Model Predictive Control with Active Learning Under Model Uncertainty: Why, When, and How

Tor Aksel N. Heirung, Joel A. Paulson, Shinje Lee, and Ali Mesbah
Dept. of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720

DOI 10.1002/aic.16180
Published online June 6, 2018 in Wiley Online Library (wileyonlinelibrary.com)

Optimal control relies on a model, which is generally uncertain because of incomplete knowledge of the system and changes in the dynamics over time. Probing the system under closed-loop control can reduce the model uncertainty through generating input-output data that is more informative than the data generated from normal operation. This paper addresses the problem of model predictive control (MPC) with active learning, with a particular focus on how incorporating probing in the control action can reduce model uncertainty. We discuss some of the central theoretical questions in this problem, and demonstrate the potential of active learning for maintaining MPC performance in the presence of uncertainty in model parameters and structure. Simulation results show that active learning is particularly beneficial when a system undergoes abrupt changes (such as the sudden occurrence of a fault) that can compromise operational safety, reliability, and profitability. © 2018 American Institute of Chemical Engineers *AIChE J*, 64: 3071–3081, 2018

Keywords: model predictive control, stochastic optimal control, active learning, passive learning, parametric uncertainty, model-structure uncertainty, Bayesian estimation

Introduction

Model predictive control[1,2] (MPC), also known as receding-horizon control, is the most widely used approach for advanced control of multivariable systems with state and input constraints.[3,4] MPC relies on a model to predict the behavior of the system. Whether data driven or based on first principles, this model is to some extent uncertain, generally because of incomplete knowledge of the system. Common sources of model uncertainty include inaccurate estimates of model parameters and unknown aspect of the model structure itself, as in the case of an unknown kinetic mechanism in a physics-based model, or the appropriate order of a data-driven model. A system may also undergo abrupt changes (such as faults and failures; see Ref. 5) or exhibit time-varying dynamics, both of which can further increase the uncertainty in the model structure.

Uncertainty in model-based control has sparked several important research directions, such as robust and stochastic control,[6,7] adaptive control,[8,9] and identification for control.[10–12] The corrective nature of feedback provides MPC with a certain degree of robustness to uncertainty. However, MPC performance degrades when feedback cannot adequately compensate for incomplete knowledge of the system. Possible consequences include excessive constraint violations and large offsets in setpoint tracking (see, e.g., Ref. 13). This has led to the development of robust MPC[14–17] (RMPC), in which model uncertainty, represented by deterministic, bounded sets, is explicitly accounted for. RMPC generally involves enforcing the state constraints with respect to all possible uncertainty realizations, including those realizations that may have a small probability of occurrence in practice. Consequently, RMPC can result in highly conservative control performance. Stochastic MPC[16–19] (SMPC) approaches can be used when the model uncertainty is described by probability distributions. By allowing the specification of a permitted probability of constraint violation, SMPC can reduce conservatism while retaining an acceptable level of robustness to uncertainty.[20]

Despite systematically accounting for model uncertainty, the performance of RMPC and SMPC strongly depends on the quality of the uncertainty descriptions. When a system undergoes changes over time, either gradual (e.g., time-varying dynamics) or abrupt (e.g., faults and malfunctions), the model and the uncertainty descriptions may no longer adequately represent the system. RMPC and SMPC are blind to the system variations in that there is no mechanism for adapting the model to the changes through adjusting its parameters, its structure, and the uncertainty description. This may cause deterioration of control performance and compromise the controller's robustness to uncertainty.

Adaptive control is an alternative approach to control under uncertainty. This approach involves adjusting the model under closed-loop control, with the goal of improving the model for control purposes.[8,9] A critical issue in adaptive control, including adaptive MPC,[21–23] is that the data generated in closed loop must be sufficiently informative for the model adaptation to be beneficial. In his seminal work,[24–28] Feldbaum was the first to recognize that in optimal control of systems with reducible model uncertainty, the control inputs must have a *probing* effect that generates informative closed-loop data for *active*

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Correspondence concerning this article should be addressed to A. Mesbah at mesbah@berkeley.edu.

learning, in addition to having a *directing* effect to control the system state. It is well known that lack of probing in adaptive control can lead to problems such as controller turn-off,[29] bursting,[30] and the loss of controllability.[31]

This article addresses the problem of MPC with active learning for systems with probabilistic uncertainty descriptions. Through a discussion of the central theoretical questions and illustrative case studies, we demonstrate the potential ability of active learning to maintain MPC performance in the presence of parameter and model-structure uncertainty. The main objective of the article is to highlight the possible benefits of active learning without in-depth treatment of the algorithmic aspects (see Ref. 32 for a recent review on this subject). We first formally introduce the optimal control problem, along with a discussion of why its exact solution is computationally intractable for practically sized systems and a brief overview of some of the main approaches to MPC with active learning. Two case studies, a continuous-stirred-tank reactor (CSTR) and a continuous bioreactor, are then used to illustrate the effects of parameter and model-structure uncertainty and how active learning can improve control performance. The article concludes with some suggestions for future research directions.

Optimal Control with Active Learning

Problem formulation

Consider nonlinear systems described in discrete time by uncertain models of the form

$$M:\begin{cases} x_{k+1}=f(x_k,u_k,\theta,w_k), \\ y_k=h(x_k,\theta,v_k) \end{cases} \tag{1}$$

where $k\in\mathbb{N}_0$ is the time index ($\mathbb{N}$ denotes the set of natural numbers, $\mathbb{N}_0=0\cup\mathbb{N}$); x_k, u_k, and y_k are the system state, input, and output, respectively; θ denotes the unknown model parameters with known initial probability distribution $p(\theta)$; $w_k\sim p(w)$ and $v_k\sim p(v)$ denote stochastic disturbances and measurement noise, respectively; and f and h denote the state and output equations. The initial system state x_0 is uncertain with known probability distribution $p(x_0)$. The distributions $p(w)$ and $p(v)$ are assumed known for all $k\in\mathbb{N}_0$, so that w_k and v_k are both sequences of independent and identically distributed random variables. All of the random variables x_0, θ, w_k, and v_i are mutually independent for all $k,i\in\mathbb{N}_0$. The system is constrained, with the control input u_k required to lie in a compact set $\mathbb{U}$, $u_k\in\mathbb{U}$, and the state x_k required to lie in a closed set $\mathbb{X}$, $x_k\in\mathbb{X}$. As the model (1) is subject to probabilistic uncertainty, the state constraints can be enforced in terms of the *chance constraint*

$$\mathrm{Pr}_k(x_k\in\mathbb{X})\geq 1-\epsilon_k \tag{2}$$

where $\mathrm{Pr}_k(\cdot)$ denotes probability given the information available at time k and ϵ_k is the permitted probability of state constraint violation. In a more general problem setting, a set of models $\mathcal{M}=\{M_1,M_2,\ldots,M_{n_m}\}$, all of the form (1), can be formulated as candidates for describing the system, with each model M_i specified by

$$M_i:=\{f^{[i]},h^{[i]},p(x_0^{[i]}),p(\theta^{[i]}),p(w^{[i]}),p(v^{[i]})\}$$

where the superscript $[i]$ distinguishes the equations, variables, and parameters in each model $M_i\in\mathcal{M}$. Note that every model in $\mathcal{M}$ can be specified with different state and output equations. Furthermore, the dimensions of the state x_k, the unknown parameters θ, the disturbances w_k, and the measurement noise v_k need not be the same across the different models. We assume that a true model of the system exists and is contained in $\mathcal{M}$ at any given time. The system can be correctly represented by different models M_i at different times, such as when some structural change occurs in the system.

