Computers and Chemical Engineering 128 (2019) 128–140

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Model predictive control with active learning for stochastic systems with structural model uncertainty: Online model discrimination

Tor Aksel N. Heirung[a], Tito L.M. Santos[a,b], Ali Mesbah[a,*]

[a] Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA
[b] Department of Electrical Engineering, Federal University of Bahia, CEP 40210-630, Salvador, Bahia, Brazil

ARTICLE INFO

Article history:
Received 4 December 2018
Revised 29 March 2019
Accepted 7 May 2019
Available online 31 May 2019

Keywords:
Predictive control
Stochastic systems
Dual control
Structural model uncertainty
Online model discrimination
Bayesian decision theory
Closed-loop fault diagnosis

ABSTRACT

Structural model uncertainty is prevalent in control design and arises from incomplete knowledge of the system or the existence of different modes of dynamic behavior, such as those arising from system faults and malfunctions. This paper addresses control of stochastic nonlinear systems using model predictive control, or MPC, under structural model uncertainty. Inspired by dual control, the MPC strategy with active learning presented here can probe the uncertain system to select, among a set of candidates, the model that best describes the observed closed-loop system data. The proposed controller involves online model selection based on estimation of the model-hypothesis probabilities and minimization of a computationally tractable measure of the predicted Bayes risk of selection error. The performance of the proposed approach is compared to that of nominal MPC with no learning, MPC with passive learning, and a robust MPC approach that systematically accounts for structural model uncertainty but has no learning mechanism. Simulation results on a nonlinear bioreactor demonstrate that active learning can have significant advantages in maintaining adequate control performance in the presence of structural uncertainty. Active learning can be particularly beneficial for improving online model discrimination and active fault diagnosis under closed-loop control.

1. Introduction

The use of multiple models has seen increasing applications in control of nonlinear systems over a wide operating range. Multiple-model control generally involves representing the system dynamics with a set of local models, designing a controller for each model, and incorporating the local controllers into a global one using an online switching method (Murray-Smith and Johansen, 1997; Azimzadeh et al., 2001; Rodriguez et al., 2003; Özkan and Kothare, 2006; Böling et al., 2007; Nandola and Bhartiya, 2008; Du and Johansen, 2014). When classic adaptive control based on a single model with parametric uncertainty can be unsatisfactory because of significant system uncertainties, multiple models can enable adaptive control under structural model uncertainty; see, e.g., Narendra and Han (2011). Uncertainty in the structure of a model can arise from incomplete knowledge of the system (such as unmodeled dynamics and disturbances), system dynamics that vary significantly over time, and the occurrence of different modes of operation that may arise from system faults and malfunctions. Accounting for significant system variations through online model adaptation in multiple-model control can ensure adequate control performance under various operating scenarios (Gopinathan et al., 1998; Kanev and Verhaegen, 2000; Maciejowski and Jones, 2003; Kuure-Kinsey and Bequette, 2009). One of the main challenges in multiple-model control is developing models that correspond to the different scenarios, some of which may be undesirable. While improved system design can circumvent a variety of such scenarios, others may occur no matter how well the system is designed and operated, in particular if their root causes are exogenous to the system.

Multiple-model control approaches generally involve switching between competing models (Murray-Smith and Johansen, 1997; Narendra and Han, 2011). That is, the control input at any given time is computed based on a single model and is not intended for learning about the model structure or discriminating between the competing models. However, when the control input affects not only the system state but possibly also the future state uncertainty, the *dual effect* is present in the system (Bar-Shalom and Tse, 1974). In this case, probing the system for information can become an important part of reducing that uncertainty. A *dual controller*, originally introduced for parametric model uncertainty (Feldbaum, 1961), naturally incorporates probing into the control input to *actively learn* about the uncertain model parameters to

* Corresponding author.
E-mail address: mesbah@berkeley.edu (A. Mesbah).

https://doi.org/10.1016/j.compchemeng.2019.05.012
0098-1354/

the extent required for achieving optimal control performance. The two aspects of dual control are often referred to as *exploitation* (i.e., making a decision based on the currently available information) and *exploration* (i.e., probing the system for learning through generating more information); see Åström and Wittenmark (1995). Although dual control is in general computationally intractable for practical systems (Wittenmark, 1995), the growing body of literature on approximate dual control has demonstrated the benefits of *active learning* under parametric uncertainty for improved control performance relative to *passive learning*, in which the controller does not increase the information content of the closed-loop data (e.g., Shouche et al., 2002; Larsson et al., 2016; Houska et al., 2017; La et al., 2017; Heirung et al., 2017b; Mesbah, 2018; Lorenzen et al., 2019).

This paper addresses the problem of simultaneous control and active learning under structural model uncertainty in the context of model predictive control, or MPC. There is a range of well-established methods for systematically accounting for model uncertainty in MPC. Among the most common are robust MPC (Bemporad and Morari, 1999; Mayne, 2014; Kouvaritakis and Cannon, 2016), which relies on deterministic bounds on the uncertainty, and stochastic MPC (Mayne, 2014; Kouvaritakis and Cannon, 2016; Mesbah, 2016), based on probabilistic uncertainty specifications. These approaches, however, generally do not involve updating the uncertainty descriptions online. Adaptive MPC, conversely, updates the system model online in response to new measurements, but the controller typically has no mechanism for active reduction of uncertainty (e.g., Tanaskovic et al., 2014; Zhu and Xia, 2016; Guay et al., 2015; Benosman, 2018).

In this work, we address the problem of MPC with active learning under structural model uncertainty (Heirung and Mesbah, 2017; Heirung et al., 2018) through an approach similar to dual control for parametric uncertainty. There are fundamental differences between the problems of reducing structural model uncertainty and learning model parameters. In particular, optimal inputs for discriminating between model structures are often ill suited for estimating parameters (Atkinson and Fedorov, 1975; Atkinson, 2008). In some cases, however, there is no clear distinction between structural and parametric model uncertainty. One example of such a case is when large variations in some quantity have more severe operational consequences than those more commonly represented as parameter drifts. Such large changes can additionally have different root causes than those of smaller magnitudes. In this case, larger variations may be better represented as significant events that are differences in kind, rather than degree, and treated as anomalous.

The primary focus of this paper is the introduction of active learning to reduce structural model uncertainty in MPC. We present a tractable formulation for MPC with active uncertainty reduction for stochastic nonlinear systems. The formulation uses a set of model hypotheses (possibly with uncertain parameters) to represent different candidates for the model structure or, alternatively, describe various operating modes for the system. The proposed MPC formulation includes a mechanism for active learning to reduce structural model uncertainty. That is, the control inputs not only direct the system state but also facilitate model discrimination based on closed-loop measurements. We accomplish this by modifying a finite-horizon optimal control problem to include a computationally tractable measure of the Bayes risk of choosing the incorrect model hypothesis (Matusita, 1971). The optimal control problem explicitly accounts for the cost of future structural model uncertainty, and its solution thus balances the costs of control and active learning. Our receding-horizon implementation of this optimal control problem relies on a Bayesian approach to recursively update the estimated model-hypothesis probabilities.

In what follows, the idea of active learning in optimal control under structural model uncertainty is first illustrated using a motivating example. Here, the optimal control problem with active learning can be solved exactly (Section 2). The subsequent section presents the control problem considered in this work and its associated challenges (Section 3), followed by the proposed solution methods and control algorithm (Section 4). A case study involving a nonlinear bioreactor with structural model uncertainty demonstrates the proposed MPC strategy with active learning in Section 5. The demonstration includes performance comparisons to nominal MPC with no learning, a robust MPC that systematically accounts for structural uncertainty but has no learning mechanism, and MPC with passive learning.

Notation: $\mathbb{Z}_{[a,b]}$ is the set of integers $\{a, a+1, \dots, b\}$. $v_k \in \mathbb{R}^n$ denotes the value of v at discrete time index k, with $v_{k+j|k}$ denoting the value of v at future time $k+j$ predicted from time k. $z_{a:b} = [z_a^\top, z_{a+1}^\top, \dots, z_b^\top]^\top$ is a vector concatenation of a sequence of values of $z_k \in \mathbb{R}^n$ from discrete time indices $k = a$ through $k = b$. With $z \in \mathbb{R}^n$ a realization of the real-valued random variable Z, $p(z)$ ($p(z\,|\,\cdot)$) denotes the (conditional) probability density function (PDF) for Z and $\mathrm{E}[Z]$ ($\mathrm{E}[Z\,|\,\cdot\,]$) denotes the (conditional) expected value of Z. $\Pr[A] \in [0, 1]$ is the probability of event A, whereas $\Pr[A\,|\,\cdot\,]$ is the conditional probability of event A.

2. Why active learning under structural model uncertainty?

Consider the finite-horizon problem of minimizing deviation from an unstable setpoint $x^* = 0$ with minimal use of input energy for a scalar, stochastic nonlinear system. Two candidate models M_0 and M_1,

$$M_0 : x_{k+1}^{[0]} := 2.25\left(x_k^{[0]}\right)^2 - 3.5x_k^{[0]} + 3 + u_k + w_k^{[0]}, \tag{1a}$$

$$M_1 : x_{k+1}^{[1]} := -0.375\left(x_k^{[1]}\right)^3 + 2.25\left(x_k^{[1]}\right)^2 - 3.5x_k^{[1]} + 3 + u_k + w_k^{[1]}, \tag{1b}$$

both predict the behavior of the system near the setpoint x^*, and in both models the state is measured directly without error. In (1), $x_k^{[i]} \in \mathbb{R}$, $u_k \in \mathbb{R}$, and $w_k \in \mathbb{R}$ denote the state, the control input, and the zero-mean Gaussian white system disturbance at time k for model $i = 0, 1$ (indicated with superscript $[i]$). One of the two models is an exact representation of the system dynamics, but it is not known which model is "true." However, the initial (or *prior*) probability of M_0 being true is known and denoted by $P_{M,0}^{[0]}$ (with $P_{M,0}^{[1]} = 1 - P_{M,0}^{[0]}$). The probabilities of each model being the exact, or the best, representation of the system can be recursively updated using Bayes' theorem with each observation of the state x_k, as discussed in Sections 3 and 4. Fig. 1 illustrates the expected successor states of the two models and suggests that in the presence of the disturbance, M_0 and M_1 are not distinguishable around the setpoint x^*.

