A modular approach to constraint satisfaction under uncertainty - with application to bioproduction systems

Yu Wang *,** **Xiao Chen** * **Elling W. Jacobsen** *,**

* *Division of Decision and Control Systems, EECS, KTH Royal Institute of Technology, Stockholm 10044, Sweden (e-mail: wang3@kth.se, xiao2@kth.se, jacobsen@kth.se).*
** *Competence Centre for Advanced Bioproduction by Continuous Processing, AdBIOPRO, Sweden.*

Abstract: The paper proposes a modular-based approach to constraint handling in process optimization and control. This is partly motivated by the recent interest in learning-based methods, e.g., within bioproduction, for which constraint handling under uncertainty is a challenge. The proposed constraint handler, called predictive filter, is combined with an adaptive constraint margin to minimize the cost of violating soft constraints due to uncertainty and disturbances. The module can be combined with any controller and is based on modifying the controller output, in a least squares sense, such that constraints are satisfied within the considered horizon. The proposed method is computationally efficient and suitable for real-time applications. The effectiveness of the method is illustrated by a simple heater example and a nonlinear and time-varying example in penicillin fed-batch production optimization.

Keywords: Constraint handling, predictive filter, adaptive constraint margin, bioproduction

1. INTRODUCTION

Recently, learning-based optimization and control methods have received increased interest in biochemical applications, mainly due to their efficient learning ability in highly uncertain systems, e.g., [Shields et al. (2021); Kim et al. (2021); Ma et al. (2021); Mehrian et al. (2018)]. Such learning algorithms are typically explorative and hence likely to violate constraints that are critical for safety or economic considerations.

In bioproduction systems, there are usually constraints on e.g., concentrations, temperature etc., and it is often optimal to operate at or close to such constraints. These constraints can be temporally violated but such violations are typically associated with significant economic costs. Due to large uncertainties in biochemical systems, high-cost constraints are likely to be violated, especially with explorative learning algorithms. To minimize the cost of such constraint violations, in the presence of uncertainty (e.g., disturbances), we propose to combine the predictive filter [Wabersich and Zeilinger (2018)], a modular approach that can ensure constraint satisfaction with arbitrary controllers, with an adaptive constraint margin.

Note that predictive filters are used in e.g., autonomous driving applications [Tearle et al. (2021)], and that the margin adaptation employed here is based on the margin optimization proposed for economic model predictive control in Trollberg et al. (2017). Our main contribution is the combination of the predictive filter and the margin adaptation such that the modular constraint handler can minimize the cost of soft constraint violations without prior knowledge on the uncertainties by learning the optimal constraint margins directly from data.

⋆ This work has been supported by the VINNOVA Competence Centre AdBIOPRO, contract 2016-05181.

Different from existing methods, the proposed method is a simple and universal module that can easily be applied to systems with arbitrary controllers or optimizers, such as learning-based algorithms. Notably, the proposed method only alters the control signal when necessary. This ensures a minimal modification of the desired closed-loop behaviour given by the preferred controller or optimizer while satisfying constraints with no prior knowledge on the uncertainties. The proposed method is computationally efficient and therefore suitable for online applications. It can adapt the optimal constraint margins according to the size of uncertainties including unmeasured disturbances, and can readily be applied in complex applications, e.g., time-varying nonlinear systems without prior knowledge on the uncertainties, as we demonstrate in Section 5.

Several methods have been proposed to handle constraints in the presence of uncertainty. If some prior knowledge on the uncertainty is available, e.g., the uncertainty belongs to a given bounded set or distribution, constraints can be handled by using e.g., robust and stochastic model predictive control (MPC) [Bayer et al. (2014, 2016); Mesbah (2016); Wu et al. (2018); Parisio et al. (2016); Lucia et al. (2014); Bayer et al. (2018); Mesbah et al. (2014)]. However, in practice, such prior knowledge on the uncertainty is rarely available. In such cases, one can handle constraints by adaptive methods [Bujarbaruah et al. (2020); Trollberg et al. (2017); Oldewurtel et al. (2013); Chachuat et al. (2008)]. All these methods rely on particular controller structures, such as MPC, and are therefore not generally applicable to control algorithms such as those based on learning methods. For such controllers, predictive filters [Tearle et al. (2021); Wabersich and Zeilinger (2018)], whose purpose is to satisfy hard safety constraints, can be applied. Predictive filters have recently also been considered for dealing with soft constraints, but then primarily to avoid infeasibility [Wabersich and Zeilinger (2021a); Wabersich et al. (2021)]. There also exist works on robust extensions of the predictive safety filter (see e.g., [Wabersich and Zeilinger (2021b)]), but these are all based on some prior knowledge of the uncertainty characteristics. Apart from the works above, we also note that in the context of learning-based control, there exist works focusing on parameter adaptation with economic considerations [Kim et al. (2021); Kordabad et al. (2021b); Gros and Zanon (2019); Kordabad et al. (2021a); Alhazmi et al. (2021); Zanon and Gros (2020); Piga et al. (2019)]. However, rather than high-cost constraint satisfaction for arbitrary systems, they consider obtaining operational conditions to optimize economical performance using learning methods. To our knowledge there exist no methods for constraint satisfaction that can be applied in systems with arbitrary controllers and adapt optimal constraint margins for soft constraints directly from data, and without prior knowledge on the uncertainty.

The rest of the paper is organized as follows. We first provide a motivating example in Section 2, then formulate the problem in Section 3. We then introduce the learning-based predictive filter in Section 4. In Section 5, we illustrate the effectiveness of the method on two biochemical applications. Finally, we provide conclusions and discussion of the work in Section 6.

