# Stochastic MPC framework for controlling the average constraint violation

Milan Korda[1], Ravi Gondhalekar[2], Frauke Oldewurtel[3] and Colin N. Jones[1]

October 21, 2013

**Abstract**

This paper considers linear discrete-time systems with additive, bounded, disturbances subject to hard control input bounds and a stochastic constraint on the amount of state-constraint violation averaged over time. The amount of violations is quantified by a loss function and the averaging can be weighted, corresponding to exponential forgetting of past violations. The freedom in the choice of the loss function makes this formulation highly flexible – for instance, probabilistic constraints or integrated chance constraints can be enforced by an appropriate choice of the loss function. For the type of constraint considered, we develop a recursively feasible receding horizon control scheme exploiting the averaged-over-time nature by explicitly taking into account the amount of past constraint violations when determining the current control input. This leads to a significant reduction in conservatism. As a simple extension of the proposed approach we show how time-varying state-constraints can be handled within our framework. The computational complexity (online as well as offline) is comparable to existing model predictive control schemes. The effectiveness of the proposed methodology is demonstrated by means of a numerical example.

## 1 Introduction

There is a significant gap between the theory of model predictive control (MPC) and its practical usage. Indeed, theoretical results on stability and recursive feasibility of MPC are available in nominal as well as robust settings [21, 28]. These results are, however, rarely used in practice: In most applications the problem of recursive feasibility is circumvented by employing soft constraints, and (robust) stability is not enforced by design but evaluated a posteriori. The main reason for this discrepancy is the excessive

[1]Milan Korda and Colin N. Jones are with the Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne, Switzerland. {milan.korda, colin.jones}@epfl.ch.

[2]Ravi Gondhalekar is with the Department of Chemical Engineering, University of California Santa Barbara, USA. gondhalekar@engineering.ucsb.edu.

[3]Frauke Oldewurtel is with the Power Systems Laboratory, Department of Electrical Engineering, Swiss Federal Institute of Technology in Zurich (ETHZ), Switzerland. oldewurtel@eeh.ee.ethz.ch.

conservatism of theoretically sound approaches compared to empirical methods, especially in a robust setting [26].

A possible remedy is for the constraint specification to be stochastic rather than robust, or, more generally, for occasional constraint violation to be allowed in a well-defined, but not necessarily probabilistic, manner. An MPC controller can then exploit the freedom to occasionally violate the constraint in order to achieve a lower objective cost, thereby reducing the conservatism.

The main theoretical challenge when we allow occasional constraint violation is recursive feasibility. Previous approaches [6, 7, 8, 9, 18, 20] considered constraints on the marginal distribution of the state, typically point-wise in time probabilistic constraints. In those works the constraints were enforced by controlling the conditional probability of constraint violations between two consecutive time instances without taking into account the past behavior of the state process. Although convenient for a receding horizon controller implementation, this necessarily limits the achievable benefits of the probabilistic specification. Indeed, even the least-restrictive[1] formulation of [18] is conservative in certain situations.

In contrast, the proposed method exploits the information on the past behavior of the state process – namely the cumulative amount of past constraint violations – when determining the current control input, and thereby is capable of enforcing stochastic properties on the closed-loop state-process as a whole. The quantity constrained in this paper is the amount of constraint violations averaged over time, where the amount is quantified by a loss function and the averaging can be weighted, allowing for exponential forgetting of past constraint violations. Since our specification is stochastic, the average amount of violations is allowed to occasionally exceed the prescribed level; we prove that under the control policy proposed this quantity converges back to the prescribed level with probability one and derive bounds on the expected return time. A simple modification of our approach allows us to handle a time-varying state-constraint, where the variation is generated by the output of a linear system driven by a fictitious bounded disturbance. The freedom in the choice of the loss function makes our approach fairly general and flexible. Indeed, probabilistic constraints (see, e.g., [18, 20]) or integrated chance constraints (see, e.g., [25]) can be readily implemented with a particular choice of the loss function.

The averaged-over-time constraints considered are as expressive and practically relevant as their traditional counterparts. For instance, this type of constraint is natural in building climate control, where comfort specifications for the room temperature prescribe the allowed amount of excursion from a given temperature band (in Kelvin hours per year) [10]. Other examples include energy-efficient datacenter cooling subject to the constraints on the number of delayed queries per unit of time (see, e.g., [17]), or performance (e.g., power output) maximization of a machine subject to fatigue constraints (see [9] for a concrete example of wind-turbine control).

Importantly, the loss function considered can also have an economic character in which case the approach presented allows imposing stochastic constraints on the running (weighted) average cost incurred. This quantity (or more precisely its asymptotic behaviour) has been recently studied extensively in a deterministic setting in the economic MPC literature; see,

[1] Least-restrictive in terms of the size of the feasible domain.

e.g., [1, 14].

The presented methodology is a generalization of the authors' previous work [19], where only probabilistic constraints and simple averaging were considered. The approach builds on the first-step stochastic invariance introduced in [18] and originally developed for nominal MPC in [12].

One of the virtues of the approach is its complete independence of the MPC problem cost function and prediction policy parametrization. In addition, the averaged-over-time constraints are enforced via *affine* constraints only. This flexibility is facilitated by the use of controlled invariant sets [2, 3], which can be parametrized either explicitly or implicitly. In either case, the computational requirements are comparable to their respective nominal and robust counterparts – in the case of explicit parametrization to the first-step nominal MPC of [11, 12]; in the case of implicit parametrization to any of the traditional robust / stochastic MPC schemes such as affine disturbance feedback [13, 18], prestabilization [20] or tubes [22] (assuming the presented approach is used with the same policy parametrization as the traditional one).

The paper is organized as follows. The problem to be solved is formulated in Section 2; a general set-based solution is described in Section 3. Section 3.2 studies in detail the convergence speed of the average amount of violations towards the prescribed level. Section 4 shows how the presented methodology can be incorporated into an MPC framework, discusses the explicit and implicit policy parametrizations (Section 4.1), and computational complexity (Section 4.2). Section 5 describes several extensions of the approach, including the time-varying constrains (Section 5.3). A numerical example is presented in Section 6.

### 1.1 Notation

Throughout the article $\mathbb{R}$ denotes the set of reals, $\mathbb{N}_{>0}$ the set of positive integers and $\mathbb{N}_{\geq 0}$ the set of nonnegative integers. Let also $x_i^j := (x_i, x_{i+1}, \ldots, x_j)$ denote a finite sequence of consecutive vectors, and $\{x_j \in \mathcal{M}\}_{j=i}^{\infty}$ an infinite sequence of vectors $(x_i, x_{i+1}, \ldots)$ with $x_j \in \mathcal{M}$ for all $j \geq i$. All random variables are defined on a common probability space with an associated probability measure $P(\cdot)$. The expectation with respect to this probability measure is denoted by $\mathbf{E}\{\cdot\}$; the conditional probability and expectation based on information available up to time $t$ are denoted by $P_t(\cdot)$ and $\mathbf{E}_t\{\cdot\}$, respectively. For a set $\mathcal{M}$ and $N \in \mathbb{N}_{>0}$, let $\mathcal{M}^N$ be the Cartesian product of the set $N$-times with itself. The indicator function of a set or a random event is denoted by $\mathbb{I}_{[\cdot]}$.

## 2 Problem statement

We consider the linear time-invariant stochastic dynamic system

$$x_{t+1} = Ax_t + Bu_t + w_t, \ t \in \mathbb{N}_{\geq 0} \tag{1}$$

with the state $x_t \in \mathbb{R}^n$, the control $u_t \in \mathbb{R}^m$, and the i.i.d. disturbance sequence $w_t \in \mathbb{R}^n$. We assume that the state $x_t$ is known at time $t$ for all $t \in \mathbb{N}_{\geq 0}$, and that the pair $(A, B)$ is

stabilizable. The control inputs are subject to hard bounds of the form

$$u_t \in \mathcal{U} \subseteq \mathbb{R}^m, \; t \in \mathbb{N}_{\geq 0} \; , \tag{2}$$

for some polyhedral input constraint set $\mathcal{U}$. The disturbance takes values in a compact polyhedron $\mathcal{W}$, i.e.,

$$w_t \in \mathcal{W} \subset \mathbb{R}^n, \quad \forall\, t \in \mathbb{N}_{\geq 0} \; .$$

We let $\boldsymbol{w}$ denote a random variable having the common distribution of $w_t$, $t \in \mathbb{N}_{\geq 0}$, that is, $P(\boldsymbol{w} \in C) = P(w_t \in C)$ for all $C \subset \mathbb{R}^n$ and $t \in \mathbb{N}_{\geq 0}$.

## 2.1 Chance constraint formulations

The essence of the paper is the handling of a single (extensions to multiple constraints are described in Section 5.1) state constraint

$$g^T x_t \leq h \; , \tag{3}$$

$g \in \mathbb{R}^n$, $h \in \mathbb{R}$, in a stochastic manner – loosely speaking, the goal is to not violate (3) very much and/or very often. The amount of violation at time $t$ is measured by $l(g^T x_t - h)$, where

$$l : \mathbb{R} \to \mathbb{R} \tag{4}$$

is a loss function. The only assumption we make on $l(\cdot)$ is the following:

**Assumption 1.** *The loss function $l(\cdot)$ is nondecreasing and lower semicontinuous.*

Note in particular that the function $l(\cdot)$ is not required to be convex, continuous or non-negative – negative values can be interpreted as a reward for not violating the constraint. The assumption that $l(\cdot)$ is nondecreasing underpins the entire proposal, as it provides a measure of distance from a state to the constraint. In contrast, the assumption of lower-semicontinuity is merely of a technical nature.

**Robust constraint** Apart from a stochastic specification (described in detail below) we impose an auxiliary robust deterministic bound on the maximum loss at each time instant:

$$l(g^T x_t - h) \leq \bar{\xi} \quad \forall\, t \in \mathbb{N}_{\geq 0} \; , \tag{5}$$

where the maximum allowed loss $\bar{\xi} > 0$ is typically prescribed by or derived from the problem specification.

**Stochastic constraint** Previous works on stochastic MPC focused predominantly on *point-wise in time* constraints of the form

$$\mathbf{E}\{l(g^T x_t - h)\} \leq \xi \quad \forall t \in \mathbb{N}_{>0} \tag{6}$$

for some $\xi \leq \bar{\xi}$. For instance, the probabilistic constraint (see, e.g., [7, 8, 18, 19])

$$P(g^T x_t > h) \leq \xi \quad \forall t \in \mathbb{N}_{>0}$$

can be encoded by taking $l(\cdot)$ equal to the indicator function of the positive real line. Similarly, the integrated chance constraint (see, e.g., [25])

$$\mathbf{E}\{\max\{g^T x_t - h, 0\}\} \leq \xi \quad \forall t \in \mathbb{N}_{>0}$$

is obtained by taking $l(\cdot) = \max\{\cdot, 0\}$.

The main difficulty when applying MPC in a stochastic setting is in guaranteeing feasibility while fully exploiting the stochastic nature of the constraint (6). In most cases, the constraint (6) is enforced via a sufficient *one-step conditional* constraint

$$\mathbf{E}\{l(g^T x_{t+1} - h) \mid x_t\} \leq \xi \ . \tag{7}$$

Since the left-hand side of this constraint is a function of $x_t$ only and the disturbance sequence is i.i.d., the decision of the controller can be based only on the current state and on the knowledge of the distribution of the disturbance $w_t$, irrespective of the past behavior of the state and/or disturbance process. That is why this constraint is amenable to invariant-set based techniques commonly used in MPC. However, the constraint (7) is far more conservative than the constraint (6) because in (7) we require the satisfaction *conditionally* on $x_t$ (i.e., for all $x_t$ that can be reached at time $t$ by the disturbance sequence under a given control policy), whereas in (6) we require the satisfaction *on average* over all realizations of the entire disturbance process up to time $t$.

In this work, rather than imposing constraints point-wise in time, such as (6) or (7), we impose constraints on the closed-loop state process as a whole. In particular, we constrain the cumulative loss averaged over time where the averaging can be weighted, corresponding to exponential forgetting of the past constraint violation. To this end, define the cumulative loss up to time $t$ by

$$v_t := \sum_{k=0}^{t} \gamma^{t-k} l(g^T x_k - h) \ , \quad t \in \mathbb{N}_{\geq 0} \ , \tag{8}$$

where the decay rate, or forgetting factor, $\gamma \in [0, 1]$ is a design parameter. Define also the running weighted sum (that will serve as a normalization factor)

$$s_t := \sum_{k=0}^{t} \gamma^{t-k} = \begin{cases} \dfrac{1-\gamma^{t+1}}{1-\gamma}, & \gamma \in [0, 1) \\ t+1 & \gamma = 1 \ . \end{cases} \tag{9}$$

The ratio $v_t/s_t$ is then the weighted average amount of violation accumulated up to time $t$ and is the main quantity of interest in this paper. In order to state the constraint imposed on the stochastic process $v_t/s_t$, define the first return time of this process to below a given level $\xi \geq 0$ as

$$\tau_t := \inf\{i \geq t \mid v_i/s_i \leq \xi\} \in \{t, t+1, \ldots\} \cup \{+\infty\} \ , \tag{10}$$

where we define the infimum of an empty set as infinity. The first return time $\tau_t$ is an integer random variable whose distribution depends on the stochastic properties of $v_t/s_t$ which in turn depend on the control policy applied to the system (1).

