

Iterative Learning Stochastic MPC with Adaptive Constraint Tightening for Building HVAC Systems

Yushen Long* Lihua Xie*

* School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, (e-mail: {yslong, elhxie}@ntu.edu.sg).

Abstract: Most of the existing stochastic model predictive control (SMPC) algorithms for systems subject to random disturbance are designed offline using the distribution information of the uncertainties. In this paper, we propose an iterative learning based MPC for systems subject to time varying stochastic constraints on states. Different from those existing offline design approaches, except for the boundedness, this algorithm does not require to know the distributions or statistics such as the covariances of the uncertainties and the parameters of the controllers are adjusted online using the observations of past state trajectories. By making use of the iterative nature of the process, pointwise in time stochastic constraints are enforced so that it can handle time-varying constraints. Under some proper assumptions, this iterative procedure is shown to be equivalent to a root-searching problem and stochastic approximation theory is applied to show that the empirical average converges to the prescribed expectation in probability. The proposed algorithm is applied to an HVAC control problem to show its effectiveness.

Keywords: Model predictive control; Iterative learning control; Stochastic approximation; Optimization.

1. INTRODUCTION

From the United Nations environment programme report Lemmet (2009), buildings account for 40 percent of energy consumption and resources and one third of greenhouse gas emissions. In Singapore, 30 percent of energy is consumed by buildings while air-conditioning systems are responsible for more than half of the total energy consumed by buildings Chua et al. (2013). Therefore, it is attractive to reduce the energy cost of buildings and one of the most promising directions is to optimize the energy efficiency of heating, ventilation and air conditioning (HVAC) systems.

One of the main control objectives of HVAC control systems is to reject disturbances which mainly consist of outdoor weather condition and indoor thermal load, and maintain certain comfort level, e.g., keep indoor temperature in a prescribed range. Conventional robust approach usually leads to a conservative control strategy since the controller always prepares for those extreme cases which only happen occasionally. On the other hand, one may consider the uncertainty as a stochastic process and enforce a chance constraint $\mathbb{P}(T(k) \in \mathbb{T}) \geq 1 - \epsilon$, where $\mathbb{P}(\cdot)$, $T(k)$ and \mathbb{T} denotes probability, indoor temperature

at time instant k and comfort range, and ϵ is a small positive constant, respectively. That is to say, we allow the temperature to go outside of the comfort region with probability less than ϵ .

Stochastic MPC has been developed extensively in recent years to handle stochastic systems subject to stochastic constraints. One of the key challenges of stochastic MPC is how to reformulate the stochastic constraint, which involves high dimensional convolution, into computable deterministic inequalities. Several approaches have been proposed to reformulate the constraints, including equivalent translation Lorenzen et al. (2017), Chebyshev's inequality Magni et al. (2009) and randomized sampling method Schildbach et al. (2014). All these approaches either require the knowledge of the uncertainty or heavy computation. In Muñoz-Carpintero et al. (2018), the chance constraint is reformulated by reinterpreting it as an average violation ratio. For instance, we define $I_{\mathbb{T}}(T) = 1$, if $T \notin \mathbb{T}$ and $I_{\mathbb{T}}(T) = 0$ otherwise. The chance constraint $\mathbb{P}(T(k) \in \mathbb{T}) \geq 1 - \epsilon$ is reformulated as $\lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} I_{\mathbb{T}}(T(i))}{N} \leq \epsilon$ in some stochastic sense. Then the constraints are adjusted online according to the difference between ϵ and $\frac{\sum_{i=0}^{N-1} I_{\mathbb{T}}(T(i))}{N}$. Although such a reformulation keeps the spirit of allowing occasional violation of the constraint, it changes the original pointwise-in-time meaning of the constraint. Due to such reformulation, this approach cannot handle time-varying constraints either.

* The work of Long and Xie is funded by the Republic of Singapore's National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore.

For a typical office building, the indoor thermal load, which is caused by occupancy and electronic devices, and the outdoor weather condition vary with a highly periodic pattern. So, the control task of a HVAC system can be considered as a repetitive task with some uncertainties. This motivates us to adopt an iterative learning control (ILC) scheme to improve system performance. In this paper, we propose a combination of MPC and ILC schemes to enforce the pointwise-in-time stochastic constraint. Compared with most existing ILC works, the learning objective here is to find some proper controller parameters so that the stochastic constraint can be satisfied. The main contribution of this paper is summarized as follows:

1) Compared with Muñoz-Carpintero et al. (2018), we enforce more general stochastic constraints on the system, which include the chance constraint as a special case. We also allow the constraints to be time-varying, e.g., different violation probability for different time instant.

2) Furthermore, in Muñoz-Carpintero et al. (2018), it is assumed that the state trajectory converges to a terminal regime along time. However, in this paper, we do not assume that the state converges either along time or along iteration. Instead, we can prove that the state trajectory converges along iteration in some stochastic sense.

The rest of the paper is organized as follows. In Section 2, the system model and control objective are presented together with a controller parameterization. In Section 3, an iterative learning based adaptive MPC is proposed. In Section 4, the convergence of the empirical ratio is analyzed. In Section 5, the proposed algorithm is applied to an HVAC system to demonstrate its effectiveness. In Section 6, conclusions are drawn.