2. MOTIVATING EXAMPLE

We first provide a simple water tank heating control example to illustrate the impact of uncertainty on constraint satisfaction, and demonstrate the need for constraint satisfaction methods that can learn the optimal constraint margins and be applied to systems with arbitrary controllers. Consider a nonlinear water tank process whose dynamics are modeled by [Trollberg et al. (2017)]

$$\begin{aligned} A\frac{dh}{dt} &= F_{in} - F_{out}, \\ \frac{dT}{dt} &= \frac{F_{in}(T_{in}-T)}{Ah} - \frac{K(hO+A)(T-T_{amb})}{CAh}, \\ T_{in} &= T_0 + (P_e+P_b)/(CF_{in}), \\ \tau_b\frac{dP_b}{dt} &= -P_b + P_b^{max}\frac{P_b^{sp}}{100}, \end{aligned} \tag{1}$$

where the states x are the water level h, the temperature in the water tank T, and the effect of combustion heater P_b; the inputs u are the tank inflow F_{in}, the effect of the electric heater P_e, and the set-point of the combustion heater given in percentage P_b^{sp}; the disturbances d are the outflow from the tank F_{out}, the water temperature at inlet T_0, and ambient temperature T_{amb}; other parameters are the specific heat of water C, the specific heat transfer coefficient from the hot part of the tank to air K, the cross-sectional area A, the circumference of the tank O, the combustion heater time constant τ_b, and the maximum effect of the combustion heater P_b^{max}. To summarize, we have

$$x = \begin{bmatrix} T \\ h \\ P_b \end{bmatrix}, \quad u = \begin{bmatrix} F_{in} \\ P_e \\ P_b^{sp} \end{bmatrix}, \quad d = \begin{bmatrix} T_0 \\ F_{out} \\ T_{amb} \end{bmatrix}.$$

Assume the nominal optimization gives an optimal operating condition $T_{\text{opt}} = 35°C$ which is also the upper limit. If the temperature T exceeds the optimal value T_{opt}, it will induce a significant economic loss, which is here motivated by biochemical applications in which the temperature usually is process critical. The ultimate goal of the proposed method is to provide constraint satisfaction for learning-based controllers that cannot handle constraints explicitly. However, for the convenience of illustration, we here apply a simple decentralized PI controller to track T_{opt}.

As shown in Fig. 1, when using the PI controller to track T_{opt}, there are as expected frequent violations of the constraint $T \leqslant T_{\text{opt}}$ that leads to large costs as illustrated by the cost L_k in the lower part of the figure. In order to handle the constraints, we apply a separate module named predictive filter [Wabersich and Zeilinger (2018)] which modifies the desired control signal if there exists a potential constraint violation within the prediction horizon; otherwise, the desired control signal is not altered. As can be seen from the figure, the constraint violation is less severe but still frequent, mainly due to various uncertainties [1] that cannot be completely characterized a priori. Therefore, it is advisable to introduce a constraint margin to reduce the impact of uncertainties. For this purpose, we introduce a constraint margin $\theta_T \in \mathbb{R}$ such that the modified constraint in the predictive filter is $T \leqslant T_{opt} - \theta_T$. As shown in Fig. 1, with $\theta_T = 0.6$, there is now no violation of T_{opt}. However, the constraint margin is conservative since the resulting T is far from the optimum T_{opt}. Thus, it is important to adapt the margin such that the combined cost of constraint violations and cost of operating away from the optimum is minimized.

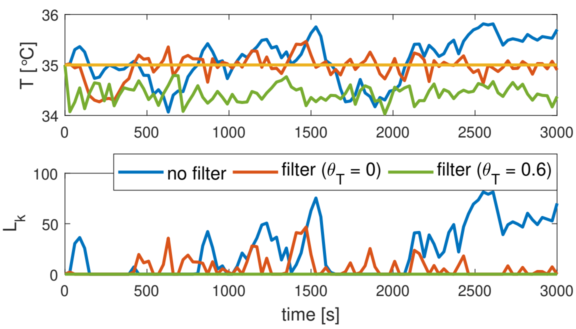

Fig. 1. The performance of PI controller (blue), PI controller with predictive filter with $\theta_T = 0$ (red) and $\theta_T = 0.6$ (green) in the presence of uncertainties including disturbances. The yellow line represents the optimal value $T_{\text{opt}} = 35°C$. The lower figure is the corresponding constraint violation cost (described later).

3. PROBLEM FORMULATION

Consider a discrete-time nonlinear process model

$$x_{k+1} = f(x_k, u_k) + \Delta_k, \tag{2}$$

where $k \in \mathbb{N}$ represents the time step of the discrete-time system; $x \in \mathbb{R}^{n_x}$ is the state vector of the plant; $u \in \mathbb{R}^{n_u}$ is the control input; $f(\cdot) : \mathbb{R}^{n_u} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ represents the nonlinear dynamics of the plant; $\Delta_k \in \mathbb{R}^{n_x}$ represents the uncertainty at time step k. A key assumption of this paper is that the magnitude and characteristics of Δ_k is unknown a priori. The system is subject to constraints

$$\begin{aligned} &g_h(x_k, u_k) \leqslant 0, h(x_k, u_k) = 0, \\ &g_e(x_k, u_k) \leqslant 0, \end{aligned} \tag{3}$$

where $g_h(x_k, u_k) \leqslant 0, h(x_k, u_k) = 0$ are hard equality and inequality constraints; $g_e(x_k, u_k) \leqslant 0$ are inequality constraints

[1] The considered uncertainties are described in Section 5.

that can be temporarily violated but with significant violation cost. Let the cost of violation of $g_e(x_k,u_k) \leqslant 0$ at sample k be

$$L_k = \lambda^T \max(0, g_e(x_k,u_k)), \tag{4}$$

where $\lambda \in \mathbb{R}^{n_\lambda}$ is a constant vector that reflects the constraint violation cost.

We here assume that the system (2) is regulated by a task-specific controller that generates a desired control signal u_k^d at sample k. Although the constraints can be integrated into the controller using e.g., MPC, we consider the case where the controller does not directly handle constraints, since we aim at solving constraint violation in general cases. To satisfy the constraints, it is necessary to modify u_k^d in case of a predicted violation. However, to maintain the closed-loop behaviour as close as possible to the desired one, the modification of u_k^d should be minimal.

Assumption 1. Since we aim at solving constraint violation problems in general cases, we do not assume linear dynamics, i.i.d. disturbance, etc. Rather we make the following reasonable assumptions:

(1) the size and characteristics of Δ_k is unknown a priori;
(2) the closed-loop system is stable;
(3) there exist at least a local optimum of the cost (5).