We consider a sequence of control laws $\mu_{0:N-1}^* = \{\mu_0, \mu_1, \dots, \mu_{N-1}\}$ such that $u_k = \mu_k(x_k, P_{M,k}^{[0]})$, with N being the length of the finite horizon. Given an initial state x_0 and a prior model probability $P_{M,0}^{[0]}$, we define the cost of using the control-law sequence $\mu_{0:N-1}$ as

$$J\left(x_0, P_{M,0}^{[0]}\right) = \mathrm{E}\left[\sum_{k=0}^{N-1} qx_k^2 + ru_k^2 + qx_N^2\right], \tag{2}$$

where q and r are positive tuning parameters and the expectation is taken over the model probabilities and the probability distributions of x_k and w_k. The optimal sequence of control laws $\mu_{0:N-1}^*$

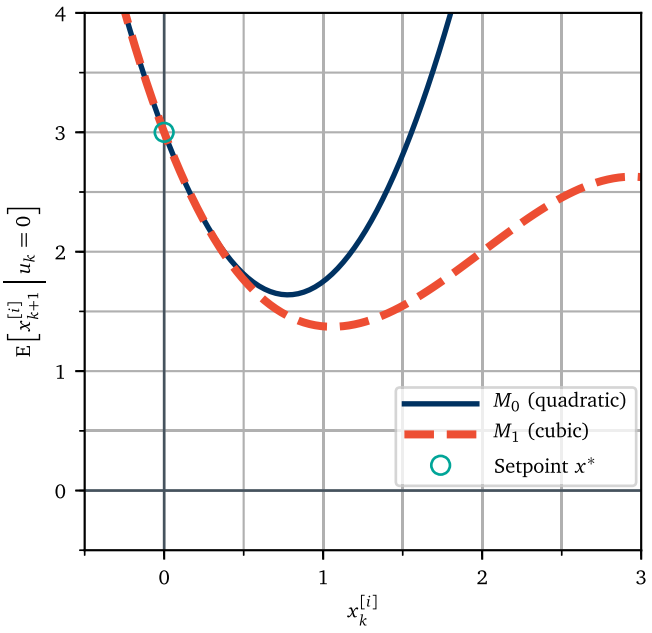

Fig. 1. Expected value of the next state $x_{k+1}^{[i]}$, given the current state $x_k^{[i]}$ and the control input $u_k = 0$ for the two models in (1).

minimizes this cost. The corresponding optimal cost, incurred by using $\mu_{0:N-1}^*$, is the solution to the optimal control problem

$$J^*\left(x_0, P_{M,0}^{[0]}\right) = \min_{\mu_{0:N-1}} J\left(x_0, P_{M,0}^{[0]}\right). \tag{3}$$

This optimal cost is also known as the *cost to go* N steps along the optimal trajectory that starts at $(x_0, P_{M,0}^{[0]})$. Solving (3) for $\mu_{0:N-1}$ numerically with dynamic programming results in a numerical representation of $\mu_{0:N-1}^*$. Dynamic programming involves formulating (3) as a Bellman equation, which relies on the principle of optimality (Bellman, 1957), and iterating backward in time from $k = N-1$ to $k = 0$. Since the state, control, disturbance, and model probability all are continuous, numerical solution involves quantizing these variables.

The model probability $P_{M,k+1}^{[0]}$ can be written as a function of $P_{M,k}^{[0]}$ and x_{k+1} using Bayes' theorem. That is, the evolution of $P_{M,k}^{[0]}$ can be predicted together with x_k, and both quantities depend on the control input u_{k-1}. Future uncertainty in the model structure, as represented by $P_{M,k}^{[0]}$, can thus be reduced in a predictable manner using the control input. Accounting for this dual effect when solving (3), which means treating both x_k and $P_{M,k}^{[0]}$ as dynamic quantities, enables active learning. We use $J_{\mathrm{A}}^*(x_0, P_{M,0}^{[0]})$ to denote the optimal cost obtained with this active-learning approach.

Conversely, solving (3) without including the equation that governs the evolution of $P_{M,k}^{[0]}$ as part of the dynamic model leads to a control law that does not anticipate that future information will influence the uncertainty. Implicitly, this control law assumes that the model probabilities will not change, or that $P_{M,k}^{[0]} = P_{M,k+k}^{[0]} \ \forall\ k \in \mathbb{Z}_{[0,N-1]}$. The control law will therefore not lead to active learning. It does, however, permit *passive* learning, since the model probabilities can be updated at every time step using the latest observation of x_k. Accordingly, using the optimal policy that results from solving (3) in this manner has the corresponding optimal cost $J_{\mathrm{P}}^*(x_0, P_{M,0}^{[0]})$.

The key question here is whether the expected cost (2) that results from using the control law with active learning, $J_{\mathrm{A}}^*(x_0, P_{M,0}^{[0]})$, is significantly different from the cost incurred when using passive learning, $J_{\mathrm{P}}^*(x_0, P_{M,0}^{[0]})$. Before a quantitative analysis of this question,

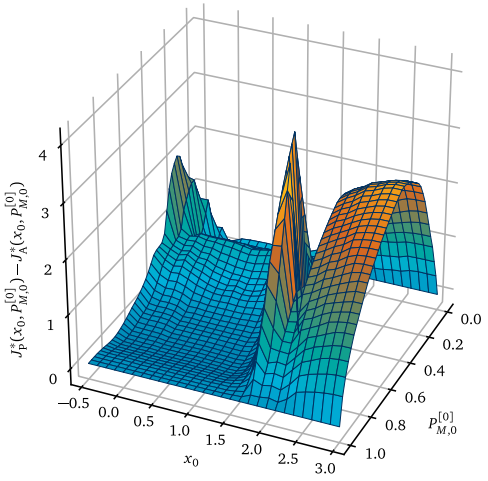

Fig. 2. Expected benefit of active learning. The figure shows the scaled difference in optimal expected cost with active and passive learning for the example problem in Section 2 with a horizon length of $N = 20$.

consider the following. First, solving $\mathrm{E}[x_{k+1}^{[i]}] = x_k^{[i]} = x^*$ for the control input gives the steady-state value $u^* = -3$ with both models. In model M_1, however, the steady-state control $u^0 = 0$ results in $\mathrm{E}[x_{k+1}^{[1]}] = x_k^{[1]} = 2$. Assume that q and r are such that operating near $x = 2$ results in a lower cost than operating near $x = x^* = 0$ if M_1 is true. The value of knowing whether the system evolves according to Models M_0 or M_1 is then high, and the optimal control law may involve probing the system to learn which of the two hypotheses is true. Qualitatively, Fig. 1 shows that close to the setpoint $x^* = 0$, the expected next state of the system is near identical regardless of which model hypothesis is true. A control policy that keeps the state in the vicinity of x^* is therefore unlikely to generate information that facilitates fast learning. It may be necessary to purposefully move the state away from x^* to speed up the learning, but this probing may come at a high cost in terms of input energy and deviation from x^*. This question of whether the cost of learning through probing is likely to generate system knowledge that the controller can use to lower the control cost is in general difficult to answer.

To quantify the improvement in cost by using control with active instead of passive learning, consider the optimal costs $J_{\mathrm{A}}^*(x_0, P_{M,0}^{[0]})$ and $J_{\mathrm{P}}^*(x_0, P_{M,0}^{[0]})$ introduced above. The benefit from active over passive learning is the difference $J_{\mathrm{P}}^*(x_0, P_{M,0}^{[0]}) - J_{\mathrm{A}}^*(x_0, P_{M,0}^{[0]})$. Fig. 2 shows this difference, scaled by the horizon length $N = 20$, for the exact numerical solutions obtained with dynamic programming. Note that the expected benefit of active learning is positive for almost the entire x_0–$P_{M,0}^{[0]}$ space. The expected benefit depends on the state and the model probabilities and is highly nonlinear in x_0 and $P_{M,0}^{[0]}$. A key insight from Fig. 2 is that there are regions of the x_0–$P_{M,0}^{[0]}$ space in which active learning has a very large expected benefit, and that in other areas there is relatively little or no expected benefit. The benefit is largest for $x_0 \geq 1.5$; the advantage of knowing the true model in this region is significant since M_0 is unstable and much more expensive to control than M_1 (see Fig. 1). When there is a high degree of certainty in which model is true, i.e., $P_{M,0}^{[0]}$ is close to zero or one, there is little benefit to active learning.

This paper considers the type of control problem discussed in this section: optimal control of stochastic nonlinear systems under structural model uncertainty using a set of model hypotheses. Since the dynamic programming approach used to solve this example problem is intractable even for moderately-sized systems, the paper presents the development of a tractable receding-horizon

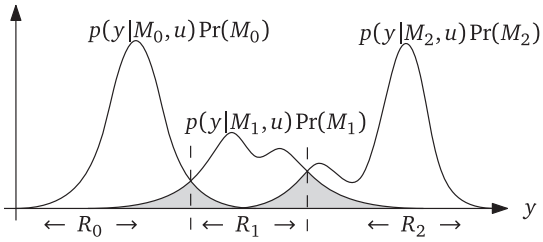

Fig. 3. Illustration of the Bayesian hypothesis selection problem (after Blackmore and Williams, 2006).

strategy with active learning for reducing structural uncertainty. Not only can the proposed algorithm learn faster, and with higher confidence, which of the models that best predicts the data, it can also improve control performance relative to robust and passive-learning MPC approaches. Before a formal presentation of the problem formulation, the next section presents a brief overview of hypothesis testing, which is central to the problem of control with multiple models.

3. Preliminaries and problem statement

3.1. Model hypothesis selection

Consider a finite set of $n_\mathrm{m}+1$ model hypotheses $\mathcal{M} := \{M_0, M_1, \ldots, M_{n_\mathrm{m}}\}$ and a prior probability $\Pr(M_i)$ for each model. The decision-theory problem of model discrimination involves classifying the measured output y, resulting from the input u, as predictions of one of the models in $\mathcal{M}$. Note that we here use u and y to denote general inputs and outputs; these variables may represent sequences such as $u_{a:b}$ and $y_{c:d}$ or the single-sample values like u_k and y_{k+1}. The Bayesian decision rule for model hypothesis selection minimizes the risk of misclassification, that is, the risk of selecting an incorrect hypothesis given u and y (Hellman and Raviv, 1970). The Bayesian decision rule can be formulated

$$\text{select } M_{i^*} \text{ such that } i^* = \arg\max_i \Pr(M_i \,|\, y, u), \tag{4}$$

where $\Pr(M_i \,|\, y, u) = p(y \,|\, M_i, u)\Pr(M_i)/p(y \,|\, u)$ by Bayes' theorem. Here, $p(y \,|\, M_i,\ u)$ is the likelihood function for observing y given hypothesis M_i and the input u and $p(y \,|\, u)$ is a normalization constant commonly referred to as the model evidence or the marginal likelihood.

Define a region R_i such that the decision rule (4) selects hypothesis M_i when the system observation y falls into R_i; that is,

$$R_i := \left\{ y : p(y \,|\, M_i, u) > p(y \,|\, M_j, u) \ \forall\ j \neq i \right\}.$$

Fig. 3 illustrates an example-set of regions and shows that $\Pr(y \in R_j,\ M_i \,|\, u)$ quantifies the probability of erroneously selecting the model hypothesis M_j when M_i is the model that best predicts the data. Hence, the probability of hypothesis-selection error Pr(error), also known as the *Bayes risk*, is (Duda et al., 2002)

$$\Pr(\text{error}) := \sum_{i=0}^{n_\mathrm{m}} \sum_{j=i+1}^{n_\mathrm{m}} \int_{R_j} p(y \,|\, M_i, u)\Pr(M_i) dy. \tag{5}$$

Eq. (5) reveals that Pr(error) depends on the input u. As evident from Fig. 3, inputs that reduce the overlap between the distributions $p(y \,|\, M_i,\ u)$ decrease Pr(error) and thus facilitate model discrimination. Note that for arbitrary distributions and in higher-dimensional space, the regions R_i may be difficult or impossible to determine (see, e.g., Duda et al., 2002). The key property of the proposed MPC algorithm developed in this paper is reduction of structural model uncertainty for a nonlinear stochastic system through reducing Pr(error) while simultaneously controlling the system.