Some remarks on notations are as follows. \mathbb{R} represents the set of real numbers and \mathbb{R}^n the n -dimensional Euclidean space. \mathbb{N} denotes the set of natural numbers. For two sets $X, Y \subset \mathbb{R}^n$, the Minkowski sum is defined as $X \oplus Y \triangleq \{x + y : x \in X, y \in Y\}$ and the Pontryagin difference is defined as $X \ominus Y \triangleq \{z | z + y \in X, \forall y \in Y\}$. The Cartesian product of N identical sets X is denoted as X^N . The transpose of a vector x is denoted as x' . A probability space is defined by the triple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space, \mathcal{F} is a σ -algebra of Ω and \mathbb{P} is a probability measure on (Ω, \mathcal{F}) . For a random variable x , $E(x)$ denotes its expectation. Let $\{\mathcal{F}_k, k \in \mathbb{T}\}$ be a filtration, where \mathbb{T} is a countable time set. The expectation of a random variable x conditioned on the variables that are \mathcal{F}_k -measurable is denoted by $E_k(x)$.

2. PROBLEM FORMULATION

Consider a linear system

$$x(k+1) = Ax(k) + Bu(k) + w(k), \quad k = 0, \dots, T-1, \quad (1)$$

where $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}^m$ is the control input and $w(k) \in \mathbb{R}^n$ is the disturbance. Its initial condition $x(0)$ is a random variable with mean value $E(x(0)) = x^0$ and bounded error support $x(0) - x^0 \in \Delta^0$. We assume that the pair (A, B) is controllable. So there exists a matrix $K \in \mathbb{R}^{m \times n}$ such that $(A + BK)$ is Schur stable. The nominal model is given by

$$\bar{x}(k+1) = A\bar{x}(k) + B\bar{u}(k) + \bar{w}(k), \quad (2)$$

with a fixed initial condition x^0 and $\bar{w}(k) = E(w(k))$. The control applied to (1) is

$$u(k) = \bar{u}(k) + K(x(k) - \bar{x}(k)). \quad (3)$$

The input is required to satisfy a hard constraint

$$u(k) \in \mathbb{U}, \quad k = 0, \dots, T-1, \quad (4)$$

where \mathbb{U} is a convex polyhedron. Divide the time horizon $[1, \dots, T]$ into several sub intervals H_1, \dots, H_s such that $\cup_{i=1, \dots, s} H_i = \{1, \dots, T\}$ and $H_i \cap H_j = \emptyset, \forall i \neq j$. For each time instant k , let p_k be the index of the time interval which contains k , i.e., $k \in H_{p_k}$. For the state, we impose a time-varying stochastic constraint

$$\frac{\sum_{k \in H_p} E(q(g_k^T x(k) - h_k))}{|H_p|} \leq v_p, \quad k = 1, \dots, s, \quad (5)$$

where $q(\cdot)$ is the loss function, g_k and h_k are given bounded vector and $v_p \in [\underline{v}, \bar{v}]$ is the average expected loss over time interval H_p .

The loss function $q(\cdot)$ satisfies the following assumption:

Assumption 2.1. $q(\cdot)$ is non-decreasing and lower-semicontinuous, and has only countable discontinuous points at most.

The objective of this paper is to design a controller to make system (1) satisfy constraints (4) and (5).

Assumption 2.2. $\bar{w}(k), k = 0, \dots, T-1$ is known. $\delta(k) \triangleq w(k) - \bar{w}(k), k = 0, \dots, T-1$ is i.i.d. with finite support Δ .

Note that we do not assume that the distribution or even the covariance of the uncertainty is known, which is different from most existing stochastic MPC approaches.

Denote $e(k) = x(k) - \bar{x}(k)$. One has

$$e(k+1) = (A + BK)e(k) + \delta(k),$$

with $e(0) = x(0) - x^0$. Constraint (4) can be rewritten as

$$\bar{u}(k) \in \mathbb{U} \ominus K\Delta(k), \quad (6)$$

where $\Delta(k) = (A + BK)^k \Delta^0 \oplus \sum_{l=0}^{k-1} (A + BK)^l \Delta$.

For constraint (5), we first consider its robust version, i.e., $x(k) \in \mathbb{X}_k$, where $\mathbb{X}_k \triangleq \{x | q(g_k^T x - h_k) \leq \underline{v}\}$. Then according to the controller parameterization (3), it can be rewritten as

$$\bar{x}(k) \in \mathbb{X}_k \ominus \Delta(k). \quad (7)$$

If constraint (7) is satisfied, then $x(k) \in \mathbb{X}_k$ for all possible realizations of $\delta(k)$. So, in this case constraint (5) is satisfied conservatively in the sense that $q(g_k^T x - h_k) \leq \underline{v}, k = 1, \dots, T$ and it implies we need to relax constraint (7) so that $\frac{\sum_{k \in H_p} E(q(g_k^T x(k) - h_k))}{|H_p|}$ can be closer to v_p . To this end, we introduce a relaxation coefficient $\beta(p_k) \leq 1$ so that constraint (7) becomes

$$\bar{x}(k) \in \mathbb{X}_k \ominus \beta(p_k)\Delta(k). \quad (8)$$

When $\beta(p_k) = 1$, constraint (8) is (7) and $x(k) \in \mathbb{X}_k$ will be satisfied almost surely and when $\beta(p_k)$ decreases, it

means that we allow the state to evolve in a larger region so that more violations may be expected.

