# On The Way To Autonomous Model Predictive Control: A distillation column simulation study<sup>1</sup>

M. Annergren \* D. Kauven \*\* C. A. Larsson \* M.G. Potters \*\*\* Q. Tran \*\*\*\* L. Özkan \*\*\*\*

\* Automatic Control Lab and ACCESS, School of Electrical Engineering, KTH, SE-100 44 Stockholm, Sweden, (e-mail: chrisla@kth.se, marann@kth.se)
\*\* RWTH Aachen University AVT - Process Systems Engineering, Turmstr. 46, 52064 Aachen, Germany, (e-mail:david.kauven@avt-rwth.aachen.de)
\*\*\* Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands, (e-mail: M.G.Potters@tudelft.nl)
\*\*\*\* Department of Electrical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, (e-mail: N.Q.Tran@tue.nl, l.ozkan@tue.nl)

Abstract Model Predictive Control (MPC) is a powerful tool in the control of large scale chemical processes and has become the standard method for constrained multivariable control problems. Hence, the number of MPC applications is increasing steadily and it is being used in application domains other than petrochemical industries. A common observation by the industrial practitioners is that success of any MPC application requires not only efficient initial deployment but also maintenance of initial effectiveness. To this end, we propose a novel high level automated support strategy for MPC systems. Such a strategy consists of components such as performance monitoring, performance diagnosis, least costly closed loop experiment design, re-identification and autotuning. This work presents the novel technological developments in each component and demonstrates them on a distillation column case study. We show that automated support strategy restores nominal performance after a performance drop is detected and takes the right course of action depending on its cause.

# 1. INTRODUCTION

Since its introduction in Richalet et al. (1976), MPC has been accepted and implemented as a standard tool to drive and maintain processes at economically optimal operating conditions. This widespread success is due to its ability to handle complex interacting systems and system constraints resulting in increase in productivity, improved product quality, safe operation and reduction in costs. Hence, the number of implementations are increasing steadily and extending to application domains other than traditional process industries. Despite the growing number of implementations and substantial benefits, MPC applications exhibit a low level of operational efficiency. One of the major reasons for this is the lack of maintenance. As quoted in Bauer and Craig (2008) "A technical shortcoming of almost all MPC systems is that if left unsupervised the performance will deteriorate over time". This performance degradation may be due to model deterioration, varying operating conditions or a change

 $^1$  The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 257059, The 'Autoprofit' project (www.fp7-autoprofit.eu).

in disturbance characteristics (Harrison and Qin (2009)). Additionally, lack of skilled process engineers who can support this kind of application is a challenge. Such factors, in most cases, lead to switching off MPC completely and returning to manual operation. Therefore, the successful implementation requires easy to use technology that can increase operational efficiency and automate maintenance of MPC systems.

This topic of maintaining operational efficiency has also been previously studied by Lee et al. (2008) and by Zhu and Padwarthan (2012). In this work, we develop this area further to a high level automated support strategy for performance monitoring, diagnosis and performance recovery. The support strategy, presented in Fig. 1, is built around the main concepts of least costly experiment design, just in-time maintenance of models and applications and automatic decision making based on an economic criteria. It brings together novel techniques in performance diagnosis and monitoring (Modén (2012a,b); Potters et al. (2012); Mesbah and Bombois (2011, 2012)), experiment design (Bombois et al. (2006); Hjalmarsson (2009); Larsson et al. (2013)), and MPC tuning (Özkan et al. (2012); Tran et al. (2012)). This paper is organized as follows. Section 2 provides preliminary information. Section 3 de-



Figure 1. Automated support strategy (The steps considered in this paper are in bold)

scribes each of the individual components of the support strategy namely, performance monitoring and diagnosis; tuning; and re-identification. In Section 4 we illustrate it on a binary distillation column case study. Results followed by conclusion and discussion are presented in Sections 5 and 6 respectively.

