraint tightening parameters Γ such that initially only conservative values are permissible, then gradually allowing less conservative values.

One of the more notable challenges when employing GPs is their poor scalability with the number of data points, as training scales cubically with the amount of data and evaluation scales quadratically [25]. This can be addressed

using different frequently encountered tools, e.g., sparse GPs [30] or networks of local experts [31].

If the MPC horizon N or the number of constraints d_c are very high, then the dimension d_γ of the constraint-tightening space Γ is potentially large, making searching for an optimizer of (27) difficult. However, this can be easily mitigated by setting

$$\Gamma = \left\{ \mathbf{D}\tilde{\gamma} \mid \tilde{\gamma} \in \tilde{\Gamma} \right\},$$

where \mathbf{D} is a matrix and $\tilde{\Gamma}$ is a low-dimensional search space, then conditioning the GP model on training inputs from $\tilde{\Gamma}$. This is exemplified in Section VI.

A further aspect of the proposed algorithm that should be considered is the initial guess for γ_0 . Though Theorem 1 requires a portion of the parameter space Γ to be explored thoroughly, convergence can potentially be sped up significantly if a good initial vector of constraint-tightening parameters γ_0 is provided, as this will decrease the size of the explored region significantly compared to a poor initial guess.

VI. NUMERICAL EXAMPLE

In this section, we aim to illustrate Theorem 1 and assess the performance of the proposed approach using a numerical experiment. Moreover, we compare our approach to different state-of-the-art approaches.

A. Simulation Setup

We consider the discrete-time system corresponding to the linearized model of a DC-DC converter [15], [32], given by

$$\mathbf{x}_{t+1} = \begin{bmatrix} 1 & 0.0075 \\ -0.143 & 0.996 \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} 4.798 \\ 0.115 \end{bmatrix} u_t + \mathbf{w}_t, \quad (32)$$

where we consider an initial state of $\mathbf{x}_0 = (0, 0)^\top$, chance constraints

$$\Pr([1 \ 0]\mathbf{x}_t \leq 0) \geq 1 - \delta, \quad \forall t \in \mathbb{N},$$

and a varying risk parameter $\delta \in [0.6, 0.01]$. The cost to be minimized is given by (4), with immediate cost

$$l(\mathbf{x}, u) = \mathbf{x}^\top \mathbf{Q} \mathbf{x} + u R u,$$

and weights

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix}, \quad R = 1.$$

We assume the control input to be bounded as $|u_t| \leq 0.2$ and consider different distributions for the stochastic disturbance \mathbf{w}_t , to be specified in the following.

For the MPC algorithms, we employ an MPC horizon time of $N = 10$ and assume to know the state matrix \mathbf{A} and the input matrix \mathbf{B} in (32). We employ a quadratic terminal cost function of $l_N(\mathbf{x}) = \mathbf{x}^\top \mathbf{P} \mathbf{x}$, where \mathbf{P} satisfies the Lyapunov function $\mathbf{P} = \mathbf{A} \mathbf{P} \mathbf{A}^\top + \mathbf{Q}$. For the backup strategy, the cost penalties are obtained by multiplying the slack variables with $c_{\text{sl}} = 10^{16}$.

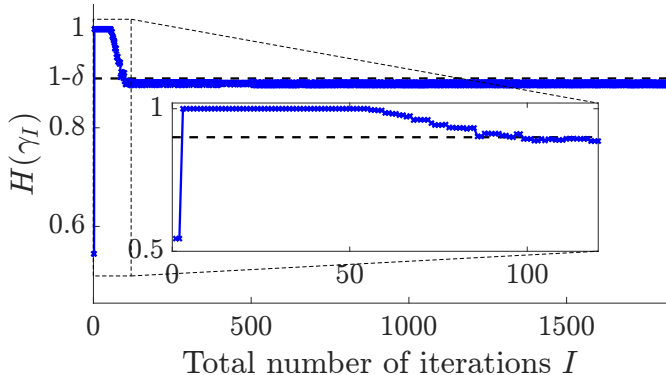


Fig. 3: Long-term probability of constraint satisfaction $H(\gamma_I)$ obtained with parameters γ_I returned by Algorithm 1 for different total number of iterations I .

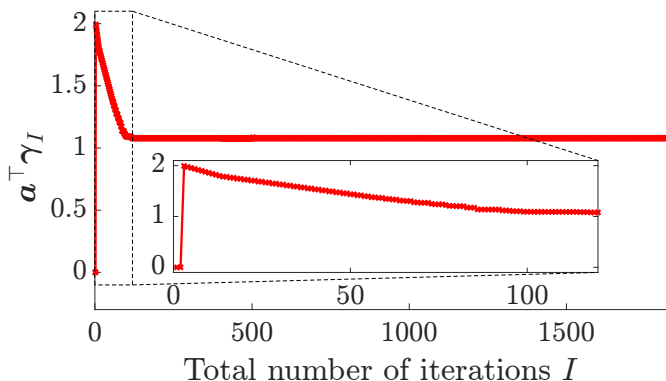


Fig. 4: Weighted sum $\mathbf{a}^\top \gamma_I$ of entries of constraint-tightening parameters γ_I returned by Algorithm 1 for different total number of iterations I .