3. ITERATIVE LEARNING CONTROLLER

Motivated by the repetitive nature of HVAC systems, we reformulate the original problem in the following sense:

Rewrite system (1) as

$$\begin{aligned} x_j(k+1) &= Ax_j(k) + Bu_j(k) + w_j(k), \\ k &= 0, \dots, T-1, \quad j \in \mathbb{N}, \end{aligned} \quad (9)$$

where j denotes the iteration index. For each iteration, the system starts from $k=0$ and $x_j(0)$ which follows the same distribution of $x(0)$.

Denote the accumulative violation ratio over the p -th time interval and the l -th iteration as $c_{l,M}(p) =$

$$\begin{cases} \frac{\sum_{k \in H_p} \sum_{j=0}^l q(g_k^T x_j(k) - h_k)}{(l+1)|H_p|}, & l \leq M-1 \\ \frac{\sum_{k \in H_p} \sum_{j=l-M+1}^l q(g_k^T x_j(k) - h_k)}{M|H_p|}, & l > M-1 \end{cases}, \text{ where } M$$

is the length of moving average window. The control objective becomes to make $c_{l,M}(p)$ converge to v_p in some sense as $l, M \rightarrow \infty$.

Assume that an initial feasible nominal trajectory $\bar{x}_0(k)$ and $\bar{u}_0(k), k=0, \dots, T-1$, which satisfy (7) and (6), is available. For the k -th time instant of the j -th iteration, we propose the following MPC problem:

Problem 1

$$\min_{\bar{U}_j(k), \epsilon(1:N_k)} J(\bar{U}_j(k)) + \frac{1}{N_k} \sum_{i=1}^{N_k} \phi(\epsilon(i))$$

subject to

$$\begin{aligned} \bar{x}_j(l+1|k) &= A\bar{x}_j(l|k) + B\bar{u}_j(l|k) + \bar{w}(l), \\ \bar{x}_j(l+1|k) + \epsilon(l-k+1) &\in \mathbb{X}_{l+1} \ominus \beta_j(p_{l+1})\Delta(l+1), \\ \bar{u}(l|k) &\in \mathbb{U} \ominus K\Delta(l), \\ l &= k, \dots, N_k, \\ \bar{x}_j(k|k) &= \bar{x}_j(k), \end{aligned}$$

where $\bar{U}_j(k) = (\bar{u}_j(k|k), \dots, \bar{u}_j(k+N_k|k))$, $\epsilon(1:N_k) = (\epsilon(1), \dots, \epsilon(N_k))$, $\epsilon(i)$ are slack variables to guarantee feasibility with a penalty term $\phi(s) = \gamma\|s\|^2$ and weight γ , $N_k = N$ when $N+k \leq T$ and $N_k = T-k$ otherwise and $J(\bar{U}_j(k))$ is the energy consumption. Denote the optimal control input of the above problem as $\bar{U}_j^*(k) = (\bar{u}_j^*(k|k), \dots, \bar{u}_j^*(k+N_k|k))$. The control to be applied to the plant is $u_j(k) = \bar{u}_j^*(k|k) + K(x_j(k) - \bar{x}_j(k))$. And the nominal state $\bar{x}_j(k)$ is updated as:

$$\bar{x}_j(k+1) = A\bar{x}_j(k) + B\bar{u}_j^*(k|k) + \bar{w}(k), \quad (10)$$

with initial condition $\bar{x}_j(0) = x^0$. The relaxation coefficient $\beta_j(p_{l+1})$ will be defined later.

Assumption 3.1. 1) There is a bounded set $\bar{\mathbb{X}} \subset \mathbb{R}^n$ such that for any iteration index $j \in \mathbb{N}$ and time instant $k=0, \dots, T$, $x_j(k) \in \bar{\mathbb{X}}$.

2) $\bar{u}_j^*(k|k)$ is a continuous function of $\beta_j(p_i)$, $i = k+1, \dots, k+N_k$ and $\bar{x}_j(k)$.

We use the following notations to collect variables over time $[0, T]$:

$$\begin{aligned} \mathbf{X}_j &\triangleq (x_j(0), \dots, x_j(T))', \\ \bar{\mathbf{X}}_j &\triangleq (\bar{x}_j(0), \dots, \bar{x}_j(T))', \\ \mathbf{W}_j &\triangleq (x_j(0) - x^0, w_j(0), \dots, w_j(T-1)), \\ \mathbf{U}_j &\triangleq (u_j(0), \dots, u_j(T-1))', \\ \bar{\mathbf{U}}_j &\triangleq (\bar{u}_j^*(0|0), \dots, \bar{u}_j^*(T-1|T-1))', \\ \mathbf{C}_{j,M} &\triangleq (c_{j,M}(1), \dots, c_{j,M}(s))', \\ \mathbf{Q}_j &\triangleq \left(\frac{\sum_{k \in H_1} q(g_k^T x_j(k) - h_k)}{|H_1|}, \dots, \frac{\sum_{k \in H_s} q(g_k^T x_j(k) - h_k)}{|H_s|} \right)', \\ \mathbf{V} &\triangleq (v_1, \dots, v_s)', \\ \mathbf{B}_j &\triangleq (\beta_j(1), \dots, \beta_j(s))', \\ \mathbf{B} &\triangleq (\beta(1), \dots, \beta(s))'. \end{aligned}$$