With these ingredients the stochastic requirement we impose reads

$$\begin{cases} \mathbf{E}_t\left\{\dfrac{v_{t+1}}{s_{t+1}}\right\} \leq \xi & \text{if} \quad \dfrac{v_t}{s_t} \leq \xi \qquad (11a) \\ \lim\limits_{k\to\infty} \dfrac{v_{\min\{t+k,\tau_t\}}}{s_{\min\{t+k,\tau_t\}}} \leq \xi & \text{if} \quad \dfrac{v_t}{s_t} > \xi\,, \qquad (11b) \end{cases}$$

where the convergence in (11b) is required to be with probability one (or "almost surely").

In words, rather than controlling the amount of constraint violation at each time separately, at each time instant $t$ we control the running (weighted) average amount of constraint violation, where the average is taken over the entire history of the state trajectory up to time $t$. The average amount of violation is controlled in such a manner that whenever it is below the prescribed level $\xi$ (Eq. (11a)), then it remains below $\xi$ in expectation at the next time instant; if the average amount of violation happens to exceed the level of $\xi$ (Eq. (11b)), we require that it converges back towards $\xi$ with probability one (either in a finite time or asymptotically depending on whether $\tau_t$ is finite or not). See Figure 1 for illustration and Section 3.2 for a detailed analysis of the convergence in (11b), for example for conditions under which the convergence occurs in a finite time.

The constraint (11), as opposed to (6), makes the dependence on the past behavior of the state process explicit, thereby allowing us to leverage the information on the past behavior when determining the current control input.

**Remark 1.** *Note that constraining the expected average loss at time $t+1$ in (11a) is just one possible choice among others. Section 5.2 describes a straightforward extension of the approach where we instead constrain the probability of exceeding the prescribed level $\xi$, i.e., impose $P_t(v_{t+1}/s_{t+1} > \xi) \leq 1-\delta$ for some $\delta \in [0,1]$.*

**Remark 2.** *The freedom in the choice of the loss function $l(\cdot)$ in (8) and the discount factor $\gamma$ makes the requirement (11) highly flexible. For instance by choosing $l(\cdot)$ equal to the indicator function of the positive real line and $\gamma = 1$ the quantity $v_t/s_t$ becomes the (simple) average* number *of constraint violations up to time $t$; this specification was considered in [19]. Note also that the loss function $l(\cdot)$ can have an economic interpretation in which case the specification (11) imposes a stochastic requirement on the (weighted) average cost incurred. This type of cost function has been recently extensively studied in the economic MPC literature; see, e.g., [1, 14].*

**Remark 3.** *(Parameters) The decay rate $\gamma \in [0,1]$ controls the speed with which the past constraint violation is forgotten – with $\gamma < 1$ we obtain exponentially decaying weights from current time $t$ to zero, whereas with $\gamma = 1$ violation is weighted equally at all time instances and therefore no forgetting occurs. If $\gamma = 0$ the constraint violation history is disregarded and the constraint (11) becomes $\mathbf{E}_t\{l(g^T x_{t+1} - h)\} \leq \xi$.*

*The allowed level of constraint violation $\xi \geq 0$ is typically derived from application requirements, but it can also be viewed as a tuning parameter for adjusting the conservativeness of the controller.*

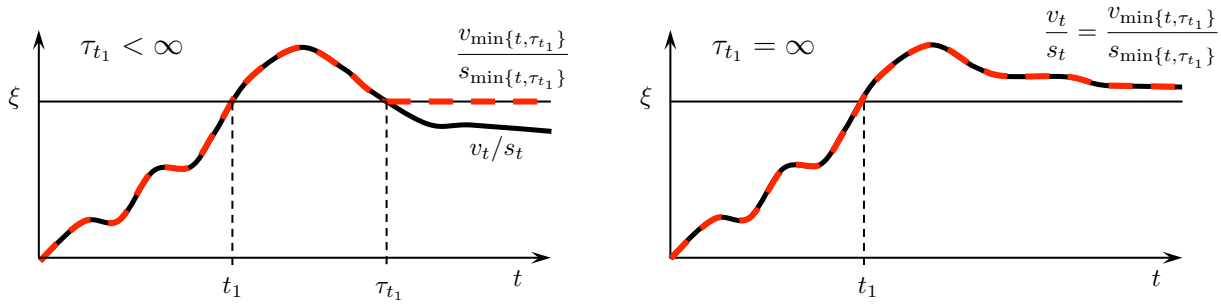


**Figure 1:** Illustration of the return time concept: At time $t_1$ the average violation process $v_t/s_t$ (solid black line) crosses the level $\xi$; the integer random variable $\tau_{t_1}$ is then the first return time of this process back to below $t_1$. The second line of specification (11) requires that the "stopped" average violation process $v_{\min\{t,\tau_{t_1}\}}/s_{\min\{t,\tau_{t_1}\}}$ (dashed red line) converges back to below $\xi$, either in finite time (left panel), corresponding to $\tau_{t_1} < \infty$, or asymptotically (right panel), corresponding to $\tau_{t_1} = \infty$.

**Remark 4.** *(Convergence speed) There is a large degree of freedom in (11b) since the speed of convergence of $v_t$ to $\xi$ is not prescribed. However, the convergence speed can be influenced by the parameters of our approach (e.g., $\gamma$) and is analyzed in detail in Section 3.2.*

# 3 Main results

In this section we present a recursively feasible receding horizon control policy under which the closed-loop state process satisfies the constraints (11) and (5). The main idea is simple: Keep track of the amount of loss accumulated in the past; if the loss is "large" (i.e., $v_t/s_t > \xi$), enforce the constraint

$$\mathbf{E}_t\{l(g^T x_{t+1} - h)\} \leq \alpha \ , \tag{12}$$

where $\alpha \leq \xi$; if it is "small" (i.e., $v_t/s_t < \xi$), loosen the constraint (12) appropriately; and at all times enforce such invariance constraints that the constraint (12) is feasible at the next time instant. The approach therefore adaptively loosens / tightens the state constraint based on the size of the loss accumulated in the past.

One can then extend the idea by also allowing loosening of the invariance constraint itself by building a family of nested one-step reachability sets around an innermost invariant set in which the existence of a control law satisfying (12) is ensured. First, we describe the simple (single-layer) version, and afterwards, in Section 3.1, we present the multi-layer one.

**Remark 5.** *The right-hand side of the constraint (12), $\alpha$, controls the speed of convergence of $v_t/s_t$ to below the prescribed violation level $\xi$; the smaller the $\alpha$, the faster the convergence (but the more conservative the controller). See Section 3.2 for details on convergence speed.*

Now we give a precise mathematical formulation of the above discussion. We start by stating the following simple Lemma whose proof is omitted for brevity:

**Lemma 1.** *Let $\nu : \mathbb{R} \to \mathbb{R}$ be a non-decreasing lower-semicontinuous function and $\omega$ a bounded scalar random variable. Then*

- *The function* $\mu \mapsto \mathbf{E}\{\nu(\mu+\omega)\}$ *is non-decreasing lower-semicontinuous*
- $\sup\{\mu \in \mathbb{R} \mid \mathbf{E}\{\nu(\mu+\omega)\} \leq a\}$ *is attained or infinite for all* $a \in \mathbb{R}$.

This Lemma is instrumental in enforcing both the robust constraint (5) and the stochastic specification (11). In particular, it follows immediately from this Lemma (applied with $\nu(\cdot) = l(\cdot)$ and $\omega = 0$) and Assumption 1 that the feasible set of the constraint (5) is a half-space given by

$$\bar{\mathcal{X}} := \{x \mid l(g^T x - h) \leq \bar{\xi}\} = \big\{x \mid g^T x \leq h + l^{-1}(\bar{\xi})\big\},$$

where

$$l^{-1}(a) := \sup\{y \in \mathbb{R} \mid l(y) \leq a\} \in [-\infty, +\infty], \quad a \in \mathbb{R}. \tag{13}$$

Note that in the extreme cases $l^{-1}(\bar{\xi}) = -\infty$ or $l^{-1}(\bar{\xi}) = +\infty$ the set $\bar{\mathcal{X}}$ is empty (and hence (5) is infeasible) or equal to $\mathbb{R}^n$ (and hence (5) is redundant), respectively.

Next we analyze the one-step conditional constraint (12). Given $x = x_t$, this constraint is equivalent to the existence of $u \in \mathcal{U}$ such that

$$\mathbf{E}\{l(g^T(Ax + Bu + w) - h)\} \leq \alpha \ , \tag{14}$$

where the expectation is over the distribution of $w$ (which coincides with the distribution of $w_t$ for all $t \in \mathbb{N}_{\geq 0}$). The constraint (14) gives rise to the first key ingredient of the proposed approach, the *stochastic feasibility set.*

**Definition 1.** *The stochastic feasibility set of the constraint (14) is*

$$\mathcal{X}_s := \big\{x \in \mathbb{R}^n \mid \exists\, u \in \mathcal{U} \text{ s.t. } \mathbf{E}\{l(g^T(Ax + Bu + w) - h)\} \leq \alpha\big\} \ .$$

In plain words, $\mathcal{X}_s$ is the set of states for which there exists an admissible input such that the conditional constraint (14) is satisfied. The subscript $s$ signifies "stochastic".

In what follows we show that $\mathcal{X}_s$ is in fact a polyhedron. To this end observe that

$$\mathbf{E}\{l(g^T(Ax + Bu + w) - h)\} = \mathbf{E}\{l(g^T(Ax + Bu) - h + g^T w)\} \leq \alpha$$

and consider the function

$$f(\mu) := \mathbf{E}\{l(\mu + g^T w)\} \ . \tag{15}$$

Note that if the random variable $g^T w$ has a probability density function $\mathrm{pdf}_{g^T w}(\cdot)$, then (15) is simply

$$f(\mu) = \int_{-\infty}^{\infty} l(\mu + y)\, \mathrm{pdf}_{g^T w}(y)\, dy \ .$$

According to the first part of Lemma 1, the function $f(\cdot)$ is non-decreasing and lower-semicontinuous. Using the second part of the same Lemma leads to the following key observation:

**Lemma 2.** *The inequality (14) is satisfied if and only if*

$$g^T(Ax + Bu) \leq h + q(\alpha) \ , \tag{16}$$

*where the function* $q : \mathbb{R} \to \mathbb{R}$ *is defined by*

$$q(\alpha) := \sup\{\mu \in \mathbb{R} \mid f(\mu) \leq \alpha\} \ . \tag{17}$$

*Proof.* Use Lemma 1 with $\nu(\cdot) = l(\cdot)$ and $\omega = g^T \boldsymbol{w}$. □

Note that the function $q(\cdot)$ is nondecreasing in the right-hand side of the constraint (14), $\alpha$, and, depending on the value of $\alpha$, it can be either negative or positive corresponding to either tightening or loosening of the constraint $g^T(Ax + Bu) \leq h$.

The function $q(\cdot)$ depends only on the distribution of the disturbance $\boldsymbol{w}$, loss function $l(\cdot)$ and on the normal vector $g$, all known in advance. Therefore, the function $q(\cdot)$ (or equivalently $f(\cdot)$) can be precomputed to virtually arbitrary precision offline, depending on the loss function $l(\cdot)$ and the distribution of $\boldsymbol{w}$ either analytically or, e.g., by Monte Carlo techniques.

Lemma 2 implies that the stochastic feasibility set $\mathcal{X}_s$ is a polyhedron.

**Corollary 1.** *If the function* $l : \mathbb{R} \to \mathbb{R}$ *is nondecreasing, then the stochastic feasibility set* $\mathcal{X}_s$ *is the polyhedron*

$$\mathcal{X}_s = \{x \in \mathbb{R} \mid \exists\, u \in \mathcal{U} \text{ s.t. } g^T(Ax + Bu) \leq h + q(\alpha)\} \ . \tag{18}$$

*Proof.* In view of Lemma 2, the representation (18) clearly holds. Consequently, $\mathcal{X}_s$ is a projection of a polyhedron and therefore also a polyhedron. □

As a result, the stochastic feasibility set $\mathcal{X}_s$ can be readily obtained from problem data via a single polyhedral projection. Note, however, that explicit computation of the set is not necessary; indeed, all sets associated with our approach can be described implicitly using a terminal constraint and suitable constraints on the predicted state trajectory, without requiring the explicit computation of polyhedral projections (see Section 4.1 for details).