B. Illustration of Theorem 1

To illustrate Theorem 1, we run Algorithm 1 multiple times for varying $I = 1, \dots, 1800$ and a fixed risk of $\delta = 0.1$. We consider process noise \mathbf{w}_t that is sampled from a uniform distribution on $[-0.14, 0.14]$. To model H , we employ a squared-exponential kernel

$$k(\gamma, \gamma') = \psi^{-1} \exp\left(-\frac{\lambda^2}{2} \|\gamma - \gamma'\|_2^2\right), \quad (33)$$

with Gaussian priors for the log hyperparameters $\log(\psi) \sim \mathcal{N}(-1, 1)$, and a modified Gaussian error function

$$s(z) = 1 + \frac{1}{\sqrt{\pi}} \int_0^z e^{-\xi^2} d\xi \quad (34)$$

as sigmoid function. To keep computations cheap while handling multiple thousands of training data points, we employ a sparse GP approximation with all different training inputs as inducing inputs [30], [33]. We consider a one-dimensional space of constraint-tightening parameters

$$\Gamma = \{\tilde{\gamma} \mathbf{1}_{d_\gamma} \mid \tilde{\gamma} \in [-1, 0.2]\}.$$

This choice of parameter space allows us to search over a one-dimensional space instead of a d_γ -dimensional one. Note Γ

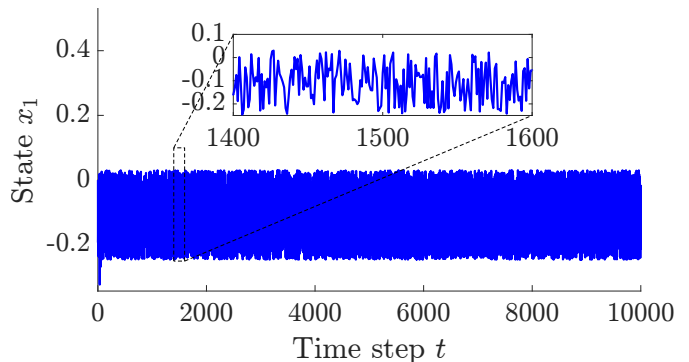
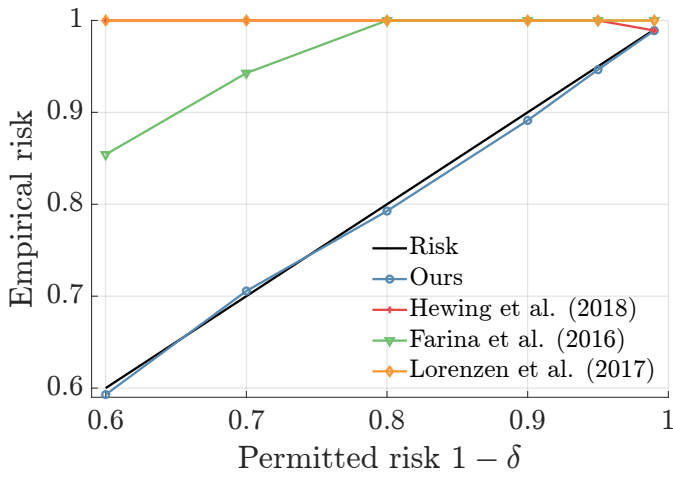


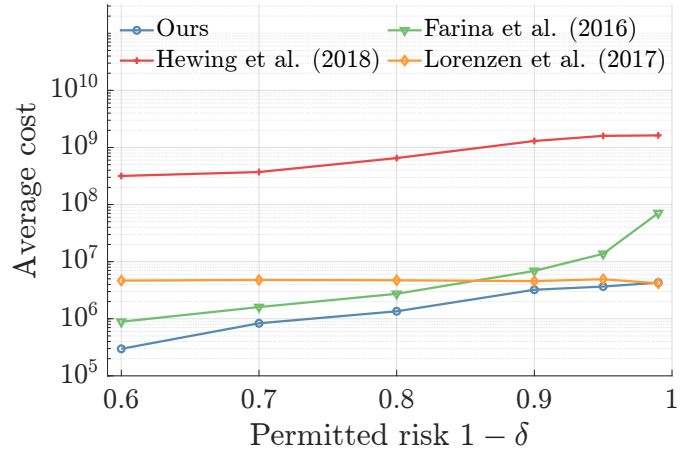
Fig. 5: State x_1 obtained for a simulation with 10^4 time steps and the constraint-tightening parameters γ_I returned by Algorithm 1 after $I = 1800$ iterations. The empirical probability of constraint satisfaction is $0.8912 < 1 - \delta - \rho_1$ for $\rho_1 = 0.01$, which is to be expected from Theorem 1.

admits γ with negative entries, allowing us to avoid unnecessary conservatism, e.g., if the unconstrained MPC algorithm already satisfies the chance constraints. For the optimization (27), we employ a weight vector of $\mathbf{a} = \mathbf{1}_{d_\gamma}$. To solve (27), we perform an exhaustive search of Γ starting at the lowest end until we find a feasible point. This is inexpensive since (27) corresponds to a one-dimensional optimization problem with a monotonically increasing cost. We set the initial constraint-tightening parameters to $\gamma_0 = \mathbf{0}$, which does not satisfy the chance constraints. The values for $T_{\text{wait}}(\mathbf{x})$ stipulated by Theorem 1 can be conservative. Instead of computing them directly, we choose $T_{\text{wait}}(\mathbf{x}) = 500$ for all I . This choice is motivated by the observation that the state x_1 converges to a seemingly stationary distribution after only a handful of steps for all starting values close to $\mathbf{0}$. We additionally set $T_{\text{col}} = 5000$ for all I and $c_{\text{rand}} = 100$, i.e., every hundredth iteration i corresponds to a random search step.