#### 2. PRELIMINARIES AND NOTATION

Let  $\mathbb{R}$ ,  $\mathbb{R}_+$  denote the field of real numbers and the set of positive reals respectively. Consider the true linear time invariant discrete time system  $\Sigma_{true}$  represented by

$$\Sigma_{true} := \{ y(t) = G_0(z)u(t) + H_0(z)e(t)$$
 (1)

and  $\Sigma_{mod}$  as the approximate model of the true system at commissioning represented by

$$\Sigma_{\rm mod} := \{ y(t) = G_{mod}(z)u(t) + H_{mod}(z)e(t)$$
 (2)

We denote the controlled system  $\Sigma_C(G_0, H_0, C(G_{mod}))$ as the closed loop system for the true plant in 1 where  $C(G_{mod})$  is defined as the MPC control law which uses  $G_{mod}$  in its algorithm. With the notation given above, we provide information on the fundamental concepts in the automated support strategy.

# 2.1 Performance measure

It is the concept that unifies the different components of the automated support strategy. Based on the outcome of this measure, necessary decisions are taken followed by the implementation of the corresponding actions. In this paper, we use a performance measure based on K key performance indicators,  $y_i$ , of the system that we want to be as close as possible to the given constraints  $b_i$ . The measure is given by

$$\mathcal{J}(t, G, H, C(G_{mod})) = \sum_{i=1}^{K} c_{1,i} P_{viol,i} + \sum_{i=1}^{K} c_{2,i} |\bar{y}_i - b_i|$$
(3)

where  $c_1$ ,  $c_2$  are user-defined weightings.  $P_{viol,i}$  and  $\bar{y}_i$ are the probability of violation and the mean value of key performance indicator  $y_i$  respectively, computed over  $N_{win}$  data points preceding time instant t.  $P_{viol,i}$  is defined as the ratio of data points exceeding constraint  $b_i$  and the total number of data points  $N_{win}$ , at time instant t. In the case of parameterized models we can equivalently consider a performance measure in the form of  $\mathcal{J}(t, \Sigma_C(G, H, C(\theta)))$  where  $\mathcal{J} \in \mathbb{R}_+$  and the control design is based on the parameter vector  $\theta$ .

The performance of the closed loop  $\Sigma_C(G, H, C(G_{mod}))$  is satisfactory when

$$\mathcal{J}(t, \Sigma_C(G, H, C(G_{mod}))) \le \beta \ \forall t \tag{4}$$

with  $\beta$  some carefully chosen, application dependent threshold. The quality of a model used in a control application will influence the performance of the control application. We use the performance measure to define the set of admissible parameters as

$$\Theta_{app}(\beta) \triangleq \{\theta : \mathcal{J}(t, \Sigma_C(G, H, C(\theta))) \le \beta\}$$
(5)

For performance diagnosis, we use the set of admissible models, defined as

 $\mathcal{D}_{adm} = \{G(z,\theta) \mid \overline{\mathcal{J}}(G(z,\theta), H_{mod}, C(G_{mod})) \leq \beta\} \quad (6)$ where  $H_{mod}$  is the disturbance model at commissioning, and  $G(z,\theta)$  represents a model with the same order as  $G_{mod}$ . Here,  $\overline{\mathcal{J}}(\cdot)$  is the time-averaged value of  $\mathcal{J}(t, G(z,\theta), H_{mod}, C(G_{mod}))$  over a period sufficiently long to obtain a static value. This set includes all plant models  $G(z,\theta)$  that have a performance below the threshold value  $\beta$ , under the original disturbance model. At commissioning, we assume that the performance is satisfactory, hence  $\mathcal{J}(t, G_0, H_0, C(G_{mod})) \leq \beta$ . In this way, we are able to distinguish between two important controlrelevant changes in a process: changes due to altered plant dynamics and disturbance characteristics changes. This will be explained in more detail in Section 3.3.

#### 2.2 Model predictive control

MPC represents a set of control strategies in which a finite time horizon optimal control problem is solved online at each time instant. Only the first value of the solution is implemented and this procedure is repeated in the next time instant. The strategy requires a model of the process to be controlled and is able to deal explicitly with MIMO plants and system constraints. A general formulation of MPC is given by

$$\min_{\{u(t+k)\}_{k=1}^{N_u}} J(t) = \sum_{k=1}^{N_y} \|y(t+k) - r(t+k)\|_{\mathbf{Q}}^2 + \sum_{k=1}^{N_u} \|\Delta u(t+k)\|_{\mathbf{R}}^2,$$
(7)

s.t. 
$$\Sigma := \begin{cases} x(t+1) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases}$$
 (8)

$$u_{min} \le u(t+k) \le u_{max}, \ k = 1, \dots, N_u$$
(9)  
$$y_{min} \le y(t+k) \le y_{max}, \ k = 1, \dots, N_u$$

$$y_{min} \ge y(\iota + \kappa) \ge y_{max}, \ \kappa = 1, \dots, N_y$$
(10)

where r(t) is a reference trajectory,  $\Delta u(t) = u(t) - u(t-1)$ is the control update,  $N_y$  and  $N_u$  are the prediction and control horizons, **Q** and **R** are output and input weighting matrices respectively.