The second key ingredient is a stochastic robust controlled invariant set (SRCI set) $\mathcal{S}$.

**Definition 2.** *A set* $\mathcal{S} \subseteq \mathcal{X}_s \cap \bar{\mathcal{X}}$ *is a stochastic robust controlled invariant set if it satisfies the following condition:*

$$\forall x \in \mathcal{S}\ \exists\, u \in \mathcal{U} \text{ s.t. } Ax + Bu + w \in \mathcal{S} \quad \forall w \in \mathcal{W}, \tag{19}$$

$$\mathbf{E}\big\{l(g^T(Ax + Bu + \boldsymbol{w}) - h)\big\} \leq \alpha \ . \tag{20}$$

In words, an SRCI set $\mathcal{S}$ is a robust controlled invariant subset of $\mathcal{X}_s \cap \bar{\mathcal{X}}$ such that the constraint (14) is satisfied under the invariance-inducing control input; by requiring that $\mathcal{S} \subset \bar{\mathcal{X}}$ we ensure that constraint (5) is also satisfied. Note that in general neither $\mathcal{X}_s \subseteq \{x \mid g^T x \leq h\}$ nor $\mathcal{S} \subseteq \{x \mid g^T x \leq h\}$. In view of Lemma 2, constraint (20) translates to the affine constraint

$$g^T(Ax + Bu) \leq h + q(\alpha).$$

We need the following assumption on $\mathcal{S}$:

**Assumption 2.** *A nonempty* polyhedral *SRCI set $\mathcal{S}$ exists and has been characterized (either explicitly or implicitly).*

The requirement of $\mathcal{S}$ being polyhedral has been imposed for the sake of computational tractability (see Section 4); all theoretical results presented in this paper hold for *any* nonempty SRCI set $\mathcal{S}$.

Now we proceed to define a quantity called *constraint leeway* that will control loosening of the one-step conditional constraint (12) (or equivalently (14)), enabling us to define a control law enforcing satisfaction of (11) in closed-loop. First, observe that (11a) is equivalent to

$$\mathbf{E}_t\{v_{t+1}\} \leq \xi s_{t+1} \ .$$

Therefore, (11a) is satisfied if

$$\begin{aligned}\xi s_{t+1} - \mathbf{E}_t\{v_{t+1}\} &= \xi s_{t+1} - \gamma v_t - \mathbf{E}_t\{l(g^T x_{t+1} - h)\} \\ &= (\gamma s_t + 1)\xi - \gamma v_t - \mathbf{E}_t\{l(g^T x_{t+1} - h)\} \\ &= \gamma(\xi s_t - v_t) + \xi - \mathbf{E}_t\{l(g^T x_{t+1} - h)\}\end{aligned}$$

is nonnegative, or equivalently

$$\mathbf{E}_t\{l(g^T x_{t+1} - h)\} \leq \gamma(\xi s_t - v_t) + \xi \ .$$

We therefore define the *constraint leeway* $\beta_t$ at time $t \in \mathbb{N}_{\geq 0}$ as

$$\beta_t := \max\left\{\gamma(\xi s_t - v_t) + \xi,\, \alpha\right\} \in [\alpha, \infty) \ . \tag{21}$$

Whenever $\gamma(\xi s_t - v_t) + \xi \geq \alpha$ (which is implied by $v_t/s_t \leq \xi$), enforcing

$$\mathbf{E}\{l(g^T x_{t+1} - h) \mid x_t\} \leq \beta_t \tag{22}$$

guarantees $\mathbf{E}_t\{v_{t+1}/s_{t+1}\} \leq \xi$, which is (11a). If, however, $\gamma(\xi s_t - v_t) + \xi < \alpha$, then enforcing (22) does not guarantee the satisfaction of $\mathbf{E}_t\{v_{t+1}/s_{t+1}\} \leq \xi$. Nevertheless, enforcing (22) at each time step does guarantee the convergence $v_t/s_t \to \xi$ with probability one (since $\alpha \leq \xi$) as required by (11b). We elaborate on this conclusion now in an instructional single-layer setting and prove it formally in a more general multi-layer setting in Theorem 1. To derive a single-layer controller satisfying (11), define an auxiliary state

$$\chi_t := \xi s_t - v_t \quad \forall\, t \in \mathbb{N}_{\geq 0} \ , \tag{23}$$

and the admissible control-input set

$$\begin{aligned}\tilde{\mathcal{U}}(x_t, \chi_t) := \big\{ & u \in \mathcal{U} \text{ s.t.} \\ & Ax_t + Bu_t + w \in \mathcal{S} \ \forall\, w \in \mathcal{W} && \text{(24a)} \\ & \mathbf{E}\{l(g^T(Ax_t + Bu_t + w) - h)\} \leq \beta_t \big\} \ , && \text{(24b)}\end{aligned}$$

where $\beta_t$ depends on the auxiliary state $\chi_t$ through (21). Note that according to Lemma 2, the constraint (24b) translates to an affine constraint.

A basic single-layer set-valued control policy under which the satisfaction of (11) is satisfied can now be defined by

$$\tilde{\kappa}(x_t, \chi_t) \in \tilde{\mathcal{U}}(x_t, \chi_t), \quad t \in \mathbb{N}_{\geq 0} \ . \tag{25}$$

In other words, at time $t$ we enforce the invariance constraint (24a), which ensures that the one-step constraint (24b) is feasible at all times whenever it is feasible at time zero (since $\beta_t \geq \alpha$); the constraint (24b) is merely an appropriately loosened version of the constraint (12), or equivalently (14).

**Remark 6.** *In (24b) it is not necessary to evaluate the function $q(\cdot)$ defined in (17) for all $\beta_t \geq \alpha$. First, since $x_t \in \bar{\mathcal{X}}$ for all $t$, it suffices to consider $\beta_t \in [\alpha, \bar{\xi}]$. Further, one can choose fixed values $\xi =: \hat{\beta}_1 < \hat{\beta}_2 < \ldots < \hat{\beta}_{n_{\hat{\beta}}} \leq \bar{\xi}$ for which the function $q(\hat{\beta}_i)$ is pre-computed and then, online, round $\beta_t < 1$ to the nearest* lower *value of $\hat{\beta}_i$.*

A formal proof of the fact that, given $\tilde{\mathcal{U}}(x_0, \chi_0) \neq \emptyset$, the closed-loop state process under the control law $u_t = \tilde{\kappa}(x_t, \chi_t)$ is well-defined at all times and satisfies the constraint (11) is given in a more general, multi-layer, setting in Theorem 1.

## 3.1 Multi-layer approach

Under the control policy (25), the invariance constraint (24a) is independent of the accumulated loss, no matter how small it is. As a consequence, this constraint can be a major, if not sole, source of conservatism for small $v_t$ (i.e., for $\beta_t \gg \alpha$).

The multi-layer approach presented in this section alleviates this by loosening the invariance constraints as the amount of past violation decreases. The idea is to construct a family of nested one-step reachability sets around the SRCI set $\mathcal{S}$. The state is then allowed to climb up if the accumulated loss is "small" and is forced to climb down if the loss gets "large"; the (possibly loosened) one-step constraint (12) is active only at the lowermost level. The approach therefore adaptively loosens or tightens not only the constraint (12) but also the invariance constraint (24b).

To this end, define the robust reachability operator of a set $\mathcal{M} \subseteq \mathbb{R}^n$ as

$$\begin{aligned} &\mathrm{Reach}(\mathcal{M}) \\ &\quad := \{x \in \mathbb{R}^n \mid \exists u \in \mathcal{U} \text{ s.t. } Ax + Bu + w \in \mathcal{M} \ \forall\, w \in \mathcal{W}\} \ . \end{aligned} \tag{26}$$

The family of nested reach sets of length $n_s$ is then given by

$$\mathcal{S}_1 := \mathcal{S}, \tag{27a}$$

$$\mathcal{S}_{k+1} := \mathrm{Reach}(\mathcal{S}_k) \cap \bar{\mathcal{X}}, \quad k \in \{1, \ldots, n_s - 1\} \ . \tag{27b}$$

**Remark 7.** *The nested property $\mathcal{S}_k \subseteq \mathcal{S}_{k+1}$ follows from the fact that $\mathcal{M} \subseteq \mathrm{Reach}(\mathcal{M})$ whenever the set $\mathcal{M}$ is robust controlled invariant and the fact that the* $\mathrm{Reach}(\cdot)$ *operator preserves the invariance.*

**Remark 8.** *When the sets $\mathcal{M}$ and $\mathcal{U}$ are polyhedral, the computation of* $\mathrm{Reach}(\mathcal{M})$ *amounts to a single polyhedral projection; consequently,* $\mathrm{Reach}(\mathcal{M})$ *is also polyhedral.*

Now we proceed to the definition of an integer random variable that will play the role of a layer index that controls to which layer (i.e., to which set $\mathcal{S}_k$) the state is permitted to move. To this end we analyze the difference $s_{t+k}\xi - v_{t+k}$. If $x_{t+1} \in \bar{\mathcal{X}}$ we have

$$\begin{aligned} s_{t+1}\xi - v_{t+1} &= \gamma(s_t\xi - v_t) + \xi - l(g^T x_{t+1} - h) \\ &\geq \gamma(s_t\xi - v_t) + \xi - \bar{\xi} \ . \end{aligned}$$

Continuing, we obtain

$$\begin{aligned} s_{t+2}\xi - v_{t+2} &= \gamma(s_{t+1}\xi - v_{t+1}) + \xi - l(g^T x_{t+2} - h) \\ &\geq \gamma^2(s_t\xi - v_t) - \gamma(\bar{\xi} - \xi) - (\bar{\xi} - \xi) \ , \end{aligned}$$

and, by induction,

$$s_{t+k}\xi - v_{t+k} \geq \gamma^k(s_t\xi - v_t) - (\bar{\xi} - \xi)\sum_{i=0}^{k-1}\gamma^i \ .$$

Consequently, as long as $x_{t+i} \in \bar{\mathcal{X}}$ for $i \in \{1, \ldots, k\}$, the requirement $s_{t+k}\xi - v_{t+k} \geq 0$ is satisfied if

$$s_t\xi - v_t \geq (\bar{\xi} - \xi)\psi(k),$$

where

$$\psi(k) := \begin{cases} \dfrac{1}{\gamma^k}\displaystyle\sum_{i=0}^{k-1}\gamma^i = \begin{cases} \dfrac{1}{\gamma^k}\dfrac{1-\gamma^k}{1-\gamma} \ , & \gamma \in (0,1) \\ k \ , & \gamma = 1 \end{cases} \\ +\infty \ , & \gamma = 0 \end{cases}$$

for all $k \geq 1$. Note that the function $\psi(\cdot)$ is nondecreasing and tends to infinity for all $\gamma \in [0, 1]$.

Define $\tilde{r}_t$ as

$$\tilde{r}_t := \max\{k \geq 1 \mid s_t\xi - v_t \geq (\bar{\xi} - \xi)\psi(k)\} \ , \tag{28}$$

with the standard convention that the maximum of an empty set is equal to $-\infty$. The layer index $r_t \in \{1, \ldots, n_s\}$ is then defined by

$$r_t := \max\{\min\{\tilde{r}_t, n_s\}, 1\}, \quad t \in \mathbb{N}_{\geq 0} \ . \tag{29}$$

In words, when $r_t = i$ for $i \geq 2$, we are guaranteed that $v_{t+k} \leq s_{t+k}\xi$ for all $k \in \{1, \ldots, i\}$ without any constraints imposed apart from $x_{t+k} \in \bar{\mathcal{X}}$. For $r_t = 1$ (i.e., at the lowermost layer of the layered structure) two situations can occur depending on the specific value of $v_t$: Either $v_{t+1} \leq s_{t+1}\xi$ for all $w_t \in \mathcal{W}$ provided that $x_{t+1} \in \bar{\mathcal{X}}$, or there may exist a disturbance realization $w_t \in \mathcal{W}$ such that $v_{t+1} > s_{t+1}\xi$. This motivates us to define a multi-layer control law which at a time $t$ imposes a robust inclusion of the state $x_{t+1}$ in the layer $\mathcal{S}_{r_t}$ and, in addition, at the lowermost level (i.e., for $r_t = 1$) imposes the same one-step expectation constraint (with tightness controlled by $\beta_t$) as in the single-layer approach.