For every γ_I , we approximate $H(\gamma_I)$ by measuring the average rate of constraint satisfaction over a simulation of 5000 steps. The progress of $H(\gamma_I)$ with I can be seen in Figure 3, that of the cost function $\mathbf{a}^\top \gamma_I$ can be seen in Figure 4. Since $\gamma_0 = \mathbf{0}$ does not satisfy the chance constraints, none of the data initially used to train the GP model \hat{H} corresponds to a feasible point. This causes the optimization problem (27) to be infeasible. Hence, the constraint-tightening parameters γ_i generated by Algorithm 1 are sampled from a uniform distribution on Γ , which in turn is reflected in γ_I if I is chosen small. However, after a handful of random searches, a vector γ_I of constraint-tightening parameters is found that strictly satisfies the chance constraints. From there, the constraint-tightening parameters γ_I returned by Algorithm 1 gradually decrease until the chance constraints are approximately satisfied with equality, after which γ_I stays approximately constant. This is because \hat{H} approximates H accurately if enough data is available. If we choose $I > 100$, then the final vector of constraint-tightening parameters γ_I will have converged to a value that satisfies the chance constraints up to a margin of $\rho_1 < 0.01$, which is to be expected by Theorem 1. The state x_1 for a sample simulation with the constraint-tightening

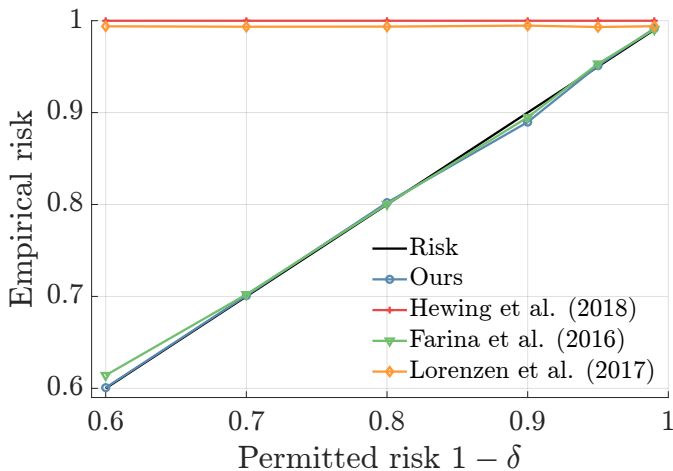


(a) Empirical rate of constraint satisfaction for single simulation with 2×10^5 time steps.

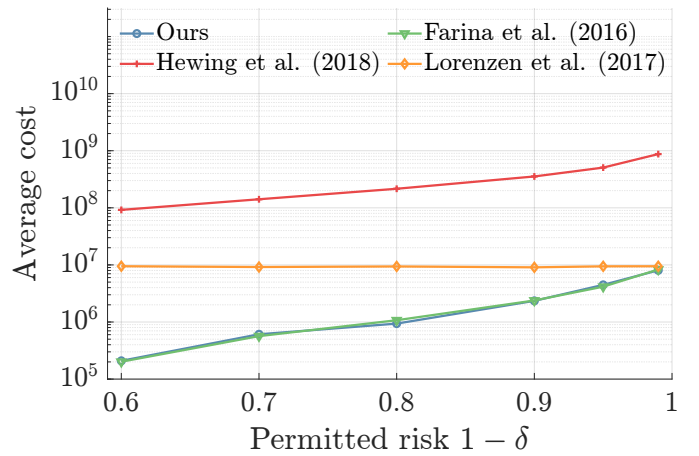


(b) Average cost for single simulation with 2×10^5 time steps.

Fig. 6: Empirical rate of constraint satisfaction and average cost for uniformly distributed uncertainties on $[-0.14, 0.14]$.



(a) Empirical rate of constraint satisfaction for single simulation with 2×10^5 time steps.



(b) Average cost for single simulation with 2×10^5 time steps.

Fig. 7: Empirical rate of constraint satisfaction and average cost for zero-mean Gaussian uncertainties on with variance 0.08.

parameters γ_I obtained with Algorithm 1 and $I = 1800$ is shown in Figure 5.

C. Comparison with Existing Approaches

We now compare our approach to [8], [14], [15], three different state-of-the-art approaches for computing constraint-tightening parameters of analytic reformulations of SMPC problems. The approach presented in [8] computes constraint-tightening parameters by determining a $1 - \delta$ credible interval for the process noise distribution and adjusting the constraints accordingly over the prediction horizon. Similarly, the approach of [14], [16], [34] leverages the process noise distribution to compute probabilistic reachable sets, guaranteeing that the chance constraints are satisfied in a closed-loop fashion. The method of [15] employs a sampling-based approach to compute the constraint-tightening parameters. Furthermore, it includes a stochastic constraint on the first step to ensure the recursive feasibility of the optimization problem. We note that

we do not employ a gain matrix, as suggested by [14], [15]. However, their approaches are also applicable without a gain matrix, provided that the system matrix \mathbf{A} is stable, which holds in this setting. Our method employs the same setup as in Section VI-B, except for the total number of iterations, which we set to $I = 150$. We consider the same input constraints and backup control policy for all approaches.

We consider two different settings, one where the entries of the process noise are uniformly distributed on $[-0.14, 0.14]$, and one where they are normally distributed with mean zero and standard deviation 0.08. We compute constraint-tightening parameters for seven different risk parameters

$$\delta \in \{0.6, 0.7, 0.8, 0.9, 0.95, 0.99\}.$$

We simulate each setting once over 20000 time steps. We additionally compare the resulting average cost for each approach.

The results for the uniformly distributed process noise entries can be seen in Figure 6 and those for the Gaussian distributed process noise entries in Figure 7. Our approach

satisfies the chance constraints for all δ up to a small margin of $\rho_1 < 0.01$, which is to be expected by Theorem 1. Furthermore, it satisfies the constraints *tightly* up to a margin of 0.01, which indicates that γ_l converged to a local minimum of (27). This is also reflected in the average cost, which is lower than that of the other approaches. This difference is particularly accentuated for the methods of [8] and [14] in the setting with uniformly distributed uncertainties due to their use of Chebyshev inequalities, which are conservative. In the case of Gaussian distributed uncertainties, the approach of [8] performs similarly to our approach, which is to be expected since the corresponding constraint-tightening parameters tightly satisfy the chance constraints.