According to the designed algorithm, it is clear that $\bar{\mathbf{X}}_j$ is determined by \mathbf{B}_j and \mathbf{X}_j is determined by $\bar{\mathbf{X}}_j$ and \mathbf{W}_j . Therefore, \mathbf{Q}_j , which is a function of \mathbf{X}_j , is essentially a function of \mathbf{B}_j and \mathbf{W}_j . So, we assume that there exists some function from \mathbf{B} to \mathbf{Q}_j which satisfies the following assumption:

Assumption 3.2. There exists a non-increasing continuous function $\sigma_p : \mathcal{B} \rightarrow \mathbb{R}_+$, with $\mathcal{B} = \{\beta : \underline{\beta} \leq \beta \leq 1\}$ and $\beta_p^* \in \mathcal{B}$, for $p = 1, \dots, s$ such that:

$$\begin{aligned} \sigma_p(\beta) &> v_p, \text{ if } \beta < \beta_p^* \\ \sigma_p(\beta) &= v_p, \text{ if } \beta = \beta_p^* \\ \sigma_p(\beta) &< v_p, \text{ if } \beta > \beta_p^* \end{aligned}$$

and $\lim_{l \rightarrow \infty} \frac{\sum_{j=0}^{l-1} \mathbf{Q}_j(\mathbf{B})}{l} \rightarrow h(\mathbf{B}) \triangleq (\sigma_1(\beta(1)), \dots, \sigma_s(\beta(s)))'$ in probability.

Note that we do not assume that $\bar{\mathbf{X}}_j$ converges along the learning procedure while in Muñoz-Carpintero et al. (2018), it is made as an assumption. In the next section, the convergence of $\bar{\mathbf{X}}_j$ along the learning procedure will be proved.

Consider the relaxation coefficient $\beta_j(p)$ updated as

$$\beta_j(p) = \Pi_{\mathcal{B}}[\beta_{j-1}(p) + \eta_{j-1}y_{j-1}(p)], \quad (11)$$

where $y_{j-1}(p) = c_{j-1,M}(p) - v_p$ and η_{j-1} is a sequence of positive numbers satisfying $\lim_{j \rightarrow \infty} \eta_j = 0$, $\sum_{j=0}^{\infty} |\eta_j - \eta_{j+1}| < \infty$ and $\sum_{j=0}^{\infty} \eta_j = \infty$ and $\Pi_{\mathcal{B}}[\cdot]$ denotes the projection to \mathcal{B} .

The stochastic MPC with adaptive constraint tightening is summarized as follows:

1) For $j=0$, initialize the algorithm with $\beta_0(p_k) = 1$ and feasible trajectory $\bar{x}_0(k)$, collect the value of loss $q(g_k^T x_0(k) - h_k)$, $k=1, \dots, T$. Then set $j=1$ and $k=0$.

2) At the beginning of iteration $j \geq 1$, reset the nominal initial state to x^0 , update $\beta_j(1), \dots, \beta_j(s)$ according to (11).

3) At each time instant $0 \leq k < T$ of iteration $j > 1$, solve Problem 1, apply $u_j(k) = \bar{u}_j^*(k|k) + K(x_j(k) - \bar{x}_j(k))$ to the plant, update $\bar{x}_j(k+1)$ according to (10) and collect $\bar{x}_j(k+1)$ and $I_k(x_j(k+1))$. At $k = T$, reset $k = 0$ and $j = j + 1$, go to 2).

4. CONVERGENCE ANALYSIS

In this section, we are going to show that under some mild conditions, the accumulated violation ratio $c_{l,M}(p)$ will converge to v_p in some stochastic sense as $l, M \rightarrow \infty$. The main idea is to treat the update law (11) as a root-searching procedure of equation $h(\mathbf{B}) = \mathbf{V}$. The main approach follows the similar line of Kushner and Yin (2003) and Muñoz-Carpintero et al. (2018). The difference between this work and Muñoz-Carpintero et al. (2018) is that we apply the results in Kushner and Yin (2003) in a multi-dimension case and interpret the system dynamics along iteration instead of along time.

For completeness, we briefly introduce the Stochastic Approximation (SA) Kushner and Yin (2003) first. Consider an equation $h(\beta) = 0$. A classical problem in SA that will be shown to be closely related to this work is to find the solution of $h(\beta) = 0$, which is denoted as β^* , when $h(\cdot)$ is not known explicitly. Suppose that there is a sequence of input $\beta(k)$ and a Markov process $\xi(k)$ and the measurement of $h(\beta(k))$, which is denoted as $y(k)$, satisfying that $E_k(y(k)) = \tilde{h}(\beta(k), \xi(k))$, where $\tilde{h}(\beta(k), \xi(k))$ represents the expectation of the noisy measurement depending on current input $\beta(k)$ and internal state $\xi(k)$. To find the root of this equation in a given bounded set \mathcal{B} , consider the following algorithm:

$$\beta(k+1) = \Pi_{\mathcal{B}}[\beta(k) + \eta(k)y(k)], \quad (12)$$

where $\eta(k)$ is the step size.

The convergence of (12) has been studied in Kushner and Yin (2003) and Kushner and Vázquez-Abad (1996), where weak convergence of $\beta(k)$ to β^* has been proved. In Muñoz-Carpintero et al. (2018), a simplified result leading to convergence in probability has been proposed. In what follows, we will use this simplified result.