Now we can define a set-valued multi-layer control policy. To this end define for all $t \in \mathbb{N}_{\geq 0}$ the sets

$$\begin{aligned}\mathcal{U}(x_t, \chi_t) := \{u \in \mathcal{U} \text{ s.t.} \\ & Ax_t + Bu_t + w \in \mathcal{S}_{r_t} \ \forall\, w \in \mathcal{W}, && \text{(30a)} \\ & \mathbf{E}\{l(g^T(Ax_t + Bu_t + w) - h)\} \leq \beta_t\} \ , && \text{(30b)}\end{aligned}$$

and

$$\Pi := \{(x_t, \chi_t) \mid \mathcal{U}(x_t, \chi_t) \neq \emptyset\} \ , \tag{31}$$

where $\beta_t$ and $r_t$ depend on the auxiliary state $\chi_t = \xi s_t - v_t$ through (21) and (29). The multi-layer control policy is then defined by

$$\kappa(x_t, \chi_t) \in \mathcal{U}(x_t, \chi_t), \quad t \in \mathbb{N}_{\geq 0} \ . \tag{32}$$

The Remark 6 about the single-layer policy (25) applies also to this control policy. Note also that, since $\mathcal{S}_1 = \mathcal{S}$, we recover the single-layer control policy for $n_s = 1$. Finally, observe that the constraint (30b) is redundant for $\beta_t \geq \bar{\xi}$ since then (30b) is implied by (30a) since $\mathcal{S}_{r_t} \subseteq \bar{\mathcal{X}}$; in particular (30b) is redundant whenever $r_t \geq 2$.

Now we can state and prove our main result:

**Theorem 1.** *Under the control law $u_t = \kappa(x_t, \chi_t)$ the following holds:*

*I. If $x_0 \in \mathcal{S}$, then $(x_0, \chi_0) \in \Pi$ (initial feasibility).*

*II. If $(x_t, \chi_t) \in \Pi$, then $(x_{t+1}, \chi_{t+1}) \in \Pi$ (recursive feasibility).*

*III. If $(x_0, \chi_0) \in \Pi$, then $x_t$ satisfies the constraint (11) (closed-loop constraint satisfaction).*

*Proof.* I. At time zero we have $r_t = 1$, and feasibility of (30a) and (30b) then follows from the definition of $\mathcal{S} = \mathcal{S}_1$.

II. Consider first $r_t = 1$. Given feasibility at time $t$, we know that the state at time $t+1$ will be in $\mathcal{S}_1 = \mathcal{S}$. Therefore, by definition of $\mathcal{S}$, the invariance constraint (30a) as well as the one-step conditional constraint (30b) will be feasible at time $t+1$.

Next, consider $r_t \geq 2$. Then, by definition of $r_t$ we have $(s_t - v_t) \geq (\bar{\xi} - \xi)(1+\gamma)/\gamma^2$, which implies that $\gamma(s_t - v_t) + \xi \geq \bar{\xi} + (\bar{\xi} - \xi)/\gamma \geq \bar{\xi}$, and hence $\beta_t \geq \bar{\xi}$ (by definition of $\beta_t$). Therefore in this case the constraint (30b) is satisfied whenever the invariance constraint (30a) is (since $\mathcal{S}_i \subset \bar{\mathcal{X}}$ for all $i \in \{1, \ldots, n_s\}$). Next, for $r_t \geq 2$ we know that the state at time $t+1$ will be in $\mathcal{S}_{r_t} = \text{Reach}(\mathcal{S}_{r_t - 1}) \cap \bar{\mathcal{X}}$. Feasibility of the invariance constraint (30a) at time $t+1$ follows since $r_t$ can only, by construction, decrease by at most one between two consecutive time instants (i.e., $r_{t+1} \geq r_t - 1$) since the maximal violation is $\bar{\xi}$.

III. Consider first $t \in \mathbb{N}_{>0}$ such that $v_t/s_t \leq \xi$. In that case we have, by definition of $\beta_t$, that

$$\begin{aligned}\mathbf{E}_t\{v_{t+1}\} &= \gamma v_t + \mathbf{E}_t\{l(g^T x_{t+1} - h)\} \leq \gamma v_t + \beta_t \\ &= \gamma v_t + \gamma(\xi s_t - v_t) + \xi = \xi(\gamma s_t + 1) = \xi s_{t+1}\end{aligned}$$

as desired by (11a).

Consider now a time instant $t \in \mathbb{N}_{>0}$ such that $v_t/s_t > \xi$ and let $\tau_t$ defined in (10) be the first return time of $v_t/s_t$ below $\xi$. Define further the process

$$\eta_k := -\chi_{t+k} = v_{t+k} - s_{t+k}\xi, \quad k \in \mathbb{N}_{\geq 0}, \tag{33}$$

and the stopped process $\eta_k^\tau := \eta_{\min\{k,\tau\}}$, where $\tau := \tau_t - t = \inf\{k \geq 0 \mid v_{t+k}/s_{t+k} \leq \xi\}$ is a stopping time. Now we show that $\eta_k^\tau$ is a supermartingale (i.e., the conditional expectation of its increments is non-positive). Indeed,

$$\begin{aligned}
&\mathbf{E}_{t+k}\{\eta_{k+1}^\tau - \eta_k^\tau\} = \\
&\quad = \mathbf{E}_{t+k}\{(\gamma\eta_k + l(g^T x_{t+k+1} - h) - \xi - \eta_k)\mathbb{I}_{[\tau>k]}\} \\
&\qquad + \mathbf{E}_{t+k}\{(\eta_\tau - \eta_\tau)\mathbb{I}_{[\tau\leq k]}\} \\
&\quad = \mathbf{E}_{t+k}\{(\gamma - 1)\eta_k + l(g^T x_{t+k+1} - h) - \xi\}\mathbb{I}_{[\tau>k]} \\
&\quad = \mathbf{E}_{t+k}\{l(g^T x_{t+k+1} - h) - \xi\}\mathbb{I}_{[\tau>k]} + (\gamma - 1)\eta_k\mathbb{I}_{[\tau>k]} \\
&\quad \leq 0 \ ,
\end{aligned}$$

where $\mathbb{I}_{[\cdot]}$ is the indicator function. Here the equalities follow from basic properties of the conditional expectation operator, and the inequality follows from the fact that $\tau > k$ implies $\mathbf{E}_{t+k}\{l(g^T x_{t+k+1} - h)\} = \beta_{t+k} = \alpha \leq \xi$ and $\eta_k \geq 0$. Therefore $\eta_k^\tau$ is a supermartingale bounded in expectation, and so by Doob's martingale convergence theorem (e.g., [29, Theorem 11.5]) it converges with probability one to some finite random variable $\eta_\infty^\tau$.

For those events on which $\tau_t < \infty$, (11b) is clearly satisfied. Therefore we only need to show that whenever $\tau_t = \infty$ (or $\tau = \infty$) we have $v_{t+k}/s_{t+k} \to \xi$. In the remainder of the proof we distinguish two cases.

First, assume $\gamma = 1$. Then $s_{t+k} \to \infty$ as $k \to \infty$ for $\gamma = 1$, and we have

$$\begin{aligned}
\lim_{k\to\infty} \frac{v_{t+k}}{s_{t+k}} - \xi &= \lim_{k\to\infty} \frac{v_{t+k} - \xi s_{t+k}}{s_{t+k}} = \lim_{k\to\infty} \frac{\eta_k}{s_{t+k}} \\
&= \lim_{k\to\infty} \frac{\eta_k^\tau}{s_{t+k}} = \lim_{k\to\infty} \frac{\eta_\infty^\tau}{s_{t+k}} = 0
\end{aligned}$$

with probability one as desired. Here we used the fact that $\eta_k = \eta_k^\tau$ on the event $\{\tau = \infty\}$.

Second, assume $\gamma \in [0, 1)$. Notice that for $v_{t+k}/s_{t+k} \to \xi$ to hold it is sufficient to have $\eta_k \to 0$. Noting that $\eta_k^\tau = \eta_k$ on the event $\{\tau = \infty\}$ and that $\eta_k^\tau$ converges, we only need to show

$$P\Big(\{\tau = \infty\} \cap \bigcup_{i=1}^{\infty} \{\inf_{k\geq 0}(\eta_k^\tau) > 1/i\}\Big) = 0 \ .$$

In words, this equation requires that whenever $\tau = \infty$, the process $\eta_k^\tau$ (which is equal to $\eta_k$ on the event $\{\tau = \infty\}$) crosses from above any positive level with probability one. Using Boole's inequality, a sufficient condition for this to hold is

$$P(\tau(i) = \infty) = 0 \quad \forall\, i \in \mathbb{N}_{>0} \ ,$$

where $\tau(i) := \inf\{k \geq 0 \mid \eta_k^\tau \leq 1/i\}$ is the first return time of $\eta_k^\tau$ below the level $1/i$. To show that $P(\tau(i) = \infty) = 0$ we over-bound the trajectories of $\eta_k^\tau$ by a random walk with a drift. By Assumption 1, the function $l(\cdot)$ is nondecreasing; thus

$$\begin{aligned} l(g^T x_{t+k+1} - h) &= l(g^T(Ax_{t+k} + Bu_{t+k}) - h + g^T w_{t+k}) \\ &\leq l(\bar{\mu} + g^T w_{t+k}) \ , \end{aligned}$$

where

$$\bar{\mu} := \max\{\mu \in \mathbb{R} \mid \mathbf{E}_{t+k}\{l(\mu + g^T w_{t+k})\} \leq \alpha\} = q(\alpha) \ ,$$

where the quantile function $q(\cdot)$ is defined in (17) and the equality follows by the i.i.d. assumption on $\{w_t\}$. Consequently, on the event $\{\tau = \infty\}$, we have

$$\eta_{k+1}^\tau = \gamma\eta_k^\tau + l(g^T x_{t+k+1} - h) - \xi \leq \gamma\eta_k^\tau + l(\bar{\mu} + g^T w_{t+k}) - \xi \ .$$

Therefore, on $\{\tau = \infty\}$, the trajectories of $\eta_k^\tau$ are bounded from above by the trajectories of the AR(1) process

$$X_{k+1} = \gamma X_k + z_k, \quad X_0 = \eta_0 > 0 \ ,$$

where $z_k := l(\bar{\mu} + g^T w_{t+k}) - \xi$ is an i.i.d. innovation with nonpositive mean. Furthermore, on $\{\tau(i) = \infty\}$ we have $\eta_k^\tau > 1/i$ and therefore the trajectories of $X_k$ (and hence $\eta_k^\tau$) are over-bounded by the random walk with a drift

$$Y_{k+1} = Y_k + z_k - (1-\gamma)/i, \quad Y_0 = X_0 = \eta_0 > 0 \ .$$

The drift of this random walk, $\mathbf{E}\{z_k\} - (1-\gamma)/i$, is strictly negative and bounded away from zero since $\mathbf{E}\{z_k\} \leq 0$ and $(1-\gamma)/i > 0$; this implies that the expected return time below $1/i$ is finite and as a result also $\mathbf{E}_t\{\tau(i)\} < \infty$ which implies $P(\tau(i) = \infty) = 0$. This finishes the proof. □

## 3.2 Convergence speed

In this section we discuss the speed of convergence of the average amount of violation $v_t/s_t$ below the prescribed level $\xi$ as required by (11b). In particular we explore its dependence on $\alpha$, the right-hand side of the conditional constraint (12), or equivalently (14), which is enforced whenever $v_t/s_t > \xi$.

Assuming $v_t/s_t > \xi$, the object of interest is the return time

$$\tau_t = \inf\{i \geq t \mid v_i/s_i \leq \xi\} \ ,$$

and in particular its expected value $\mathbf{E}_t\{\tau_t\}$. In order to study this object define consider the process $\eta_k$ defined in (33). Clearly, studying the return time of $v_t/s_t$ to below $\xi$ is equivalent to studying the return time of $\eta_k$ to below zero. Following the arguments laid out in part III of the proof of Theorem 1, as long as $v_{t+k}/s_{t+k} > \xi$, the trajectories of $\eta_k$ are bounded from above by the trajectories of the AR(1) process

$$X_{k+1} = \gamma X_k + z_k, \quad X_0 = \eta_0 > 0 \ , \tag{34}$$

where $z_k = l(\bar{\mu} + g^T w_{t+k}) - \xi$ is an i.i.d. innovation with nonpositive mean and $\bar{\mu} = q(\alpha)$, where the quantile function $q(\cdot)$ is defined in (17) and $\alpha$ is the right hand side of the constraint (12). Therefore

$$\tau_t \leq \tau + t \ ,$$

where

$$\tau := \inf\{k \geq 0 \mid X_k \leq 0\} \ .$$

Summarizing, the analysis of $\tau_t$ boils down to the analysis of $\tau$, the return (or first passage) time of the AR(1) process $X_k$ below zero. The distribution of $\tau$ is typically studied using Monte Carlo simulation. Simulating this system is straightforward and accurate estimates of the expected return time or its whole distribution can be obtained at low computation cost. Nevertheless, here we provide a flavor of existing analytic results.

The most recent result is due to Novikov and Kordzakhia [24] and applies to the case of $\gamma \in [0, 1)$.