VII. CONCLUSION

We have presented a binary regression approach for choosing the constraint-tightening parameters of the analytic reformulation of a stochastic optimal control problem. We have provided sufficient conditions under which a finite number of system samples can approximate the long-term rate of constraint satisfaction. Our algorithm leverages these conditions to compute constraint-tightening parameters that provably satisfy the chance constraints up to an arbitrarily low error using a GP binary regression model. In numerical simulations, our approach yielded constraint-tightening parameters that tightly satisfied chance constraints, allowing it to outperform state-of-the-art approaches in average cost. Potential future research directions include obtaining a convergence rate for the proposed method and experimenting with different binary regression tools, e.g., logistical or probit regression approaches.

APPENDIX - PROOF OF THEOREM 1

To show that the constraint-tightening parameters returned by our algorithm converge to values that respect the chance constraints, we first need to show that the GP-based binary regression model is increasingly accurate as more data is employed to train it. To this end, we employ the following result, which applies to arbitrary data sets.

Lemma 2: Let Assumptions 1 to 4 hold, and consider a sequence of data sets, increasing in size, where the training inputs are fixed but potentially different between sets, i.e.,

$$\tilde{\mathcal{D}}_j = \{\tilde{\gamma}_{j,n}, \tilde{y}_{j,n}\}_{n=1,\dots,j}, \quad j \in \mathbb{N}$$

where

$$\tilde{\gamma}_{j,1}, \dots, \tilde{\gamma}_{j,j} \in \Gamma$$

and $\tilde{y}_{j,1}, \dots, \tilde{y}_{j,j}$ are Bernoulli random variables with

$$\tilde{y}_{j,n} = \mathbb{I}_{\mathbb{R}^d_c}(\mathbf{h}(\mathbf{x})), \quad \mathbf{x} \sim \pi_{\tilde{\gamma}_{j,n}}.$$

Here $\pi_{\tilde{\gamma}_{j,n}}$ denotes the unique stationary measure for the closed-loop system (8) using $\tilde{\gamma}_{j,n}$. Let \tilde{H}_j denote the posterior GP conditioned on $\tilde{\mathcal{D}}_j$. Then, for every $\rho_1, \rho_2 > 0$, there exists a $J \in \mathbb{N}$, such that

$$\Pr\left(\frac{1}{j} \sum_{n=1}^j |H(\tilde{\gamma}_{j,n}) - \tilde{H}_j(\tilde{\gamma}_{j,n})| > \rho_1 \mid \tilde{\mathcal{D}}_j\right) < \rho_2.$$

holds for all $j \geq J$.

Proof: The proof follows straightforwardly from Lemma 1 and [26, Theorem 2]. ■

Lemma 2 implies that if the measurements $\tilde{y}_{j,1}, \dots, \tilde{y}_{j,j}$ are sampled from the stationary distributions $\pi_{\tilde{\gamma}_{j,n}}$ for arbitrary but fixed training inputs $\tilde{\gamma}_{j,1}, \dots, \tilde{\gamma}_{j,j} \in \Gamma$, then the accuracy of the GP model grows in expectation with the data size j . A direct consequence is that we obtain convergence as the data size grows, independently of the training input locations.

Lemma 3: Let Assumptions 1 to 4 hold. Then, for every $\rho_1, \rho_2 > 0$, there exists a $J \in \mathbb{N}$, such that

$$\Pr\left(\frac{1}{j} \sum_{n=1}^j |H(\tilde{\gamma}_{j,n}) - \tilde{H}_j(\tilde{\gamma}_{j,n})| > \rho_1 \mid \tilde{\mathcal{D}}_j\right) < \rho_2.$$

holds for all $j \geq J$ and all data sets

$$\tilde{\mathcal{D}}_j = \{\tilde{\gamma}_{j,n}, \tilde{y}_{j,n}\}_{n=1,\dots,j}, \quad j \in \mathbb{N}$$

with training inputs

$$\tilde{\gamma}_{j,1}, \dots, \tilde{\gamma}_{j,j} \in \Gamma,$$

Bernoulli random variables

$$\tilde{y}_{j,n} = \mathbb{I}_{\mathbb{R}^d_c}(\mathbf{h}(\mathbf{x})), \quad \mathbf{x} \sim \pi_{\tilde{\gamma}_{j,n}},$$

where $\pi_{\tilde{\gamma}_{j,n}}$ denotes the unique stationary measure for the closed-loop system (8) using $\tilde{\gamma}_{j,n}$, and \tilde{H}_j denotes the posterior GP conditioned on $\tilde{\mathcal{D}}_j$.

Proof: Assume the contrary is true. Then, there exist scalars $\rho_1, \rho_2 > 0$ such that, for every $J \in \mathbb{N}$, there exists a $j \geq J$ and a sequence of training inputs $\tilde{\gamma}_{j,1}, \dots, \tilde{\gamma}_{j,j} \in \Gamma$, such that

$$\Pr\left(\frac{1}{j} \sum_{n=1}^j |H(\tilde{\gamma}_{j,n}) - \tilde{H}_j(\tilde{\gamma}_{j,n})| > \rho_1 \mid \tilde{\mathcal{D}}_j\right) \geq \rho_2,$$

where

$$\tilde{\mathcal{D}}_j = \{\tilde{\gamma}_{j,n}, \tilde{y}_{j,n}\}_{n=1,\dots,j}, \quad (35)$$

$$\tilde{y}_{j,n} = \mathbb{I}_{\mathbb{R}^d_c}(\mathbf{h}(\mathbf{x})), \quad \mathbf{x} \sim \pi_{\tilde{\gamma}_{j,n}}, \quad (36)$$

and \tilde{H}_j denotes the posterior GP conditioned on $\tilde{\mathcal{D}}_j$. Since

$$\tilde{\gamma}_{j,1}, \dots, \tilde{\gamma}_{j,j}$$

is fixed for every j and j can be made arbitrarily large, this contradicts Lemma 2. ■

Lemma 3 implies pointwise convergence of the GP model with the number of data, provided that the data is sampled from the stationary distributions. We now show that Lemma 1 can be leveraged to recover guarantees similar to those of Lemma 3 for the GP model trained on data collected over a finite horizon, provided that the horizon is sufficiently long.