3.2. Problem formulation

Consider a stochastic system $\mathcal{S}$ with discrete-time dynamics

$$\mathcal{S}: \begin{cases} x_{k+1} = f_k(x_k, u_k, w_k), \\ y_k = h_k(x_k, v_k), \end{cases} \tag{6}$$

where k is the time index; $x_k \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$, and $y_k \in \mathbb{R}^{n_y}$ are the system state, control input, and output, respectively; $w_k \in \mathbb{R}^{n_w}$ is the system disturbance; $v_k \in \mathbb{R}^{n_v}$ is the measurement noise; and $f_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \to \mathbb{R}^{n_x}$ and $h_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_y}$ describe the (possibly) nonlinear system dynamics and the measurements, respectively. Because of the structural uncertainty in modeling f_k and h_k and the potential changes in the dynamics over time, we consider a finite set of models denoted by $\mathcal{M} := \{M_0, M_1, \ldots, M_{n_\mathrm{m}}\}$. Here, each model M_i represents a hypothesis for describing the system $\mathcal{S}$ with dynamics of the form

$$M_i: \begin{cases} x_{k+1}^{[i]} = f^{[i]}\left(x_k^{[i]}, u_k, \theta^{[i]}, w_k^{[i]}\right), \\ y_k^{[i]} = h^{[i]}\left(x_k^{[i]}, \theta^{[i]}, v_k^{[i]}\right). \end{cases} \tag{7}$$

Each of the models may have uncertain parameters $\theta^{[i]}$, and the superscript $[i]$ is used to distinguish the functions and variables associated with each model $M_i \in \mathcal{M}$. Note that the functions and variables associated with the true system $\mathcal{S}$ in (6) have no superscript $[i]$.

The sequences $w_k^{[i]}$ and $v_k^{[i]}$ are both independent realizations of identically distributed random variables with *known* probability distributions $p(w^{[i]})$ and $p(v^{[i]})$. The initial condition $x_0^{[i]}$ and parameter vector $\theta^{[i]}$ in model i are unknown, and are modeled as time-invariant probabilistic uncertainties with known joint PDF $p(x_0^{[i]}, \theta^{[i]})$. A hypothesis M_i contains all the information necessary for specifying model i;

$$M_i := \left\{ f^{[i]}, h^{[i]}, p\left(x_0^{[i]}, \theta^{[i]}\right), p\left(w_k^{[i]}\right), p\left(v_k^{[i]}\right) \right\}. \tag{8}$$

The true system $\mathcal{S}$ is not necessarily contained in the set of model candidates $\mathcal{M}$. All models in $\mathcal{M}$ and the system $\mathcal{S}$ have the same control inputs u_k and measurements y_k; that is, the respective dimensions of these signals, n_u and n_y, are identical across all models in $\mathcal{M}$. Note that the numbers of states, unknown parameters, and stochastic disturbance and noise elements may vary between models and the system $\mathcal{S}$.

The control inputs u_k in (6) are constrained to a polytope U:

$$u_k \in U := \{u_k \,|\, A_u u_k \leq b_u\}, \tag{9}$$

where $A_u \in \mathbb{R}^{n_\mathrm{c} \times n_u}$, $b_u \in \mathbb{R}^{n_\mathrm{c}}$, and n_c is the number of input constraints. The state constraints are specified as a set of n_p inequality constraints,

$$g^{[i]}\left(x_k^{[i]}\right) \leq 0, \qquad \forall\, i \in \{1, \ldots, n_\mathrm{m}\}, \tag{10}$$

where $g^{[i]} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_\mathrm{p}}$ defines the inequalities.[1] However, since the system evolves as a stochastic process that is observed through noisy measurements, the constraints (10) cannot be enforced directly. We implement these constraints probabilistically in the proposed MPC, as discussed below.

As demonstrated in the motivating example in Section 2, active learning under model uncertainty can be a central component in optimal control. The optimal control problem we consider is thus twofold: (i) regulating the system dynamics, and (ii) active learning that reduces the structural model uncertainty. Inspired

[1] If the number of state constraints n_p vary between the models, which may be the case if the models have different state dimension n_x, $g^{[i]}$ is defined accordingly for every model.

by *explicit* methods for dual control with parametric uncertainty (Wittenmark, 1995; Filatov and Unbehauen, 2000; Mesbah, 2018), this paper formulates an explicit cost for the probability of future model-selection error that is minimized along with the stage and terminal costs. Including the cost of reducing the structural uncertainty in the optimal control problem (or OCP) enables probing of the system to determine which model hypothesis in (8) that best describes the observed closed-loop dynamics.

For the stochastic nonlinear system $\mathcal{S}$ in (6) with structural model uncertainty, we state the OCP considered in this paper formally as follows.

Problem 1 (Optimal control of stochastic systems with active learning under structural model uncertainty)**.** At sampling time k, determine the optimal input sequence $u^*_{k:k+N-1}$ for the system $\mathcal{S}$ in (6) that minimizes a cost function consisting of the terminal cost $\ell_N(x_N)$, the stage cost $\ell_k(x_k, u_k)$ over the horizon $\mathbb{Z}_{[k,\,k+N-1]}$, and the probability of model-selection error Pr(error) in (5), given the model set $\mathcal{M}$ in (8), while satisfying the state constraints (10) with specified confidence, and the input constraints (9).

The solution to the OCP in Problem 1 involves a balance between controlling the system based on current knowledge and reducing the structural uncertainty at every sampling time k to improve control performance. Solving Problem 1 on a receding horizon thus leads to an MPC strategy with active learning. The MPC applies only the first element of the optimal input sequence, $u^*_{k|k}$, to the system $\mathcal{S}$ at every discrete time k and re-solves the optimal control problem at the subsequent sampling time incorporating new information from the system measurements.

Remark 1. Problem 1 does not specify how the terminal and stage costs, the state constraints, and probability of model-selection error are evaluated. For example, this general formulation permits implementation choices such as whether to enforce the state constraints for all models in the set $\mathcal{M}$ or only the one that has the highest probability at any given time. The choices made in this paper are discussed in the following section.

The goal of this paper is to present a tractable formulation for MPC with active learning for structural model uncertainty. We discuss some of the main considerations in receding-horizon implementation of the OCP in Problem 1 in the following.

Computation of the probability of model-selection error: Evaluating the probability of model-selection error Pr(error) in (5) requires computing multivariate integrals over the probability distribution $p(y_{k+1:k+N} \,|\, M_i, u_{k:k+N-1})$ over the decision regions depicted in Fig. 3. In general, no closed-form expression for $p(y_{k+1:k+N} \,|\, M_i, u_{k:k+N-1})$ exists, rendering exact evaluation of Pr(error) intractable (Blackmore and Williams, 2006). The primary challenge in solving the OCP in Problem 1 is deriving a computationally tractable approximation for Pr(error) given the model set $\mathcal{M}$.

Online estimation of model probabilities: Receding-horizon implementation of Problem 1 also requires online estimation of the conditional model probabilities $\Pr(M_i \,|\, y_k)$ at every sampling time k to update the probability of model-selection error Pr(error) based on the output y_k. This problem is in its general form referred to as multiple-model estimation (Ackerson and Fu, 1970). This is an intractable problem that suffers from exponential growth in complexity over time; see Li and Bar-Shalom (1996) for an overview and discussion of different solution approaches.

Uncertainty propagation: Solving Problem 1 requires efficient propagation of the probabilistic model uncertainty in x_0 and θ as well as the stochastic disturbances w_k and noise v_k through the nonlinear models $M_i \in \mathcal{M}$. The probability distribution of the model outputs $p(y_{k+1:k+N} \,|\, M_i, u_{k:k+N-1})$, or their statistical moments, are required for evaluating the cost function of the OCP. The probabilistic state information can also be used to enforce the state constraints (10) with a specified probability of satisfaction. In general, however, joint propagation of parametric uncertainty and stochastic noise poses a significant challenge to optimal control for nonlinear systems (Paulson and Mesbah, 2017). The widely-used random-sampling approaches to uncertainty propagation (Cowles and Carlin, 1996; Caflisch, 1998; Kantas et al., 2009) are primarily developed for propagating one source of uncertainty (i.e., either time-invariant parametric uncertainty or time-varying stochastic disturbances) and can be computationally prohibitive for online optimization. Moreover, these methods are generally ill-suited for gradient-based optimization since they do not lend themselves to accurate and precise gradient computation.

The proposed MPC strategy involves approximating the probability distribution $p(y_{k+1:k+N} \,|\, M_i, u_{k:k+N-1})$ in terms of its moments. As shown in the next section, this enables approximating the OCP in Problem 1 with a deterministic surrogate that is amenable to online solution. Established methods for moment-based approximation of probability distributions include *linearization* (Ljung, 1979) and the *unscented transform* (Julier and Uhlmann, 1997; 2004). For conceptual and notational simplicity, we here approximately solve Problem 1 through approximating the first two moments of the probability distributions using linearization. Other methods for approximating $p(y_{k+1:k+N} \,|\, M_i, u_{k:k+N-1})$, such as the unscented transform, are equally well-suited for our proposed approach. The unscented transform is particularly useful when the system dynamics exhibit a high degree of nonlinearity, or when computation of the model Jacobians is computationally prohibitive.

4. Approximate solution to optimal control with active learning

This section first presents the methods used to address the challenges associated with solving the OCP in Problem 1 as discussed above. We subsequently present an algorithm for MPC with active learning for reducing structural model uncertainty based on a deterministic surrogate for Problem 1.

4.1. Joint propagation of probabilistic model uncertainty and stochastic disturbances

The uncertainty propagation approach adopted in this work is based on augmenting the state vector $x_k^{[i]}$ with the unknown time-invariant parameters $\theta^{[i]}$ in the model Eq. (7). Omitting the model-index superscript $[i]$, denote the augmented state vector by $z_k = [x_k^\top,\, \theta_k^\top]^\top$, where the parameters θ_k have constant dynamics $\theta_{k+1} = \theta_k$. The augmented state vector then evolves according to

$$z_{k+1} = f_z(z_k, u_k, w_k) = \begin{bmatrix} f(x_k, u_k, \theta_k, w_k) \\ \theta_k \end{bmatrix}. \tag{11a}$$

Similarly, the output equation takes the form

$$y_k = h_z(z_k, v_k), \tag{11b}$$

where $f_z : \mathbb{R}^{n_z} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \to \mathbb{R}^{n_z}$ and $h_z : \mathbb{R}^{n_z} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_y}$, $n_z = n_x + n_\theta$, and $p(z_0) = p(x_0, \theta)$ is the joint probability distribution of the initial augmented state.

There exists no closed-form method for propagation of probabilistic uncertainties through a general nonlinear system (Chen, 2003). This work relies on propagating the moments of the probability distributions of the augmented state and output in (11). The moment-based method used here for joint propagation of probabilistic model uncertainty and stochastic disturbances is based on linearization of the augmented system dynamics (11) and approximation of the probability distributions of the state and

output as multivariate Gaussian distributions. Linearization is a well-established method for uncertainty propagation in recursive Bayesian estimation, resulting in the extended Kalman filter, or EKF, for state estimation in nonlinear systems (e.g., Ljung, 1979; Simon, 2006).