**Lemma 3** (Novikov & Kordzakhia)**.** *Let $\gamma \in [0, 1)$ and*[2] *$P(z_0 < 0) > 0$. Then*

$$\mathbf{E}\{\tau\} = \frac{1}{\log(1/\gamma)} \int_0^\infty (\mathbf{E}e^{-uX_\tau} - e^{-u\eta_0}) e^{-\phi(u)} \frac{1}{u} \, \mathrm{d}u < \infty \ , \tag{35}$$

*where $\phi(u) = \sum_{k=0}^\infty \log \mathbf{E}e^{-\lambda^k u z_0}$.*

*Proof.* This is a restatement of the [23, Eq. (6)] originally proved in [24, Theorem 3] under slightly stronger assumptions. The results of [23] apply to the problem of the first passage of a given barrier from below, but our problem is that of passing the barrier from above. However, since our barrier is equal to zero, the results of [23] apply to the symmetrical version of our problem with $X_k$ replaced by $-X_k$ and $z_k$ by $-z_k$. This leads to (35) after verifying technical assumptions: Since $z_0 = l(\bar{\mu} + g^T \boldsymbol{w}) - \xi$ is bounded (by Assumption 1 and compactness of $\mathcal{W}$) we have

$$\mathbf{E}\{(-z_0)^-\} < \infty \ \text{ and } \ \log \mathbf{E}e^{-uz_0} < \infty \quad \forall\, u \in [0, \infty) \ .$$

Under these technical assumptions and $P(z_0 < 0) > 0$, the results of [23] ensure that $\mathbf{E}\{\tau\} < \infty$ and (35) holds. □

This Lemma provides easy-to-evaluate upper and lower bounds on the expected return time.

**Corollary 2.** *Let $\gamma \in [0, 1)$ and $P(z_0 < 0) > 0$. Then*

$$\frac{1}{\log(1/\gamma)} \int_0^\infty (1 - e^{-u\eta_0}) e^{-\phi(u)} \frac{1}{u} \mathrm{d}u \leq \mathbf{E}\{\tau\}, \tag{36a}$$

$$\mathbf{E}\{\tau\} \leq \frac{1}{\log(1/\gamma)} \int_0^\infty (e^{u\bar{z}} - e^{-u\eta_0}) e^{-\phi(u)} \frac{1}{u} \, \mathrm{d}u \ , \tag{36b}$$

*where $\phi(u) = \sum_{k=0}^\infty \log \mathbf{E}e^{-\lambda^k u z_0}$ and $\bar{z} = \xi - \min_{w \in \mathcal{W}} l(q(\alpha) + g^T w)$*

[2] Due to the i.i.d. assumption the distribution of $z_k$ is equal to that of $z_0$ for all $k \in \mathbb{N}_{\geq 0}$. In view of $\mathbf{E}\{z_0\} \leq 0$, the assumption $P(z_0 < 0) > 0$ is equivalent to $z_0$ not being identically zero with probability one in which case the AR(1) process (34) becomes deterministic and the convergence to zero occurs only asymptotically.

*Proof.* Follows from (35) by noticing that $\bar{z} \geq -X_\tau \geq 0$. □

**Remark 9.** *Note that since* $\mathbf{E}\{\tau\}$ *only upper bounds* $\mathbf{E}\{\tau_t\} - t$*, the upper bound (36b) is a valid upper bound on* $\mathbf{E}\{\tau_t\} - t$*, whereas the lower bound (36a) may not be.*

Now we turn to the somewhat simpler case of $\gamma = 1$, where the AR(1) process (34) becomes a random walk. It is well known that when $\mathbf{E}\{z_k\} \geq 0$, then the expected return time to zero starting from a positive value is infinite. Assume therefore that $\alpha < \xi$, where $\alpha$ is the right-hand side of the constraint (12); since $\mathbf{E}\{z_k\} \leq \alpha - \xi$ this ensures that $\mathbf{E}\{z_k\} < 0$. Then we have the following characterization of the return time:

**Lemma 4.** *Let* $\gamma = 1$ *and* $\alpha < \xi$*. Then*

$$\mathbf{E}\{\tau\} = \frac{\eta_0 - \mathbf{E}\{X_\tau\}}{\hat{z}} \ , \tag{37}$$

*where* $\hat{z} = -\mathbf{E}\{l(q(\alpha) + g^T \boldsymbol{w})\} + \xi$*.*

*Proof.* By one of the assumptions of the Lemma and since $z_k$ are i.i.d. we have $-\hat{z} = \mathbf{E}\{z_k\} \leq \alpha - \xi < 0$. Therefore $X_k$ is random walk with a strictly negative drift and since $X_0 = \eta_0 > 0$, we have $\mathbf{E}\{\tau\} < \infty$. We will show that the process $M_k := X_k + \hat{z}k$ is a martingale. Indeed, using the i.i.d. property of $z_k$ we have

$$\mathbf{E}_k\{X_{k+1} + \hat{z}(k+1)\} = X_k + \mathbf{E}\{z_k\} + \hat{z}(k+1) = X_k + \hat{z}k \ ,$$

verifying the claim. Next, we have

$$|M_{k+1} - M_k| = |z_k + (\xi - \alpha)| = \sup_{w \in \mathcal{W}} |l(\bar{\mu} + g^T w) - \alpha| < \infty$$

by Assumption 1 and compactness of $\mathcal{W}$. This verifies the assumptions of the optional stopping theorem (e.g., [29, Theorem 10.10. (iii)]), which then applied to $M_k$ yields

$$\eta_0 = \mathbf{E}\{M_0\} = \mathbf{E}\{M_\tau\} = \mathbf{E}\{X_\tau + \hat{z}\tau\} \ ,$$

which gives (37). □

This leads to the following upper and lower bounds on the expected return time $\mathbf{E}\{\tau\}$:

**Corollary 3.** *Assume* $\gamma = 1$ *and* $\alpha < \xi$*, then*

$$\frac{\eta_0}{\hat{z}} \leq \mathbf{E}\{\tau\} \leq \frac{\eta_0 + \bar{z}}{\xi - \alpha} \ , \tag{38}$$

*where* $\hat{z} = -\mathbf{E}\{l(q(\alpha) + g^T \boldsymbol{w})\} + \xi$ *and* $\bar{z} = \xi - \min_{w \in \mathcal{W}} l(q(\alpha) + g^T w)$*.*

*Proof.* Follows from (35) by noticing that $-\bar{z} \leq X_\tau \leq 0$ and $\hat{z} \geq \xi - \alpha$. □

Remark 9 applies also to the bounds (38).

**Remark 10.** *The bounds (36) and (38) quantify the effect of the parameter $\alpha$ on the speed of convergence of the violation process $v_t/s_t$ back to below $\xi$. As already pointed out in Remark 5, lower values of $\alpha$ ensure faster converge but, on the other hand, lead to a tighter constraint set of the problem (1) and are therefore likely to result in a worse closed-loop cost performance. In contrast, higher values of $\alpha$ allow for a slower convergence of $v_t/s_t$ to below $\xi$ (i.e., allow for more constraint violation) and therefore are likely to lead to a better closed-loop cost performance.*

**Remark 11.** *Once again we stress that the bounds (36) and (38) provide easy-to-obtain information on the expected value of the return time $\tau_t$ of $v_t/s_t$ to below $\xi$. Information on the* entire distribution *can be obtained (at a slightly higher but still entirely tractable) computational cost using Monte Carlo simulation of the AR(1) process $X_k$ whose return time to below zero provides an upper bound on $\tau_t$.*

# 4 Implementation

In this section we discuss how the general theory developed in previous sections can be employed within an MPC framework. The set valued control policy (32) gives rise to a generic MPC problem

**Problem 1.**

$$\text{minimize}\,\{J \mid u_t \in \mathcal{U}_t(x_t, \chi_t)\}\ ,$$

where the cost function $J$ is completely free to choose (as far as constraint satisfaction is concerned), as well as is the prediction horizon and the policy parametrization with respect to which the cost function $J$ is minimized.

By Lemma 2, the one-step expectation constraint translates to an affine constraint, and therefore for each value of $t$, the constraint set of Problem 1 is a polyhedron (by Assumption 2), and hence if $J$ is convex in the decision variables of the problem (i.e., the policy parametrization), then Problem 1 is convex. Computational aspects of this problem are discussed in the remainder of this section.

## 4.1 Parametrization of SRCI sets

A crucial step for the application of Problem 1 is the parametrization of a family of SRCI sets $\mathcal{S}_k$ (27).

### 4.1.1 Explicit parametrization

One way of parametrizing the family of SRCI sets is an explicit construction of an SRCI set $\mathcal{S}$ (Definition 2) and a subsequent application of the Reach$(\cdot)$ operator (Eq. (26) and Remark 8). It is desirable that the set $\mathcal{S}$ (and hence all $\mathcal{S}_k$) be large, preferably *maximal*. However, the computation of maximum (stochastic) (robust) controlled invariant sets is

known to be difficult in larger dimensions, as polyhedral projections are required, and the maximal set may not be polyhedral [2, 3]. Nevertheless, there are effective algorithms for the computation of *polyhedral* controlled invariant under-approximations of these sets which do not require polyhedral projections; see, e.g., [4, 5, 27].

With a family of polyhedral SRCI sets $S_{r_t} = \{x \mid g_{r_t,j}^T x \leq h_{r_t,j}, j = 1, \ldots, n_{r_t}\}$, the inclusion $u_t \in \mathcal{U}(x_t, \chi_t)$ is enforced direcly as

$$\begin{aligned} & g_{r_t,,j}^T(Ax_t + Bu_t) + \max_{w \in \mathcal{W}} g_{r_t,j}^T w \leq h_{r_t,j}, \quad j = 1, \ldots, n_{r_t} \\ & g^T(Ax_t + Bu_t) \leq h + q(\beta_t), \end{aligned}$$

where the two constraints are equivalent to (30a) and (30b), respectively. The maximum on the left-hand side of the first constraint can be precomputed offline; for simple $\mathcal{W}$ (e.g., box or ellipsoid) analytically, for a polytopic $\mathcal{W}$ by solving a linear program.

### 4.1.2 Implicit parametrization

The computation of maximum controlled invariant sets (or large under-approximations thereof) can be avoided if a family of SRCI sets $\mathcal{S}_k$ is parametrized *implicitly*. The implicit inclusion to a family of SRCI sets can be achieved via the traditional MPC with a terminal invariant set [21, 28]; this procedure is now briefly described.

For the traditional MPC scheme, at a time $t$ and on a prediction horizon $N$, the control input predictions $u_{t+k}$ are given by an explicit policy parametrization (the decision variable) for $k \in \{0, \ldots, N-1\}$, and by a fixed terminal controller for $k \geq N$. Let the explicit policy (in general a causal state-sequence feedback or, equivalently, causal disturbance feedback) be $\pi := (\pi_0, \ldots, \pi_{N-1})$, that is,

$$u_{t+k} = \pi_k(x_t^{t+k}), \quad k \in \{0, \ldots, N-1\} \ ,$$

and let the terminal state-feedback controller be $\kappa_f$, that is,

$$u_{t+k} = \kappa_f(x_{t+k}), \quad k \geq N \ .$$

The constraint satisfaction is enforced explicitly along the prediction horizon through constraints on the policy parametrization $\pi$ and implicitly beyond the prediction horizon by constraining the terminal state $x_{t+N}$ to a positively invariant set associated with the terminal controller $\kappa_f$. Specifically, the terminal set $\mathcal{X}_{\kappa_f}$ employed is a subset of $\mathcal{X}_s \cap \bar{\mathcal{X}}$ such that

$$Ax + B\kappa_f(x) + w \in \mathcal{X}_{\kappa_f} \qquad \forall w \in \mathcal{W} \ , \tag{39a}$$

$$g^T(Ax + B\kappa_f(x)) \leq h + q(\alpha), \tag{39b}$$

$$\kappa_f(x) \in \mathcal{U} \tag{39c}$$

is satisfied for all $x \in \mathcal{X}_{\kappa_f}$.

At a time $t$, given $x_t$ and $v_t$ (and hence $r_t$ and $\beta_t$), we wish to ensure that $u_t \in \mathcal{U}_t(x_t, \chi_t)$. The one-step expectation constraint (30b) remains unchanged. The invariance constraint (30a)

is enforced implicitly as follows. By definition of $\mathcal{S}_k$ and $\mathcal{S}$, the inclusion $x_{t+1} \in \mathcal{S}_{r_t}$ for all $w \in \mathcal{W}$ is ensured if, first,

$$\mathbf{E}\{l(g^T x_{t+r_t+i} - h) \mid x_{t+r_t+i-1}, x_t, v_t\} \leq \alpha \tag{40}$$

is satisfied for a given $(x_t, v_t)$, all $i \in \mathbb{N}_{>0}$ and all possible $x_{t+r_t+i-1}$ generated by all possible $w_t^{t+r_t+i-2} \in \mathcal{W}^{r_t+i-1}$ under the given policy $\pi$ and the terminal controller $\kappa_f$, and, second, if

$$x_{t+i} \in \bar{\mathcal{X}} \quad \forall\, i \in \mathbb{N}_{>0}\ . \tag{41}$$

The constraint (40) is enforced explicitly for $i \in \{1, \ldots, N - r_t\}$ by constraints on $\pi$ and implicitly for $i > N - r_t$ by requiring that $x_{t+N} \in \mathcal{X}_{\kappa_f}$ for all $w_t^{t+N-1}$ under $\pi$. The constraint (41) is a standard robust constraint and is enforced explicitly for $i \in \{1, \ldots, N-1\}$ and implicitly for $i \geq N$ by requiring that $x_{t+N} \in \mathcal{X}_{\kappa_f}$ for all $w_t^{t+N-1}$.