The state x_k is entirely or partially observed through the measurements y_k. Let the input and output data available at time k be denoted by $\mathcal{Y}_k:=\{y_k,y_{k-1},\ldots,y_0,u_{k-1},u_{k-2},\ldots,u_0\}$, with $\mathcal{Y}_0:=y_0$. The state and the unknown parameters can be lumped together to form the augmented state vector $z_k^\top=[x_k^\top,\ \theta^\top]$. The conditional probability distribution of z_k, given $\mathcal{Y}_k$, is known as the *hyperstate* $\xi_k:=p(z_k\,|\,\mathcal{Y}_k)$. As the model (1) represents a Markov process (see, e.g., Ref. 33), the hyperstate ξ_k is a Bayesian *posterior* at time k, whereas the *prior* at time k is $\xi_{k-1}=p(z_{k-1}\,|\,\mathcal{Y}_{k-1})$. To determine the posterior, the prior can be used to predict the hyperstate $\xi_{k|k-1}$. The predicted and posterior hyperstates are collectively propagated through the Bayesian recursion[34]

$$\xi_{k|k-1}=\int p(z_k\,|\,z_{k-1},u_{k-1})\cdot\xi_{k-1}\,dz_{k-1} \tag{3a}$$

$$\xi_k=\frac{p(y_k\,|\,z_k)\cdot\xi_{k|k-1}}{\int p(y_k\,|\,z_k)\cdot\xi_{k|k-1}\,dz_k} \tag{3b}$$

starting from $\xi_{0|-1}:=p(z_0)$, which is defined by $p(x_0)$ and $p(\theta)$. Note that the Bayesian filter (3) simplifies greatly when the model (1) is linear, all model parameters are known, and w_k and v_k are Gaussian sequences. In this case, $p(z_k\,|\,z_{k-1},u_{k-1})$ and $p(y_k\,|\,z_k)$ are Gaussian for all k, and (3) reduces to the well-known Kalman filter.[35] In the general case, however, the right-hand sides of (3a) and (3b) have no closed-form solution and must be approximated.[34]

It is clear from (3a) that the control input u_{k-1} directly affects $\xi_{k|k-1}$ and thus the hyperstate ξ_k, which quantifies the uncertainty associated with the state and parameters. Our knowledge of the system can therefore depend on the control input. When this is the case, the control input can be used to reduce the model uncertainty through actively learning about the system. Informally, *active learning* is possible when the control input affects the covariance, or other higher-order central moments, of the augmented state z_k.[36] In this case, the control input affects both the system state and the uncertainty, which is referred to as *dual effect*.[36] In contrast, the system is *neutral* when the control input almost surely affects only the mean of z_k (the first-order central moment) and not its uncertainty. The presence of the dual effect is necessary for active learning as it enables *probing* the system for information that can reduce uncertainty. Note that both *passive* and active learning use the information $\mathcal{Y}_k$ to reduce model uncertainty, typically through some form of Bayesian recursion (3). However, probing is what differentiates a controller with active learning from one that passively uses the input-output data, with no ability to increase the information content.

Here, we consider optimal control of systems described by (1) with reducible model uncertainty. Define a cost function in terms of a terminal cost $\ell_N(x_N)$ and a stage cost $\ell_k(x_k,u_k)$ over a time horizon $0\leq k\leq N-1$. With the sequence of control inputs $\pi_k:=\{u_k,u_{k+1},\ldots,u_{N-1}\}$, the cost from time k to time N is

$$J_k(\xi_k,\pi_k)=\mathrm{E}_k\left[\sum_{j=k}^{N-1}\ell_j(x_j,u_j)+\ell_N(x_N)\right] \tag{4}$$

where $\mathrm{E}_k[\cdot]$ denotes the conditional expectation with respect to $\xi_k=p(z_k\,|\,\mathcal{Y}_k)$, $p(w)$, and $p(v)$ for all $j\geq k+1$. Consider some control policy $\mu_k(\xi_k)$ such that $u_k=\mu_k(\xi_k)$. Let π_k^* denote the policy sequence that results in the optimal cost $J_k^*(\xi_k)$. The sequence of optimal costs can then be written recursively as

$$J_k^*(\xi_k)=\min_{u_k} \mathrm{E}_k[\ell_k(x_k,u_k)+J_{k+1}^*(\xi_{k+1})], \quad k=0,1,\ldots,N-1 \tag{5}$$

with $J_N^*(\xi_N)=\mathrm{E}_N[\ell_N(x_N)]$. The optimal cost $J_k^*(\xi_k)$ is also known as the *cost-to-go* from time k to time N, starting from ξ_k. The recursive equation (5) is the *Bellman equation*, which results from Bellman's principle of optimality.[37] From the Bellman equation, it is clear that as the cost function and optimality are defined in terms of the hyperstate, the uncertainty and how it is affected by the control input are implicitly included in the optimal control problem. That is, unless the system is neutral so that the control input cannot affect the uncertainty, the optimal sequence of control policies inherently probe the system when advantageous. This problem of stochastic optimal control under model uncertainty is formulated formally as follows.

Problem 1. (Stochastic optimal control with active learning). *For the horizon* $0 \le k \le N$, *determine the optimal sequence of control policies* π_0^*

$$\pi_0^* := \underset{\pi_0}{\operatorname{argmin}} J_0(\xi_0, \pi_0) \tag{6}$$

subject to a model M of the form (1) or a model set $\mathcal{M}$, *the probability distributions p(w) and p(v), the state chance constraint (2), and the input constraint* $u_k \in \mathbb{U}$.

Problem 1 was first posed by Feldbaum[24–28] for unconstrained systems with parametric model uncertainty only. The problem is often referred to as the *dual control problem* as the dual effect and thus active learning is naturally accounted for.

There is no analytic solution to the Bellman equation except for specific simple cases, such as the linear-quadratic regulator, or LQR.[38] In general, the Bellman equation can in principle be solved numerically, backward in time, using dynamic programing. However, dynamic programing suffers from exponential complexity growth in the dimensions of the state, input, and uncertainty and is therefore intractable as a solution strategy for larger problems, an issue known as the *curse of dimensionality*. This has led to the development of approximate methods for solving the Bellman equation, broadly referred to as approximate dynamic programming.[39,40]

In the following, we demonstrate the benefit from active learning on a scalar system, for which the dual control problem can be solved almost exactly using dynamic programing.

Illustrative example

Consider two scalar, linear models M_1 and M_2 of the form

$$x_{k+1}=ax_k+bu_k+w_k$$

with the stochastic sequence w_k being zero-mean, white Gaussian noise and the state x_k observed directly without noise. The models have different parameters $\theta^\top=[a, b]$; Model M_1 is stable but expensive to control (low input gain), while model M_2 is unstable but cheap to control (large input gain). Under the assumption that either M_1 or M_2 is a true representation of the system, the model uncertainty arises from not knowing which one is true. This leads to a simple dual control problem that can be solved numerically using dynamic programing. Despite its simplicity, the problem demonstrates the benefit of probing the system to actively reducing uncertainty. The optimal control problem is minimizing a cost function of the form $J_k(\xi_k)=\mathrm{E}_k[\sum_{j=k}^{N-1}(qx_j^2+ru_j^2)+qx_N^2]$ with $q,\ r > 0$; see the Supporting Information for details. In the following, P_k and $1-P_k$ denote the probabilities that models M_1 and M_2 are true, respectively.

The benefit from solving the optimal control problem with active learning (Problem 1), so that the dual effect is accounted for, can be quantified in various ways. Arguably, the most insight can be gained from comparing the optimal costs for the solutions with active and passive learning. Using dynamic programing to determine an optimal control policy with active learning requires including the equations that describe the dual effect for the system. In this problem, the dual effect is accounted for by including the Bayesian recursion (3) in the dynamic model, which relates the probabilities P_k and $1-P_k$ to the observations of the state, which in turn depend on the control input. Solving the Bellman equation (5) backward in time, from $k=N$ to $k=0$, accounting for how the control input affects both the state and the model probabilities, enables the control policy to anticipate the learning outcome from its actions. That is, the optimal control policy with active learning accounts for how future state observations affect the model uncertainty. In contrast, if the dual effect is not included in the model when solving the Bellman equation, meaning the effect of future state observations on uncertainty is unaccounted for, the resulting control policy cannot anticipate learning. In other words, the solution is determined under the implicit assumption that the model probabilities remain constant over time; that is, $P_{k+1}=P_k$ for all $0 \le k \le N-1$. The optimal control policy can therefore only result in passive learning.