Problem 1. Under Assumption 1, use u_k with the minimal modification of u_k^d to satisfy all hard constraints while minimizing the long-term cost of violating $g_e(x_k,u_k) \leqslant 0$, defined as

$$J = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \bar{L}_k, \tag{5}$$

with

$$\bar{L}_k = L_k + P_k, \tag{6}$$

where L_k is the cost of violating the constraint, as defined in (4), and P_k reflects the cost of operating away from the optimal operational conditions. The purpose of P_k is to avoid over-conservative constraint margin and make a trade-off between constraint satisfaction and optimal operating conditions.

In this work, we propose a method based on the predictive filter [Wabersich and Zeilinger (2018)] combined with online learning of constraint margins to solve Problem 1.

4. LEARNING-BASED PREDICTIVE FILTER

In the absence of uncertainty, the high-cost constraint satisfaction problem can be solved using a safety predictive filter [Tearle et al. (2021); Wabersich and Zeilinger (2018)]. In the presence of uncertainty, one can employ robust extensions of the safety predictive filter [Wabersich and Zeilinger (2021b)] for the constraint satisfaction problem, if a priori knowledge on the uncertainty magnitude and characteristics are available. However, such a priori knowledge on the uncertainty is usually not available in practice.

To solve Problem 1, we here propose an adaptive predictive filter framework, as shown in Fig. 2. We introduce a constraint margin $\theta \in \mathbb{R}^{n_\lambda}$ for the high-cost constraints, i.e., $g_e(x_k,u_k) + \theta \leqslant 0$, which is learned online directly from available data using the constraint adaptor. Given a learned θ, the constraint satisfaction is then ensured with the minimal modification of the desired closed-loop behaviour by the predictive filter. Note that although one can also introduce some margins on e.g., setpoint of the PI controller in Section 2, the predictive filter provides a more robust and less conservative margin due to its predictive ability.

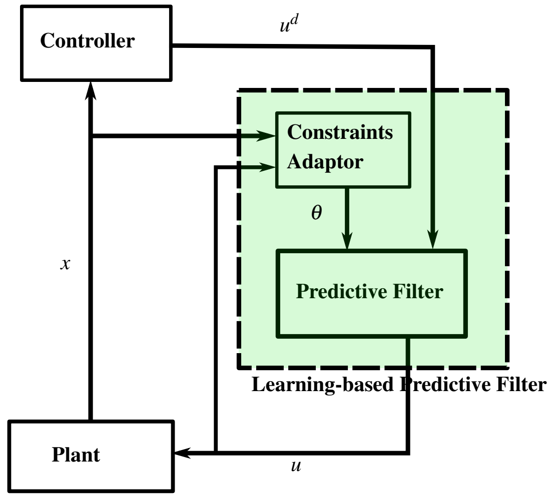

Fig. 2. The learning-based predictive filter framework

4.1 Predictive filter

Given a constraint margin θ, we apply a predictive filter to improve constraint satisfaction $g_e(y_k,u_k) + \theta \leqslant 0$, for arbitrary desired control signals generated by a task-specific controller. This is repeated at every sample.

The predictive filter [Wabersich and Zeilinger (2018)] applies a mechanism that is independent of a task-specific objective. The basic idea is that, given a desired control signal and the current state measurement or estimate, we first determine whether it will satisfy all constraints, including $g_e(x_k,u_k) + \theta \leqslant 0$, within the given horizon. If all constraints are satisfied within the horizon, the desired control signal is applied directly. If a constraint violation is predicted, we obtain the closest control signal sequence that satisfy the constraints within the horizon, and apply the first control signal in the modified sequence to the plant. Formally, given a constraint margin θ, the predictive filter is given by

$$\begin{aligned} \min_{u_{i|k}} \quad & \|u_k^d - u_{0|k}\|_2 \\ \text{s.t.} \quad & \text{for all } i = 0,\ldots,N: \\ & x_{0|k} = x_k, \\ & x_{i+1|k} = f(x_{i|k}, u_{i|k}), \\ & g_h(x_{i|k}, u_{i|k}) \leqslant 0, h(x_{i|k}, u_{i|k}) = 0, \\ & g_e(x_{i|k}, u_{i|k}) + \theta \leqslant 0, \end{aligned} \tag{7}$$

where u_k^d is the desired control generated by the task-specific controller at time step k; N is the prediction horizon of the predictive filter; the subscript $i|k$ represents the i^{th} step ahead prediction where the predictive filter is initialized at time step k. The predictive filter generates the optimal control sequence $\{u_{i|k}^*\}$. Then, the first optimal input signal $u_{0|k}^*$ in the optimal control sequence is applied to the plant. If all the constraints can be satisfied within the prediction horizon, the control signal u_k^d is directly applied to the plant, since $u_k^d = u_{0|k}^*$; otherwise, the desired control signal u_k^d is modified to $u_{0|k}^*$ that is closest to u_k^d in the sense of 2-norm and ensures constraint satisfactions within the horizon. By repeatedly applying the predictive filter at each time step, the satisfaction of the constraints can be guaranteed [Tearle et al. (2021); Wabersich and Zeilinger (2018)].

In bioproduction applications, hard constraints can usually be enforced by physical saturations. Motivated by this, we here focus on learning the optimal constraint margin for high-cost soft constraints. Also note that we assume that the modification of the desired input through the predictive filter is bounded in a way such that the closed-loop system is stable, for any feasible choice of θ. We refer the interested readers to e.g., Wabersich and Zeilinger (2018) and Wabersich and Zeilinger (2021b) for discussions on e.g., feasibility of the predictive filter.

The remaining problem is then to determine the optimal constraint margin θ that minimizes the impact of uncertainties on constraint violations without being conservative. We next apply the constraint adaptor to learn the optimal constraint margin.