Lemma 4.1. Muñoz-Carpintero et al. (2018) The stochastic process $\beta(k)$ defined in (12) converges to β^* in probability if:

1) There is a memory process $\{\xi(k) \subset \mathcal{E} : k \in \mathbb{T}\}$ for some set \mathcal{E} , a filtration $\{\mathcal{F}_k : k \in \mathbb{T}\}$, where \mathcal{F}_k measures all the information used to obtain $\{\beta(k), \xi(k), y(i) : i < k, \beta(0)\}$, and a measurable continuous bounded function $\tilde{h}(\cdot, \cdot)$ satisfying:

$$E_k(y(k)) = \tilde{h}(\beta(k), \xi(k)).$$

2) $\eta(k) \rightarrow 0$, $\eta(k) > 0$, $\sum_{k=0}^{\infty} |\eta(k) - \eta(k+1)| < \infty$ and $\sum_{k=0}^{\infty} \eta(k) = \infty$.

3) The sequence $\{(y(k), \beta(k), \xi(k)) : k \in \mathbb{T}, k < \infty\}$ is bounded.

4) There is a transition function $P(\cdot, \cdot|\beta)$ such that $P(\xi, \cdot|\beta)$ is weakly continuous in (β, ξ) and satisfies

$$\mathbb{P}(\xi(k+1) \in \cdot | \xi(i), \beta(i), i \leq k) = P(\xi(k), \cdot | \beta(k)). \quad (13)$$

Note that $P(\cdot, \cdot|\beta)$, for a fixed β , defines a Markov chain and the associated random variables are denoted as $\xi_{\beta}(k)$.

5) For a fixed β and any initial condition $\xi_{\beta}(0)$,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} E(\tilde{h}(\beta, \xi_{\beta}(i))) = h(\beta).$$

6) The differential equation

$$\dot{\beta} = h(\beta) + z, \quad z(t) \in -C_{\mathcal{B}}(\beta(t)),$$

where $z(t)$ is the minimum force needed to keep β in \mathcal{B} , has a unique stationary limit point β^* .

Based on the above lemma, now we are ready to show that \mathbf{B}_j and $\frac{\sum_{l=0}^{k-1} \mathbf{Q}_l}{k}$ converge to \mathbf{B}^* and \mathbf{V} in probability, respectively. The key idea is to treat the iterative learning process as a dynamic system along iteration index j instead of along time index k .

Denote $\mathbf{B}^* \triangleq (\beta_1^*, \dots, \beta_s^*)'$.

Proposition 4.1. Suppose that $\sigma_p(\cdot)$, $p = 1, \dots, s$ satisfy Assumption 3.2. Then the solution of the differential equation $\dot{\mathbf{B}} = g(\mathbf{B}) + z$, where $g(\mathbf{B}) = h(\mathbf{B}) - \mathbf{V}$, $z(t) \in -C_{\mathcal{B}^s}(\mathbf{B}(t))$ is the minimum force needed to keep \mathbf{B} in \mathcal{B}^s , converges to its unique stationary point \mathbf{B}^* if $\mathbf{B}(0) \in \mathcal{B}^s$.

Proof. By Assumption 3.2, for any $p = 1, \dots, s$, if $\beta_p(0) \in \mathcal{B}$, then $\beta_p(t) \in \mathcal{B}$, $\forall t > 0$. Therefore, the force term $z(t) = 0$, $\forall t > 0$. Consider $V(t) = \frac{1}{2}(\mathbf{B} - \mathbf{B}^*)'(\mathbf{B} - \mathbf{B}^*)$. Some simple calculations lead to that

$$\begin{aligned} \dot{V}(t) &= \dot{\mathbf{B}}'(\mathbf{B} - \mathbf{B}^*) \\ &= (h(\mathbf{B}) - \mathbf{V})'(\mathbf{B} - \mathbf{B}^*) \\ &= \sum_{p=1}^s (\sigma_p(\beta_p(t)) - v_p)(\beta_p(t) - \beta_p^*) \end{aligned}$$

Assumption 3.2 implies that $(\sigma_p(\beta_p(t)) - v_p)(\beta_p(t) - \beta_p^*) < 0$, $\forall \beta_p(t) \neq \beta_p^*$. Then by Lyapunov stability we have the desired result.

The system dynamics along the iteration index j can be rewritten as follows:

$$\begin{aligned} \bar{\mathbf{X}}_{j+1} &= g(\mathbf{B}_{j+1}), \\ \mathbf{X}_{j+1} &= f(\mathbf{W}_{j+1}, \mathbf{B}_{j+1}), \end{aligned} \quad (14)$$

where $\mathbf{B}_{j+1} = \Pi_{\mathcal{B}^s}[\mathbf{B}_j + \eta_j \mathbf{Y}_j]$, where $\mathbf{Y}_j = \mathbf{C}_{j,M} - \mathbf{V}$. By Assumption 3.1 we know that $g(\cdot)$ and $f(\cdot, \cdot)$ are bounded and continuous.

Proposition 4.2. If Assumption 2.2, 3.1 and 3.2 are satisfied, then \mathbf{B}_j , $\frac{\sum_{l=0}^{k-1} \mathbf{Q}_l}{k}$ and $\bar{\mathbf{X}}_j$ converge to \mathbf{B}^* , \mathbf{V} and $g(\mathbf{B}^*)$ in probability, respectively.