Here, the expected benefit from active learning is quantified in terms of the difference between the optimal costs $J_0^*(\xi_0)$ obtained by solving (5) with and without accounting for the dual effect. Denote the optimal costs with active and passive learning by $J_{0,\mathrm{A}}^*(\xi_0)$ and $J_{0,\mathrm{P}}^*(\xi_0)$, respectively. The difference $\Phi(\xi_0) := (J_{0,\mathrm{P}}^*(\xi_0)-J_{0,\mathrm{A}}^*(\xi_0))/N$ quantifies the per-stage improvement in the expected cost of optimal control with active learning relative to that of optimal control with passive learning. Figure 1 shows $\Phi(\xi_0)$ as a function of the initial hyperstate ξ_0, which consists of the initial state x_0 and the initial probability P_0 that Model M_1 is true. The horizon N is chosen sufficiently long so that the solution to the Bellman equation does not change if N is increased further; that is, the solutions are the infinite-horizon control policies. Figure 1 shows that active learning has no value over passive learning along $P_0=0$ and $P_0=1$. This is because these regions of the hyperstate correspond to absolute certainty in which model is true, implying no learning is possible. In the vicinity of $x_0=0$, there are two reasons the benefit from active learning is small. First, $(x_k, u_k)=(0,0)$ is an equilibrium for both models, so the cost function remains fairly small in a neighborhood of this point. Second, even with passive learning, the correct model will eventually be identified at a low cost as the state is near the origin. Hence, even though the correct model can be determined faster in the vicinity of $x_0=0$ with active learning, the benefit of reduced model uncertainty does not outweigh the cost of probing in this region of the hyperstate space. Further away from $x_0=0$, however, the control effort required to bring the state to the origin is significant. Moreover, the consequence of not knowing which model is correct is much greater when the uncertainty is higher (i.e., P_0 is not close to 0 or 1). Accordingly, the optimal control policy with active learning

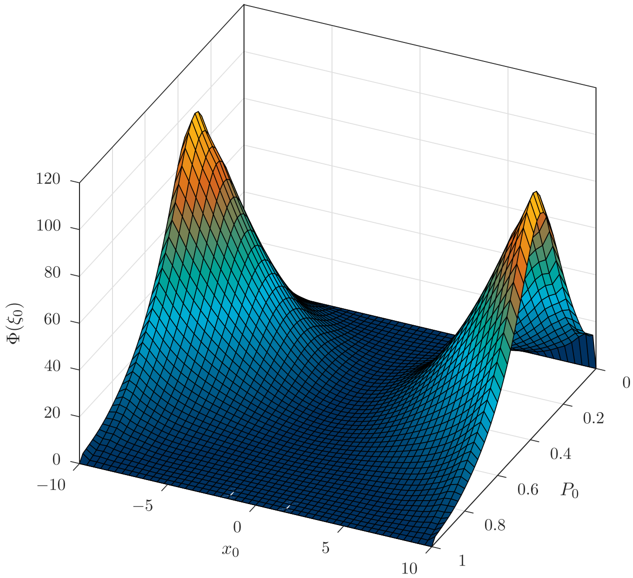

Figure 1. The expected per-stage value of optimal control with active learning relative to optimal control with passive learning, $\Phi(\xi_0)$, as a function of the hyperstate ξ_0 defined in terms of the initial state x_0 and the initial probability P_0 for Model M_1.

[Color figure can be viewed at wileyonlinelibrary.com]

significantly outperforms the one with passive learning when x_0 is far from the origin and P_0 is close to 0.5. Note that the benefit from active learning is always nonnegative as the controller only probes the system when beneficial in expectation.

While the degree to which active learning is beneficial, if at all, depends on a variety of factors, including operating region, how much reduced uncertainty can improve the performance, and the potential risks associated with probing. However, this example demonstrates that even for a very simple system with model uncertainty, optimal control with active learning can have significant advantages for the control performance. Solving the Bellman equation numerically with dynamic programing is tractable in this case as the hyperstate has dimension two, the control input is scalar, and there are only two models. Practically sized problems, however, require significantly more computational resources, and the exponential growth in problem complexity renders it intractable to determine the exact solution to the Bellman equation.

Tractable solutions to the optimal control problem with active learning

Exact solutions to certain simple instances of Problem 1 are available in the literature,[41,42] yet determining the sequence of optimal control policies that account for the dual effect through dynamic programing is generally not tractable.* The computational challenges associated with Problem 1 have motivated the development of methods capable of solving practically sized problems of optimal control with active learning. These methods, commonly known as approximate dual control, can be classified as implicit and explicit approaches to dual control.[32,45] Broadly, implicit dual control involves obtaining an approximate solution to the Bellman equation (5), which has the learning component implicitly included through the hyperstate as discussed above.[46,47] Explicit dual control, conversely, generally refers to replacing Problem 1 with a surrogate optimization problem that accounts for a measure of reducible model uncertainty.[48–53] Thus, explicit methods incorporate some form of probing into the optimal control problem explicitly through approximation of Problem 1.

MPC with active learning is a growing area of research.[32] The various approaches in the literature primarily differ in how probing is introduced in the optimal control problem. One class of methods relies on a constraint added to the optimal control problem so that the closed-loop data generated by the MPC be sufficiently informative for maintaining the model.[54,55] A potential disadvantage of this approach is that the resulting probing of the system, necessary to generate the informative data, may be excessive and lead to a reduction in control performance. Another class of methods modifies the cost function of the optimal control problem to include a term that quantifies the model uncertainty.[50,51] The MPC then balances the control cost, based on the uncertain model, with probing for active learning. Some of the key challenges in these approaches include tuning the controller to achieve an appropriate balance between the control and probing features of the inputs, and ensuring that the controller probes the system only when beneficial.[32] Note that implicit dual control methods generally do not suffer from these challenges, often at the expense of computational complexity.

While outside the scope of this article, MPC with active learning is not only relevant for model uncertainty. In output-feedback MPC,[56–59] the controller does not have direct access to the system state for feedback, but rather a measured subset affected by noise. An output-feedback MPC generally relies on a state estimator, commonly based on a Bayesian framework like (3), to estimate the state. For nonlinear systems, the quality of the state estimate depends on the control action. Furthermore, the state in a nonlinear system may be only locally observable, in which case the controller must take special care to prevent the loss of observability. In these situations, probing can be used to improve the state estimate and in turn improve the control performance.[52,60,61]

The remainder of this article demonstrates the importance of active learning in the context of MPC under model uncertainty. We present two case studies, one with parameter and one with model-structure uncertainty, in which tractable surrogates for Problem 1, inspired by explicit dual control, are solved in a receding-horizon fashion. Without advocating particular approaches to MPC with active learning, we illustrate the importance of active learning as a general strategy for improving MPC performance when the model uncertainty is significant.

MPC with Active Learning for Parametric Uncertainty

This section presents a case study in which a change in the process, represented by an abrupt reduction in a model parameter, can cause significant deterioration of the control performance. The results demonstrate how online model adaptation using informative closed-loop data, generated by a controller with active learning, can mitigate control performance loss

*While an in-depth discussion of state chance constraints is beyond the scope of this article, this form of constraints are intractable in optimization[43,44] and contribute to the complexity of Problem 1.

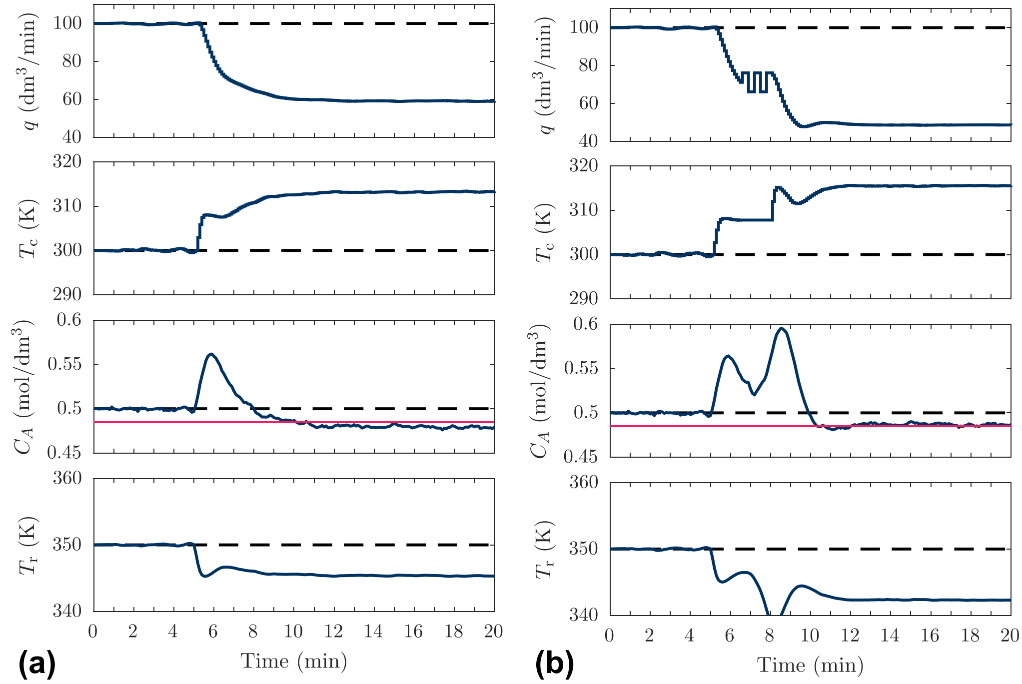

Figure 2. Closed-loop performance of (a) standard MPC in Strategy 1 and (b) MPC with open-loop model adaptation in Strategy 2 under parametric model uncertainty.