#### 2.3 Prediction error identification

Consider a discrete time, linear time-invariant dynamic system described by a model on innovations form of

$$\Sigma := \begin{cases} x(t+1) = A(\theta)x(t) + B(\theta)u(t) + K(\theta)e(t), \\ y(t) = C(\theta)x(t) + e(t) \end{cases}$$
(11a)

where  $\theta \in \mathbb{R}^{n_{\theta}}$  is the same as the parameter vector in (5),  $x(t) \in \mathbb{R}^{d}$  is the state vector,  $u(t) \in \mathbb{R}^{m}$  is the input,  $y(t) \in \mathbb{R}^{p}$  the output, and  $e(t) \in \mathbb{R}^{p}$  the innovations: a zero-mean, random process with covariance matrix  $\mathbb{E}\left\{e(t)e^{T}(s)\right\} = \Lambda_{e}\delta_{t,s}$ . We assume that there exists a vector  $\theta_{o}$  which corresponds to the true system parameters. In system identification, we want to find the best possible estimate, according to some measure, of the true system parameters  $\theta_{o}$ . For this purpose, we use the prediction error method with quadratic cost and  $\hat{\theta}_{N}$ denotes the estimate resulting from N samples of input– output data,  $Z_{N} = \{u(t), y(t)\}_{t=1}^{N}$ , (Ljung, 1999).

For an unbiased estimator the Cramér-Rao inequality provides a lower bound on the covariance matrix for the parameter estimation error. Given data collected from time m to time n, this bound is given by the inverse of the Fisher information matrix  $\mathcal{I}_m^n(\theta_o)$ , where the latter is defined by

$$\mathcal{I}_{m}^{n}(\theta_{o}) = \sum_{t=m}^{n} \mathbb{E}\left\{\psi(t,\theta)\Lambda_{e}^{-1}\psi^{T}(t,\theta)\right\}\Big|_{\theta=\theta_{o}}, \qquad (12)$$

$$\psi_i(t,\theta) = \frac{\mathrm{d}\hat{y}(t)}{\mathrm{d}\theta_i} \tag{13}$$

$$\psi(t,\theta) = \left[\psi_1(t,\theta) \cdots \psi_{n_\theta}(t,\theta)\right]^T \tag{14}$$

where  $\hat{y}(t)$  is the one-step-ahead predictor. Under very general conditions Ljung and Caines (1979), it holds that

$$\mathcal{I}_m^n(\theta_o)^{1/2}(\hat{\theta}_N - \theta_o) \in As\mathcal{N}(0, I)$$
(15)

meaning that the random variable  $\mathcal{I}_m^n(\theta_o)^{1/2}(\hat{\theta}_N - \theta_o)$  converges in distribution to the normal distribution with zero mean and covariance matrix *I*. This implies that

$$\hat{\theta}_N \in \mathcal{E}_{id}(\alpha) \triangleq \left\{ \theta : \left[\theta - \theta_o\right]^T \mathcal{I}_m^n(\theta_o) \left[\theta - \theta_o\right] \le \chi_\alpha^2(n_\theta) \right\}$$
(16)

with probability  $\alpha$  as  $N = n - m \to \infty$ . Here  $\chi^2_{\alpha}(n_{\theta})$  is the  $\alpha$ -percentile of the  $\chi^2$ -distribution with  $n_{\theta}$  degrees of freedom. We call  $\mathcal{E}_{id}(\alpha)$  the *identification ellipsoid*. The result in (16) is crucial for the re-identification procedure described in Section 3.5.