Lemma 4: Let Assumptions 1 to 4 hold. Then, for every $\rho_1, \rho_2 > 0$, there exists a $J \in \mathbb{N}$, such that

$$\Pr\left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_j(\gamma_n)| > \rho_1 \mid \mathcal{D}_j\right) < \rho_2.$$

holds for all data sets

$$\mathcal{D}_j = \{\gamma_n, y_n\}_{n=1,\dots,j}$$

with $j \geq J$, where the training inputs $\gamma_1, \dots, \gamma_j \in \Gamma$ of \mathcal{D}_j are identical to those of \mathcal{D}_{j-1} up to the j -th training input, and Bernoulli random variables

$$y_n = \mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x}_{n,T_n})),$$

where \mathbf{x}_{n,T_n} corresponds to the state after T_n steps under the closed-loop dynamics (8) given γ_n and arbitrary initial state $\mathbf{x}_{n,0} \in \mathcal{X}$, and the number of steps satisfy

$$T_n \geq \frac{1}{-\log(\varphi)} (\log(\vartheta V(\mathbf{x}_{n,0})) + j \log(2)), \quad (37)$$

with φ and ϑ as in Lemma 1. Here $\hat{H}_j \in \Gamma$ denotes the posterior GP conditioned on $\mathcal{D}_j = \{\gamma_n, y_n\}_{n=1, \dots, j}$.

Proof: Choose arbitrary $\rho_1 > 0$ and $\rho_2 > 0$. Note that, since y_n are Bernoulli random variables, we can write

$$\begin{aligned} & \Pr \left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_j(\gamma_n)| > \rho_1 \right) \\ &= \sum_{m=1}^{j^2} P_m \mathbb{I}_{\mathbb{R}_+} \left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_n(\gamma_n | \mathbf{y}_m)| - \rho_1 \right), \end{aligned}$$

where $\hat{H}_n(\gamma_n | \mathbf{y}_m)$ denotes the GP model conditioned on the measurements $\mathbf{y}_m \in \{0, 1\}^n$, the vectors $\mathbf{y}_1, \dots, \mathbf{y}_{j^2}$ correspond to all possible different j -dimensional vectors containing zeros and ones, and

$$P_m = \prod_{n=1}^j p_n^m,$$

with

$$p_n^m = \begin{cases} \mathbb{E} \left[\mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x}_{n,T_n})) \right] & \text{if } y_{m,n} = 1 \\ 1 - \mathbb{E} \left[\mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x}_{n,T_n})) \right] & \text{if } y_{m,n} = 0, \end{cases}$$

where $y_{m,n}$ denotes the n -th entry of the vector \mathbf{y}_m . Consider now the data set $\tilde{\mathcal{D}}_j = \{\gamma_n, \tilde{y}_n\}_{n=1, \dots, j}$ obtained by sampling from the stationary distribution of the closed-loop system given γ_n , i.e.,

$$\tilde{y}_n = \mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x})), \quad \mathbf{x} \sim \pi_{\gamma_n}.$$

By applying the same procedure as above, we obtain

$$\begin{aligned} & \left| \Pr \left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_j(\gamma_n)| > \rho_1 \mid \mathcal{D}_j \right) \right. \\ & \quad \left. - \Pr \left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_j(\gamma_n)| > \rho_1 \mid \tilde{\mathcal{D}}_j \right) \right| \\ &= \sum_{m=1}^{j^2} |P_m - \tilde{P}_m| \mathbb{I}_{\mathbb{R}_+} \left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_n(\gamma_n | \mathbf{y}_m)| - \rho_1 \right), \end{aligned}$$

where

$$\tilde{P}_m = \prod_{n=1}^j \tilde{p}_n^m,$$

and

$$\tilde{p}_n^m = \begin{cases} \mathbb{E}_{\pi_{\gamma}} \left[\mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x})) \right] & \text{if } y_{m,n} = 1 \\ 1 - \mathbb{E}_{\pi_{\gamma}} \left[\mathbb{I}_{\mathbb{R}^{d_c}}(\mathbf{h}(\mathbf{x})) \right] & \text{if } y_{m,n} = 0. \end{cases}$$

From Lemma 1, it follows that

$$|\tilde{p}_n^m - p_n^m| \leq \vartheta V(\mathbf{x}_{n,0}) \varphi^{T_n}. \quad (38)$$

By inserting (29) into (38), we obtain $|\tilde{p}_n^m - p_n^m| \leq 2^{-j}$, implying

$$\begin{aligned} |P_m - \tilde{P}_m| &\leq \tilde{P}_m \max \left\{ 1 - (1 - 2^{-j})^j, (1 + 2^{-j})^j - 1 \right\} \\ &= \tilde{P}_m \left((1 + 2^{-j})^j - 1 \right) \leq \tilde{P}_m. \end{aligned}$$

By employing Lemma 3, we obtain, for j sufficiently large

$$\begin{aligned} & \Pr \left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_j(\gamma_n)| > \rho_1 \mid \mathcal{D}_j \right) \\ &\leq \Pr \left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_j(\gamma_n)| > \rho_1 \mid \tilde{\mathcal{D}}_j \right) \\ &\quad + \left| \Pr \left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_j(\gamma_n)| > \rho_1 \mid \mathcal{D}_j \right) \right. \\ &\quad \left. - \Pr \left(\frac{1}{j} \sum_{n=1}^j |H(\gamma_n) - \hat{H}_j(\gamma_n)| > \rho_1 \mid \tilde{\mathcal{D}}_j \right) \right| \\ &\leq (1 + 1) \frac{\rho_2}{2} = \rho_2. \end{aligned}$$

Lemma 4 states that if we wait long enough between constraint-tightening parameter updates and data collection, we obtain a model that faithfully captures the long-term probability of constraint satisfaction over a finite set of points in Γ . This allows us to prove our main theorem.