In principle there are no restrictions on the policy parametrization $\pi$ and the terminal controller $\kappa_f$ as long as they are "compatible" in the sense that the shifted solution is feasible at the next time instant. In fact, most of the widely used robust and stochastic MPC parametrizations such as affine disturbance feedback [13, 18], pre-stabilization [20] or tubes [22] can be used and give rise to *affine* constraints only.

As a concrete instance we detail the implicit parametrization for the most general of the these policies, the affine disturbance feedback. For this control policy the predicted input sequence $u_{t+k}$ for $k \in \{0, \ldots, N-1\}$ can be written in a compact form

$$\boldsymbol{u} = \underbrace{\begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_{N-1} \end{bmatrix}}_{\boldsymbol{\theta}} + \underbrace{\begin{bmatrix} 0 & 0 & \ldots & 0 \\ \Theta_{1,1} & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \\ \Theta_{N-1,1} & \ldots & \Theta_{N-1,N-1} & 0 \end{bmatrix}}_{\boldsymbol{\Theta}} \boldsymbol{w}\ , \tag{42}$$

where

$$\boldsymbol{u} = [u_t^T, \ldots, u_{t+N-1}^T]^T, \quad \boldsymbol{w} = [w_t^T, \ldots, w_{t+N-1}^T]^T$$

are the predicted control and disturbance sequences, respectively. The terminal controller is an affine state feedback

$$u_{t+k} = K_s x_{t+k} + d_s, \quad k \geq N,$$

such that all eigenvalues of $A + BK_s$ are strictly less than one in magnitude.

The constraint (30b) translates (in view of Lemma 2) to

$$g^T(Ax_t + B\theta_0) \leq h + q(\beta_t)\ , \tag{43}$$

where $q(\cdot)$ is defined in (17). The constraint (43) is an affine constraint on $\theta_0$, the first component of $\boldsymbol{\theta}$.

For $i \in \{1, \ldots, N - r_t\}$ the constraint (40) reads

$$\begin{aligned} g^T(A^{r_t+i}x_t + \mathcal{B}_{r_t+i}\boldsymbol{\theta}) + \max_{\boldsymbol{w} \in \mathcal{W}^N} g^T(\mathcal{B}_{r_t+i}\boldsymbol{\Theta} + A\mathcal{C}_{r_t+i-1})\boldsymbol{w} \\ \leq h + q(\alpha) \qquad \forall\, i \in \{1, \ldots, N - r_t\}\ , \end{aligned} \tag{44}$$

where the expression on the left-hand side is the prediction of $g^T x_{t+r_t+i}$ under the affine disturbance feedback policy starting from $x_t$, and

$$\mathcal{B}_k = [A^{k-1}B, \ldots, B, 0, \ldots, 0],\ \mathcal{C}_k = [A^{k-1}, \ldots, I, 0, \ldots, 0].$$

Note that the structure of the matrices $\mathcal{B}_{r_t+i}$, $\mathcal{C}_{r_t+i-1}$ and $\boldsymbol{\Theta}$ ensures that only the disturbances $w_t^{t+r_t+i-2}$ contribute to the second term on the first line of (44). Consequently, the maximum on the left-hand side can be taken only over $w_t^{t+r_t+i-2} \in \mathcal{W}^{r_t+i-1}$. When the matrix $\boldsymbol{\Theta}$ is fixed in advance (which encompasses prestabilization and tubes), the maximum can be evaluated offline, in which case the constraint (44) becomes affine in the decision variable $\boldsymbol{\theta}$. For many cases of practical interest (e.g., $\mathcal{W}$ ellipsoid or a box) the maximum has an analytical expression; for a polytopic $\mathcal{W}$ the computation of the maximum boils down to solving a linear program (LP). When $\boldsymbol{\Theta}$ is left free online as a decision variable, the constraint translates to a second order cone constraint for an ellipsoidal $\mathcal{W}$; for a polytopic $\mathcal{W}$ it translates (using LP duality theory, introducing a vector of Lagrange multipliers $\boldsymbol{\lambda}$) to the following affine constraints in the decision variables $(\boldsymbol{\theta}, \boldsymbol{\Theta}, \boldsymbol{\lambda})$:

$$\begin{aligned}
&g^T(A^{r_t+i}x_t + \mathcal{B}_{r_t+i}\boldsymbol{\theta}) + h_w^T\boldsymbol{\lambda} \leq h + q(\alpha)\ , \\
&G_w^T\boldsymbol{\lambda} = (\mathcal{B}_{r_t+i}\boldsymbol{\Theta} + A\mathcal{C}_{r_t+i-1})^T g\ , \\
&\boldsymbol{\lambda} \geq 0, \qquad \forall\, i \in \{1, \ldots, N - r_t\}\ ,
\end{aligned}$$

where $G_w$ and $h_w$ are such that $\mathcal{W}^N = \{w \mid G_w w \leq h_w\}$.

For $i > N - r_t$ the constraint (40) is enforced by requiring that $x_{t+N} \in \mathcal{X}_{\kappa_f}$ for all $\boldsymbol{w} \in \mathcal{W}^N$. Given $\mathcal{X}_{\kappa_f} = \{x \mid g_{f,j}x \leq h_{f,j}, j = 1, \ldots, n_f\}$, this is equivalent to

$$g_{f,j}^T(A^N x_t + \mathcal{B}_N\boldsymbol{\theta}) + \max_{\boldsymbol{w}\in\mathcal{W}^N} g_{f,j}^T(\mathcal{B}_N\boldsymbol{\Theta} + A\mathcal{C}_N)\boldsymbol{w} \leq h_{f,j}\ , \tag{45}$$

for all $j \in \{1, \ldots, n_f\}$. This constraint is then handled in exactly the same fashion as (44).

Finally, the constraint (41) is enforced (by construction of $\mathcal{X}_{\kappa_f}$) via (45) for $i \geq N$, and for $i \in \{1, \ldots, N-1\}$, given $\bar{\mathcal{X}} = \{x \mid \bar{g}_j x \leq \bar{h}_j, j = 1, \ldots, \bar{n}\}$, by

$$\bar{g}_j^T(A^i x_t + \mathcal{B}_i\boldsymbol{\theta}) + \max_{\boldsymbol{w}\in\mathcal{W}^N} \bar{g}_j^T(\mathcal{B}_i\boldsymbol{\Theta} + A\mathcal{C}_i)\boldsymbol{w} \leq \bar{h}_j\ , \tag{46}$$

for all $j \in \{1, \ldots, \bar{n}\}$. This constraint is again handled in the same fashion as (44).

To summarize, an MPC problem with affine disturbance feedback policy parametrization minimizes the cost function $J$ subject to the constraints (43), (44), (45) and (46). Imposing these constraints enforces the inclusion $u_t \in \mathcal{U}(x_t, \chi_t)$ and therefore Theorem 1, and all subsequent results, hold.

At this point it should be stressed that exactly the same steps (e.g., evaluation or dualization of the maximum of an affine function over the disturbance set) need to be taken when setting up a standard robust MPC problem with affine disturbance feedback as policy parametrization, and the same holds for other policy parametrizations; see, e.g., [13].

## 4.2 Computational complexity discussion

This section briefly compares the computational complexity of the presented approaches with existing MPC formulations.

If the SRCI sets are parametrized explicitly, then the *offline* complexity is governed by the computation of maximum controlled invariant sets (or under-approximations thereof) and as such is comparable to the nominal first-step MPC of [11, 12]. The *online* computational requirements are then governed by the type of the cost function $J$, policy parametrization and by the complexity of the SRCI sets employed.

If the SRCI sets are parametrized implicitly, then the *offline* computational complexity is determined by the policy parametrization $\pi$ and the terminal controller $\kappa_f$, the former being also the main determinant of the *online* complexity along with the type of the cost function $J$. Most importantly, online as well as offline computational requirements are analogous to those of the traditional robust and stochastic MPC schemes with the same parametrization, be it affine disturbance feedback [13, 18], prestabilization [20], or tubes [22].

It should be stressed that with explicitly parametrized SRCI sets as well as with implicit parametrization (as long as $\mathcal{W}$ is polytopic) using polyhedral terminal set and any of the above-mentioned policies, the constraint set of Problem 1 is *polyhedral* and hence the class of the problem (e.g., quadratic / linear program) is not altered by introducing the constraint (11). It is also very common to consider an ellipsoidal terminal set; in that case quadratic constraints are introduced, analogously to nominal MPC. Similarly, for an ellipsoidal disturbance set $\mathcal{W}$ and the affine disturbance feedback policy, the constraint (44) becomes a second-order cone constraint analogously to the robust MPC of [13].

# 5 Extensions

## 5.1 Multiple constraints

This section describes how multiple constraints can be handled.

### 5.1.1 Individual constraints

First we treat the case of multiple individual constraints of the form

$$g_i^T x \leq h_i, \quad i \in \{1, \ldots, n_g\} \,, \tag{47}$$

to each of which we associate a loss function $l_i : \mathbb{R} \to \mathbb{R}$ satisfying Assumption 1. Then we define individual cumulative violations $v_t^i$ as in (8) with $l$, $g$ and $h$ replaced by $l_i$, $g_i$ and $h_i$, respectively; and the normalization factors $s_t^i$ as in (9) with $\gamma$ replaced by $\gamma_i$. The constraint (11) is then imposed for each $i \in \{1, \ldots, n_g\}$ individually, possibly with the right-hand side $\xi$ replaced by $\xi_i$. This constraint is then enforced in a completely analogous fashion. The stochastic feasibility set is defined as the intersection of the individual stochastic feasibility sets associated to each constraint; similarly the SRCI set is defined

as in Definition 2 but now the second line is required to hold for all $i \in \{1, \dots, n_g\}$. The individual constraint leeways $\beta_i$ are defined as in (21) with $\gamma$, $\xi$ and $\alpha$ replaced by $\gamma_i$, $\xi_i$ and $\alpha_i$. The nested family of sets (27) is defined in exactly the same way. As a layer index $r_t$ we use the *minimum* of individual layer indices defined in (29) with $\tilde{r}_t$ replaced by $\tilde{r}_t^i$ (where all quantities defining $\tilde{r}_t$ in (28) are now indexed by $i$). Therefore the layered structure is shared by all the constraints, allowing the state to move to an upper layer only if all layer indices allow to do so. The multilayer control law is then defined by

$$\kappa(x_t, \chi_1, \dots, \chi_{n_g}) \in \mathcal{U}(x_t, \chi_1, \dots, \chi_{n_g}) \ ,$$

where $\chi_i = \xi_i s_t^i - v_t^i$ and $\mathcal{U}(x_t, \chi_1, \dots, \chi_{n_g})$ is defined as in (30) with (30b) enforced for each constraint with $g$, $h$, $l$ and $\beta_t$ replaced by $g_i$, $h_i$, $l_i$ and $\beta_t^i$. All theoretical results hold for each constraint individually with proofs being almost verbatim copies.

### 5.1.2 Joint constraint

Second, we describe how to handle a joint constraint

$$\|Gx_t + h\|_\infty \leq 1,$$

violation of which is given by $l(\|Gx_t + h\|_\infty - 1)$, where the loss function $l : \mathbb{R} \to \mathbb{R}$ satisfies Assumption 1. In order to apply the previously developed techniques all we need to show is that $\mathbf{E}_t\{l(\|Gx_{t+1} + h\|_\infty - 1)\} \leq \alpha$ translates to an affine constraint on $u_t$ (while possibly introducing some conservatism). To this end write

$$\begin{aligned}
&\mathbf{E}\big\{l(\|G(Ax + Bu + w) + h\|_\infty - 1)\big\} \\
&\quad \leq \mathbf{E}\big\{l(\|Gw\|_\infty + \|G(Ax + Bu) + h\|_\infty - 1)\big\} \\
&\quad = \hat{f}(\|G(Ax + Bu) + h\|_\infty - 1) \ ,
\end{aligned}$$

where the function $\hat{f}(\mu) := \mathbf{E}\{l(\|Gw\|_\infty + \mu)\}$ is known in advance given the distribution of $w$. Here the inequality in the above computation follows from the triangle inequality, which is in general strict, introducing a degree of conservatism. Satisfaction of $\mathbf{E}_t\{l(\|Gx_{t+1} + h\|_\infty - 1)\} \leq \alpha$ is then implied by

$$\|G(Ax_t + Bu_t) + h\|_\infty - 1 \leq \hat{q}(\alpha) \ , \tag{48}$$

where $\hat{q}(\alpha) := \{\mu \in \mathbb{R} \mid \hat{f}(\mu) \leq \alpha\}$.