4.2 Constraint adaptor

The constraint adaptor learns the optimal constraint margin θ that minimizes (5) based on online measurements. We here consider the finite-difference stochastic approximation (FDSA) [Spall (1998, 2005); Kushner and Clark (2012)] to adapt the constraint margin θ. Note that the proposed framework is general and the learning method can be replaced by other, possibly more effective, methods. Below we briefly outline the FDSA-based adaptation method. More details can be found in Trollberg et al. (2017).

Assume there exist at least a local optimum of J defined in (5) with respect to θ, and that J is well defined for any feasible choice of θ. The FDSA method seeks the local optimum of J by approximating the gradient $\partial J/\partial\theta$ by its finite-difference approximation denoted as DJ. More specifically, at each iteration of FDSA, we component-wise perturb the applied constraint margin in the predictive filter, then approximate the gradient $\partial J/\partial\theta$ based on the online measurements of (6) corresponding to the perturbed margins. The constraint margin is then adapted along the direction of the gradient approximation DJ. The local optimum of J can be obtained by applying the FDSA recursively. We next formally introduce the FDSA method.

At the j^{th} iteration of FDSA update, the l^{th} component of the finite-difference approximation of the gradient is given by

$$[DJ(\theta_j,c_j)]_l = \frac{J^+_{j,l}-J^-_{j,l}}{2c_j}, \tag{8}$$

where θ_j is the constraint margin value at the j^{th} iteration; c_j is the perturbation size on θ to approximate the gradient using finite difference; $J^+_{j,l}$ and $J^-_{j,l}$ are the long-term constraint violation cost (5) with $\theta_j + c_j e^l$ and $\theta_j - c_j e^l$ as constraint margins in the predictive filter, respectively, with e^j as the standard basis vector, and can be approximated by averaging the online measurements of (6) over large samples.

Similar to gradient descent, the FDSA method recursively updates θ based on

$$\theta_{j+1} = \theta_j + a_j DJ(\theta_j,c_j), \tag{9}$$

where a_j is the step size along the gradient approximation $DJ(\theta_j,c_j)$. The convergence of θ can be guaranteed under some conditions on the step size sequences $\{a_j\}$ and $\{c_j\}$, [Spall (1998, 2005)]:

(1) $c_j \to 0$ as $j \to \infty$, to diminish the error in the finite-difference approximation of the gradient over time;
(2) $\lim_{n\to\infty}\sum_{j=1}^{n} a_j = \infty$, to avoid premature convergence;
(3) $\lim_{n\to\infty}\sum_{j=1}^{n} a_j c_j < \infty$ and $\lim_{n\to\infty}\sum_{j=1}^{n} a_j^2/c_j^2 < \infty$, to reduce the adaptation rate by ensuring the step sizes decrease over time.

Note that the FDSA-based method does not require gradient information of J to update θ, but instead adapts θ directly from online measurements. As we will show in Section 5, the FDSA can converge relatively fast to the optimal θ. It implies that the optimal margin can be learned efficiently in terms of computational burden. Therefore, the proposed method is suitable for online applications.

We remark that although the proposed predictive filter resembles a simple MPC, it focuses solely on constraint satisfaction for arbitrary desired controllers. Also, the proposed framework is general and not limited by specific choices of the learning method.

5. EXAMPLES

To demonstrate the effectiveness of the proposed method, we first revisit the nonlinear water tank control example in Section 2. Then, we consider a nonlinear time-varying example in fed-batch penicillin fermentation optimization.

5.1 Water tank control - motivating example revisited

Uncertainties We consider the following uncertainties in the process [Trollberg et al. (2017)]. The disturbance d is governed by

$$\dot{d} = \begin{bmatrix} -1/\tau_T & 0 & 0 \\ 0 & -1/\tau_F & 0 \\ 0 & 0 & 0 \end{bmatrix} d + d_0 + \nu_d, \tag{10}$$

where the time constants are $\tau_T = 20s$, $\tau_F = 20s$; the nominal value is $d_0 = [15\ \ 1\ \ 25]^T$; the stochastic part

$$\nu_d \sim [U(-0.5,0.5)\ \ \beta(B(1,0.5)-0.5)\ \ U(-0.5,0.5)]^T,$$

($B(1,0.5)$ is the Bernoulli distribution) keeps constant over the sampling interval and is updated at each sampling instance. We introduce 30% relative input uncertainties, i.e., the actual inputs to the plant are given by $F_{in}(1+\nu_F)$, $P_b(1+\nu_b)$, $P_b^{sp}(1+\nu_{b,sp})$ where $\nu_F, \nu_b, \nu_{b,sp} \sim U(-0.3,0.3)$.

Note that we consider a number of uncertainties on a nonlinear system, including modelling error, discretization error, disturbances and input uncertainties. The size and characteristics of these are assumed unknown a priori. Hence, constraint handling is hard using existing methods.

Constraints We consider the following constraints:

$$g_h(x,u) = \begin{bmatrix} h-20 \\ F_{in}-10 \\ P_e-200 \\ P_b^{sp}-100 \\ 0-h \\ 0.01-F_{in} \\ 0-P_b^{sp} \\ 0-P_e \end{bmatrix} \leqslant 0,\ \ g_e(x,u) = T-35 \leqslant 0, \tag{11}$$

where $g_h(x,u)$ contains hard constraints on the states and the inputs, while $g_e(x,u)$ represents high-cost soft constraints. We assume the input and state hard constraints are enforced by physical saturations. We therefore only adapt the constraint margins of the high-cost soft constraints $g_e(x,u)$ in the predictive filter. The objective is to minimize the impact of the uncertainties on the high-cost constraint violations, i.e., $g_e(x,u) > 0$, with a minimal modification of the desired control signal generated by the PI controller. Obviously, the constraints above cannot be handled using the simple PI controller.