Proof. This proof is to check all conditions in Lemma 4.1 as in Muñoz-Carpintero et al. (2018). Consider $j \geq M - 1$.

1) According to the definition, for $j \geq 1$, \mathbf{Q}_j is a function of \mathbf{X}_j and consequently, a function of \mathbf{W}_j and \mathbf{B}_j , so we denote it as $\mathbf{Q}_j(\mathbf{W}_j, \mathbf{B}_j)$. On the other hand, we know that $\mathbf{C}_{j,M} = \frac{\sum_{l=j-M+1}^j \mathbf{Q}_l(\mathbf{W}_l, \mathbf{B}_l)}{M}$. Denote $\bar{\mathbf{Q}}_j(\mathbf{B}_j) = E(\mathbf{Q}_j(\mathbf{W}_j, \mathbf{B}_j)|\mathbf{B}_j)$ and $\zeta_j(\mathbf{W}_j, \mathbf{B}_j) = \mathbf{Q}_j(\mathbf{W}_j, \mathbf{B}_j) - \bar{\mathbf{Q}}_j(\mathbf{B}_j)$. Clearly, $E(\zeta_j(\mathbf{W}_j, \mathbf{B}_j)|\mathbf{B}_j) = 0$. Now we can define the memory process $\xi_j = [\mathbf{Q}_{j-1}, \dots, \mathbf{Q}_{j-M+1}]$ and the function $\tilde{h}(\cdot, \cdot)$ in Lemma 4.1 is given by

$$\tilde{h}(\mathbf{B}_j, \xi_j) = \frac{\bar{\mathbf{Q}}_j(\mathbf{B}_j) + \sum_{l=j-M+1}^{j-1} \mathbf{Q}_l}{M} - \mathbf{V},$$

and

$$\mathbf{Y}_j = \tilde{h}(\mathbf{B}_j, \xi_j) + \frac{\zeta_j(\mathbf{W}_j, \mathbf{B}_j)}{M}.$$

Since the sub- σ -algebra \mathcal{F}_{j-1} collects all available information after the $(j-1)$ -th iteration, which includes \mathbf{B}_j , we have $E_{j-1}(\frac{\zeta_j(\mathbf{W}_j, \mathbf{B}_j)}{M}) = 0$ and $\frac{\zeta_j(\mathbf{W}_j, \mathbf{B}_j)}{M}$ is a martingale difference.

The boundedness of $\tilde{h}(\cdot, \cdot)$ directly follows the definition of $\bar{\mathbf{Q}}_j$. To show its continuity, we only need to show that

$$\bar{\mathbf{Q}}_j(\mathbf{B}_j) = \left(\frac{\sum_{k \in H_1} E(q(g_k^T x_j(k) - h_k)|\mathbf{B}_j)}{|H_1|}, \dots, \frac{\sum_{k \in H_s} E(q(g_k^T x_j(k) - h_k)|\mathbf{B}_j)}{|H_s|} \right),$$

is continuous.

By the designed algorithm, $x_j(k+1) = Ax_j(k) + B(\bar{u}_j^*(k|k) + K(x_j(k) - \bar{x}_j(k))) + w_j(k)$. Thus, given \mathbf{B}_j , \mathbf{X}_j can be expressed as

$$\begin{aligned} \mathbf{X}_j &= Lx^0 + P\bar{\mathbf{U}}_j + S\bar{\mathbf{X}}_j + Q\mathbf{W}_j, \\ &= Lx^0 + P\bar{\mathbf{U}}_j + Sg(\mathbf{B}_j) + Q\mathbf{W}_j, \end{aligned}$$

where L, P, S and Q are matrices with proper dimension. Note that $\bar{\mathbf{U}}_j$ is also a continuous function of \mathbf{B}_j . So, every $x_j(k)$ is a continuous function of \mathbf{B}_j and $q(g_k^T x_j(k) - h_k)$ is a semicontinuous function of \mathbf{B}_j with at most countable discontinuous points. Therefore, $E(q(g_k^T x_j(k) - h_k)|\mathbf{B}_j)$ is continuous with respect to \mathbf{B}_j since this expectation is essentially an integral of \mathbf{W}_j depending on the parameter \mathbf{B}_j .

2) This condition is satisfied by design.

3) \mathbf{Y}_j is bounded since $\tilde{h}(\cdot, \cdot)$ and \mathbf{Q}_j are bounded. By Assumption 3.1, $\bar{\mathbf{X}}_j$ is bounded so ξ_j is bounded as well. \mathbf{B}_j is bounded due to the projection operator.

4) Note that $\xi_{j+1} = [\mathbf{Q}_j, \dots, \mathbf{Q}_{j-M+2}]$. We have shown that \mathbf{Q}_j is a function of \mathbf{W}_j and \mathbf{B}_j so the distribution of \mathbf{Q}_j is determined by \mathbf{B}_j and the distribution of \mathbf{W}_j . $\bar{\mathbf{X}}_j$ is also a function of \mathbf{B}_j . Therefore, the distribution of ξ_{j+1} is completely determined by ξ_j and \mathbf{B}_j since $[\mathbf{Q}_{j-1}, \dots, \mathbf{Q}_{j-M+2}]$ are included in ξ_j . So we have shown that (13) is satisfied.