From top to bottom, the plots show the control inputs q and T_c and the measured states C_A and T_r. The black dashed lines and the red solid lines represent the setpoints and constraints, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

from the parametric model uncertainty. Consider a CSTR with two states, two inputs, and two uncertain model parameters.[62] The process dynamics are[†]

$$\frac{dC_A}{dt}=\frac{q}{V}(C_{A,\text{in}}-C_A)-k_0\exp\left(\frac{-E}{RT_\text{r}}\right)C_A \tag{7a}$$

$$\frac{dT_\text{r}}{dt}=\frac{q}{V}(T_\text{in}-T_\text{r})-\frac{\Delta H}{\rho c_\text{p}}k_0\exp\left(\frac{-E}{RT_\text{r}}\right)C_A+\frac{UA}{\rho c_\text{p}V}(T_\text{c}-T_\text{r}) \tag{7b}$$

where C_A is the concentration of species A, with the inlet concentration $C_{A,\text{in}}$; T_r is the reaction temperature, with the inlet temperature T_in; q is the manipulated volumetric flow rate; T_c is the manipulated coolant temperature; k_0 is the kinetic constant of the reaction; E is the activation energy; R is the gas constant; ΔH is the reaction heat; U is the heat transfer coefficient; A is the heat transfer area; ρ is the density; and c_p is the heat capacity. Both state variables C_A and T_r are measured and are subject to additive zero-mean Gaussian white noise with known variance. However, the model is uncertain as the kinetic constant k_0 and the reaction heat ΔH are unknown, but have known initial probability distributions $p(k_0)$ and $p(\Delta H)$.

A discrete-time state-space model of the form (1), with parametric uncertainty in $\theta^\top=[k_0,\ \Delta H]$, is used to design an MPC for the CSTR. The control objective is to maintain a specified productivity $\bar{q}\bar{C}_A$, while avoiding product quality loss by an unacceptably low concentration of A. This control objective translates to keeping C_A at a desired setpoint with minimal control effort while ensuring C_A is above a specified lower bound at all times. We formulate this optimal control problem as

$$\min_{\{u_{k+j}\}_{j=0}^{N-1}} \ J \tag{8a}$$

$$\text{subject to}\ \ x_{k+j+1}=f(x_{k+j},u_{k+j},\theta_k) \tag{8b}$$

$$40\ \text{dm}^3/\text{min}\le q_{k+j}\le 120\ \text{dm}^3/\text{min} \tag{8c}$$

$$290\ \text{K}\le T_{\text{c},k+j}\le 320\ \text{K} \tag{8d}$$

$$0.485\ \text{mol/dm}^3\le C_{A,k+j+1} \tag{8e}$$

$$j=0,1,\ldots,N-1$$

where

$$J=\sum_{j=0}^{N-1}\|x_{k+j}-\bar{x}\|_Q^2+\|u_{k+j}-\bar{u}\|_R^2+\|x_{k+N}-\bar{x}\|_Q^2$$

with $(\bar{x},\bar{u})=(\bar{C}_A,\ \bar{T}_\text{r},\ \bar{q},\ \bar{T}_\text{c})$ denoting the desired operating point and $\|s\|_M^2:=s^\top Ms$. The parameters used in the model and in the MPC are listed in the Supporting Information. Note that the lowest acceptable value of C_A, 0.485 mol/dm^3, is only 3% below the target concentration $\bar{C}_A=0.5$ mol/dm^3. The optimal control problem (8) is solved in a receding-horizon manner at every measurement sampling instant k given the state x_k, which is estimated by an extended Kalman filter.

Four MPC strategies are considered for the CSTR (7) with parametric model uncertainty. Strategy 1 uses a standard MPC with no form of parameter estimation; that is, the uncertain model parameters θ_k in the optimal control problem (8) are at all times equal to their nominal values determined at the MPC commissioning. In Strategy 2, once a noticeable drop in the control performance is observed, an open-loop identification

[†]For clarity of presentation, a deterministic process with no disturbances is considered in this case study.

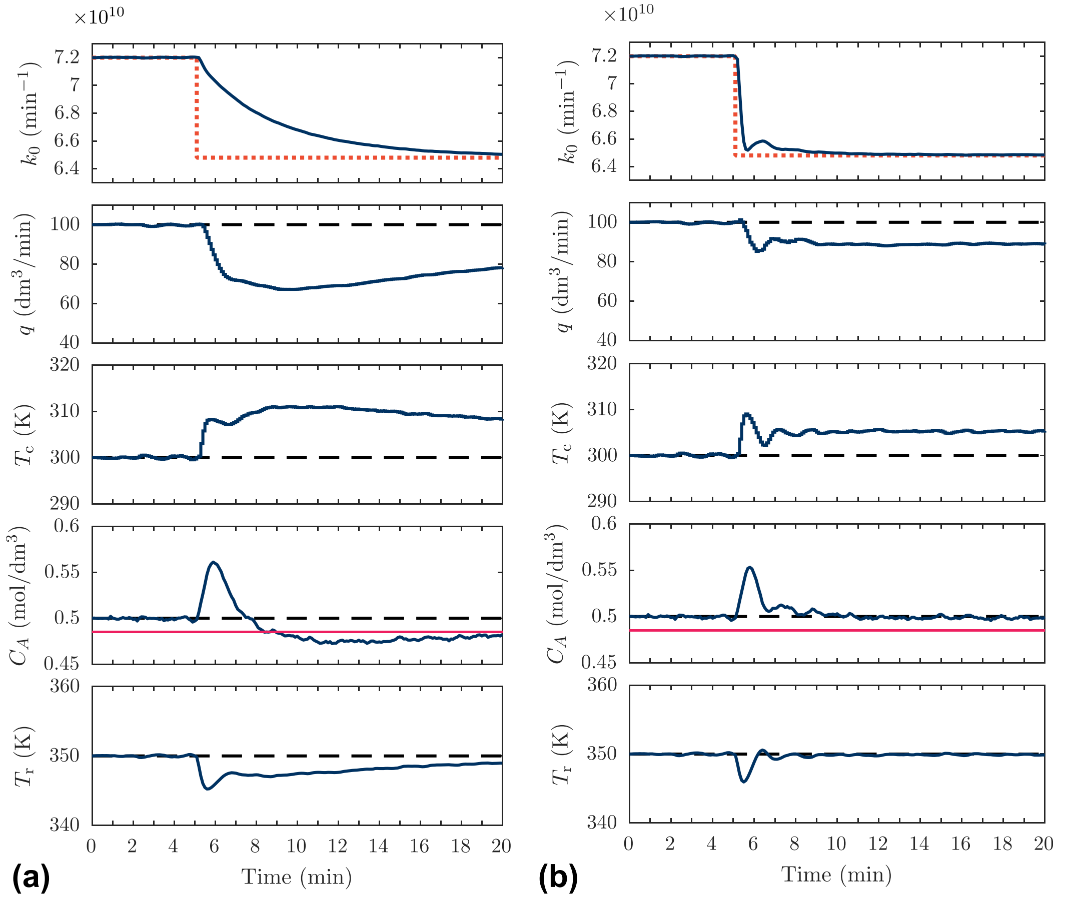

Figure 3. Closed-loop performance of (a) MPC with passive learning in Strategy 3 and (b) MPC with active learning in Strategy 4 under parametric model uncertainty.