Results To achieve the objective, we apply the learning-based predictive filter described in Section 4. In the constraint adaptor, we introduce $\theta_T \in \mathbb{R}$ as the constraint margin of the high-cost constraints $g_e(x,u)$. The desired control signal $u^d \in \mathbb{R}^3$ is generated by the decentralized PI controller with input saturations. In the predictive filter, we apply the linearized model around the given setpoint with sampling time $T_s = 30s$. The prediction horizon is chosen to $N = 5$. The optimal constraint margin θ_T is learned online using the FDSA method. For each θ_T value considered in the gradient approximation step, we measure the violation cost with penalty term (6) where $\lambda = 100$ and $P_k = \max(0, T_{opt} - T)$ from 200 samples, and then take the average of these costs as an approximation of (5) for the corresponding θ_T value in the gradient approximation step. As shown in Fig. 3 (in blue), θ_T quickly converges to $\theta_T^* = 0.26$ with a low violation cost within few iterations, starting from $\theta_{T,0} = 0$ where the constraint violation cost is relatively high. Note that when using the PI controller without constraint handling (in yellow, Fig. 3), the violation cost is around 6 times larger than the case when using the predictive filter with $\theta_{T,0} = 0$. This shows that constraint handling as such is important under uncertainty.

When running the process, the violation cost should be monitored constantly. If the violation cost is close to 0 for a period of time, the current margin θ_T may no longer be optimal due to the change of the uncertainties. Therefore, a new optimal θ_T should be learned by resetting the step size sequences in (9). For example, in Fig. 3 (in red), after reducing the size of the uncertainties by 30%, a new optimal margin $\theta_T = 0.2$ is learned, starting from the previous optimal margin $\theta_T^* = 0.26$. This shows that the proposed method has the ability to adjust optimal constraint margin according to a change in the uncertainties.

The proposed method provides the minimal modification on the control signal provided by the controller in order to satisfy constraints. To illustrate this, we compare the performance of the decentralized PI controller with and without the learning-based predictive filter in the presence of the same uncertainties. As shown in Fig. 4 (a), the actual temperature seldom exceeds $T_{opt} = 35°C$ using the learning-based predictive filter with $\theta_T^* = 0.26$, while the violation is severe and quite frequent using the PI controller only. Note that, due to the minimal modification on the desired input, the temperature, when using the learning-based predictive filter, is still around $T_{opt} = 35°C$ with a similar behaviour as in the case with the PI controller only.

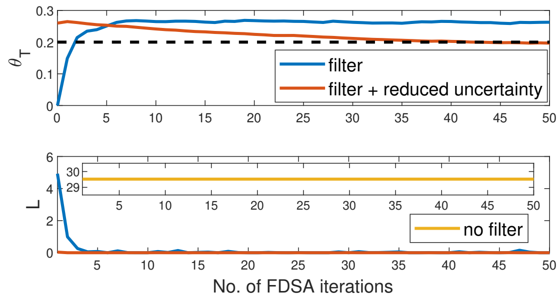

Fig. 3. The trajectory of constraint margin θ_T and constraint violation cost measurement L (obtained by averaging (4) of 200 samples for each θ_T) during learning iterations using the FDSA method in the water tank example.

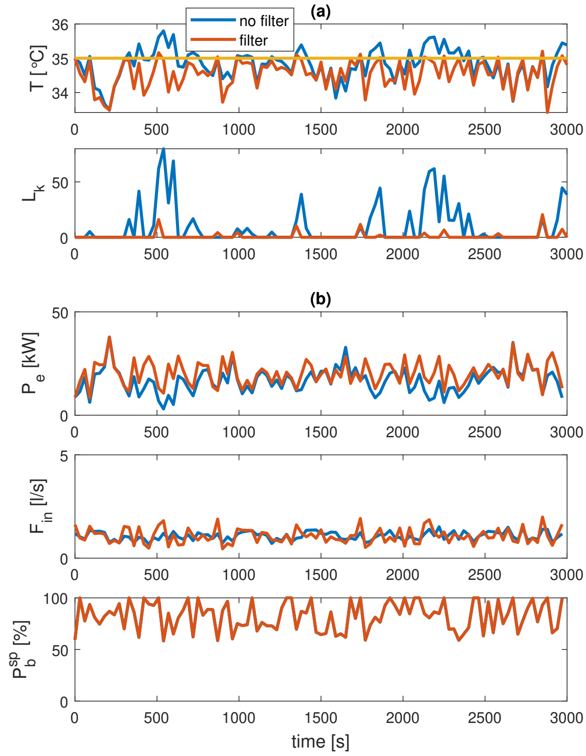

Fig. 4. Comparison of the decentralized PI controller with(red)/without(blue) the learning-based predictive filter with the same unknown uncertainties. The predictive filter is with the converged constraint margin $\theta_T^* = 0.26$. Fig. (a) shows the temperature setpoint tracking and constraint satisfaction performance; Fig. (b) shows the corresponding input signals. The solid yellow line represents the constraint $T_{\text{opt}} = 35°C$.

5.2 Application to fed-batch penicillin fermentation

Fed-batch fermentation is common for production in the pharmaceutical industry, in which substrates are supplied to the bioreactor without outflow during operation [Bonvin (1998)]. An important problem in fed-batch fermentation is how to minimize undesired side reactions and thereby improve quality of the product. To minimize undesired side reactions, constraints on the maximum substrate concentrations should be imposed [San and Stephanopoulos (1989)]. At the same time, there should be constraints on the minimum substrate concentrations to ensure cell growth and production of the desired product [Srinivasan et al. (2003)]. This is a challenging task due to the fact that available models are highly uncertain and the size and characteristic of the uncertainties are usually not fully known.

We consider the constraint satisfaction problem in fed-batch penicillin fermentation with several sources of uncertainty. The objective is to ensure substrate constraint satisfaction, while maximizing the penicillin concentration in the bioreactor at the end of the fermentation. We employ a nominal optimal feeding strategy proposed as in Srinivasan et al. (2003) and combine it with the proposed predictive filter and adaptive constraint margin. Compared with the water tank example, the main difference is that the plant is time-varying and a nonlinear model is employed in the predictive filter.