Note that for arbitrary set $\mathcal{X} \in \mathbb{R}^{n(T+1)}$, the set $\mathcal{X} \oplus (-Lx^0 + P\bar{\mathbf{U}}_j + Sg(\mathbf{B}_j))$ varies continuously with respect to \mathbf{B}_j . As a consequence, $\mathbb{P}(Q\mathbf{W}_j \in \mathcal{X} \oplus (-Lx^0 + P\bar{\mathbf{U}}_j +$

$Sg(\mathbf{B}_j))|\mathbf{B}_j)$ is a continuous function of \mathbf{B}_j and $\mathbb{P}(\mathbf{X}_j \in \mathcal{X}|\mathbf{B}_j)$ is also continuous.

For the continuity, one needs to note that for an arbitrary set \mathcal{A} , $\mathbb{P}(\mathbf{Q}_j \in \mathcal{A}|\mathbf{B}_j)$ can be rewritten as $\mathbb{P}(\mathbf{X}_j \in \mathcal{X}|\mathbf{B}_j)$ for some set \mathcal{X} . Furthermore, $[\mathbf{Q}_{j-1}, \dots, \mathbf{Q}_{j-M+2}]$ are included in ξ_j . Combining the above facts together leads to that $P(\xi_j, \cdot|\mathbf{B}_j)$ is continuous.

5) This condition is satisfied by Assumption 3.2.

6) It has been shown in Proposition 4.1.

Therefore, by Lemma 4.1 and (14), \mathbf{B}_j and $\bar{\mathbf{X}}_j$ converge to \mathbf{B}^* and $g(\mathbf{B}^*)$ in probability.

Assumption 3.2 implies that $\lim_{l \rightarrow \infty} \frac{\sum_{j=0}^{l-1} \mathbf{Q}_j(\mathbf{B}^*)}{l} \rightarrow \mathbf{V}$ in probability. Note that \mathbf{Q}_j is continuous with respect to \mathbf{B}_j . Consider $\frac{\sum_{j=0}^l \mathbf{Q}_j(\mathbf{B}_j) - \mathbf{Q}_j(\mathbf{B}^*)}{l+1}$. Since \mathbf{B}_j converges to \mathbf{B}^* in probability, for any given probability p and ϵ , there exists a large enough integer \bar{l} such that $|\mathbf{Q}_l(\mathbf{B}_l) - \mathbf{Q}_l(\mathbf{B}^*)| < \epsilon$ with probability larger than p for all $l \geq \bar{l}$. Therefore, $\frac{\sum_{j=0}^l |\mathbf{Q}_j(\mathbf{B}_j) - \mathbf{Q}_j(\mathbf{B}^*)|}{l+1} = \frac{\sum_{j=0}^{\bar{l}-1} |\mathbf{Q}_j(\mathbf{B}_j) - \mathbf{Q}_j(\mathbf{B}^*)|}{l+1} + \frac{\sum_{j=\bar{l}}^l |\mathbf{Q}_j(\mathbf{B}_j) - \mathbf{Q}_j(\mathbf{B}^*)|}{l+1}$ where the first term can be arbitrarily small by choosing large enough \bar{l} and the second term is less than ϵ with probability higher than p , which means that $\frac{\sum_{j=0}^l \mathbf{Q}_j}{l}$ converges to $\lim_{l \rightarrow \infty} \frac{\sum_{j=0}^{l-1} \mathbf{Q}_j(\mathbf{B}^*)}{l}$ in probability and consequently converges to \mathbf{V} in probability.

5. NUMERICAL EXAMPLE

In this section, we consider the following HVAC system from Muñoz-Carpintero et al. (2018) and Ma et al. (2011):

$$x(k+1) = Ax(k) + Bu(k) + V(\bar{w}(k) + \delta(k)),$$

where $A = \begin{bmatrix} 0.762 & 0.231 \\ 0.013 & 0.988 \end{bmatrix}$, $V = \begin{bmatrix} 0.0069 & 0.0004 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} -3.181 \\ 0 \end{bmatrix}$, $x = (x_1, x_2)^T$ with x_1 being the air temperature of the room and x_2 the temperature of a slow-dynamics mass (floor, walls and furniture), the control input is the cool air mass flow injected into the room, $\bar{w}(k)$ is the deterministic part of the disturbance with \bar{w}_1 the ambient temperature and \bar{w}_2 indoor thermal load and $\delta(k)$ is the stochastic part of the disturbance. δ_1 follows a truncated standard normal distribution over $[-4, 4]$, and δ_2 is 100ν with ν following a truncated normal distribution over $[-4, 4]$. The initial condition $x_1(0)$ and $x_2(0)$ are random variables uniformly distributed in $[24.3, 24.7]$ and $[23.8, 24.2]$, respectively. $k = 0, \dots, 47$ represents the sampling time instants of one day where sampling interval is 30 minutes.

In this example we consider two kinds of constraints. The first one is the chance constraint which aims to regulate the violation frequency and the second one is the integrated chance constraint which aims to regulate the violation magnitude and frequency.