From top to bottom, the plots show the estimates (solid blue) and true values (dashed red) of k_0, the control inputs q and T_c and the measured states C_A and T_r. In the plots of C_A, the red solid lines represent the constraint. The black dashed lines represent the setpoints. [Color figure can be viewed at wileyonlinelibrary.com]

experiment is performed to reidentify the model parameters offline and, accordingly, update the model in the optimal control problem (8). The parameter estimate resulting from the one-time reidentification is used at all subsequent times. Strategy 3 uses the same MPC formulation as the above strategies, except the uncertain model parameters are estimated online at every sampling time using an extended Kalman filter and are accordingly updated in the optimal control problem (8). Note that the online adaptation of the model in Strategy 3 relies on the closed-loop data generated by the feedback action of the controller, with no effort to increase the amount of information. This is therefore an MPC strategy with passive learning. Strategy 4 involves modifying the optimal control problem (8) to incorporate some form of probing in the control inputs to increase the information content of the closed-loop data used for the online parameter estimation with the extended Kalman filter. The controller adds probing to the input q that consists of steps of magnitude 1 dm^3/min when the prediction error is above a given threshold. Thus, this is an MPC strategy with active learning.

The performance of these four MPC strategies is evaluated when a change in the process at time 5 min causes an abrupt drop in the reaction constant k_0 from 7.2×10^{10} to 6.48×10^{10} min^{-1}, a reduction of 10%. The closed-loop performance of Strategies 1 and 2 are shown in Figure 2. Strategy 1 is incapable of keeping C_A close to the setpoint $\bar{C}_A$ after the drop in k_0, with C_A significantly below its lowest acceptable value (see Figure 2a). In Strategy 2, the controller is switched off 1.5 min after the drop in k_0 to perform an open-loop identification experiment from time 6.5 to 8.0 min. The identification experiment involves five steps of magnitude ±5.0 dm^3/min in the flow rate q (see Figure 2b).[‡] The open-loop data are used for offline least-squares estimation of the model parameters, with the estimate of k_0 within 2.4% of its true value. Despite using the reidentified model in the MPC after time 8.0 min, Figure 2b shows that the process has already drifted away from the setpoint (shown by a black dashed line) and the MPC is unable to restore the control performance in terms of adequate setpoint tracking. Furthermore, the MPC is only partially capable of meeting the product quality requirement on C_A, with minor but frequent violations of the constraint (8e) (shown by the red solid line).

The effect of online estimation of the uncertain model parameters on the MPC performance is presented in Figure 3. Figure 3a shows the performance of the MPC with passive learning in Strategy 3. After the abrupt change in k_0 at time 5.0 min, online parameter estimation enables adapting the model and reducing the uncertainty. However, the estimate of k_0 converges slowly to its true value, leading to a large offset in setpoint tracking and more importantly constraint violation for C_A. The slow convergence of k_0 can be attributed to the

‡For an overview of systematic approaches to optimal experiment design for parameter estimation, see Refs. 11 and 12.

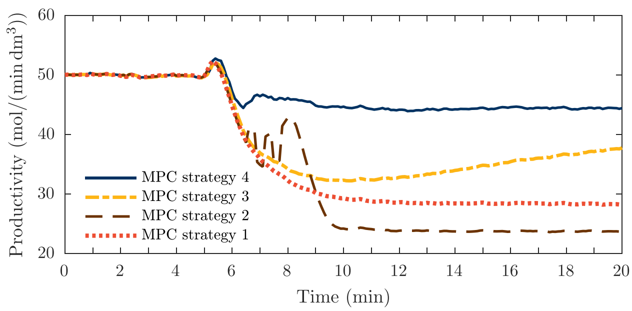

Figure 4. Productivity under standard MPC (Strategy 1), MPC with offline model adaptation (Strategy 2), MPC with passive learning (Strategy 3), and MPC with active learning (Strategy 4).

The oscillations visible in Strategies 2 and 4 result from the probing. [Color figure can be viewed at wileyonlinelibrary.com]

controller taking no active steps to improve the learning. That is, Strategy 3 incorporates only passive learning from the data generated by the feedback action of the MPC. Figure 3b demonstrates the performance of the MPC with active learning in Strategy 4. After detecting an increase in the model-output prediction error, the MPC with active learning increases its measure of model uncertainty by resetting the covariance in the extended Kalman filter to a large value. To reduce the model uncertainty, the controller then adds a probing signal to the control with the purpose of causing the process to generate more informative input-output data for parameter estimation. As a result, the estimate of k_0 quickly converges to its true value. This in turn enables the MPC to maintain C_A close to the setpoint and avoid constraint violation. The results shown in Figures 2 and 3 clearly demonstrate the superior performance of the MPC with active learning in effectively dealing with the parametric model uncertainty and consequently mitigating the control performance loss.

Figure 4 contrasts the performance of the four MPC strategies in terms of maintaining the desired productivity. Note that this productivity is closely related to the economics of the operation. While some productivity loss from the process change is unavoidable, the figure shows that the MPC with active learning (Strategy 4) results in the least loss. The standard MPC (Strategy 1) and the MPC with offline model adaptation (Strategy 2) lead to the highest losses among the four strategies. This is because the MPC in these strategies do not benefit from any form of online model adaptation to reduce model uncertainty. Note that Strategy 2 performs worse than Strategy 1 as a result of the excitation during open-loop operation, which takes the process so far from the operating point that the controller is unable to bring the process back. While the MPC with passive learning (Strategy 3) results in a lower loss in productivity, the model adaptation is still inadequate relative to the MPC with active learning (Strategy 4), which probes the process for active uncertainty reduction. This case study clearly illustrates the potential advantages of integrating MPC and probing for active learning to handle abrupt process changes.

MPC with Active Learning for Model-Structure Uncertainty

The focus of this section is the problem of uncertainty in the structure of the model used to control the process. This problem is posed using a set of candidate models $\mathcal{M}$, which is assumed to contain a model whose structural form matches that of the process at any given time. This type of model structure uncertainty may arise when the process can transition between different modes of operation in an unpredictable manner or can undergo abrupt structural changes. A typical form of abrupt change that can be captured by a change in model structure is process faults and failures, making the framework presented here particularly suitable for MPC with active fault detection and diagnosis.[5,63–65] Under this type of uncertainty, the probing is introduced for discriminating between multiple model candidates. The problem of MPC with active learning for model-structure uncertainty has received considerably less attention than the formulation that considers uncertain parameters.[66] In the following, we demonstrate the possible benefits of MPC with active learning under an uncertain model structure by letting the controller inject probing signals for improved discrimination between multiple models.

We consider a continuous bioreactor where the control objective is to maximize the process productivity, defined as the concentration multiplied by the dilution rate. The process dynamics are[67,68]

$$dX=(-DX+\mu X)dt+\sigma_X dw_X(t) \tag{9a}$$

$$dS=(D(S_f-S)-\frac{1}{Y_{X/S}}\mu X)dt+\sigma_S dw_S(t) \tag{9b}$$

$$dP=(-DP+(\alpha\mu+\beta)X)dt+\sigma_P dw_P(t) \tag{9c}$$

Here, X, S, and P are the concentrations of biomass, substrate, and product. The dilution rate D is the control input ($u(t)=D(t)$), and the volume is kept constant by ensuring the volumetric inlet and outlet flows are identical. The substrate concentration in the inlet feed is denoted by S_f, $Y_{X/S}$ is the yield of biomass per substrate consumed, α and β are the yield parameters, μ is growth rate of biomass, and $w_X(t)$, $w_S(t)$, and $w_P(t)$ are independent, zero-mean unit-variance Weiner processes scaled by standard deviations σ_X, σ_S, and σ_P.

Here, we consider three process models: the nominal model with saturation/Monod kinetics, a product inhibition model, in which the growth rate decreases with the product concentration,[67] and a model with saturation kinetics combined with a drop in the substrate inlet feed concentration. That is,

$$\mu=\frac{\mu_{max}S}{K_M+S} \quad \text{in Models } M_1 \text{ and } M_2 \tag{10a}$$

and

$$\mu=\frac{\mu_{max}(1-P/P_m)S}{K_M+S} \quad \text{in Model } M_3 \tag{10b}$$

where μ_{max} denotes the maximum growth rate, K_M is an affinity constant, and P_m is the maximum production rate. Models M_1 and M_2 can be interpreted as representing structural model uncertainty, with two different growth-model hypotheses that may be valid for different growth conditions. The third model represents a disturbance or fault that may occur during operation but is difficult to identify under regular feedback control:

$$S_f^{[2]}=0.3S_f^{[1]} \tag{11}$$

The model parameters, initial conditions, and operating conditions are listed in Table S3 in the Supporting Information.