We now prove the first part of Theorem 1, which states that the final optimization step in Algorithm 1 is feasible with high probability.

Lemma 5: Let Assumptions 1 to 4 hold. Then, for any $\rho_1 > 0$ and $\rho_2 \in (0, 1)$, there exists a $I \in \mathbb{N}$, such that (28) is feasible with probability at least $1 - \rho_2$.

Proof: Since the kernel k is continuously differentiable and Γ is compact, Assumption 3 implies that the function H is Lipschitz continuous on Γ . This means there exists a set of non-zero measure Γ_c , such that $H(\gamma) > 1 - \delta + \frac{\varepsilon_{\text{feas}}}{2}$ holds for all $\gamma \in \Gamma_c$, where $\varepsilon_{\text{feas}}$ is as in Assumption 2. Let $\gamma_i, i = 1, \dots, I$ denote the sequence of constraint-tightening parameters generated by Algorithm 1. Define the finite set

$$\Gamma_{I,c} = \{\gamma_m \mid m \leq I, \gamma_m \in \Gamma_c\}$$

and recall that whenever $i = n c_{\text{rand}}$, $n \in \mathbb{N}$, the constraint-tightening parameter vector γ_i is sampled from a uniform distribution on Γ . Hence, by the law of large numbers, with probability one

$$\lim_{I \rightarrow \infty} \frac{|\Gamma_{I,c}|}{I} \geq \frac{\text{vol}(\Gamma_c)}{\text{vol}(\Gamma)} > 0. \quad (39)$$

Hence, due to Assumption 2, for I large enough, we have

$$\Pr \left(\left| \frac{|\Gamma_{I,c}|}{I} - \frac{\text{vol}(\Gamma_c)}{\text{vol}(\Gamma)} \right| \geq \frac{\text{vol}(\Gamma_c)}{2 \text{vol}(\Gamma)} \right) \leq \frac{\rho_2}{2}.$$

Note that, if (28) is infeasible, then

$$\hat{H}_I(\gamma) < 1 - \delta \leq H(\gamma) - \frac{\varepsilon_{\text{feas}}}{2}$$

holds for all $\gamma \in \Gamma_c$, which in turn implies

$$\frac{1}{I} \sum_{\gamma \in \Gamma_c} |H(\gamma) - \hat{H}_I(\gamma)| \geq \frac{|\Gamma_{I,c}| \varepsilon_{\text{feas}}}{2I}. \quad (40)$$

By employing Lemma 4 and I large enough, we can bound the probability that (28) is infeasible as

$$\begin{aligned} & \Pr \left(\hat{H}_I(\gamma) < 1 - \delta \quad \forall \gamma \in \{\gamma_1, \dots, \gamma_I\} \right) \\ & \leq \Pr \left(\frac{1}{|\Gamma_{I,c}|} \sum_{\gamma \in \Gamma_{I,c}} |H(\gamma) - \hat{H}_I(\gamma)| > \frac{|\Gamma_{I,c}| \varepsilon_{\text{feas}}}{2I} \mid \mathcal{D}_j \right) \\ & = \Pr \left(\left(\frac{1}{|\Gamma_{I,c}|} \sum_{\gamma \in \Gamma_{I,c}} |H(\gamma) - \hat{H}_I(\gamma)| > \frac{|\Gamma_{I,c}| \varepsilon_{\text{feas}}}{2I} \right) \right. \\ & \quad \cup \left(\left| \frac{|\Gamma_{I,c}|}{I} - \frac{\text{vol}(\Gamma_c)}{\text{vol}(\Gamma)} \right| \geq \frac{\text{vol}(\Gamma_c)}{2\text{vol}(\Gamma)} \right) \mid \mathcal{D}_j \Big) \\ & \quad + \Pr \left(\left(\frac{1}{|\Gamma_{I,c}|} \sum_{\gamma \in \Gamma_{I,c}} |H(\gamma) - \hat{H}_I(\gamma)| > \frac{|\Gamma_{I,c}| \varepsilon_{\text{feas}}}{2I} \right) \right. \\ & \quad \cup \left(\left| \frac{|\Gamma_{I,c}|}{I} - \frac{\text{vol}(\Gamma_c)}{\text{vol}(\Gamma)} \right| < \frac{\text{vol}(\Gamma_c)}{2\text{vol}(\Gamma)} \right) \mid \mathcal{D}_j \Big) \\ & \leq \frac{\rho_2}{2} + \Pr \left(\left(\frac{1}{|\Gamma_{I,c}|} \sum_{\gamma \in \Gamma_{I,c}} |H(\gamma) - \hat{H}_I(\gamma)| > \frac{|\Gamma_{I,c}| \varepsilon_{\text{feas}}}{2I} \right) \right. \\ & \quad \cup \left(\left| \frac{|\Gamma_{I,c}|}{I} - \frac{\text{vol}(\Gamma_c)}{\text{vol}(\Gamma)} \right| < \frac{\text{vol}(\Gamma_c)}{2\text{vol}(\Gamma)} \right) \mid \mathcal{D}_j \Big) \\ & \leq \frac{\rho_2}{2} + \Pr \left(\left(\frac{I}{|\Gamma_{I,c}|} \frac{1}{I} \sum_{n=1}^I |H(\gamma_n) - \hat{H}_I(\gamma_n)| > \frac{|\Gamma_{I,c}| \varepsilon_{\text{feas}}}{2I} \right) \right. \\ & \quad \cup \left(\left| \frac{|\Gamma_{I,c}|}{I} - \frac{\text{vol}(\Gamma_c)}{\text{vol}(\Gamma)} \right| < \frac{\text{vol}(\Gamma_c)}{2\text{vol}(\Gamma)} \right) \mid \mathcal{D}_j \Big) \\ & \leq \frac{\rho_2}{2} \\ & \quad + \Pr \left(\frac{1}{I} \sum_{n=1}^I |H(\gamma_n) - \hat{H}_I(\gamma_n)| > \frac{\varepsilon_{\text{feas}}}{8} \left(\frac{\text{vol}(\Gamma_c)}{\text{vol}(\Gamma)} \right)^2 \mid \mathcal{D}_j \right) \\ & \leq \rho_2, \end{aligned}$$