A first-order Taylor series approximation of the augmented state equations in (11a) with respect to the state and disturbance results in

$$z_{k+1} \approx f_z\big(\hat{z}_k, u_k, \bar{w}\big) + A_k\big(z_k - \hat{z}_k\big) + E_k(w_k - \bar{w}), \tag{12}$$

where

$$A_k := \left.\frac{\partial f_z}{\partial z_k}\right|_{\hat{z}_k, u_k, \bar{w}}, \quad E_k := \left.\frac{\partial f_z}{\partial w_k}\right|_{\hat{z}_k, u_k, \bar{w}},$$

$\hat{z}_k \approx \bar{z}_k = \mathrm{E}[z_k]$ is the approximated mean of the augmented state, and $\bar{w} = \mathrm{E}[w_k]$. The evolution of $\hat{z}_k$ is described by

$$\hat{z}_{k+1} = f_z(\hat{z}_k, u_k, \bar{w}), \tag{13}$$

starting from $\hat{z}_0 = \bar{z}_0$, which has a known distribution. Hence, (12) can be written in terms of the deviation variables $\Delta z_k = z_k - \hat{z}_k$ and $\Delta w_k = w_k - \bar{w}$, leading to the linear time-varying system

$$\Delta z_{k+1} = A_k \Delta z_k + E_k \Delta w_k. \tag{14}$$

Note that since we do not linearize the augmented state equations in (11a) with respect to u_k, the input enters (14) directly through Δz_{k+1}, which is a function of $f_z(\hat{z}_k, u_k, \bar{w})$. The covariance of the augmented state, denoted by Σ_{z_k}, can now be approximated as $\hat{\Sigma}_{z_k} = \mathrm{E}[\Delta z_k \Delta z_k^\top]$. Using (14), we arrive at the Lyapunov-type recursive expression

$$\hat{\Sigma}_{z_{k+1}} = A_k \hat{\Sigma}_{z_k} A_k^\top + E_k \Sigma_w E_k^\top, \quad \text{with} \quad \hat{\Sigma}_{z_0} = \Sigma_{z_0}, \tag{15}$$

for the evolution of the approximate covariance. Similarly, the approximate mean and covariance of the model output in (11b) evolve as

$$\hat{y}_k = h_z\big(\hat{z}_k, \bar{v}\big), \tag{16a}$$

$$\hat{\Sigma}_{y_{k+1}} = C_k \hat{\Sigma}_{z_k} C_k^\top + M_k \Sigma_v M_k^\top, \tag{16b}$$

where

$$C_k := \left.\frac{\partial h_z}{\partial z_k}\right|_{\hat{z}_k, \bar{v}}, \quad M_k := \left.\frac{\partial h_z}{\partial v_k}\right|_{\hat{z}_k, \bar{v}}.$$

To predict the future evolution of the first two moments of the augmented state and output in (11) over the prediction horizon k to $k+N$ while capturing the cross-correlation of moments in time, we collect the moment propagation equations for all future sampling times into one set of equations. The predicted sequence of augmented state deviations from time k to $k+N$ can be written as $\Delta z_{k:k+N} = A\Delta z_k + E\Delta w_{k:k+N}$, where

$$A = \begin{bmatrix} I \\ A_1 \\ A_1A_2 \\ \vdots \\ \prod_{j=1}^{N-1} A_j \\ \prod_{j=1}^{N} A_j \end{bmatrix},$$

$$E = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ E_0 & 0 & & & 0 \\ A_1E_0 & E_1 & \ddots & & \vdots \\ A_2A_1E_0 & A_2E_1 & E_2 & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ \prod_{j=1}^{N-1} A_jE_0 & \prod_{j=2}^{N-1} A_jE_1 & \cdots & A_{N-1}E_{N-2} & E_{N-1} \end{bmatrix}.$$

Note that each A_j and E_j both depend on the predicted input $u_{k+j|k}$ from the sequence $u_{k:k+N}$. The covariances of the augmented state from time k to $k+N$ evolve as

$$\hat{\Sigma}_{z_{k:k+N}} = A\hat{\Sigma}_{z_k}A^\top + E\,\mathrm{diag}\,(\Sigma_w, \ldots, \Sigma_w)E^\top.$$

Similarly, using (16), we collect the equations for predicting the output moments over the prediction horizon k to $k+N$ into

$$\Delta y_{k:k+N} = C\Delta y_k + M\Delta v_{k:k+N}, \tag{17a}$$

$$\hat{\Sigma}_{y_{k:k+N}} = C\hat{\Sigma}_{z_{k:k+N}}C^\top + M\,\mathrm{diag}(\Sigma_v, \ldots, \Sigma_v)M^\top, \tag{17b}$$

where $C = \mathrm{diag}(C_j, \ldots, C_N)$ and $M = \mathrm{diag}(M_j, \ldots, M_N)$. Note that the symmetric matrix $\hat{\Sigma}_{y_{k:k+N}}$ in (17b) captures the cross-correlation of the predicted output covariances in time.

4.2. Tractable approximation of the probability of model-selection error

For the nonlinear setting considered in this work, there exists no closed-form expression for evaluating the Bayes risk of hypothesis-selection error Pr(error) in the cost function of the OCP in Problem 1. Assuming that the output of the model hypotheses in the set $\mathcal{M}$ has a multivariate Gaussian distribution, an upper bound (Matusita, 1971; Blackmore and Williams, 2006) can be derived for Pr(error). Given the information available at time k, let $P_{M,k}^{[i]} := \Pr(M_i \,|\, y_{0:k}, u_{0:k-1})$ be the probability of model M_i being the best, or most appropriate, representation of the system. For the $n_\mathrm{m}+1$ model hypotheses in $\mathcal{M}$, Pr(error) can be bounded as

$$\Pr(\text{error}) \le P_\mathrm{UB},$$

where the bound P_UB over the prediction horizon $k+1$ to $k+N$ is defined in terms of pairwise *Bhattacharyya bounds* as

$$P_\mathrm{UB} := \sum_{i=0}^{n_\mathrm{m}} \sum_{j=i+1}^{n_\mathrm{m}} \sqrt{P_{M,k}^{[i]} P_{M,k}^{[j]}} \times \exp\left(-d_\mathrm{B}\left(y_{k+1:k+N}^{[i]}, \Sigma_{y_{k+1:k+N}}^{[i]}, y_{k+1:k+N}^{[j]}, \Sigma_{y_{k+1:k+N}}^{[j]}\right)\right) \tag{18a}$$

with

$$d_\mathrm{B}\big(y^{[i]}, \Sigma^{[i]}, y^{[j]}, \Sigma^{[j]}\big) := \frac{1}{4}\big(y^{[i]} - y^{[j]}\big)^\top \big[\Sigma^{[i]} + \Sigma^{[j]}\big]^{-1}\big(y^{[i]} - y^{[j]}\big) + \frac{1}{2}\ln\left(\frac{|\Sigma^{[i]} + \Sigma^{[j]}|}{2\sqrt{|\Sigma^{[i]}|\,|\Sigma^{[j]}|}}\right). \tag{18b}$$

The quantity d_B is known as the *Bhattacharyya distance* between the predicted distributions of the output of models i and j, assuming that the distributions are Gaussian and thus described by the predicted mean and covariance of the output y; see (17). We here use the bound (18a) as a tractable surrogate for the probability of model-selection error Pr(error) in the cost function of Problem 1. While the bound P_UB is not guaranteed to hold in the case of nonlinear models or non-Gaussian distributions, the goal of the proposed approach is not to quantify the probability of model-selection error. Rather, the goal is to lower this probability to the extent that it benefits control performance in expectation. If reducing the bound P_UB does not sufficiently improve the control performance, there is a variety of metrics that quantify the overlap or

similarity between distributions that can be employed. Gibbs and Su (2002) provide an extensive overview of various such metrics, including the Hellinger distance, the Kullback-Leibler divergence, the Kolmogorov metric, and the Wasserstein metric. However, these metrics do not directly correspond to (a bound on) the probability of selecting the wrong hypothesis, the reduction of which is the objective in the proposed MPC strategy.

4.3. Model probability estimation and model selection

As evident from (18), the bound P_{UB} depends on the model probabilities $P_{M,k}^{[i]}$. Thus, online estimation of the model probabilities is important for evaluating P_{UB} when solving the OCP in Problem 1 on a receding horizon. We use Bayesian recursion to estimate $P_{M,k}^{[i]}$ for each of the model hypothesis in the set $\mathcal{M}$ given the measurements y_k.

Let the prediction error of each model M_i be defined as

$$e_k^{[i]} = y_k - \mathrm{E}\left[x_k^{[i]} \mid y_{0:k-1}, u_{0:k-1}, M_i\right]. \tag{19}$$

When the model outputs are sufficiently separated, the model prediction error $e_k^{[i]}$ is small for the model i that currently best predicts the data generated by the true system $\mathcal{S}$, relative to the prediction error for the other models; cf. Fig. 1. That is, the error (19) is a measure of how well each model predicts the observed system outputs at time k. The prediction errors $e_k^{[i]}$ can be used to estimate the model probabilities $P_{M,k}^{[i]}$ recursively using Bayes' theorem (e.g., see Simon, 2006). The likelihood of measuring y_k under model hypothesis i is proportional to the probability of observing y_k conditioned on model M_i and the past measurements $y_{0:k-1}$; i.e.,

$$L_k^{[i]} \propto p(y_k \mid y_{0:k-1}, M_i).$$

Determining the density on the right-hand side of the above expression involves evaluation of multivariate integrals for which no closed-form solution exists for general nonlinear model equations. In this work, we approximate the likelihood $L_k^{[i]}$ in terms of the first two moments of y_k, such that the right-hand side density takes the form of a multivariate Gaussian distribution. That is, we approximate the likelihood as

$$p(y_k \mid y_{0:k-1}, M_i) \approx \exp\left(-\frac{1}{2}\left(e_k^{[i]}\right)^\top \Sigma_\nu^{-1} e_k^{[i]}\right). \tag{20}$$

Eq. (20) incorporates that the components of y_k with large measurement noise covariance have a relatively smaller contribution to the likelihood. Based on the likelihood of observing the measurements under each of the model hypotheses, we can now use Bayes' theorem to estimate the probability of each of the models best predicting the data. Let $P_{M,k}^{[i]}$ and $P_{M,k-1}^{[i]}$ denote the posterior and prior probabilities of model i. The Bayesian recursion approximates the posterior probability of each model i as

$$P_{M,k}^{[i]} = \frac{P_{M,k-1}^{[i]} L_k^{[i]}}{\sum_{j=0}^{n_{\text{m}}} P_{M,k-1}^{[j]} L_k^{[j]}}. \tag{21}$$

With the model probabilities determined, the Bayesian decision rule selects the model i_k^* that has the largest posterior probability at time k; i.e.,

$$i_k^* := \arg\max_i P_{M,k}^{[i]}.$$

At times when there are two or more models that have the highest and near-highest probabilities, using i_k^* for control can lead to frequent switching between models. To avoid this behavior, the proposed MPC strategy chooses the control model i_k^{M} according to

$$i_k^{\text{M}} = \begin{cases} i_k^* & \text{if } i_k^* = i_{k-j}^* \ \forall\ j \in \{1, 2, \ldots, N_{\text{S}}\}, \\ i_{k-1}^{\text{M}} & \text{otherwise.} \end{cases} \tag{22}$$

Here, N_{S} is the number of sampling times the most probable model i_k^* must remain the same for it to be selected as the control model. The value of N_{S} must be chosen on a case-by-case basis, with particular attention to the variance of the stochastic disturbances and noise as well as the tuning of the state estimators.

Remark 2. When a sequence of system observations is significantly different from the predictions of a particular model hypothesis, the corresponding model probability quickly approaches zero. However, note that if the scenario corresponding to that model hypothesis subsequently occurs, the model probability would increase very slowly despite high likelihoods, since the likelihood is multiplied by a near-zero prior probability. Thus, the controller can only adapt slowly to the change in scenario, potentially leading to degradation of the control performance. To avoid this issue, we introduce a lower bound $P_{\min}$ on the probabilities $P_{M,k}^{[i]}$, similar to the approach of Aufderheide and Bequette (2003), and renormalize to ensure that $\sum_{i=0}^{n_{\text{m}}} P_{M,k}^{[i]} = 1$.