Note that the identification ellipsoid in (16) depends on the Fisher information matrix defined in (12) and on the unknown true parameter vector  $\theta_o$ . Instead of evaluating the expected value in the expression of the Fisher information matrix, we approximate it by

$$I_m^n(\theta_o) = \sum_{t=m}^n \psi(t,\theta) \Lambda_e^{-1} \psi^T(t,\theta) \bigg|_{\theta=\theta_o}$$
(17)

In fact, it holds under fairly mild conditions that

$$\lim_{n \to \infty} \frac{1}{n} I_m^n(\theta_o) = \lim_{n \to \infty} \frac{1}{n} \mathcal{I}_m^n(\theta_o)$$
(18)

Instead of using the true parameter vector  $\theta_o$ , we use an initial estimate  $\hat{\theta}_{init}$ . This estimate can be obtained from the model used at commissioning or from an identification experiment that is much cheaper than the one performed during re-identification.

# 3. AUTOMATED SUPPORT STRATEGY AND ITS COMPONENTS

The automated support strategy consists of four main components written in bold as shown in Fig. 1. These are performance monitoring, performance diagnosis, closed loop identification and automated tuning.

Performance monitoring is a crucial step and is discussed in Section 3.2. Once the performance drop is detected, the performance diagnosis tool is called upon (Section 3.3). This tool distinguishes if the performance drop is due to a plant change or a disturbance characteristics change. Subsequent actions, being re-identification and tuning, are discussed in Section 3.5 and Section 3.4 respectively.

#### 3.1 Performance Monitoring

This component uses the performance measure defined in (3) to decide whether the closed loop system  $\Sigma_C(G_{mod}, H_{mod}, C(G_{mod}))$  is satisfactory or not. The satisfactory performance is the case when  $\overline{\mathcal{J}}(\Sigma_C(G_0, H_0, C(G_{mod}))) \leq \beta$ 

### 3.2 Performance Monitoring

This component uses the performance measure defined in (3) to decide, using definition (4), whether the performance of the closed-loop system  $\Sigma_C(G_{mod}, H_{mod}, C(G_{mod}))$  is satisfactory or not.

### 3.3 Performance Diagnosis

This component is triggered when a performance drop is observed and used to accurately distinguish between a performance drop due to a plant change or due to a disturbance change. Base layer problems such as jammed or leaking valves are excluded. The diagnosis problem decides between the following hypotheses when a performance drop is observed:

$$\mathcal{H}_0: G_0(z) \in \mathcal{D}_{adm} \mathcal{H}_1: G_0(z) \notin \mathcal{D}_{adm},$$
 (19)

where  $\mathcal{H}_0$  ( $\mathcal{H}_1$ ) corresponds to a performance drop caused by a change in disturbance characteristics (plant change). To be able to discriminate between the two hypotheses stated in (19), we identify the unknown true system (1) in closed-loop operation with the existing MPC controller  $C(G_{mod})$ . To this end, an external signal  $r(t) \in \mathbb{R}^m$  is added to the input of the system:  $u(t) = u_{MPC}(t) + r(t)$  where  $u_{MPC}(t)$  is the input signal computed by the MPC controller. By applying the excitation signal r(t) for  $(t = 1, \ldots, N)$  to the closed-loop system and measuring the signals  $Z_N = \{u(t), y(t) | t = 1, ..., N\}$ , a model  $\{G(z, \hat{\theta}_N), H(z, \hat{\theta}_N)\}$  of the true system is identified using prediction error identification.

The identified model  $G(z, \hat{\theta}_N)$  is then exploited to choose between hypotheses  $\mathcal{H}_0$  or  $\mathcal{H}_1$ . We recall that the identification experiment for the diagnostic tool should be cheap. Consequently,  $\hat{\theta}_N$  has a large uncertainty compared to the size of  $\mathcal{D}_{adm}$ . To accommodate for the large uncertainty of  $\hat{\theta}_N$  inherent to a cheap experiment, we make use of its uncertainty ellipsoid (16). We know that  $\theta_0$  lies in  $\mathcal{E}_{id}$ with probability  $\alpha$ . Thus, for a given  $\alpha$ , we can compute the fraction  $\kappa$  of transfer function  $G(\theta)$ , with  $\theta \in \mathcal{E}_{id}$ , that lie inside  $\mathcal{D}_{adm}$ .