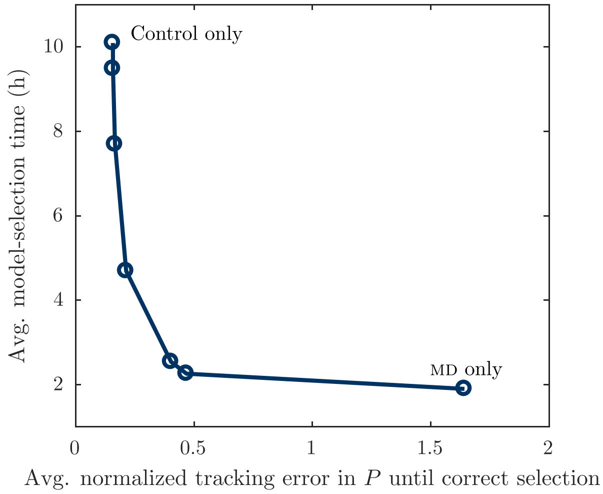

Figure 5. Pareto analysis for choosing an appropriate value of the model-discrimination weight, which increases from left to right in the figure from a pure passive-learning control strategy to a pure model discrimination type strategy (disregarding the stage cost for control).

[Color figure can be viewed at wileyonlinelibrary.com]

The parameter μ_{max} is assumed unknown and estimated together with the states, using an extended Kalman filter with the substrate and product concentrations both measured; $y_k=[S_k, P_k]^\top+v_k$. Note that there is no active learning of this parameter in the algorithm used in the results presented below.

The objective in the MPC is to minimize the stage cost $(P^{[i_k^*]}_{k+j}-\bar{P})^2+(u_{k+j}-\bar{D})^2$, where i_k^* is the index of the model with the highest probability. The terminal cost is $(P^{[i_k^*]}_{k+N}-\bar{P})^2$. The optimal operating point $(\bar{P},\bar{D})$ for maximizing productivity is determined offline.[68] The goal of active learning here is to faster and with higher confidence identify the structure that represents the current process behavior, while simultaneously maintaining production until intervention, if necessary, restores the process. We consider a scenario in which the process initially evolves according to Model M_1. At time $t=3$ h, a change occurs and the process starts evolving according to Model M_2. If and when the change is detected, the process is restored through intervention so that it again evolves according to Model M_1 after 0.5 h. A Bayesian recursion of the form (3) is used to compute the model probabilities P_M online.

The algorithm adopted here for MPC with active learning under model-structure uncertainty is presented in detail in Ref. 69, and involves actively probing the process for discriminating between a set of model hypotheses. The MPC algorithm uses the Bayesian decision rule for hypothesis selection to minimize the risk of selecting the wrong model, given the input and output data recorded from the process up until the current time.[70] This risk does not have a closed-form expression for general nonlinear models, and must therefore be approximated. Thus, active learning under model-structure uncertainty is incorporated in the MPC by explicitly augmenting the control cost with a weighted term that approximates an upper bound on the selection-error risk. The approximate bound[71] is computed online by linearizing the model around a state trajectory and predict the first two statistical moments of the state. The associated weight, referred to as a model-discrimination weight, is a tuning parameter in the control design. An appropriate value for this weight is chosen through Monte-Carlo simulation of the process with different values of the weight, as discussed in the following.

When explicitly augmenting a cost function for control with a term that provides some measure of model uncertainty, there is a form of trade-off between the control cost and the cost associated with reducing model uncertainty. Minimizing the augmented cost function is to balance minimization of the control cost, based on the current potentially wrong model, with minimization of the uncertainty, which improves that model. Here, the trade-off between reducing the selection-error risk to improve the model and minimizing the control cost based on the current model, which is potentially wrong structurally, is captured with the model-discrimination weight in the cost function of the MPC. We determine this weight for best overall performance through Pareto analysis. Figure 5 shows a Pareto front that demonstrates the trade-off between detecting a structural change and the tracking error in product concentration during the detection phase. The ordinate shows the average model-selection time, defined as the time elapsed between the occurrence of the structural change and correct model selection. The abscissa is the average closed-loop squared tracking error in product concentration, normalized through dividing by the number of sampling intervals elapsed between the structural change and selection of the corresponding model. Hence, this provides a measure of how much the product concentration tends to deviate at any given sampling time as a consequence of increasing the weight of probing for active learning. The top-left point in Figure 5 corresponds to a weight of zero, which means the controller makes no attempt to reduce the probability of error in the model selection. The bottom-right point corresponds to the other extreme: the stage and terminal cost for control are removed from the cost function and the controller only minimizes the selection-error risk. In this case, the algorithm makes no attempt at keeping P_k and D_k close to their desired values. For the values in between, the weight increases from left to right. Slightly increasing the weight from zero leads to a large improvement in the model-selection time with minimal increase in tracking error. This shows that a relatively small adjustment to the operating strategy can result in large gains in uncertainty reduction with negligible change in the standard control cost. Note that the potential gains from early detection of the structural change are not included in Figure 5. Thus, this figure does not imply any overall trade-off between reducing model-selection error and improving productivity.

A value of the model-discrimination weight that balances the stage cost and the probability of model selection error is used in the MPC algorithm with active learning to induce probing after a sudden drop in productivity signals a possible change. That is, the weight goes from zero to its preselected value, enabling the controller to probe the process for discrimination between the models and select the model that best describes the recorded measurements. This phase, with probing for determining the model with the correct structure, lasts until the probability of a model that was previously considered to not represent the process has the largest probability for a predefined amount of time. In other words, when one model retains the highest probability for a sufficiently long time, this model is selected by controller.

Figure 6 shows a simulation that compares the evolution of the probabilities P_M for the three model hypotheses with passive (a zero weight on model discrimination) and active

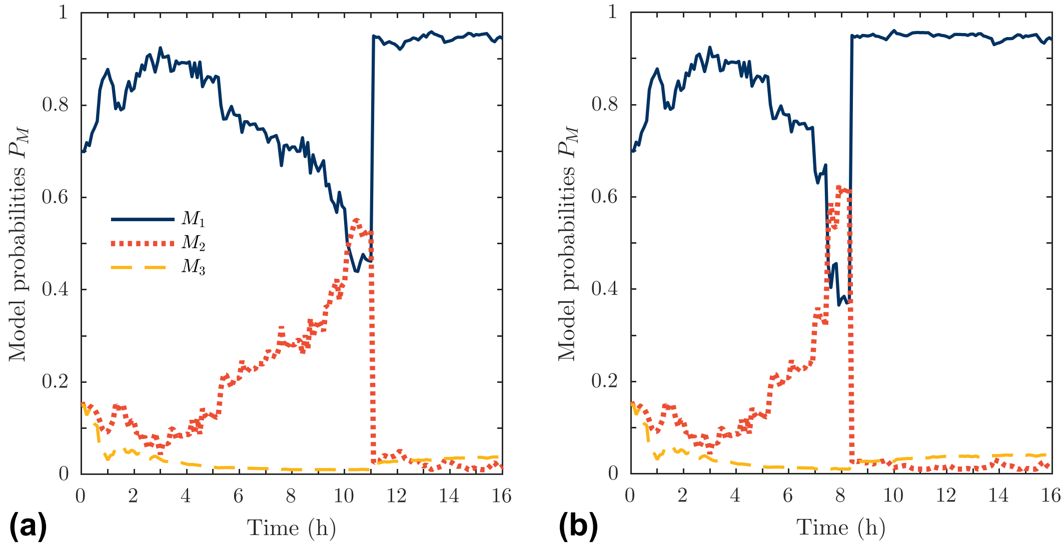

Figure 6. Comparison of model probabilities for the three model hypotheses using MPC with (a) passive and (b) active learning.

Active learning reduces the time required to select the correct model by 36% relative to passive learning. The jump in probability for Model M_1 occurs as the system is restored to its original structure. [Color figure can be viewed at wileyonlinelibrary.com]

(a positive weight) learning. Some time after the new model for the structural change is correctly selected the process is restored to its original structure. In this particular case, going from passive to active learning reduces the time to correctly identify the change in the process from 7.6 to 4.9 h, which is a 36% reduction.