Denote the set of time instants of working hours as H_{work} and that of non-working hours as H_{off} . The chance con-

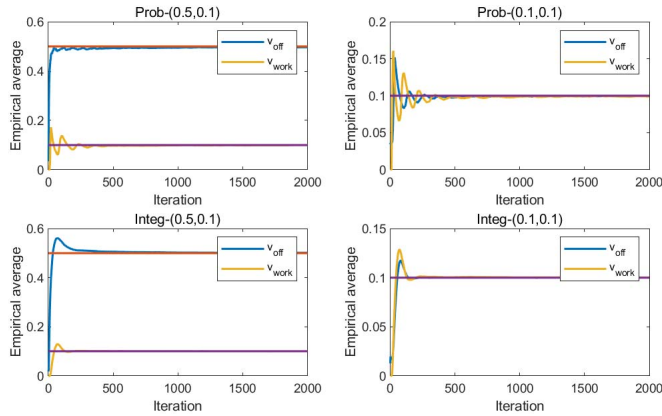


Fig. 1. Evolution of empirical average loss

straint (Prob) is given by $v_{\text{work}} \triangleq \frac{\sum_{k \in H_{\text{work}}} \mathbb{P}(x(k) \geq 24)}{|H_{\text{work}}|}$ and $v_{\text{off}} \triangleq \frac{\sum_{k \in H_{\text{off}}} \mathbb{P}(x(k) \geq 24)}{|H_{\text{off}}|}$. The integrated chance constraint (Integ) is given by $v_{\text{work}} \triangleq \frac{\sum_{k \in H_{\text{work}}} E(\max(x(k) - 24, 0))}{|H_{\text{work}}|}$ and $v_{\text{off}} \triangleq \frac{\sum_{k \in H_{\text{off}}} E(\max(x(k) - 24, 0))}{|H_{\text{off}}|}$.

We test the proposed algorithm for $(v_{\text{off}}, v_{\text{work}}) = (0.1, 0.1)$ and $(0.5, 0.1)$ for both types of constraint. In the following, we use 'Integ' and 'Prob' as the abbreviation of the integrated chance constraint and probabilistic constraint, respectively. Then the four combinations of different kinds of constraints and expected losses are denoted as Integ – $(0.1, 0.1)$, Integ – $(0.5, 0.1)$, Prob – $(0.1, 0.1)$ and Prob – $(0.5, 0.1)$, respectively.

In the 4 cases the constraint of the control input is $0 \leq u \leq 1$. The objective function is $\sum_{i=k}^{k+N-1} \bar{u}(i|k)$ with $N = 7$ as defined in Problem 1. The initialization of the tightened constraints are calculated using the tube based approach in Mayne et al. (2005). The step size η_j in (12) is given by $\eta_j = \frac{20}{j+20}$ and $M = 50$. In Fig. 1, the evolution of the empirical average for working and non-working hours along the iteration index is shown.

Finally, we directly apply the relaxation coefficients found by the proposed algorithm with time-varying and uniform stochastic constraints to the system for 400 iterations to compare their mean costs. The performance improvement compared with the robust approach in Mayne et al. (2005) is shown in Table 1. Note that the approach in Muñoz-Carpintero et al. (2018) leads to the result of Prob – $(0.1, 0.1)$.

Table 1. Average cost and performance improvement of different approaches

Constraint type- $(v_{\text{off}}, v_{\text{work}})$	Mean Cost	Improvement
Integ-(0.5, 0.1)	6.8131	21.02%
Integ-(0.1, 0.1)	7.6413	11.42%
Prob-(0.5, 0.1)	7.9691	7.62%
Prob-(0.1, 0.1)	8.0690	6.47%
Robust constraint	8.6268	0%

6. CONCLUSION

In this paper, an iterative learning based approach has been proposed for MPC subject to time varying stochastic constraints. Without assuming the knowledge of the distribution function of the uncertainty, an adaptive updating law has been designed to relax and tighten the constraints according to past realizations of the state trajectories. Under some proper assumptions, the learning procedure has been shown to be equivalent to a root-searching problem and stochastic approximation theory has been applied to ensure convergence. A numerical example on HVAC system control has shown the effectiveness of the proposed algorithm.

REFERENCES

Chua, K.J., Chou, S.K., Yang, W.M., and Yan, J. (2013). Achieving better energy-efficient air conditioning – A review of technologies and strategies. *Applied Energy*, 104, 87–104.

Kushner, H. and Vázquez-Abad, F. (1996). Stochastic approximation methods for systems over an infinite horizon. *SIAM Journal on Control and Optimization*, 34(2), 712–756.

Kushner, H.J. and Yin, G.G. (2003). *Stochastic Approximation and Recursive Algorithms and Applications*. Springer.

Lemmet, S. (2009). Building and climate change: summary for decision maker. Technical report, United Nations Environment Programme.

Lorenzen, M., Dabbene, F., Tempo, R., and Allgöwer, F. (2017). Constraint-tightening and stability in stochastic model predictive control. *IEEE Transactions on Automatic Control*, 62(7), 3165–3177.

Ma, Y., Anderson, G., and Borrelli, F. (2011). A distributed predictive control approach to building temperature regulation. In *Proceedings of the 2011 American Control Conference*, 2089–2094.

Magni, L., Pala, D., and Scattolini, R. (2009). Stochastic model predictive control of constrained linear systems with additive uncertainty. In *2009 European Control Conference (ECC)*, 2235–2240.

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

Muñoz-Carpintero, D., Hu, G., and Spanos, C.J. (2018). Stochastic model predictive control with adaptive constraint tightening for non-conservative chance constraints satisfaction. *Automatica*, 96, 32 – 39.

Schildbach, G., Fagiano, L., Frei, C., and Morari, M. (2014). The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations. *Automatica*, 50(12), 3009 – 3018.