Inspired by Tyler and Morari (1996), and using the idea described above, we can now define a more careful decision rule. The hypothesis  $\mathcal{H}_1$  will only be chosen as the correct hypothesis if the identified model  $G(z, \hat{\theta}_N)$  is not only outside  $\mathcal{D}_{adm}$ , but is far outside  $\mathcal{D}_{adm}$ . With far, we mean that  $\kappa \ll 1$ . The decision rule becomes

$$G(z, \hat{\theta}_N) \in \mathcal{D}_{adm} \Rightarrow \text{choose} \quad \mathcal{H}_0$$
  

$$G(z, \hat{\theta}_N) \notin \mathcal{D}_{adm}, \text{ but close to } \mathcal{D}_{adm} \Rightarrow \text{choose} \quad \mathcal{H}_0.$$
  

$$G(z, \hat{\theta}_N) \notin \mathcal{D}_{adm} \text{ and far outside } \mathcal{D}_{adm} \Rightarrow \text{choose} \quad \mathcal{H}_1.$$
  
(20)

The numerical procedure of the above computation is as follows. Once an identified model  $G(\hat{\theta}_N)$  has been obtained, we randomly pick a large number of  $\theta$ 's from its associated set  $\mathcal{E}_{id}$ , where  $\alpha$  is chosen by the user. For each one of these  $\theta$ 's, we verify via (6) whether  $G(\theta) \in \mathcal{D}_{adm}$ . This allows us to calculate the fraction  $\kappa$ , which is subsequently used to deduce which of the hypotheses is correct according to (20).

The data  $Z_N$  obtained from the cheap identification experiment is passed on to the next steps in the support strategy. This information is used as a prior, and keeps the overall costs as low as possible.

#### 3.4 Automated Tuning for MPC

The tuning component in the automated support strategy can be either used at commissioning or when new tuning is required in the course of MPC automatic performance maintenance. The latter is implemented when the diagnosis test detects a disturbance change or a new model is identified because the performance drop is due to the changes in the plant dynamics. Only the weighting matrices  $\mathbf{Q}$  and  $\mathbf{R}$  in (7) are considered as tuning parameters. We also assume that the constraints are inactive and the disturbance energy distribution has low-pass characteristics. The auto-tuning method is based on the impact of modeling uncertainty on closed-loop performance as described in Tran et al. (2012). Without any modeling errors, increasing the closed loop bandwidth of the system results in a decrease in the output variance. On the contrary, in the presence of modeling errors, increasing the bandwidth of the closed-loop system from some frequency onwards will lead to an increase in the output variance.

This relation is illustrated in Fig.2 on the quadruple-tank system from (Johansson (1997)).



Figure 2. Relation between closed-loop bandwidth and output variance.

This relation shows that there exists a specific bandwidth  $(\omega_{ds})$  where the balance between nominal performance and robustness is achieved. Hence, the aim of the auto-tuning method is to find this bandwidth. The MPC auto-tuning method consists of the following steps:

- (1) Initialization: Determine a certain initial bandwidth based on disturbance characteristics and available information on modeling errors. Find the corresponding MPC tuning parameters. Fix the prediction horizon so as to cover the main dynamics of the open-loop system and the control horizon based on computational capacity.
- (2) Automation: Increase the bandwidth. Find the corresponding MPC tuning parameters.
- (3) Monitoring: Monitor the output variance. Compare it with the previous tuning. If the variance decreases, keep increasing the bandwidth. If the output variance increases, the previous tuning is assumed to be optimal.

In the auto-tuning procedure, we need to relate the selected bandwidth to the tuning parameters of MPC. The closed-loop bandwidth is determined by the weighting matrices  $\mathbf{Q}$  and  $\mathbf{R}$  as shown in Tran et al. (2012). In the unconstrained case, the control input of MPC can be considered as a state feedback (Maciejowski (2002)) and its sensitivity and complimentary sensitivity functions can be computed. The cut-off frequencies of the sensitivity function define the bandwidth of the closed-loop system which determines the capacity of rejecting disturbances. To find the tuning parameters  $\mathbf{Q}$  and  $\mathbf{R}$  that give a certain desired bandwidth, we formulate the following optimization problem:

$$\min_{\mathbf{Q},\mathbf{R}} (\omega_{ds} - \omega_{cs}(\sigma_i(\mathbf{Q},\mathbf{R})))^2$$
(21)