The overall objective in the optimal control problem is to maximize productivity, defined as $P_k D_k$. Figure 7 shows the evolution of process productivity, comparing three control strategies: (i) a standard MPC with a static model structure and no updating of the model-structure probabilities (no learning), (ii) a controller identical to (i) but with a Bayesian estimator that updates the probabilities of the model candidates (passive learning), and (iii) a controller identical to (ii) but with a term added to the cost function that represents the risk of selecting the wrong model (inducing probing for active learning). The controller with no learning never recovers from the structural change as it makes no attempt to identify its occurrence. The controller with passive learning performs better, but not as well as the approach with active learning. The active learning is thus instrumental in bringing back the productivity since it facilitates selecting the new correct model faster.

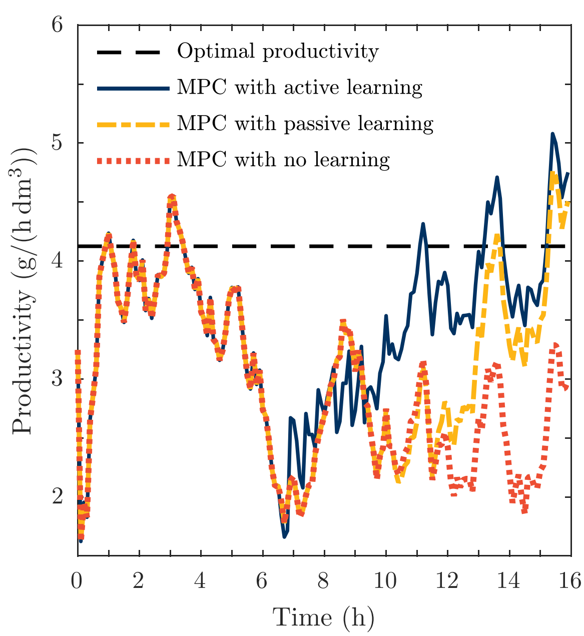

Figure 7. Comparison of productivity $P_k D_k$ in the cases of standard MPC with no learning, passive learning, and active learning.

The dashed black line represents the optimal productivity. [Color figure can be viewed at wileyonlinelibrary.com]

The extent to which productivity is improved by MPC with active learning relative to passive learning and no learning is summarized in Table 1. The productivity is determined through 20 Monte-Carlo simulations that each span 16 h. The MPC with active learning increases average productivity by 14% over passive learning and by 27% over no learning. This demonstrates that probing for active reduction of model-structure uncertainty has the potential to significantly improve control performance on average.

Future Research Directions

This article demonstrates that MPC with active learning under parametric or structural model uncertainty can mitigate significant loss in control performance from unanticipated sudden changes in the system. This control approach can advance the area of active fault diagnosis and fault-tolerant control.[66] Fast and high-confidence diagnosis of incipient faults under closed-loop control can prevent severe failures and facilitate safe and graceful degradation until appropriate intervention can take place. This can be of great importance in a wide range of applications, in particular where safety is imperative.

Table 1. Productivity per Sampling Instant, Averaged over 20 Monte-Carlo Simulations, for the MPC with Active Learning, Passive Learning, and No Learning

Active learning	Passive learning	No learning
3.59	3.16	2.83

However, theoretical and practical aspects of MPC with active learning for fault-tolerant control remain open and further research is required.

In the chemical process industry, optimal operation is commonly related to economic objectives. Here, MPC is generally implemented in a hierarchical control structure, where the "optimal" setpoints to the MPC layer are determined by *real-time optimization* (RTO) of an economic cost function.[72] The way to best incorporate active learning into this hierarchical control structure is an open question. MPC with active learning will generally improve the quality of the uncertain model (when possible and necessary), but the setpoints previously computed by the RTO may no longer be economically optimal for the updated model. In other words, active learning will improve tracking and constraint satisfaction at the MPC layer, but this may not necessarily lead to the best economic outcome of the operation. Yet, there may still be significant advantages to using the model obtained with active learning, in particular in terms of constraint satisfaction, until new economically optimal setpoints are recomputed by the RTO using the adapted model.

Combining the RTO and MPC layers in the hierarchical control structure has been an active area of research in recent years. One promising approach is economic MPC (EMPC), which directly optimizes the economics of the system subject to the dynamic model.[73,74] Incorporating active learning in EMPC is a potentially promising area for future research, as this formulation would be able to determine the true economic cost and benefit of improving the model. In this way, the EMPC with active learning could select an operation mode that optimally balances the "cost of learning" and the expected economic gains from improving the model.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 1705706.

Literature Cited

1. Morari M, Lee JH. Model predictive control: past, present and future. *Comput Chem Eng.* 1999;23(4–5):667–682.
2. Rawlings JB, Mayne DQ, Diehl M. *Model Predictive Control: Theory and Design*, 2nd ed. Madison, WI: Nob Hill Publishing, 2017.
3. Qin SJ, Badgwell TA. A survey of industrial model predictive control technology. *Control Eng Pract.* 2003;11(7):733–764.
4. Forbes MG, Patwardhan RS, Gopaluni RB. Model predictive control in industry: challenges and opportunities. In: *Proceedings of the IFAC International Symposium on Advanced Control of Chemical Processes*. Whistler, BC: The International Federation of Automatic Control, 2015:532–539.
5. Blanke M, Kinnaert M, Lunze J, Staroswiecki M. *Diagnosis and Fault-Tolerant Control*, 2nd ed. Berlin, Germany: Springer, 2006.
6. Zhou K, Doyle JC. *Essentials of Robust Control*. Upper Saddle River, NJ: Prentice Hall, 1999.
7. Åström KJ. *Introduction to Stochastic Control Theory*. New York: Academic Press, 1970.
8. Bodson M, Sastry SS. *Adaptive Control: Stability, Convergence and Robustness*. Englewood Cliffs, NJ: Prentice Hall, 1994.
9. Åström KJ, Wittenmark B. *Adaptive Control*, 2nd ed. Reading, MA: Addison-Wesley, 1995.
10. Van den Hof PMJ, Schrama JP. Identification and control—closed-loop issues. *Automatica.* 1995;31(12):1751–1770.
11. Gevers M. Identification for control: from the early achievements to the revival of experiment design. *Eur J Control.* 2005;11(4–5):335–352.
12. Hjalmarsson H. From experiment design to closed-loop control. *Automatica.* 2005;41(3):393–438.
13. Grimm G, Messina MJ, Tuna SE, Teel AR. Examples when nonlinear model predictive control is nonrobust. *Automatica.* 2004;40(10):1729–1738.
14. Bemporad A, Morari M. Robust model predictive control: a survey. In: Garulli A, Tesi A, editors. *Robustness in Identification and Control, Vol.* 245. Berlin, Germany: Springer, 1999:207–226.
15. Mayne DQ, Seron MM, Raković SV. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica.* 2005;41(2):219–224.
16. Mayne DQ. Model predictive control: recent developments and future promise. *Automatica.* 2014;50(12):2967–2986.
17. Kouvaritakis B, Cannon M. *Model Predictive Control: Classical, Robust and Stochastic*. London: Springer, 2016.
18. Cannon M, Kouvaritakis B, Raković SV, Cheng Q. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Trans Automat Control.* 2011;56(1):194–200.
19. Mesbah A. Stochastic model predictive control: an overview and perspectives for future research. *IEEE Control Syst.* 2016;36(6):30–44.
20. Heirung TAN, Paulson JA, O'Leary J, Mesbah A. Stochastic model predictive control—how does it work? *Comput Chem Eng.* DOI: 10.1016/j.compchemeng.2017.10.026
21. Mayne DQ, Michalska H. Adaptive receding horizon control for constrained nonlinear systems. In: *Proceedings of the IEEE Conference on Decision and Control*. San Antonio, TX: IEEE, 1993:1286–1291.
22. Adetola V, Guay M. Adaptive model predictive control for constrained nonlinear systems. In: *Proceedings of the IFAC World Congress, 2007*. Seoul, Korea: The International Federation of Automatic Control, 2008:1946–1951.
23. Tanaskovic M, Fagiano L, Smith RS, Goulart PJ, Morari M. Adaptive model predictive control for constrained linear systems. In: *Proceedings of the European Control Conference*. Zürich, Switzerland, 2013:382–387.
24. Feldbaum AA. Dual-control theory. I. *Autom Remote Control.* 1961; 21:874–880.
25. Feldbaum AA. Dual control theory. II. *Autom Remote Control.* 1961; 21:1033–1039.
26. Feldbaum AA. The theory of dual control. III. *Autom Remote Control.* 1961;22:1–12.
27. Feldbaum AA. The theory of dual control. IV. *Autom Remote Control.* 1961;22:109–121.
28. Feldbaum AA. *Optimal Control Systems*. New York, NY: Elsevier, 1965.
29. Wieslander J, Wittenmark B. An approach to adaptive control using real time identification. *Automatica.* 1971;7(2):211–217.
30. Anderson BDO. Adaptive systems, lack of persistency of excitation and bursting phenomena. *Automatica.* 1985;21(3):247–258.
31. Mareels I, Polderman JW. *Adaptive Systems: An Introduction*. Basel, Switzerland: Birkhäuser, 1996.
32. Mesbah A. Stochastic model predictive control with active uncertainty learning: a survey on dual control. *Annu Rev Control.* DOI: 10.1016/j.arcontrol.2017.11.001
33. Puterman ML. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ: Wiley, 2005.
34. Simon D. *Optimal State Estimation*. Hoboken, NJ: Wiley, 2006.
35. Kalman RE. A new approach to linear filtering and prediction problems. *J Basic Eng.* 1960;82(1):35–45.
36. Bar-Shalom Y, Tse E. Dual effect, certainty equivalence, and separation in stochastic control. *IEEE Trans Automat Control.* 1974;19(5):494–500.
37. Bellman R. *Dynamic Programming*. Princeton, NJ: Princeton University Press, 1957.
38. Kalman RE. Contributions to the theory of optimal control. *Bol Soc Mat Mex.* 1960;5:102–119.
39. Powell WB. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Hoboken, NJ: Wiley, 2011.
40. Bertsekas DP. *Dynamic Programming and Optimal Control: Approximate Dynamic Programming*, 4th ed., *Vol.* 2. Belmont, MA: Athena Scientific, 2012.
41. Sternby J. A simple dual control problem with an analytical solution. *IEEE Trans Automat Control.* 1976;21(6):840–844.
42. Åström KJ, Helmersson A. Dual control of an integrator with unknown gain. *Comput Math Appl.* 1986;12(6):653–662.
43. Ben-Tal A, El Ghaoui L, Nemirovski A. *Robust Optimization*. Princeton, NJ: Princeton University Press, 2009.
44. Bertsimas D, Brown DB, Caramanis C. Theory and applications of robust optimization. *SIAM Rev.* 2011;53(3):464–603.