Remark 3. We here make the conventional choice (Murray-Smith and Johansen, 1997; Narendra and Han, 2011) of using only one model i_k^{M} for evaluating the control objective and the state constraints (discussed below) and switching between the models as appropriate. Alternatively, one can use two or more models in determining the control cost and in constraint enforcement, for instance by weighting the output variables (Kuure-Kinsey and Bequette, 2010) and constraints by the model probabilities used in their evaluation. Using such a weighted approach can enable smoother control profiles, but it can also result in more conservative performance as well as increased computational cost owing to a larger number of constraints.

4.4. State and parameter estimation

Our proposed approach to MPC with active learning does not depend on a specific type of estimation algorithm. Hence, both for simplicity of presentation and for its symmetry with the proposed propagation approach presented in Section 4.1, we here use the EKF over more involved alternatives such as the unscented Kalman filter (UKF; Julier and Uhlmann, 2004), particle-based filters (e.g., Arulampalam et al., 2002), and moving-horizon estimation (Robertson et al., 1996). The EKF uses the model Jacobians to propagate the mean and covariance of z_k using (13) and (15), respectively (Simon, 2006).[2] The prediction step of the EKF for every model M_i consists of

$$\hat{z}_{k|k-1}^{[i]} = f_z^{[i]}\left(\hat{z}_{k-1|k-1}^{[i]}, u_{k-1}, \bar{w}^{[i]}\right), \tag{23a}$$

$$\hat{\Sigma}_{z_{k|k-1}}^{[i]} = A_{k-1}^{[i]} \hat{\Sigma}_{z_{k-1|k-1}}^{[i]} \left(A_{k-1}^{[i]}\right)^\top + E_{k-1}^{[i]} \Sigma_w^{[i]} \left(E_{k-1}^{[i]}\right)^\top. \tag{23b}$$

Once the system measurements y_k are observed, the posterior distribution of z_k is updated according to

$$\hat{z}_{k|k}^{[i]} = \hat{z}_{k|k-1}^{[i]} + K_k^{[i]}\left(y_k - h_z^{[i]}\left(\hat{z}_{k|k-1}^{[i]}, \bar{v}^{[i]}\right)\right), \tag{24a}$$

$$\hat{\Sigma}_{z_{k|k}}^{[i]} = \hat{\Sigma}_{z_{k|k-1}}^{[i]} K_k^{[i]} \hat{\Sigma}_{y_{k|k-1}}^{[i]} K_k^{[i]}, \tag{24b}$$

where the Kalman gain is defined by

$$K_k^{[i]} = \hat{\Sigma}_{z_{k|k-1}, y_{k|k-1}}^{[i]} \left(\hat{\Sigma}_{y_{k|k-1}}^{[i]}\right)^{-1}.$$

[2] Alternatively, state and parameter estimation can be performed separately, for instance using a dual-estimation approach (see, e.g., Wan and Nelson, 1996.)

4.5. State constraint implementation

A rigorous treatment of various approaches to enforcing and implementing the state constraints (10) is beyond the scope of this paper. We here employ the approach of selecting a parameter vector $c \in \mathbb{R}^{n_p} \geq 0$ for constraint back-off (see, e.g., Farina et al., 2016; Heirung et al., 2017a; Paulson and Mesbah, 2018; Koller et al., 2018). The proposed MPC enforces the state constraints for the expected-value state predictions from the current control model i_k^{M} at time k. That is, we implement

$$g^{[i_k^{\mathrm{M}}]}\left(\hat{x}_{k+j+1|k}^{[i_k^{\mathrm{M}}]}\right) \leq -c, \quad \forall\ j \in \mathbb{Z}_{[0,N-1]}. \tag{25}$$

Note that since the predicted (augmented) state statistics are computed to evaluate P_{UB} (cf. Sections 4.1–4.2), specifying satisfaction probabilities for the individual constraints in (10) is a trivial extension of the proposed MPC strategy.

4.6. Algorithm for MPC with active learning for structural model uncertainty

We now formulate a deterministic surrogate for the OCP in Problem 1 using the methods discussed above. The control objective is evaluated over the horizon k to $k+N$ and consists of the bound $P_{\mathrm{UB}}(u_{k:k+N-1})$ in (18a), which provides a tractable approximation of the probability of model-selection error, as well as the stage cost $\ell_j\left(\hat{x}_{j|k}^{[i_k^{\mathrm{M}}]}, u_{j|k}\right)$ and the terminal cost $\ell_{k+N}\left(\hat{x}_{k+N|k}^{[i_k^{\mathrm{M}}]}\right)$, both for the nominal state trajectory predicted with the current control model $M^{[i_k^{\mathrm{M}}]}$.[3] That is, we define the control objective

$$\begin{aligned} J(u_{k:k+N-1}) = &\sum_{j=k}^{k+N-1} \ell_j\left(\hat{x}_{j|k}^{[i_k^{\mathrm{M}}]}, u_{j|k}\right) + \ell_{k+N}\left(\hat{x}_{k+N|k}^{[i_k^{\mathrm{M}}]}\right) \\ &+ r_{\mathrm{B}} P_{\mathrm{UB}}(u_{k:k+N-1}), \end{aligned} \tag{26}$$

where $r_{\mathrm{B}} \geq 0$ is a (user-specified) scalar weight. The first two terms in the objective (26) represent the performance predicted with the current control model chosen by the selection rule (22). The third term is the cost associated with model-selection error, the reduction of which induces a probing action for lowering the structural model uncertainty.

The tractable surrogate OCP for Problem 1 can now be stated as

$$\min_{u_{k:k+N-1}} \ J(u_{k:k+N-1}) \tag{27a}$$

subject to

$$\hat{z}_{k+j+1|k}^{[i]} = f_z^{[i]}\left(\hat{z}_{k+j|k}^{[i]}, u_{k+j|k}, \bar{w}^{[i]}\right), \qquad \forall\ j \in \mathbb{Z}_{[0,N-1]},\ \forall\ M_i \in \mathcal{M}, \tag{27b}$$

$$\hat{\Sigma}_{z_{k:k+N}}^{[i]} = A^{[i]} \hat{\Sigma}_{z_{k|k}}^{[i]} (A^{[i]})^\top + E^{[i]}\, \mathrm{diag}\left(\Sigma_w^{[i]}, \ldots, \Sigma_w^{[i]}\right)\left(E^{[i]}\right)^\top, \qquad \forall\ M_i \in \mathcal{M}, \tag{27c}$$

$$\hat{y}_{k+j|k}^{[i]} = h_z\left(\hat{z}_{k|k}^{[i]}, \bar{v}^{[i]}\right), \qquad \forall\ j \in \mathbb{Z}_{[1,N]},\ \forall\ M_i \in \mathcal{M}, \tag{27d}$$

$$\hat{\Sigma}_{y_{k:k+N}}^{[i]} = C^{[i]} \hat{\Sigma}_{z_{k:k+N}}^{[i]} \left(C^{[i]}\right)^\top + M^{[i]}\, \mathrm{diag}\left(\Sigma_v^{[i]}, \ldots, \Sigma_v^{[i]}\right)\left(M^{[i]}\right)^\top, \qquad \forall\ M_i \in \mathcal{M}, \tag{27e}$$

$$u_{k+j|k} \in U, \qquad \forall\ j \in \mathbb{Z}_{[0,N-1]}, \tag{27f}$$

$$g^{[i_k^{\mathrm{M}}]}\left(\hat{x}_{k+j+1|k}^{[i_k^{\mathrm{M}}]}\right) \leq -c, \qquad \forall\ j \in \mathbb{Z}_{[0,N-1]}, \tag{27g}$$

$$w_{k+j}^{[i]} \sim p_{w^{[i]}},\ v_{k+j}^{[i]} \sim p_{v^{[i]}},\ z_{k|k}^{[i]} \sim p_{z_k^{[i]}}, \qquad \forall\ j \in \mathbb{Z}_{[0,N-1]},\ \forall\ M_i \in \mathcal{M}, \tag{27h}$$

[3] For clarity, the dependence of P_{UB} on the input sequence is indicated explicitly here.

with the state estimate $\hat{z}_{k|k}^{[i]}$ and its covariance $\hat{\Sigma}_{z_{k|k}}^{[i]}$ given by the estimator equation (24) for all models $M_i \in \mathcal{M}$. Here, the constraints (27b) and (27c) predict the approximate state mean and covariance (cf. (13) and (15)) for each of the models $M_i \in \mathcal{M}$, whereas (27d) and (27e) predict the corresponding output means and covariances (cf. (16)); the constraints (27f) and (27g) specify the input and state constraints, the latter for the current control model i_k^{M} (cf. (25)); and (27h) specifies the given probability distributions for the disturbances, measurement noise, and augmented state.

The proposed MPC strategy with active learning for structural model uncertainty involves solution of the deterministic optimal control problem (27) at every measurement sampling time k on a receding horizon. This requires online estimation of the mean and covariance of the augmented state (i.e., $\hat{z}_{k|k}^{[i]}$ and $\hat{\Sigma}_{z_{k|k}}^{[i]}$), as well as the model probabilities $P_{M,k}^{[i]}$. We summarize the proposed MPC algorithm as follows.

0. Initialize at time $k = 0$: specify $p_{z_0^{[i]}}$ and $P_{M,0}^{[i]}$ for all models.
1. At time k, obtain measurements y_k and update $\hat{z}_{k|k}^{[i]}$ and $\hat{\Sigma}_{z_{k|k}}^{[i]}$ for all models using (24).
2. Update all model probabilities $P_{M,k}^{[i]}$ using (21), and use (22) to select i_k^{M}.
3. Solve the OCP (27) to obtain the optimal control input sequence $u_{k:k+N-1}^*$.
4. Implement the control input $u_{k|k}^*$.
5. Predict $\hat{z}_{k+1|k}^{[i]}$ and $\hat{\Sigma}_{z_{k+1|k}}^{[i]}$ for all models using (23).
6. Set $k \leftarrow k+1$ and go to step 1.

Remark 4. The number of variables in (27) depends on its implementation. A straight-forward formulation results in Nn_u control variables, $Nn_z n_{\mathrm{m}}$ augmented-state mean variables, $Nn_z(n_z+1)n_{\mathrm{m}}/2$ variables in the symmetric state covariance matrices, and $N(N+1)n_y(n_y+1)n_{\mathrm{m}}/4$ variables in the full symmetric output covariance matrix (17b). The largest number of variables thus comes from this last matrix, necessary to evaluate Pr(error).

5. Case study

We here apply our proposed approach to optimizing productivity in a bioreactor case study adapted from Agrawal et al. (1989) and Henson and Seborg (1992).

5.1. Problem description with multiple model hypotheses

In the bioreactor, biomass and a substrate react to form a product; their respective concentrations are X, S, and P. The control input is the dilution rate $u = D$ and the reactor volume is kept constant by ensuring the volumetric inlet and outlet flows are identical. In the true system and in the model hypotheses, the states $x = [X,\ S,\ P]^\top$ evolve according to discrete-time equations of the forms (6) and (7), respectively, with

$$f(x,u,w) = \begin{bmatrix} (-DX + \mu X)\Delta t + X \\ \left(D(S_{\mathrm{f}} - S) - \dfrac{1}{Y_{X/S}}\mu X\right)\Delta t + S \\ (-DP + (\alpha\mu + \beta)X)\Delta t + P \end{bmatrix} + w, \tag{28}$$

where $w = [w_X,\ w_S,\ w_P]^\top$ is a vector of independent realizations of zero-mean unit-variance Gaussian variables scaled by their respective standard deviations σ_X, σ_S, and σ_P, and Δt is the sample time. Further, S_{f} is the substrate concentration in the inlet feed, $Y_{X/S}$ is the yield of biomass per substrate consumed, α and β are the yield parameters for the production of P, and μ is the biomass growth rate. The substrate and product concentrations are both

Table 1
Nominal parameters and operating conditions of the continuous bioreactor (Agrawal et al., 1989; Henson and Seborg, 1992).