We adapt the fed-batch penicillin fermentation model presented in Srinivasan et al. (2003); Lucia and Engell (2013); Bajpai and Reuss (1981)

$$\begin{aligned} \frac{dX}{dt} &= \mu(S)X - \frac{F}{V}X, \\ \frac{dS}{dt} &= -\frac{\mu(S)}{Y_X} - \frac{\gamma X}{Y_p} + \frac{F}{V}(S_{in} - S), \\ \frac{dP}{dt} &= \gamma X - \frac{F}{V}P \\ \frac{dV}{dt} &= F, \end{aligned} \tag{12}$$

where $\mu(S) = \frac{\mu_m S}{K_m + S + (S^2/K_i)}$; X is the biomass concentration; P is the product concentration; V is the volume of the bioreactor; μ_m, K_m, K_i, γ are kinetic parameters; Y_x and Y_p are yield coefficients; F is the feed rate; S_{in} is the inlet substrate concentration.

We consider F as the control input, i.e., $u = F$. The objective is to ensure the inequality constraint $S \in [S_{min}, S_{max}]$, i.e.,

$$g_e(x,u) = \begin{bmatrix} S - S_{max} \\ S_{min} - S \end{bmatrix} \leqslant 0 \tag{13}$$

is satisfied for arbitrary feeding strategy u. The imposed upper bound $S_{max} = 0.6$ g/l is to minimize undesired side reactions, while the imposed lower bound $S_{min} = 0.2$ g/l is to ensure cell growth and production of the desired product. By ensuring constraint satisfaction, we improve product quality and hence the economic profit in the presence of uncertainty, including disturbances. We here consider the growth phase in fed-batch fermentation such that the biomass concentration reaches the desired value X_{max}. We assume that the desired feeding strategy is given by the nominal optimal nonlinear feedback control proposed in Srinivasan et al. (2003)

$$u_d = \frac{V}{S_{in} - S}\left(\frac{1}{Y_x}\mu(S)X - \frac{1}{Y_p}\gamma X\right)\Big|_{S=\sqrt{K_i K_m}}, \text{ as } X \leqslant X_{max}, \tag{14}$$

where $X_{max} = 3.7g/l$. The initial operating conditions and parameters are summarized in Appendix B. We consider similar uncertainties as in Srinivasan et al. (2003):

(1) Y_X is a time-varying parameter that varies every 2 hours within the range [0.3,0.5];
(2) S_{in} is normally distributed with mean $200g/l$ and standard deviation $25g/l$, which is assumed to change very 2 hours;
(3) we also introduce 10% relative input uncertainties on u that vary between samples, i.e., the actual input is given by $(1 + \nu_u)u$ where $\nu_u \sim U(-0.1, 0.1)$.

We assume the characteristics and size of the uncertainties above are unknown a priori.

Results To ensure constraint satisfaction without prior knowledge on the uncertainties and with minimal modification of the nominal optimal feeding strategy, we apply the learning-based predictive filter. We introduce $\theta = [\theta_{upper} \; \theta_{lower}]^T \in \mathbb{R}^2$ as the constraint margin of the high-cost constraints $g_e(x,u)$. The desired feeding strategy $u^d \in \mathbb{R}$ is given by (14). In the predictive filter, we apply the nominal nonlinear model (12) with sampling time $T_s = 2h$. The prediction horizon is chosen as $N = 5$.

The optimal constraint margin θ is learned online using the FDSA method. For each θ value considered in the gradient approximation step, we measure the violation cost (4) with $\lambda = [1 \; 1]^T$ from 20 parallel fed-batch experiments using the optimal feeding strategy (14) together with the corresponding predictive filter, and then take the average of these costs as an approximation of (5) for the corresponding θ value in the gradient approximation step. As shown in Fig. 5, by using the FDSA method, a value in the region of minimum violation cost $\theta^* = [0.08 \; 0.11]^T$ is learned online starting from the initial guess $\theta_0 = [0.02 \; 0.02]^T$, where the constraint violation cost is relatively high. Although it is relatively time-consuming to learn the optimal margin θ in this example, we remark that the experiments can be conducted in parallel and, as shown in Fig. 6, θ converges to a region of low violation cost within only 10 iterations. Note that, without constraint handling (in yellow, Fig. 6), the violation cost is around 10 times larger than with the case using the predictive filter with θ_0. This demonstrates the importance of constraint handling.

Fig. 7 shows the comparison of the optimal feeding strategy (14) with and without the learning-based predictive filter in the presence of the same uncertainties. For the first 50 hours, the input is the same as the nominal feeding strategy without predictive filter, since there is no potential constraint violation within the horizon. After 50 hours, the substrate concentration exceeds the given upper limit $S_{\max} = 0.6g/l$ in the case without the predictive filter, while the predictive filter modifies the feeding strategy such that the potential constraint violation is avoided. Notably, both penicillin and biomass concentrations are very close to those for the case without predictive filter, due to the minimal modification of the desired input. This implies that, by adding the learning-based predictive filter module, the optimization performance is not altered significantly from the desired one, but the constraints with high violation cost due to production of undesired byproducts can be satisfied.

We stress that this example serves to demonstrate the generality and potential of the proposed method for more complex applications. While the case with many fed-batch bioreactors with similar behavior may not be realistic, our ultimate goal is to provide constraint satisfaction for explorative learning based methods in continuous bioproduction which is fully realistic with the proposed method.

6. CONCLUSION AND DISCUSSION

In this paper we proposed a universal modular approach to constraint handling, based on a predictive filter combined with constraint margin adaptation to minimize the cost of soft constraint violations. The main advantage of the proposed method is that, as a universal module, it can be easily applied to a system with arbitrary controller and can ensure constraint satisfaction with minimal modification of the desired closed-loop performance. This provides a simple approach to enable systems to complete tasks using arbitrary controllers or optimizers, such as learning-based algorithms, while satisfying constraints. We illustrated the potential of the proposed method using a water tank example and a fed-batch penicillin optimization example. The efficiency of the constraint margin adaptation and constraint satisfaction, while largely maintaining the desired closed-loop behaviour, demonstrates the potential of the proposed method for application in complex control and optimization systems that do not explicitly handle constraints.