Constraint (48) translates to a finite number of affine constraints and therefore all the previously developed techniques can be readily applied, replacing the constraint (16) with (48).

**Remark 12.** *In [19] the authors investigated the special case of a joint chance constraint specification with $l(\cdot)$ being the indicator function of the positive real line and no discounting, in which case $v_t$ captures the total* number *of excursions outside a given polytopic constraint set up to a time $t$. In this special case it is possible to use a confidence region of the disturbance to enforce that the expected loss incurred by the successor state falls below a given level, avoiding the use of the triangle inequality; see [19] for details.*

**Remark 13.** *Instead of the infinity norm specifying the joint chance constraint we could use any function satisfying the triangle inequality (e.g., any norm) without changing the theoretical results. However, not all such functions allow for a computationally tractable characterization of the invariant sets associated to (48).*

## 5.2 Other constraint specifications

This section discusses another possible stochastic specification that our framework can be readily extended to. The motivation for this is to demonstrate the flexibility of the approach provided by the freedom in the choice of the function $\nu(\cdot)$ in Lemma 1. In particular we show how to enforce the probabilistic constraint

$$P_t\left\{\frac{v_{t+1}}{s_{t+1}} > \xi\right\} \leq 1-\delta \quad \text{if} \quad \frac{v_t}{s_t} \leq \xi \,, \tag{49}$$

where $\delta \in [0,1]$ is a parameter. For $v_t/s_t > \xi$ the approach remains unchanged. All we need to do is to show that, given $x_t$ and $v_t$, this constraint translates to an affine constraint on $u_t$. To this end write

$$\begin{aligned}
&P_t\left\{\frac{v_{t+1}}{s_{t+1}} > \xi\right\} = \\
&= P_t\left(\gamma v_t + l(g^T(Ax_t + Bu_t + w_t) - h) > s_{t+1}\xi\right) \\
&= P_t\left(g^T w_t + g^T(Ax_t + Bu_t) - h > l^{-1}(s_{t+1}\xi - \gamma v_t)\right) \\
&= \mathbf{E}_t\left\{\mathbb{I}_{++}[g^T(Ax_t + Bu_t + w_t) - h - l^{-1}(s_{t+1}\xi - \gamma v_t)]\right\} \,,
\end{aligned}$$

where the first equality follows from Lemma 1 with $\nu(\cdot) = l(\cdot)$ and $\omega = 0$, $l^{-1}(\cdot)$ is defined in (13) and $\mathbb{I}_{++}[\cdot]$ is the indicator function of the positive real line. Since $\mathbb{I}_{++}[\cdot]$ is non-decreasing and lower-semicontinuous, invoking Lemma 1 again with $\nu(\cdot) = \mathbb{I}_{++}[\cdot]$ and $\omega = g^T w_t$ implies that

$$P_t\left\{\frac{v_{t+1}}{s_{t+1}} > \xi\right\} \leq 1-\delta$$

is equivalent to

$$g^T(Ax_t + Bu_t) - h \leq \bar{q}(l^{-1}(s_{t+1}\xi - \gamma v_t)) \,, \tag{50}$$

where the function $\bar{q}(a) := \sup\{\mu \mid P(g^T \boldsymbol{w} + \mu > a) \leq 1-\delta\}$ is known in advance given the knowledge of $\boldsymbol{w}$. Therefore the constraint (49) indeed translates to an affine constraint, as before with a time-varying right-hand side; thus the previously developed techniques can be readily applied with only minor modifications. In particular note that $v_t/s_t \leq \xi$ implies that $s_{t+1}\xi - \gamma v_t = \gamma(\xi s_t - v_t) + \xi \geq \xi$. Therefore the tightest right-hand side of (50) that we can encounter if $v_t/s_t \leq \xi$ is $\bar{q}(l^{-1}(\xi))$. Consequently the stochastic feasibility set $\mathcal{X}_s$ and the lowest-level SRCI set $\mathcal{S}$ are designed with respect to the constraint $g^T(Ax_t + Bu_t) - h \leq \bar{q}(l^{-1}(\xi))$ in addition to the constraint (20). The nested family of sets $\mathcal{S}_i$ (27) and the layer index $r_t$ (29) are defined in exactly the same way as before. Finally, the control law is defined by the inclusion $\kappa(x_t, \chi_t) \in \bar{\mathcal{U}}(x_t, \chi_t)$, where $\chi_t = \xi s_t - v_t$ (as

in (23)) and

$$\begin{aligned}\bar{\mathcal{U}}(x_t,\chi_t) := \big\{ & u \in \mathcal{U} \text{ s.t. } Ax_t + Bu_t + w \in \mathcal{S}_{r_t}\ \forall\, w \in \mathcal{W}, \\ & \mathbf{E}\{l(g^T(Ax_t + Bu_t + w) - h)\} \le \alpha \qquad \text{if } v_t/s_t > \xi, \\ & g^T(Ax_t + Bu_t) - h \le \bar{q}(l^{-1}(\gamma\chi_t + \xi)) \ \text{ if } v_t/s_t \le \xi \big\}\ .\end{aligned}$$

All theoretical results established for the control law (30) hold for this control law as well (after changing (11a) for (49)).

## 5.3 Time-varying constraints

In this section we outline how the presented approach can be extended to handle a time-varying right-hand side of the constraint (3). We assume that the right-hand side of the constraint $g^T x \le h$ suffers an unmeasurable disturbance modelled as the output of an auxiliary dynamical system driven by a fictitious disturbance. This simple extension provides the presented approach with another degree of modelling flexibility. For instance, in building climate control the time-variation of the constraint may come from allowing the occupants to change the room thermostat set-point within a pre-specified band around a given temperature (e.g., around 21°C). In general, it is impossible to track an abrupt (e.g., a step-wise) change of the constraint; therefore, the set-point change request is pre-filtered by a pre-specified (i.e., known to the MPC controller) linear system, the dynamics of which are a design parameter influencing the transient behavior (e.g, the speed of the response to the request). See Section 6 for a concrete example from building climate control.

Define the auxiliary single output system

$$\tilde{x}_{t+1} = \tilde{A}\tilde{x}_t + \tilde{B}\tilde{w}_t, \quad \tilde{y}_t = \tilde{C}\tilde{x}_t\ , \tag{51}$$

where the fictitious disturbance $\tilde{w}$ is constrained to lie in a compact polyhedron $\tilde{\mathcal{W}}$. The dimensions of $\tilde{x}$ and $\tilde{w}$ and the matrices $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ are free modeling parameters with the only requirement that $\tilde{C}$ has only one row. Then, the constraint

$$\frac{1}{s_t}\mathbf{E}\Big\{\sum_{k=0}^{t}\gamma^{t-k} l(g^T x_k - h - \tilde{y}_k)\Big\} \le \xi,\ \forall\, t \in \mathbb{N}_{\ge 0}, \{\tilde{w}_t \in \tilde{\mathcal{W}}\}_{t=0}^{\infty} \tag{52}$$

can be enforced by applying the previously developed theory to the augmented system with the state given by $\hat{x}_t := [x_t^T,\ \tilde{x}_t^T]^T$ and the constraint vector $\hat{g} := [g^T, -\tilde{C}]^T$. Note that we do not assume any distribution over the fictitious disturbance $\tilde{w}_t$ but rather treat it robustly; as a result, provided the setup is feasible (i.e., the set $\mathcal{S}$ nonempty and $\hat{x}_0 \in \mathcal{S}$), we can track any variation of the right-hand side that can be generated by the auxiliary system (51) with the input $\tilde{w}_t \in \tilde{\mathcal{W}}$.

# 6 Numerical example

As a numerical example we consider a simple building climate control problem adapted from [15]. The discrete-time system matrices are

$$A = \begin{bmatrix} 0.7534 & 0.2009 & 0.0358 \\ 0.0633 & 0.9349 & 0.0014 \\ 0.0500 & 0.0063 & 0.8671 \end{bmatrix}, B = \begin{bmatrix} 1.7 \\ 0.1 \\ 0.1 \end{bmatrix},$$

where the control input is the input power (in MW) and the states are respectively the room air temperature, interior-wall surface temperature and exterior-wall core temperature, all in degrees Celsius. This system was obtained by discretization with the sampling period 30 min. The control input is constrained to $\mathcal{U} = [0, 0.2]$. The i.i.d. disturbance sequence $w_t$ is given by $w_t = D\omega_t$, where

$$D = \begin{bmatrix} 0.0497 & 0.0434 & 0.0502 \\ 0.0018 & 0.0162 & 0.0019 \\ 0.3833 & 0.0013 & 0.0016 \end{bmatrix}$$

and $\omega_t$ is a vector of independent zero-mean normal random variables with standard deviations equal to 1/3 truncated to the interval $[-1, 1]$. The support of the disturbance $w_t$ is therefore equal to $\mathcal{W} = \{D\omega \mid ||\omega||_\infty \leq 1\}$. The components of the disturbance $z_t$ correspond to (scaled versions of) outside air temperature, solar radiation and internal heat sources, respectively. We consider a single comfort constraint $x_1 \geq T_{\min}$, where $T_{\min}$ is allowed to vary between $[20, 22]$ degrees Celsius in a manner compliant with the dynamics of the auxiliary system (51) given by

$$\tilde{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0.76 \end{bmatrix}, \tilde{B} = \begin{bmatrix} 0 \\ 0.3 \end{bmatrix}, \tilde{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \tilde{\mathcal{W}} = [-0.8, 0.8].$$

This constraint variation may be considered as a degree of freedom given to the occupants of the building allowing them to adjust the room temperature according to their needs although not permitting gross deviations for the 21 °C standard set-point.

We consider three constraint specifications summarized in Table 1 featuring a probabilistic constraint as well as an integrated chance constraint. The proposed policies were compared to the robust MPC of [13], modified to handle the time-varying constraint as outlined in Section 5.3. As a prediction policy we used the affine disturbance feedback with the prediction horizon $N = 8$ for all specifications compared. The cost function minimized was the sum of the (nonnegative) control inputs over the prediction horizon which corresponds to the minimization of the energy expenditure. We consider the multi-layer policy with six layers; the SRCI sets $\mathcal{S}_i$, $i = 1, \ldots, 6$ were computed explicitly with $\mathcal{S}_1 = \mathcal{S}$ being the maximum SRCI set. To compare the policies we carried out 100 Monte Carlo simulations, each 50 h (100 time steps) long. At time step 20, we assume a jump of the constraint from 21 °C to 22 °C. Table 3 and Figure 2 summarize the results, showing that our formulation fully exploits the freedom given by the constraint specification and brings about a significant cost improvement compared to the standard robust MPC.

In addition we investigated the return time distribution for the "Integ0.95" specification using the AR(1) process (34). Figure 3 shows kernel estimates (see, e.g., [16]) of the

probability density function of the return time and Table 2 compares the Monte Carlo estimates with the analytic bounds of Corollary 3. We can see that the lower bound provided by the Corollary is relatively tight, whereas the upper bound is rather loose. Closer investigation revealed that this is because the distribution of $-X_\tau$ is tightly concentrated around zero with a light but relatively long positive tail resulting in a high discrepancy between the expected value of $-X_\tau$ and its maximum value used in the bound.

# 7 Conclusion

This paper presented a framework to handle the constraints on the amount of state-constraint violations averaged over time, where the average amount of violations can be bounded either in expected value or robustly. The key ingredient of the proposed approach is the explicit incorporation of the information on the past behavior of the state process into the decision on the current control input. The approach significantly reduces the conservatism of previous stochastic MPC formulations, which is confirmed by a numerical example. The computational requirements (both online and offline) are comparable to those of the recursively feasible nominal MPC with a first-step constraint of [11, 12] in the case of explicit parametrization of controlled invariant sets, or comparable to those of popular, existing, robust MPC schemes in the case of implicit parametrization (assuming the same parametrization for the presented and for the robust approach).

# References


[1] D. Angeli, R. Amrit, and J. B. Rawlings. On average performance and stability of economic model predictive control. *IEEE Transactions on Automatic Control*, 57(7):1615–1626, 2012.

[2] F. Blanchini. Set invariance in control. *Automatica*, 35(11):1747–1767, 1999.

[3] F. Blanchini. *Set-Theoretic Methods in Control*. Birkhäuser Boston, first edition, 2007.