Variable	Nominal value	Unit
$Y_{X/S}$	0.4	g/g
α	2.2	g/g
β	0.2	h^{-1}
μ_{max}	0.48	h^{-1}
K_m	1.2	g/L
S_f	20	g/L
P_m	50	g/L
Δt	0.1	h
σ_X	0.090	g/L
σ_S	0.013	g/L
σ_P	0.112	g/L
σ_{v_S}	0.03	g/L
σ_{v_P}	1.00	g/L
P_{min}	0.01	–
N_S	5	–

measured;

$$y_k = [S_k,\ P_k]^\top + v_k,$$

where $v_k \in \mathbb{R}^2$ and $v_k \sim \mathcal{N}(0, \mathrm{diag}(\sigma_{v_S}^2, \sigma_{v_P}^2))$.

We consider three models: $\mathcal{M} = \{M_0,\ M_1,\ M_2\}$. The nominal M_0 model with saturation/monod kinetics; a model M_1 with saturation kinetics combined with a drop in the substrate inlet feed concentration, and a product inhibition model M_2, in which the growth rate decreases with the product concentration (Agrawal et al., 1989). The kinetic models can be interpreted as representing structural model uncertainty, with the following two growth hypotheses that may be valid under different conditions;

$$\mu = \frac{\mu_{max} S}{K_m + S} \qquad \text{in models } M_0 \text{ and } M_1,$$

and

$$\mu = \frac{\mu_{max}(1 - P/P_m)S}{K_m + S} \qquad \text{in Model } M_2,$$

where μ_{max} denotes the maximum growth rate, K_m is an affinity constant, and P_m is the maximum production rate. Model M_1 represents a fault causing a significant disturbance to the nominal operation of the bioreactor, which may be difficult to identify under regular feedback control:

$$S_f^{[1]} = (1 - 0.25)S_f.$$

This fault and the random fluctuations in substrate concentration share certain similarities, but the fault prevents acceptable operation and must be addressed through intervention.

We consider a scenario in which the true system initially evolves according to Model M_0. At time $t = 3\,h$, a change occurs and the bioreactor starts evolving according to the structure of Model M_1. If and when the change is detected, the process is returned to structurally evolve according to Model M_0 after 0.5 h through some form of intervention. Table 1 lists the parameter values used to represent the true system and the operating conditions for the system. The initial conditions are $X_0 = 7.78\,g/L$, $S_0 = 0.55\,g/L$, and $P_0 = 27.50\,g/L$.

We *assume* that all model parameters are known, with no uncertainty specified in θ for any model. However, we introduce some plant–model mismatch by using wrong values for several parameters in all models used in the controllers. Relative to the true values listed in Table 1, all models in $\mathcal{M}$ have values of α, μ_{max}, and σ_P that are all 10% smaller, while β, K_m, σ_X, and σ_S are 10% larger. The model M_1 contains an additional parameter error: the 25% reduction in S_f is a 22% reduction in the true system if this fault occurs.

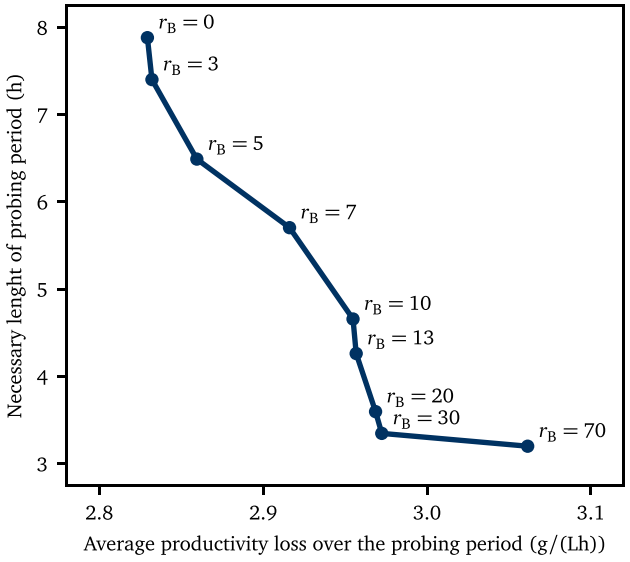

Fig. 4. Simulation-based tuning of the parameter r_B, with the top-left point corresponding to no active learning. The ordinate shows the average time elapsed between the controller initiates probing for active learning and correct selection of the model that corresponds to the structural change. The abscissa shows the average per-hour productivity loss, relative to the optimal steady-state productivity, over the corresponding period of time.

We determine the desired operating point offline under the assumption that the system evolves according to Model M_0. The product concentration should stay as close as possible to the desired value $P_d = 27.5$ g/L and must be between 25 g/L and 30 g/L.

The lower and upper bounds on the product concentration are enforced for the expected-value predictions with a back-off parameter $c_1 = 0.5$ g/L. The dilution rate (the control) is constrained to $0.015\,h^{-1} \le u_k \le 0.8\,h^{-1}$, and its desired value corresponding to P_d is $D_d = 0.15\ h^{-1}$. The stage cost is thus formulated as the reference-tracking objective

$$\ell_j\left(\hat{x}_{j|k}^{[i_k^M]}, u_{j|k}\right) = \left(\hat{P}_{j|k}^{[i_k^M]} - P_d\right)^2 + \left(u_{j|k} - D_d\right)^2$$

for all j on the prediction horizon. The terminal cost is $\ell_{k+N}\left(\hat{x}_{k+N|k}^{[i_k^M]}\right) = \left(\hat{P}_{k+N|k}^{[i_k^M]} - P_d\right)^2$. In all simulations reported below, the prediction horizon is $N = 8$ since there is no marginal benefit to increase it beyond this value.

5.2. MPC with active learning

The stage cost in (26) is the tracking error in P for the current control model $M_{i_k^M}$, whereas the bound P_{UB} quantifies the structural model uncertainty. The trade-off between the two objectives, minimizing the tracking error for the most probable model and active learning for correct identification of a potential structural change, thus explicitly appears in the objective function. The relative importance of these two goals is specified with the weight r_B. Turning off the active-learning feature by setting r_B to zero increases the risk of productivity loss from a structural change that is not identified. Conversely, increasing r_B to a value so high that active learning is the dominant objective results in a large tracking error, which leads to productivity loss. That is, a large r_B can result in probing that disrupts the operation to an extent that causes performance loss greater than the gain from faster learning. This is demonstrated for this case study in Fig. 4, which shows the necessary length of time for which active learning is required, plotted against the average reduction in productivity over this period of time.

An appropriate value for r_B can be determined through a tuning procedure like the one we describe next. The results are obtained through simulating the closed-loop system over a set of different r_B values ranging from 0 to 70, with 100 Monte Carlo simulations for each value of r_B.

The ordinate in Fig. 4 shows the average time elapsed between the controller initiates probing for active learning and correct selection of the model that corresponds to the structural change. The abscissa shows the average productivity loss, relative to the optimal steady-state productivity $P_d D_d$, over the length of time for which the probing is performed. The top-left point in Fig. 4, $r_B = 0$, corresponds to the controller making no attempt to reduce the probability of error in model selection. Increasing r_B from 0 leads to a large improvement in the time taken to select the correct model, with minimal increase in productivity loss. This shows that a relatively small adjustment to the operating strategy can result in large gains in uncertainty reduction with minor immediate loss of performance. Note that the curve in Fig. 4 shows a short-term cost of active learning, but it does not show the potential gain in productivity that is realized after the learning. That is, the trade-off discussed in the previous paragraph does not imply an overall loss in control performance. Also note that there is minimal benefit to increasing r_B past 30, and that setting $r_B > 70$ does not further expedite the learning.

Based on the insight from Fig. 4, we choose $r_B = 30$ for the objective function of the optimal control problem (27) solved in the simulation results presented below. This value of r_B balances the productivity loss and the probability of selection error and is used for structural change identification after a sudden drop in productivity signals that a change may have occurred. Further simulations show that $r_B = 30$ works well across a range of scenarios of structural change.

5.3. Benchmark MPC controllers

We compare the performance of our proposed MPC strategy with active learning to three other control strategies: (i) nominal MPC with no learning, (ii) robust MPC that accounts for structural model uncertainty, stochastic disturbances, and measurement noise through tightening the state constraints, and (iii) MPC with passive learning where the model is updated at any given time, but the controller has no probing feature to actively learn about the process. All four MPC strategies use the EKF discussed in Section 4.4 for state estimation.

The nominal MPC strategy uses M_0 as the prediction model and enforces expected-value state constraints for M_0 with the back-off parameter c_1 discussed above. In this approach, the uncertainty arises from not knowing which model best predicts the process data as well as from the stochastic disturbances and measurement noise. To increase robustness to this uncertainty, we use the constraint-tightening approach of Limón et al. (2002, 2005) for robust MPC. To ensure robust constraint satisfaction while using M_0 as the control model, we determine the bounds on the additive disturbances and the prediction errors through 1000 Monte Carlo simulations using each of the different models to represent the process. With these simulations we determine upper bounds for the estimate error of the constrained state P_k, $|P_k - \hat{P}_{k|k}|$, and the one-step-ahead prediction error $\|K(y_k - C\hat{x}_{k|k})\|_\infty$, both of which hold for all system realizations. With a Lipschitz constant $L_c = 1.1185$ for constraint tightening and after adding a 5% margin to each of these bounds, a prediction horizon of $N = 3$ is used in the robust MPC strategy.

The MPC with passive learning is identical to our proposed approach with active learning except r_B is set to zero.

All tuning parameters are identical in all four controllers except as noted above, with the shorter prediction horizon N in the robust MPC, a nonzero r_B for active-learning MPC, and updating the control-model structure in the two controllers with learning. Hence, only the MPC with active learning predicts the covariances in (27c) and (27e) as well as the other quantities for more than one model.

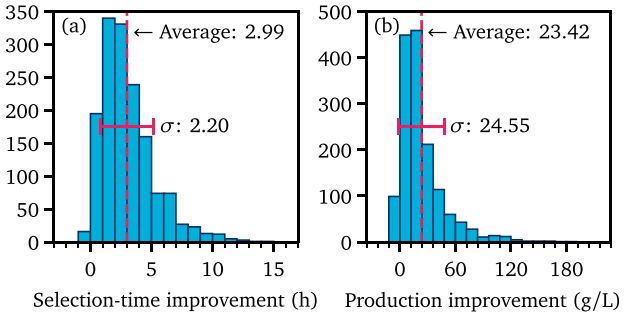

Fig. 5. Histograms showing improvements that result from MPC with active learning over MPC with passive learning in terms of (a) selection time of the correct model hypothesis and (b) production over the selection time obtained with passive learning.

5.4. Simulation results

We implement all four MPC strategies the by solving the associated special cases of the OCP (27) using IPOPT 3.12.5 (Wächter and Biegler, 2006) under MATLAB with explicit functions for the gradients. In this implementation, all variables are treated as decision variables. The problem can also be solved in the reduced space of the control inputs, and the gradients can be determined using automatic differentiation with tools such as CASADI (Andersson et al., 2019).

Evaluating the control performance on 1500 Monte Carlo simulations reveals the proposed approach with active learning outperforms the three other controllers, with the passive-learning MPC being second best. For these simulations, the CPU time for solving each instance of the optimal control problem (which with active learning has 872 variables) on a standard laptop is on average 0.167 s (standard deviation 0.023 s) for MPC with passive learning and 0.301 s (standard deviation 0.126 s) for MPC with active learning. Fig. 5 compares these two strategies and quantifies the two main advantages of MPC with active learning: faster selection of the correct model hypothesis and, as a result, improved mitigation of productivity loss.