where the second inequality is due to the union bound. \blacksquare

Proof of Theorem 1: Since Lemma 5 establishes the feasibility of (28) with arbitrarily high probability, to prove Theorem 1 we only need to show that the corresponding solution satisfies the chance constraints up to an arbitrarily small margin. To this end, we show that

$$|\hat{H}_I(\gamma) - H(\gamma)| < \rho_1 \quad \forall \gamma \in \{\gamma_1, \dots, \gamma_I\} \quad (41)$$

holds with arbitrarily high probability $1 - \rho_2$, which implies the desired result.

Recall that Algorithm 1 collects $T_{\text{col}} = c_{\text{col}} I$ data points before updating γ_i . Hence, if there exists a $\gamma \in \{\gamma_1, \dots, \gamma_I\}$, such that $|\hat{H}_I(\gamma) - H(\gamma)| \geq \rho_1$, then

$$\sum_{n=1}^I |\hat{H}_I(\gamma_n) - H(\gamma_n)| \geq T_{\text{col}} \rho_1 = c_{\text{col}} I \rho_1$$

holds. By applying the union bound, we obtain, for I high enough,

$$\begin{aligned} & \Pr \left(\exists \gamma \in \{\gamma_1, \dots, \gamma_I\} : |\hat{H}_I(\gamma) - H(\gamma)| \geq \rho_1 \right) \\ & \leq \Pr \left(\frac{1}{I} \sum_{n=1}^I |\hat{H}_I(\gamma_n) - H(\gamma_n)| \geq c_{\text{col}} \rho_1 \right) \leq \rho_2. \end{aligned}$$

\blacksquare

- [1] F. Oldewurtel, C. N. Jones, A. Parisio, and M. Morari. Stochastic model predictive control for building climate control. *IEEE Transactions on Control Systems Technology*, 22(3):1198–1205, May 2014.
- [2] A. T. Schwarm and M. Nikolaou. Chance-constrained model predictive control. *AIChE Journal*, 45(8):1743–1752, 1999.
- [3] I. Jurado, P. Millán, D. Quevedo, and F. R. Rubio. Stochastic mpc with applications to process control. *International Journal of Control*, 88(4):792–800, 2015.
- [4] R. Kumar, M. J. Wenzel, M. J. Ellis, M. N. ElBsat, K. H. Drees, and V. M. Zavala. A stochastic model predictive control framework for stationary battery systems. *IEEE Transactions on Power Systems*, 33(4):4397–4406, July 2018.
- [5] Y. Jiang, C. Wan, J. Wang, Y. Song, and Z. Y. Dong. Stochastic receding horizon control of active distribution networks with distributed renewables. *IEEE Transactions on Power Systems*, 34(2):1325–1341, March 2019.
- [6] A. Mesbah. Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems*, 36(6):30–44, Dec 2016.
- [7] D. Mayne. Robust and stochastic model predictive control: Are we going in the right direction? *Annual Reviews in Control*, 41:184–192, 2016.
- [8] M. Farina, L. Giulioni, and R. Scattolini. Stochastic linear model predictive control with chance constraints – a review. *Journal of Process Control*, 44(Supplement C):53 – 67, 2016.
- [9] G. Schildbach, L. Fagiano, C. Frei, and M. Morari. The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations. *Automatica*, 50(12):3009–3018, 2014.
- [10] G.C. Calaforo and L. Fagiano. Robust model predictive control via scenario optimization. *IEEE Transactions on Automatic Control*, 58(1):219–224, Jan 2013.
- [11] L. Blackmore, M. Ono, A. Bektassov, and B.C. Williams. A probabilistic particle-control approximation of chance-constrained stochastic predictive control. *IEEE Transactions on Robotics*, 26(3):502–517, June 2010.
- [12] L. Blackmore, M. Ono, and B. C. Williams. Chance-constrained optimal path planning with obstacles. *IEEE Transactions on Robotics*, 27(6):1080–1094, 2011.
- [13] M. Farina, L. Giulioni, L. Magni, and R. Scattolini. An approach to output-feedback mpc of stochastic linear discrete-time systems. *Automatica*, 55:140–149, 2015.
- [14] L. Hewing and M. Zeilinger. Stochastic model predictive control for linear systems using probabilistic reachable sets. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 5182–5188. IEEE, 2018.
- [15] M. Lorenzen, F. Dabbene, R. Tempo, and F. Allgoewer. Constraint-tightening and stability in stochastic model predictive control. *IEEE Transactions on Automatic Control*, 62(7):3165–3177, July 2017.
- [16] L. Hewing and M. Zeilinger. Scenario-based probabilistic reachable sets for recursively feasible stochastic model predictive control. *IEEE Control Systems Letters*, 4(2):450–455, 2020.
- [17] J. Paulson and A. Mesbah. Nonlinear model predictive control with explicit backoffs for stochastic systems under arbitrary uncertainty. *IFAC-PapersOnLine*, 51(20):523–534, 2018.
- [18] E. Bradford, L. Imsland, and E. A. del Rio Chanona. Nonlinear model predictive control with explicit back-offs for gaussian process state space models. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 4747–4754. IEEE, 2019.
- [19] E. Bradford, L. Imsland, D. Zhang, and E. A. del Rio Chanona. Stochastic data-driven model predictive control using Gaussian processes. *Computers & Chemical Engineering*, 139:106844, 2020.