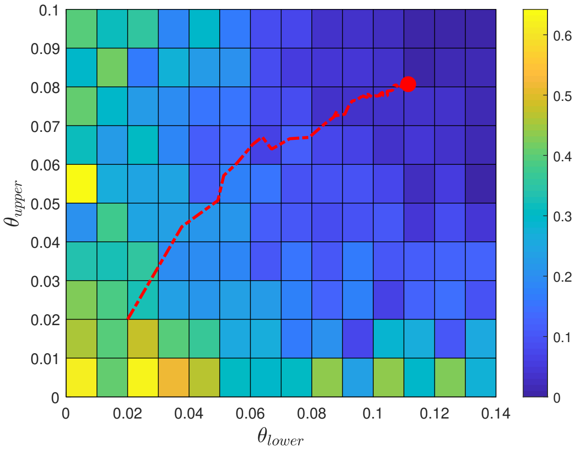

Fig. 5. The relation between constraint violation cost J and θ values in the fed-batch penicillin fermentation example. The constraint violation cost J for each θ is obtained by averaging (6) of 50 parallel fed-batch experiments using the feeding strategy (14) with the corresponding predictive filter, and represented by different colors. The dash line represents the trajectory of θ during online learning process from 20 parallel fed-batch experiments (See Fig. 6), and the red dot represents the converged θ using the FDSA method.

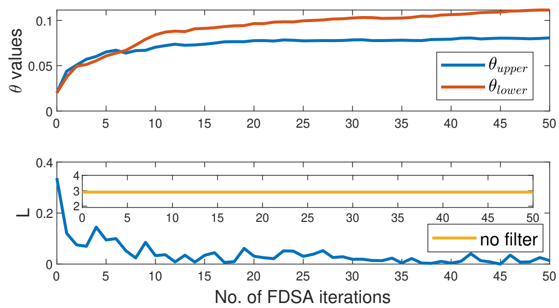

Fig. 6. The trajectory of constraint margin θ and violation cost measurement L (obtained by averaging (4) of 20 fed-batch experiments) during learning iterations using the FDSA method in the fed-batch penicillin fermentation example.

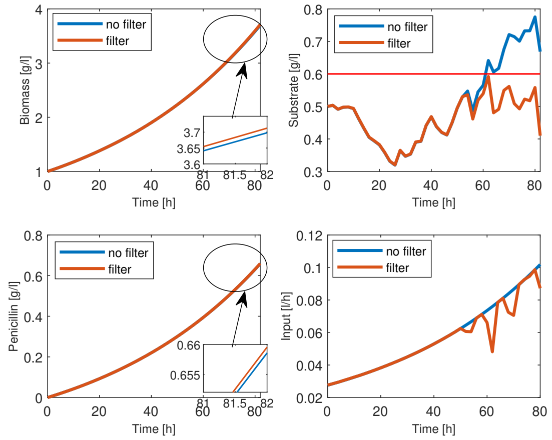

Fig. 7. Comparison of the optimal feeding strategy with(red)/without(blue) the learning-based predictive filter with the same uncertainties. The converged constraint margin $\theta^* = [0.08\ 0.11]^T$. The solid red line represents the upper bound for substrate concentration $S_{\max} = 0.6g/l$. The case with violation against the lower bound is omitted here.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Lukas Hewing and Prof. Melanie Zeilinger for their valuable feedback. Y. Wang and E.W. Jacobsen have been supported by the VINNOVA Competence Centre AdBIOPRO, contract 2016-05181. The support is gratefully acknowledged.

REFERENCES

Alhazmi, K., Albalawi, F., and Sarathy, S.M. (2021). A reinforcement learning-based economic model predictive control framework for autonomous operation of chemical reactors. *arXiv preprint arXiv:2105.02656*.

Bajpai, R. and Reuss, M. (1981). Evaluation of feeding strategies in carbon-regulated secondary metabolite production through mathematical modeling. *Biotechnology and Bioengineering*, 23(4), 717–738.

Bayer, F.A., Lorenzen, M., Müller, M.A., and Allgöwer, F. (2016). Robust economic model predictive control using stochastic information. *Automatica*, 74, 151–161.

Bayer, F.A., Müller, M.A., and Allgöwer, F. (2014). Tube-based robust economic model predictive control. *Journal of Process Control*, 24(8), 1237–1246.

Bayer, F.A., Müller, M.A., and Allgöwer, F. (2018). On optimal system operation in robust economic MPC. *Automatica*, 88, 98–106.

Bonvin, D. (1998). Optimal operation of batch reactors—a personal view. *Journal of Process Control*, 8(5-6), 355–368.

Bujarbaruah, M., Zhang, X., Tanaskovic, M., and Borrelli, F. (2020). Adaptive stochastic MPC under time-varying uncertainty. *IEEE Transactions on Automatic Control*, 66(6), 2840–2845.

Chachuat, B., Marchetti, A., and Bonvin, D. (2008). Process optimization via constraints adaptation. *Journal of Process Control*, 18(3-4), 244–257.

Gros, S. and Zanon, M. (2019). Data-driven economic NMPC using reinforcement learning. *IEEE Transactions on Automatic Control*, 65(2), 636–648.

Kim, J.W., Park, B.J., Oh, T.H., and Lee, J.M. (2021). Model-based reinforcement learning and predictive control for two-stage optimal control of fed-batch bioreactor. *Computers & Chemical Engineering*, 107465.

Kordabad, A.B., Cai, W., and Gros, S. (2021a). MPC-based reinforcement learning for economic problems with application to battery storage. *arXiv preprint arXiv:2104.02411*.

Kordabad, A.B., Esfahani, H.N., Lekkas, A.M., and Gros, S. (2021b). Reinforcement learning based on scenario-tree MPC for ASVs. *arXiv preprint arXiv:2103.11949*.

Kushner, H.J. and Clark, D.S. (2012). *Stochastic approximation methods for constrained and unconstrained systems*, volume 26. Springer.

Lucia, S., Andersson, J.A., Brandt, H., Diehl, M., and Engell, S. (2014). Handling uncertainty in economic nonlinear model predictive control: A comparative case study. *Journal of Process Control*, 24(8), 1247–1259.