Fig. 5(a) shows the average improvement in the speed of learning using active over passive learning. We quantify this through improvement in selection time, which is the time elapsed between the structural change occurring and the controller selecting corresponding model. On average, the correct model is selected 2.99 h faster with active learning, with a standard deviation of 2.20 h. There is a large number of cases for which the improvement is more than one standard deviation above the average, and crucially there is only a small number of cases where passive learning slightly outperforms active learning. In these few instances, the realizations of the stochastic noise and disturbances are such that the passive strategy of not actively improving the learning yield a better outcome, despite the active approach being significantly better in expectation and on average. Fig. 5(b) compares the difference in total production during the selection time relative to the passive learning case (that is, the total productivity with passive learning over this time span is subtracted from the total productivity with active learning over the *same* time span). This is thus a metric for the production improvement obtained using active instead of passive learning. The improvement in production is distributed similarly to the improvement in selection time. The average

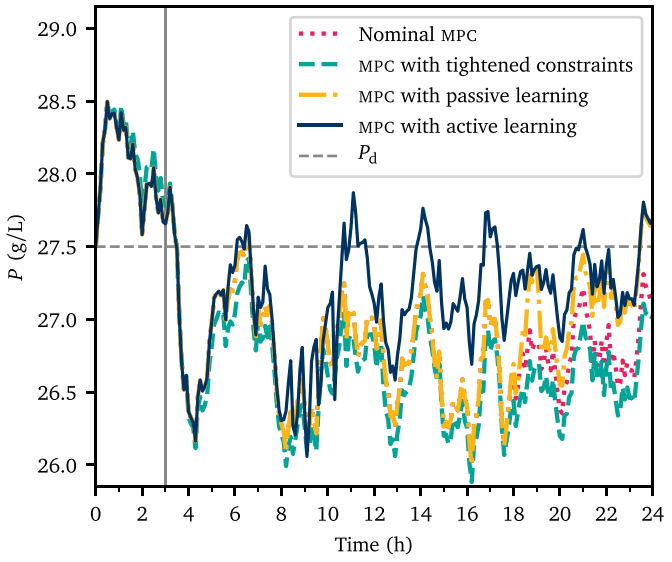

Fig. 6. Comparison of setpoint tracking in product concentration P using nominal MPC with no learning, robust MPC with no learning, MPC with passive learning ($r_B = 0$), and MPC with active learning ($r_B = 30$). The time of the structural change at 3 h is indicated with a vertical line.

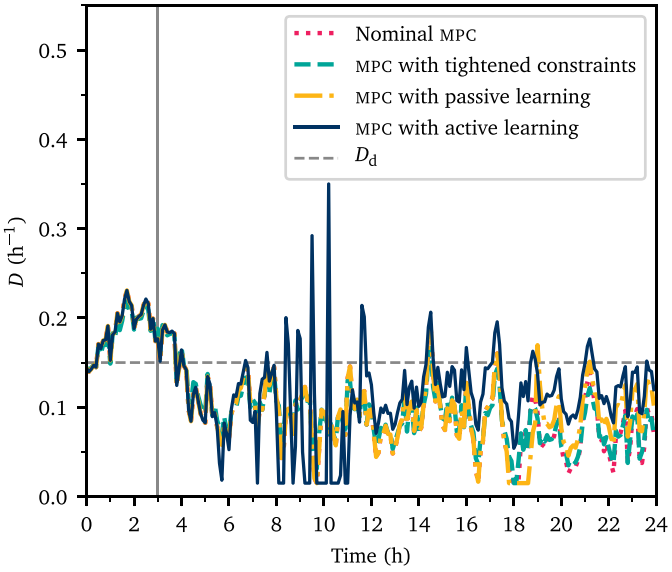

Fig. 7. Comparison of input profiles D using nominal MPC with no learning, robust MPC with no learning, MPC with passive learning ($r_B = 0$), and MPC with active learning ($r_B = 30$). The time of the structural change at 3 h is indicated with a vertical line.

production improvement of 23.42 g/L is significant and corresponds to 5.6 h of production at the optimal productivity $P_d D_d$. There is also here a large number of cases where the production improvement is larger than one standard deviation of 24.55 g/L and only a few cases in which using active learning reduces the production relative to using passive learning. These simulation results illustrate the importance of active learning for effectively restoring the MPC performance in the presence of structural model uncertainty and consequently mitigating losses in the process performance.

Figs. 6–8, show representative simulations for all four strategies under the same realization of disturbances and measurement noise, illustrating their different performance. The figures show, respectively, setpoint tracking in product concentration P, input profiles D, and productivity profiles PD.

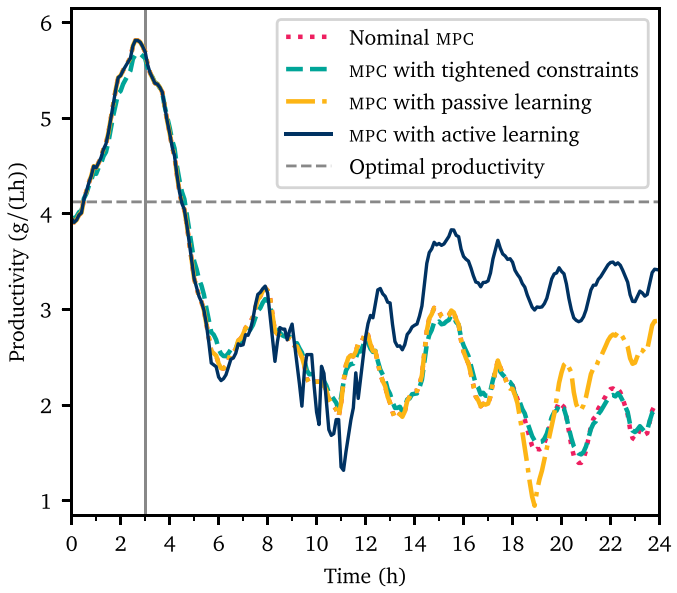

Fig. 8. Comparison of productivity PD using nominal MPC with no learning, robust MPC with no learning, MPC with passive learning ($r_B = 0$), and MPC with active learning ($r_B = 30$). The time of the structural change at 3 h is indicated with a vertical line.

Fig. 6 shows how the nominal MPC results in significant tracking error in P, a consequence of not accounting for uncertainty. This is the cause of the large productivity drop shown in Fig. 8. Fig. 6 also illustrates how the robust MPC does not significantly improve performance over the nominal one. There is minimal improvement in tracking error and no substantial change in the control input profile, apparent from Fig. 7. Accordingly, the productivity drop shown in Fig. 8 is unacceptable also when using the robust MPC. Together, these figures provide representative demonstrations of how neither the nominal MPC nor the robust MPC recovers from the structural change since neither has a mechanism to detect that the change occurred. Both controllers lower the dilution rate to maintain the product concentration near its optimal steady-state value. The reduced inlet substrate concentration and the maintenance of this product concentration together cause a slow depletion of substrate in the reactor, which necessitates reducing the dilution rate further before the process settles around a lower steady-state productivity.

Fig. 7 shows how the proposed MPC with active learning uses the dilution rate to probe the process and to reduce the probability of selecting the wrong model. In this case, the probing results in a significant improvement in mitigating the productivity loss; see Fig. 8. Averaged over the 1500 Monte Carlo simulations, the integral of the squared deviation from optimal productivity for the nominal, the robust, and the with passive- and active-learning MPC strategies are 88.7 $g^2/(L^2h)$, 82.0 $g^2/(L^2h)$, 79.4 $g^2/(L^2h)$, and 66.9 $g^2/(L^2h)$, respectively. By this metric, all MPC strategies that account for the process uncertainty in some way improve the process performance relative to the nominal MPC, with the MPC with active learning being superior to the others.

To identify the occurrence of the structural change, the MPC strategies with passive and active learning both estimate the model probabilities. These two strategies differ in the value of r_B, which as noted above is set to zero for passive learning. Fig. 9 compares the estimated probabilities for the three model hypotheses that result from MPC with passive and active learning, and it illustrates how the average improvement in selection time shown in Fig. 5 is achieved. These profiles correspond to the same uncertainty realization shown in Figs. 6–8. Fig. 9 shows that the probability of the correct model hypothesis, $P^{[1]}_{M,k}$, increases faster as a result of the

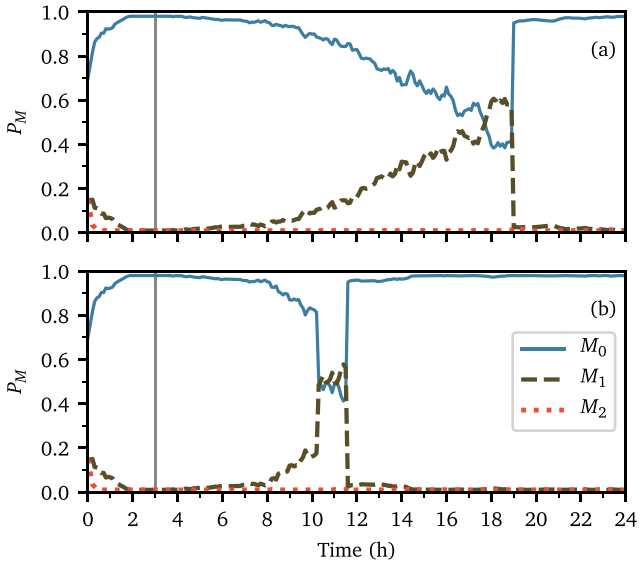

Fig. 9. Comparison of model probabilities for the three model hypotheses for MPC with passive (a) and active (b) learning. The time of the structural change at 3 h is indicated with a vertical line.

probing induced by active learning. In this particular case, the controller with active learning selects the correct model at time 11.1 h, 8.1 h after the structural change occurred. Conversely, with passive learning the controller does not select the correct model until time 18.5 h, 15.5 h after the change. That is, the active learning reduces the time elapsed before correct selection by 47.7 %.

6. Conclusions

This work adopts the concept of explicit dual control to develop a computationally tractable approach to MPC with active learning that increases the information content of closed-loop system data for reducing structural model uncertainty. The proposed approach involves improved hypothesis selection from a set of candidate models through minimizing an approximate bound on the Bayes risk of model-selection error. Simulation results with a nonlinear bioreactor case study demonstrate that MPC with active learning can lead to significant improvement in the control performance relative to nominal MPC with no learning, a robust MPC that accounts for structural model uncertainty but has no learning mechanism, and MPC with passive learning. MPC with active learning can create new avenues for online model discrimination and active fault diagnosis under closed-loop control (e.g., see Blanke et al., 2006; Heirung and Mesbah, 2019).

The computationally tractable approach to MPC with active learning developed in this paper relies on several approximations. In particular, we approximate the state probability distributions and likelihoods using the first two moments and use an approximate bound on the Bayes risk of model-selection error. Future work involves analyzing the theoretical implications of these approximations and identifying potential sources of significant conservatism. The development of a tighter bound on the Bayes risk of model-selection error and analysis of the convergence properties of the proposed approach are topics of future papers.

References

Ackerson, G.A., Fu, K.S., 1970. On state estimation in switching environments. IEEE Trans. Automat. Contr. 15 (1), 10–17.

Agrawal, P., Koshy, G., Ramseier, M., 1989. An algorithm for operating a fed-batch fermentor at optimum specific-growth rate.. Biotechnol. Bioeng. 33 (1), 115–125.

Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B., Diehl, M., 2019. Casadi: a software framework for nonlinear optimization and optimal control. Math. Program. Comput. 11 (1), 1–36.

Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T., 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. Signal Process. 50 (2), 174–188.

Åström, K.J., Wittenmark, B., 1995. Adaptive Control, 2nd ed. Addison-Wesley.

Atkinson, A.C., 2008. DT-optimum designs for model discrimination and parameter estimation. J. Stat. Plann. Inference 138, 56–64.

Atkinson, A.C., Fedorov, V.V., 1975. The design of experiments for discriminating between two rival models. Biometrika 62 (1), 57–70.

Aufderheide, B., Bequette, B.W., 2003. Extension of dynamic matrix control to multiple models. Comput. Chem. Eng. 27 (8–9), 1079–1096.

Azimzadeh, F., Galán, O., Romagnoli, J.A., 2001. On-line optimal trajectory control for a fermentation process using multi-linear models. Comput. Chem. Eng. 25 (1), 15–26.

Bar-Shalom, Y., Tse, E., 1974. Dual effect, certainty equivalence, and separation in stochastic control. IEEE Trans. Automat. Control 19 (5), 494–500.

Bellman, R., 1957. Dynamic Programming. Princeton University Press.

Bemporad, A., Morari, M., 1999. Robust model predictive control: a survey. In: Garulli, A., Tesi, A. (Eds.), Robustness in Identification and Control. In: Lecture Notes in Control and Information Sciences, 245. Springer, pp. 207–226.

Benosman, M., 2018. Model-based vs data-driven adaptive control: an overview. Int. J. Adapt. Control Signal Process. 32 (5), 753–776.

Blackmore, L., Williams, B.C., 2006. Finite horizon control design for optimal discrimination between several models. In: IEEE Conference on Decision and Control, San Diego, CA, pp. 1147–1152.

Blanke, M., Kinnaert, M., Lunze, J., Staroswiecki, M., 2006. Diagnosis and Fault-Tolerant Control, 2nd ed. Springer.

Böling, J.M., Seborg, D.E., Hespanha, J.P., 2007. pH Multi-model control of a simulated pH neutralization process. Control Eng. Pract. 15 (6), 663–672.

Caflisch, R.E., 1998. Monte Carlo and quasi-Monte Carlo methods. Acta Numerica 7, 1–49.

Chen, Z., 2003. Bayesian filtering: from Kalman filters to particle filters, and beyond. Statistics 182 (1), 1–69.

Cowles, M.K., Carlin, B.P., 1996. Markov chain Monte Carlo convergence diagnostics: a comparative review. J. Am. Stat. Assoc. 91 (434), 883–904.

Du, J., Johansen, T.A., 2014. A gap metric based weighting method for multimodel predictive control of MIMO nonlinear systems. J. Process Control 24 (9), 1346–1357.

Duda, R.O., Hart, P.E., Stork, D.G., 2002. Pattern Classification, 2nd ed. Wiley.

Farina, M., Giulioni, L., Scattolini, R., 2016. Stochastic linear model predictive control with chance constraints – A review. J. Process Control 44, 53–67.

Feldbaum, A.A., 1961. Dual-control theory. I. Automat. Remote Control 21 (9), 874–880.

Filatov, N.M., Unbehauen, H., 2000. Survey of adaptive dual control methods. IEE Proc. Control Theory Appl. 147 (1), 118–128.

Gibbs, A.L., Su, F.E., 2002. On choosing and bounding probability metrics. Int. Stat. Rev. 70 (3), 419–435.

Gopinathan, M., Bošković, J.D., Mehra, R.K., Rago, C., 1998. A multiple model predictive scheme for fault-tolerant flight control design. In: IEEE Conference on Decision and Control, Tampa, FL, pp. 1376–1381.

Guay, M., Adetola, V., DeHaan, D., 2015. Robust and adaptive model predictive control of nonlinear systems. Inst. Eng. Technol..

Heirung, T.A.N., Mesbah, A., 2017. Stochastic nonlinear model predictive control with active model discrimination: a closed-loop fault diagnosis application. In: IFAC World Congress, Toulouse, France, pp. 16504–16509.

Heirung, T.A.N., Mesbah, A., 2019. Input design for active fault diagnosis. Annual Rev. Control in press.

Heirung, T.A.N., Paulson, J.A., Lee, S.J., Mesbah, A., 2018. Model predictive control with active learning under model uncertainty: why, when, and how. AIChE J. 64 (8), 3071–3081.

Heirung, T.A.N., Paulson, J.A., O'Leary, J., Mesbah, A., 2017a. Stochastic model predictive control — how does it work? Comput. Chem. Eng. 114, 158–170.

Heirung, T.A.N., Ydstie, B.E., Foss, B., 2017b. Dual adaptive model predictive control. Automatica 80, 340–348.

Hellman, M.E., Raviv, J., 1970. Probability of error, equivocation, and the Chernoff bound. IEEE Trans. Inf. Theory 16 (4), 368–372.

Henson, M.A., Seborg, D.E., 1992. Nonlinear control strategies for continuous fermenters. Chem. Eng. Sci. 47 (4), 821–835.

Houska, B., Telen, D., Logist, F., Van Impe, J., 2017. Self-reflective model predictive control. SIAM J. Control Optim. 55 (5), 2959–2980.

Julier, S.J., Uhlmann, J.K., 1997. A new extension of the Kalman filter to nonlinear systems. In: Aerospace/Defense Sensing, Simulation and Controls, Orlando, FL, pp. 182–193.

Julier, S.J., Uhlmann, J.K., 2004. Unscented filtering and nonlinear estimation. Proc. IEEE 92 (3), 401–422.

Kanev, S., Verhaegen, M., 2000. Controller reconfiguration for non-linear systems. Control Eng. Pract. 8 (11), 1223–1235.

Kantas, N., Maciejowski, J.M., Lecchini-Visintini, A., 2009. Sequential Monte Carlo for model predictive control. In: Magni, L., Raimondo, D.M., Allgöwer, F. (Eds.), Nonlinear Model Predictive Control – Towards New Challenging Applications. In: Lecture Notes in Control and Information Sciences, 384. Springer, pp. 263–273.

Koller, R.W., Ricardez-Sandoval, L.A., Biegler, L.T., 2018. Stochastic back-off algorithm for simultaneous design, control, and scheduling of multiproduct systems under uncertainty. AIChE Journal 64 (7), 2379–2389.

Kouvaritakis, B., Cannon, M., 2016. Model Predictive Control: Classical, Robust and Stochastic. Springer.
Kuure-Kinsey, M., Bequette, B.W., 2009. Multiple model predictive control of nonlinear systems. In: Magni, L., Raimondo, D.M., Allgöwer, F. (Eds.), Nonlinear Model Predictive Control. In: Lecture Notes in Control and Information Sciences, 384. Springer, pp. 153–165.
Kuure-Kinsey, M., Bequette, B.W., 2010. Multiple model predictive control strategy for disturbance rejection. Indus. Eng. Chem. Res. 49, 7983–7989.
La, H.C., Potschka, A., Schlöder, J.P., Bock, H.G., 2017. Dual control and online optimal experimental design. SIAM J. Sci. Comput. 39 (4), 640–657.
Larsson, C.A., Ebadat, A., Rojas, C.R., Bombois, X., Hjalmarsson, H., 2016. An application-oriented approach to dual control with excitation for closed-loop identification. Eur. J. Control 29, 1–16.
Li, X.-R., Bar-Shalom, Y., 1996. Multiple-model estimation with variable structure. IEEE Trans. Automat. Control 41 (4), 478–493.
Limón, D., Alamo, T., Camacho, E.F., 2002. Input-to-state stable MPC for constrained discrete-time nonlinear systems with bounded additive uncertainties. In: IEEE Conference on Decision and Control, Las Vegas, NV, pp. 4619–4624.
Limón, D., Bravo, J.M., Alamo, T., Camacho, E.F., 2005. Robust MPC of constrained nonlinear systems based on interval arithmetic. IEE Proc. Control Theory Appl. 152 (3), 325–332.
Ljung, L., 1979. Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. IEEE Trans. Control Syst. Technol. 24 (1), 36–50.
Lorenzen, M., Cannon, M., Allgöwer, F., 2019. Robust MPC with recursive model update. Automatica 103, 461–471.
Maciejowski, J.M., Jones, C.N., 2003. MPC fault-tolerant flight control case study: flight 1862. In: IFAC Fault Detection, Supervision and Safety of Technical Processes, pp. 119–124.
Matusita, K., 1971. Some properties of affinity and applications. Annal. Inst. Stat. Math. 23, 137–155.
Mayne, D.Q., 2014. Model predictive control: recent developments and future promise. Automatica 50 (12), 2967–2986.
Mesbah, A., 2016. Stochastic model predictive control: an overview and perspectives for future research. IEEE Control Syst. 36 (6), 30–44.
Mesbah, A., 2018. Stochastic model predictive control with active uncertainty learning: a survey on dual control. Annual Rev. Control 45, 107–117.
Murray-Smith, R., Johansen, T.A., 1997. Multiple Model Approaches To Nonlinear Modelling And Control. CRC Press.
Nandola, N.N., Bhartiya, S., 2008. A multiple model approach for predictive control of nonlinear hybrid systems. J. Process Control 18 (2), 131–148.
Narendra, K.S., Han, Z., 2011. The changing face of adaptive control: the use of multiple models. Annu. Rev. Control 35 (1), 1–12.
Özkan, L., Kothare, M.V., 2006. Stability analysis of a multi-model predictive control algorithm with application to control of chemical reactors. J. Process Control 16 (2), 81–90.
Paulson, J.A., Mesbah, A., 2017. An efficient method for stochastic optimal control with joint chance constraints for nonlinear systems. Int. J. Robust Nonlinear Control. In press.
Paulson, J.A., Mesbah, A., 2018. Nonlinear model predictive control with explicit backoffs for stochastic systems under arbitrary uncertainty. In: IFAC Nonlinear Model Predictive Control Conference, Madison, WI, pp. 523–534.
Robertson, D.G., Lee, J.H., Rawlings, J.B., 1996. A moving horizon-based approach for least-squares estimation. AIChE J. 42 (8), 2209–2224.
Rodriguez, J.A., Romagnoli, J.A., Goodwin, G.C., 2003. Supervisory multiple regime control. J. Process Control 13 (2), 177–191.
Shouche, M.S., Genceli, H., Nikolaou, M., 2002. Effect of on-line optimization techniques on model predictive control and identification (MPCI). Comput. Chem. Eng. 26 (9), 1241–1252.
Simon, D., 2006. Optimal State Estimation. John Wiley & Sons.
Tanaskovic, M., Fagiano, L., Smith, R.S., Morari, M., 2014. Adaptive receding horizon control for constrained MIMO systems. Automatica 50 (12), 3019–3029.
Wächter, A., Biegler, L.T., 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math. Progr. Ser. A 106 (1), 25–57.
Wan, E.A., Nelson, A.T., 1996. Dual Kalman filtering methods for nonlinear prediction, smoothing, and estimation. Neural Inf. Process. Syst. 793–799.
Wittenmark, B., 1995. Adaptive dual control methods: an overview. In: IFAC Symposium on Adaptive Systems in Control and Signal Processing, Budapest, Hungary, pp. 67–72.
Zhu, B., Xia, X., 2016. Adaptive model predictive control for unconstrained discrete–time linear systems with parametric uncertainties. IEEE Trans. Automat. Control 61 (10), 3171–3176.