**Table 1:** Parameters of the control policies investigated. The loss function $l(\cdot)$ corresponds to a probabilistic constraint for the first specification and to an integrated chance constraint for the second and third. Different levels of allowed violation $\xi$ were chosen because of the very different character of the two specifications (see Figure 2 for simulation results). Note also the very different meaning of the maximum violation per one time step $\bar{\xi}$: for the probabilistic constraint it is naturally 1 (constraint is violated) but for the integrated chance it is user-specified 0.5°C (maximum allowed *amount* of constraint violation).

| Policy | Probab0.95 | Integ0.95 | Integ1.0 |
|---|---|---|---|
| loss function $l(x)$ | $\mathbb{I}_{[x>0]}$ | $\max\{x, 0\}$ | $\max\{x, 0\}$ |
| allowed violation $\xi$ | 0.2 | 0.1 | 0.1 |
| maximum violation $\bar{\xi}$ | 1 | 0.5°C | 0.5°C |
| forgetting factor $\gamma$ | 0.95 | 0.95 | 1.0 |

**Table 2:** Estimates of the mean return time for the "Integ0.95" specification and four different values of the initial overshoot $100(v_t/s_t - \xi)/\xi$. A Monte Carlo estimate using $10^5$ trajectories of the AR(1) process (34) is denoted by $\hat{\tau}$, the lower and upper bounds (36) then by $\tau_{\text{lb}}$ and $\tau_{\text{ub}}$. We assumed $t \gg 0$ and therefore $s_t \approx 1/(1-\gamma)$, so the initial condition of the AR(1) process was overshoot/$(100(1-\gamma))$.

| overshoot | 1 % | 10 % | 50 % | 100 % |
|---|---|---|---|---|
| $\tau_{\text{lb}}$ [h] | 3.5 | 18.1 | 33.4 | 40.0 |
| $\hat{\tau}$ [h] | 6.4 | 20.5 | 35.7 | 42.5 |
| $\tau_{\text{ub}}$ [h] | 57.3 | 71.9 | 87.2 | 93.8 |

**Table 3:** Relative cost improvement of the proposed policies with respect to the robust MPC measured as the average over 100 Monte Carlo simulations of the energy consumed during each 50 h long realization.

| Policy | Probab0.95 | Integ0.95 | Integ1.0 |
|---|---|---|---|
| Cost improvement | 8.8 % | 19.0 % | 20.0 % |


[4] F. Blanchini, S. Miani, and C. Savorgnan. Dynamic augmentation and complexity reduction of set-based constrained control. In *17th IFAC World Congress*, pages 14324–14329, 2008.

[5] T. B. Blanco, M. Cannon, and B. D. Moor. On efficient computation of low-complexity controlled invariant sets for uncertain linear systems. *International Journal of Control*, 83(7):1339–1346, 2010.

[6] M. Cannon, B. Kouvaritakis, and D. Ng. Probabilistic tubes in linear stochastic model predictive control. *Systems & Control Letters*, 58(10-11):747–753, 2009.

[7] M. Cannon, B. Kouvaritakis, S. V. Raković, and Q. Cheng. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Transactions on Automatic Control*, 56(1):194–200, 2011.

[8] M. Cannon, B. Kouvaritakis, and X. Wu. Model predictive control for systems with stochastic multiplicative uncertainty and probabilistic constraints. *Automatica*, 45(1):167–172, 2009.

[9] M. Cannon, B. Kouvaritakis, and X. Wu. Probabilistic Constrained MPC for Multiplicative and Additive Stochastic Uncertainty. *IEEE Transactions on Automatic Control*, 54(7):1626–1632, 2009.

[10] EN 15251:2007. Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics, 2008.

[11] R. Gondhalekar and J. Imura. Least-restrictive move-blocking model predictive control. *Automatica*, 46(7):1234–1240, 2010.

[12] R. Gondhalekar, J. Imura, and K. Kashima. Controlled invariant feasibility – A general approach to enforcing strong feasibility in MPC applied to move-blocking. *Automatica*, 45(12):2869–2875, 2009.

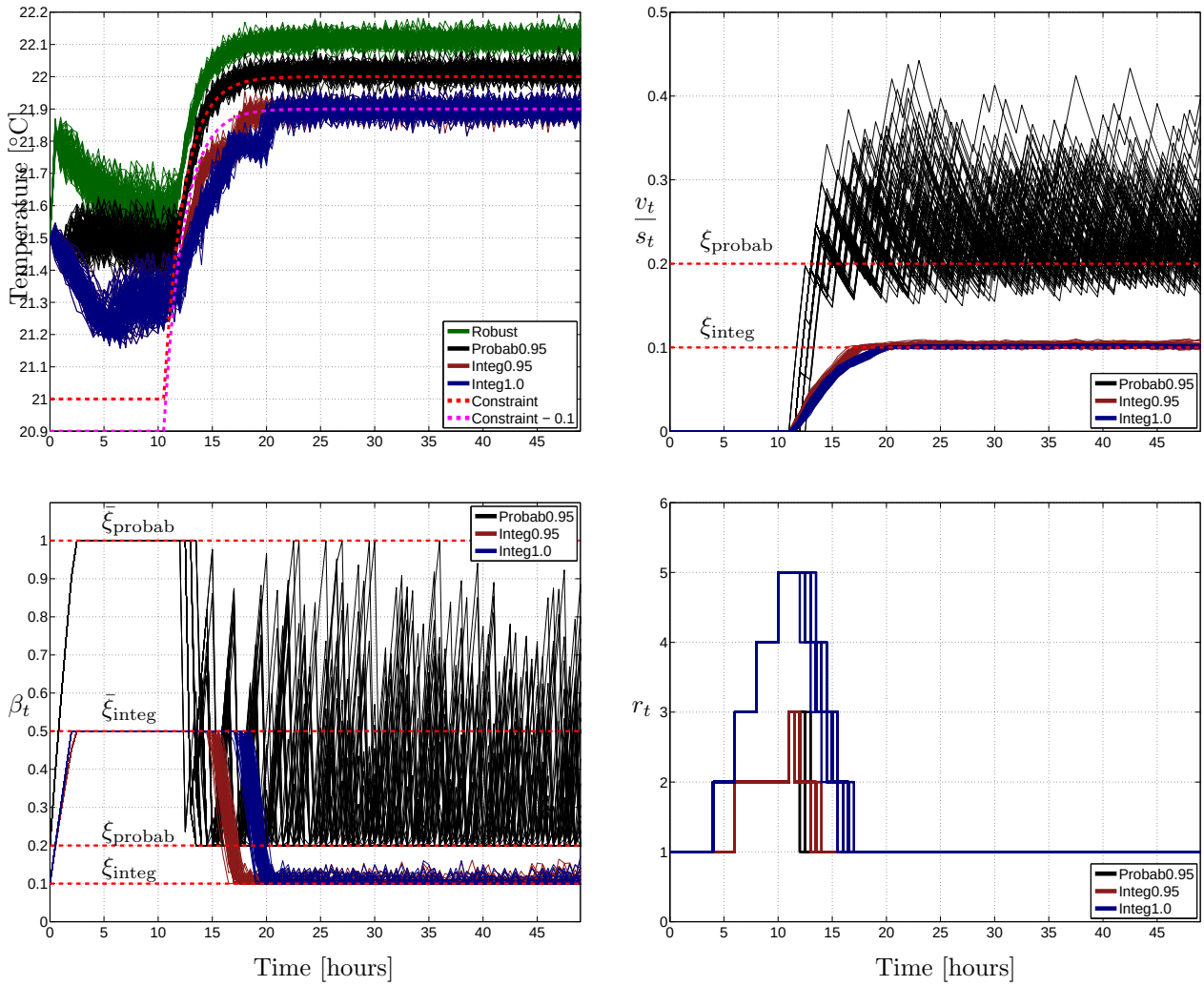


**Figure 2:** One hundred sample paths of the room temperature, the average number of violations $v_t/s_t$, the right-hand side of the one-step conditional constraint $\beta_t$ and the layer index $r_t$. The simulation was carried out for three sets of parameters described in Table 1. The freedom of the stochastic specification is most exploited during the transient when temperature constraint changes because of the previously accumulated non-violations. Note in particular that the transient for Integ0.95 is faster than that for Integ1.0 since the previous non-violations are gradually forgotten. This is confirmed by the trajectories of the constraint leeway $\beta_t$ and the layer index $r_t$ where $\beta_t$ returns to 0.1 and $r_t$ to 1 sooner for Integ0.95. Finally note the very different character of the trajectories for $\beta_t$ and $v_t/s_t$ for Probab0.95 which is due to the very different nature of the constraint – with Probab0.95 all violations incur the same loss (equal to one) whereas with the "Integ" specifications the loss grows linearly with the violation size. This is also reflected in the different profiles of the room temperature – the "Integ" specifications are allowed to persistently violate by $\xi_{\text{inter}} = 0.1°\text{C}$ and indeed in stationarity the mean temperature for both of these two specifications is close to 21.9°C, whereas the mean of the "Probab0.95" specification is close to 22.0°C; to emphasize this distinction we show in red the time evolution of the constraint and in magenta the time evolution of the constraint shifted down by 0.1°C.


[13] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski. Optimization over state feedback policies for robust control with constraints. *Automatica*, 42(4):523–533, 2006.

[14] L. Grüne. Economic receding horizon control without terminal constraints. *Automatica*,

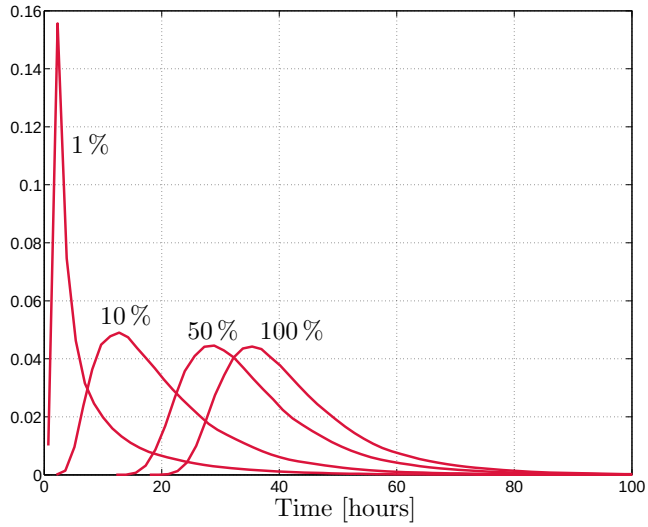


**Figure 3:** Kernel estimates of the probability density functions of the return time for the "Integ0.95" specification for initial overshoots $100(v_t/s_t - \xi)/\xi$ of 1 %, 10 %, 50 % and 100 %. The mean values of the return times with their estimates are summarized in Table 2.


49(3):725–734, 2013.

[15] M. Gwerder and J. Tödtli. Predictive control for integrated room automation. In *REHVA World Congress for Building Technologies – CLIMA*, 2005.

[16] M. Hollander and D. A. Wolfe. *Nonparametric Statistical Methods*. Wiley, second edition, 1999.

[17] J. Kim, M. Ruggiero, and D. Atienza. Free cooling-aware dynamic power management for green datacenters. *2012 International Conference on High Performance Computing & Simulation (HPCS)*, pages 140–146, 2012.

[18] M. Korda, R. Gondhalekar, J. Cigler, and F. Oldewurtel. Strongly feasible stochastic model predictive control. In *Conference on Decision and Control*, pages 7381–7386, 2011.

[19] M. Korda, R. Gondhalekar, C. N. Jones, and F. Oldewurtel. Stochastic model predictive control: Controlling the average number of constraint violations. In *Conference on Decision and Control*, 2012.

[20] B. Kouvaritakis, M. Cannon, S. V. Raković, and Q. Cheng. Explicit use of probabilistic distributions in linear predictive control. *Automatica*, 46(10):1719–1724, 2010.

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

[22] D. Q. Mayne, M. M. Seron, and S. V. Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219–224, 2005.

[23] A. Novikov. On distributions of first passage times and optimal stopping of AR(1) sequences. *Theory of Probability and Its Applications*, 53(3):419–429, 2009.

[24] A. Novikov and N. Kordzakhia. Martingales and first passage times of AR(1) sequences. *Stochastics: An international Journal of Probability and Stochastic Processes*, 80:197–210, 2008.

[25] A. Prékopa. *Stochastic Programming*. Springer, first edition, 1995.

[26] S. J. Qin and T. A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11:733–764, 2003.

[27] S. V. Raković, E. C. Kerrigan, D. Q. Mayne, and K. I. Kouramas. Optimized robust control invariance for linear discrete-time systems: Theoretical foundations. *Automatica*, 43(5):831–841, 2007.

[28] J. B. Rawlings and D. Q. Mayne. *Model Predictive Control: Theory and Design*. Nob Hill Publishing, first edition, 2009.

[29] D. Williams. *Probability with Martingales*. Cambridge University Press, first edition, 1991.