### 3.5 Re-identification of the model

The re-identification component is a closed loop identification method which is referred to as MPC-X (MPC with excitation) and is presented in Larsson et al. (2013). The excitation signal requirements are introduced as an extra constraint in the MPC formulation. In this section, we explain the identification and outline the underlying least costly identification concepts. Application set The previously defined application set (5) is typically non convex and difficult to work with. Therefore, we approximate the application set with an ellipsoid given by

 $\Theta_{app}(\beta) \approx \mathcal{E}_{app}(\beta) \triangleq \left\{ \theta : \theta^T A \theta + 2b^T \theta + c \leq 0 \right\}.$ (22) The ellipsoid (22) is not necessarily centered at  $\theta_o$ . The ellipsoid could come from a Taylor approximation of  $\mathcal{J}$ evaluated at  $\theta_o$  or from fitting it to a number of points  $\theta$ fulfilling the inequality in (5). The approximation method chosen determines the matrix A, the vector b and the scalar c in (22).

Applications oriented input design The objective of applications oriented input design is to provide an input signal which, when used in an identification experiment, gives a model that fulfills any performance requirements on the application when used in the control design, e.g., as the model in MPC. To achieve this, we require

$$\mathcal{E}_{id} \subseteq \mathcal{E}_{app}.\tag{23}$$

This inclusion implies that the estimated model parameters will lie in the set of acceptable parameters with the specified probability  $\alpha$ . By the S-procedure, the inclusion (23) holds if and only if there exists a  $\tau > 0$ , such that

$$\begin{bmatrix} I_1^N(\hat{\theta}_{init}) & -I_1^N(\hat{\theta}_{init})\hat{\theta}_{init} \\ -\hat{\theta}_{init}^T I_1^N(\hat{\theta}_{init}) & \hat{\theta}_{init}^T I_1^N(\hat{\theta}_{init})\hat{\theta}_{init} - \chi_{\alpha}^2(n_{\theta}) \end{bmatrix} \succeq \tau \begin{bmatrix} A & b \\ b^T & c \end{bmatrix}$$
(24)

as shown in Boyd and Vandenberghe (2003).

Receding horizon implementation and MPC-X We include the (24) as an additional constraint in the receding horizon implementation of MPC given in (7) for a plant described as in 11 where we assume that u(t + k) = u(t + k - 1) for  $k > N_u$ . This formulation of the MPC is referred as MPC-X, non-convex and typically difficult to solve. Therefore we make a convex relaxation based on LMI lifting and dropping a troublesome rank constraint on the resulting variables. The relaxation is detailed in Larsson et al. (2013) and based on the works of Manchester (2010); Luo et al. (2010).

We make two comments regarding the MPC-X formulation. First, even though the formulation allows for different control and prediction horizons, we will only consider the case  $N_u = N_y$ . In the (most common) case of  $N_y > N_u$ some assumption on the inputs beyond the control horizon has to be made. Typically they are set to be constant or zero, neither choice is suitable for the case where excitation is desired. Second, since the information matrix is constructed as the sum of  $N_u$  rank 1 matrices,  $N_u$  has to be at least as long as the rank of the matrix on the right hand side of the excitation constraint (24).

# 4. IMPLEMENTATION OF AUTOMATED SUPPORT STRATEGY ON A DISTILLATION COLUMN CASE STUDY

A distillation column case study is selected to illustrate the automated support strategy. To this end, a simulation environment for benchmark validation provided by Lundh and Modén (2012) is used. The model of the distillation column consists of 110 trays and describes the composition profile. A detailed description of this process is given in Skogestad (1997). The reboiler and the condenser are not modelled. The model has 2 control variables (top  $(y_{top})$  and bottom purity  $(y_{bot})$ ) and 2 manipulated variables (liquid (LF) and vapor flow rates(VF)). The feed (F = 219 Kmol/min) flow with a light component composition of  $z_F = 0.65$  enters the column at the 39<sup>th</sup> tray. It is assumed that the relative volatility is  $\alpha = 1.35$  and the liquid holdup is constant at M = 30 kmol. The objective of the controller is to keep the bottom and top compositions at their set points.