45. Filatov NM, Unbehauen H. Survey of adaptive dual control methods. *IEE Proc Control Theory Appl*. 2000;147(1):118–128.
46. Lee JM, Lee JH. An approximate dynamic programming based approach to dual adaptive control. *J Process Control*. 2009;19(5): 859–864.
47. Bayard DS, Schumitzky A. Implicit dual control based on particle filtering and forward dynamic programming. *Int J Adapt Control Signal Process*. 2010;24(3):155–177.
48. Bavdekar VA, Mesbah A. Stochastic model predictive control with integrated experiment design for nonlinear systems. In: *Proceedings of the IFAC Conference on Dynamics and Control of Process Systems*. Trondheim, Norway: The International Federation of Automatic Control, 2016:49–54.
49. Bavdekar VA, Ehlinger V, Gidon D, Mesbah A. Stochastic predictive control with adaptive model maintenance. In: *Proceedings of the IEEE Conference on Decision and Control*. Las Vegas, NV: IEEE, 2016:2745–2750.
50. La HC, Potschka A, Schlöder JP, Bock HG. Dual control and online optimal experimental design. *SIAM J Sci Comput*. 2017;39(4):B640–B657.
51. Heirung TAN, Ydstie BE, Foss B. Dual adaptive model predictive control. *Automatica*. 2017;80:340–348.
52. Houska B, Telen D, Logist F, Van Impe J. Self-reflective model predictive control. *SIAM J Control Optim*. 2017;55(5):2959–2980.
53. Annergren M, Larsson CA, Hjalmarsson H, Bombois X, Wahlberg B. Application-oriented input design in system identification: optimal input design for control. *IEEE Control Syst*. 2017;37:31–56.
54. Marafioti G, Bitmead RR, Hovd M. Persistently exciting model predictive control. *Int J Adapt Control Signal Process*. 2014;28(6):536–552.
55. Larsson CA, Rojas CR, Bombois X, Hjalmarsson H. Experimental evaluation of model predictive control with excitation (MPC-X) on an industrial depropanizer. *J Process Control*. 2015;31:1–16.
56. Findeisen R, Imsland L, Allgöwer F, Foss B. State and output feedback nonlinear model predictive control: an overview. *Eur J Control*. 2003;9(2–3):190–206.
57. Imsland L, Findeisen R, Bullinger E, Allgöwer F, Foss B. A note on stability, robustness and performance of output feedback nonlinear model predictive control. *J Process Control*. 2003;13(7):633–644.
58. Goulart PJ, Kerrigan EC. Output feedback receding horizon control of constrained systems. *Int J Control*. 2007;80(1):8–20.
59. Artstein Z, Raković SV. Set invariance under output feedback: a set-dynamics approach. *Int J Syst Sci*. 2011;42(4):539–555.
60. Hovd M, Bitmead RR. Interaction between control and state estimation in nonlinear MPC. In: *Proceedings of the IFAC Dynamics and Control of Process Systems*. Cambridge, MA: The International Federation of Automatic Control, 2004:119–124.
61. Knudsen BR, Alessandretti A, Jones CN. Sensor fault tolerance in output feedback nonlinear model predictive control. In: *Proceedings of Control and Fault-Tolerant Systems*. Barcelona, Spain: IEEE, 2016:711–716.
62. Adetola V, Guay M. Robust adaptive MPC for systems with exogeneous disturbances. In: *Proceedings of the IFAC International Symposium on Advanced Control of Chemical Processes*. Istanbul, Turkey: The International Federation of Automatic Control, 2009: 255–260.
63. Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri SN. A review of process fault detection and diagnosis. Part I: quantitative model-based methods. *Comput Chem Eng*. 2003;27(3):293–311.
64. Venkatasubramanian V, Rengaswamy R, Kavuri SN. A review of process fault detection and diagnosis. Part II: qualitative models and search strategies. *Comput Chem Eng*. 2003;27(3):313–326.
65. Venkatasubramanian V, Rengaswamy R, Kavuri SN, Yin K. A review of process fault detection and diagnosis. Part III: process history based methods. *Comput Chem Eng*. 2003;27(3):327–346.
66. Heirung TAN, Mesbah A. Stochastic nonlinear model predictive control with active model discrimination: a closed-loop fault diagnosis application. *IFAC-PapersOnLine*. 2017;50(1):15934–15939.
67. Agrawal P, Koshy G, Ramseier M. An algorithm for operating a fed-batch fermentor at optimum specific-growth rate. *Biotechnol Bioeng*. 1989;33(1):115–125.
68. Henson MA, Seborg DE. Nonlinear control strategies for continuous fermenters. *Chem Eng Sci*. 1992;47(4):821–835.
69. Heirung TAN, Santos TLM, Mesbah A. Model predictive control with active learning for stochastic systems with structural model uncertainty: online model discrimination. *J Process Control*. Under revision.
70. Hellman ME, Raviv J. Probability of error, equivocation, and the Chernoff bound. *IEEE Trans Inf Theory*. 1970;16(4):368–372.
71. Blackmore L, Williams BC. Finite horizon control design for optimal discrimination between several models. In: *Proceedings of the IEEE Conference on Decision and Control*. San Diego, CA: IEEE, 2006:1147–1152.
72. Darby ML, Nikolaou M. MPC: current practice and challenges. *Control Eng Pract*. 2012;20(4):328–342.
73. Rawlings JB, Amrit R. Optimizing process economic performance using model predictive control. In: Magni L, Raimondo DM, Allgöwer F, editors. *Nonlinear Model Predictive Control, Vol. 384 of Lecture Notes in Control and Information Sciences*. Berlin, Germany: Springer, 2009:119–138.
74. Amrit R, Rawlings JB, Biegler LT. Optimizing process economics online using model predictive control. *Comput Chem Eng*. 2013;58: 334–343.

Manuscript received Sep. 30, 2017, and revision received Mar. 29, 2018.