Lucia, S. and Engell, S. (2013). Robust nonlinear model predictive control of a batch bioreactor using multi-stage stochastic programming. In *2013 European Control Conference (ECC)*, 4124–4129.

Ma, Y., Wang, Z., Castillo, I., Rendall, R., Bindlish, R., Ashcraft, B., Bentley, D., Benton, M.G., Romagnoli, J.A., and Chiang, L.H. (2021). Reinforcement learning-based fed-batch optimization with reaction surrogate model. In *2021 American Control Conference (ACC)*, 2581–2586.

Mehrian, M., Guyot, Y., Papantoniou, I., Olofsson, S., Sonnaert, M., Misener, R., and Geris, L. (2018). Maximizing neotissue growth kinetics in a perfusion bioreactor: an in silico strategy using model reduction and bayesian optimization. *Biotechnology and Bioengineering*, 115(3), 617–629.

Mesbah, A. (2016). Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems Magazine*, 36(6), 30–44.

Mesbah, A., Streif, S., Findeisen, R., and Braatz, R.D. (2014). Stochastic nonlinear model predictive control with probabilistic constraints. In *2014 American Control Conference*, 2413–2419.

Oldewurtel, F., Sturzenegger, D., Esfahani, P.M., Andersson, G., Morari, M., and Lygeros, J. (2013). Adaptively constrained stochastic model predictive control for closed-loop constraint satisfaction. In *2013 American Control Conference*, 4674–4681.

Parisio, A., Rikos, E., and Glielmo, L. (2016). Stochastic model predictive control for economic/environmental operation management of microgrids: An experimental case study. *Journal of Process Control*, 43, 24–37.

Piga, D., Forgione, M., Formentin, S., and Bemporad, A. (2019). Performance-oriented model learning for data-driven MPC design. *IEEE control systems letters*, 3(3), 577–582.

San, K.Y. and Stephanopoulos, G. (1989). Optimization of fed-batch penicillin fermentation: A case of singular optimal control with state constraints. *Biotechnology and Bioengineering*, 34(1), 72–78.

Shields, B.J., Stevens, J., Li, J., Parasram, M., Damani, F., Alvarado, J.I.M., Janey, J.M., Adams, R.P., and Doyle, A.G. (2021). Bayesian reaction optimization as a tool for chemical synthesis. *Nature*, 590(7844), 89–96.

Spall, J.C. (1998). Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on Aerospace and Electronic Systems*, 34(3), 817–823.

Spall, J.C. (2005). *Introduction to stochastic search and optimization: estimation, simulation, and control*, volume 65. John Wiley & Sons.

Srinivasan, B., Bonvin, D., Visser, E., and Palanki, S. (2003). Dynamic optimization of batch processes: II. handling uncertainty using measurements. *Computers and Chemical Engineering*, 27(1), 27–44.

Tearle, B., Wabersich, K.P., Carron, A., and Zeilinger, M.N. (2021). A predictive safety filter for learning-based racing control. *arXiv preprint arXiv:2102.11907*.

Trollberg, O., Rojas, C.R., and Jacobsen, E.W. (2017). Online constraint adaptation in economic model predictive control. In *20th IFAC World Congress*, 9065–9070. IFAC.

Wabersich, K.P., Krishnadas, R., and Zeilinger, M.N. (2021). A soft constrained MPC formulation enabling learning from trajectories with constraint violations. *IEEE Control Systems Letters*, 6, 980–985.

Wabersich, K.P. and Zeilinger, M.N. (2018). Linear model predictive safety certification for learning-based control. In *2018 IEEE Conference on Decision and Control (CDC)*, 7130–7135.

Wabersich, K.P. and Zeilinger, M.N. (2021a). Predictive control barrier functions: Enhanced safety mechanisms for learning-based control. *arXiv preprint arXiv:2105.10241*.

Wabersich, K.P. and Zeilinger, M.N. (2021b). A predictive safety filter for learning-based control of constrained nonlinear dynamical systems. *Automatica*, 129, 109597.

Wu, Z., Zhang, J., Zhang, Z., Albalawi, F., Durand, H., Mahmood, M., Mhaskar, P., and Christofides, P.D. (2018). Economic model predictive control of stochastic nonlinear systems. *AIChE Journal*, 64(9), 3312–3322.

Zanon, M. and Gros, S. (2020). Safe reinforcement learning using robust MPC. *IEEE Transactions on Automatic Control*, 66(8), 3638–3652.

Appendix A. DECENTRALIZED PI CONTROLLER DESIGN

In the water tank example, we apply the linearized model around the setpoint an operating point $T_{\text{ref}} = 35°C$, $h_{\text{ref}} = 8.5$ dm, $P_{b,\text{ref}} = 83.6$ kW. To design a decentralized PI controller, we first apply Relative Gain Array (RGA) to determine suitable pairing between inputs and states. The RGA of the linearized system is given by

$$\Lambda(G(\omega)) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \forall\omega \tag{A.1}$$

which suggests that it is suitable to control T, h, P_b using P_e, F_{in}, P_b^{sp} respectively. For each pairing, we design a PI controller such that the crossover frequency is $\omega_c = 1/200$ rad/s. In simulation, we apply the discretized decentralized PI controller obtained by discretizing the continuous-time decentralized PI controller using zero-order hold with sampling time $T_s = 30s$.

Appendix B. PARAMETERS OF FED-BATCH PENICILLIN FERMENTATION EXAMPLE

Table B.1. Nominal parameters of the bioreactor model [Srinivasan et al. (2003)].

$\mu_{m,0} = 0.02\,h^{-1}$	$K_{m,0} = 0.05\,g/l$	$K_{i,0} = 5\,g/l$
$Y_{x,0} = 0.5\,g[X]/g[S]$	$Y_{p,0} = 1.2\,g[P]/g[S]$	$\gamma_0 = 0.004\,h^{-1}$

Table B.2. Initial operating conditions of the process [Srinivasan et al. (2003)].

$X_0 = 1\,g/l$	$S_0 = 0.5\,g/l$	$P_0 = 0\,g/l$	$V_0 = 150\,l$