- [20] K. Wang and S. Gros. Recursive feasibility of stochastic model predictive control with mission-wide probabilistic constraints. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 2312–2317. IEEE, 2021.
- [21] F. Oldewurtel, D. Sturzenegger, P. M. Esfahani, G. Andersson, M. Morari, and J. Lygeros. Adaptively constrained stochastic model predictive control for closed-loop constraint satisfaction. In *2013 American Control Conference (ACC)*, pages 4674–4681, Washington, DC, USA, 2013.
- [22] D. Chatterjee and J. Lygeros. On stability and performance of stochastic predictive control techniques. *IEEE Transactions on Automatic Control*, 60(2):509–514, 2015.
- [23] P. H. Baxendale. Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *The Annals of Applied Probability*, 15(1B):700 – 738, 2005.
- [24] C. Micchelli, Y. Xu, and H. Zhang. Universal kernels. *Journal of Machine Learning Research*, 7(Dec):2651–2667, 2006.
- [25] C. E. Rasmussen and C. K. I. Williams. Gaussian processes for machine learning. 2006. *The MIT Press, Cambridge, MA, USA*, 2006.
- [26] Subhashis Ghosal and Anindya Roy. Posterior consistency of gaussian process prior for nonparametric binary regression. *The Annals of Statistics*, 34(5):2413–2429, 2006.
- [27] O. Häggström. *Finite Markov chains and algorithmic applications*. Number 52. Cambridge University Press, 2002.
- [28] X. Shen and L. Wasserman. Rates of convergence of posterior distributions. *The Annals of Statistics*, 29(3):687–714, 2001.
- [29] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, 2006.
- [30] E. Snelson and Z. Ghahramani. Sparse gaussian processes using pseudo-inputs. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems*, volume 18. MIT Press, 2005.
- [31] M. Deisenroth and J. Ng. Distributed Gaussian processes. In Francis Bach and David Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 1481–1490, Lille, France, 07–09 Jul 2015. PMLR.
- [32] M. Cannon, B. Kouvaritakis, S. V. Rakovic, and Q. Cheng. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Transactions on Automatic Control*, 56(1):194–200, Jan 2011.
- [33] C. E. Rasmussen and H. Nickisch. Gaussian processes for machine learning (gpml) toolbox. *The Journal of Machine Learning Research*, 11:3011–3015, 2010.
- [34] L. Hewing, K. Wabersich, and M. Zeilinger. Recursively feasible stochastic model predictive control using indirect feedback. *Automatica*, 119:109095, 2020.

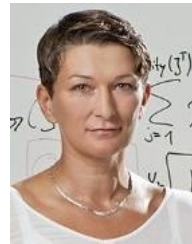


Alexandre Capone received the B.S. and M.S. degrees in mechanical engineering from RWTH Aachen University, Aachen, Germany, in 2014 and 2016, respectively. He received his Ph.D. from TUM in 2024. He was a visiting researcher at the California Institute of Technology in 2022. He is currently a postdoctoral researcher at Carnegie Mellon University. He was awarded the Springorium Commemorative coin for his studies at RWTH Aachen. His main research interests include Gaussian processes, safe learning-

based control, reinforcement learning, uncertainty quantification, and event-triggered control, with applications to human-machine interaction, autonomous driving, and robotic systems.



Tim Brüdigam received his B.Sc. and M.Sc. degree in electrical engineering from the Technical University of Munich (TUM), Germany in 2014 and 2017, respectively. In 2022, he obtained his Ph.D. degree from TUM. During his studies, he was a research scholar at the California Institute of Technology in 2016. He joined the Chair of Automatic Control Engineering at TUM as a research associate in 2017 and was a visiting researcher at the University of California, Berkeley in 2021. His main research interest lies in advancing Model Predictive Control (MPC), especially stochastic MPC, with possible application in automated driving.



Sandra Hirche (M03-SM11-F'20) received the Diplom-Ingenieur degree from Technical University Berlin in 2002 and the Doktor-Ingenieur degree from Technical University Munich (TUM), Germany, in 2005.

From 2005 to 2007 she was awarded a JSPS Postdoc scholarship at the Fujita Laboratory, Tokyo Institute of Technology, Japan. She became associate professor at TUM in 2008. Since 2013 she is TUM Liesel Beckmann Distinguished Professor and heads the Chair of Information-oriented Control in the Department of Electrical and Computer Engineering at TUM. Her main research interests include learning, cooperative, and networked control with application in human-robot interaction, haptics, multi-robot systems, and general robotics. She has published more than 200 papers in international journals, books and refereed conferences. She has received multiple awards such as the Rohde Schwarz PhD Award, the 2005 IFAC World Congress Best Poster Award, the Best Paper Award of the 2009 IEEE Worldhaptics, and the Outstanding Student Paper Award of the 2018 IEEE Conference on Decision and Control. In 2013 she has been awarded with an ERC Starting Grant on the "Control based on Human Models" and in 2019 with an ERC Consolidator Grant on "Safe data-driven control for human-centric systems".