#### 4.1 Linear input-output model

We consider a linear state space model to represent the process in MPC implementation in the form (11) where  $x \in \mathbb{R}^2$ ,  $u \in \mathbb{R}^2$ ,  $y \in \mathbb{R}^2$  are the system states, inputs and outputs respectively. The state space matrices are linearly dependent on the parameter vector  $\theta$ , meaning, each element of the matrices correspond to one element in  $\theta$ . To ensure identifiability of the linear model, we assume that we know the true elements of C in (11). The considered linear model can thus be written as

$$\Sigma_{lin} := \begin{cases} x(t+1) = \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_3 & \theta_4 \end{pmatrix} x(t) + \begin{pmatrix} \theta_5 & \theta_6 \\ \theta_7 & \theta_8 \end{pmatrix} u(t) + \\ \begin{pmatrix} \theta_9 & \theta_{10} \\ \theta_{11} & \theta_{12} \end{pmatrix} e(t), \\ y(t) = \begin{pmatrix} \hat{\theta}_{13}^{init} & \hat{\theta}_{14}^{init} \\ \hat{\theta}_{15}^{init} & \hat{\theta}_{16}^{init} \end{pmatrix} x(t) + e(t) \end{cases}$$

$$(25a)$$

MPC-X currently only handles output error systems. This means that MPC-X assumes  $\theta_9, \ldots, \theta_{12} = 0$  in (25). The input signal to be used in the identification experiment is therefore designed as if the system was output error. Despite this, we still use the obtained input signal to identify the linear model in (25), that is, to estimate  $\theta_1, \ldots, \theta_{12}$ .

# 4.2 Performance measure

For the case study, we used  $\mathcal{J}(t, \Sigma_C(G, H, C(G_{mod}))) = \sum_{i=y_{bot}, y_{top}} \{c_{1,i}P_{viol} + c_{2,i}|y_i - b_i|\}$  where we chose  $c_{1,1} = c_{1,2} = c_{2,1} = c_{2,2} = 1$ , the time window over which averaging takes place is  $T_{win} = 300$  samples, and  $b_i$  are the constraints on the top and bottom products of the column. These are set to 0.08 for  $y_{bot}$  and 0.92 for  $y_{top}$ .

Application set We need to find the application set (22). The approach here is to randomly sample parameter values  $\theta$  and to evaluate the application cost for these parameters in simulation. Based on the parameter values that give sufficiently low application cost, we can estimate A, b and c in (22). Finding these estimates can be formulated as a convex optimization problem and as such it can be solved efficiently Boyd and Vandenberghe (2003).

### 4.3 Simulation scenarios

We consider three different scenarios in the simulation study.

Nominal operations During initial identification, the variance of the control inputs are var(LF) = var(VF) =

100. The model is obtained around the desired operating point where the top composition is 0.95 and the bottom one is 0.05. The variance of the feed rate and feed composition are  $var(\Delta F) = 64$  and  $var(\Delta zF) = 2.5 \times 10^{-3}$ , respectively. These disturbance signals are then filtered by a second-order Butterworth low-pass filter with cutoff frequency 0.05. The noise on the outputs has a variance of 0.1.

*Plant change* The directionality is well known and frequently observed phenomenon in high purity distillation columns. To induce a performance drop due to a change in plant dynamics, an input uncertainty of "rotation type" is introduced, which is similar to the uncertainty shown in Skogestad and Morari (1987) and Skogestad and Morari (1988):

$$\begin{pmatrix} \Delta LF_{imp} \\ \Delta VF_{imp} \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \Delta LF_{MPC} \\ \Delta VF_{MPC} \end{pmatrix}, \quad (26)$$

where  $\Delta LF_{MPC}$  and  $\Delta VF_{MPC}$  are the control inputs computed by the MPC, i.e. the deviation of the liquid and vapor flow from their nominal values;  $\Delta LF_{imp}$  and  $\Delta VF_{imp}$  are the control inputs implemented on the plant;  $\phi$  is the rotation angle of the inputs. In the following,  $\phi = 0$  at the commissioning phase, and then a rotation of  $\phi = -\pi/5$  is used give a plant change. The explanation of the performance drop is given in more detail in Tran et al. (2012) and partly in Özkan et al. (2012). It was proved that the model uncertainty limits the MPC's ability to restore nominal performance.

We run a simulation of 1,600 samples in four phases of 400 samples each:

- (1) Normal operation using commissioning model.
- (2) Operation under plant change.
- (3) Re-identification phase.
- (4) Operation using new model from phase 3) in MPC.

The performance index is calculated for each phase. To evaluate the performance of the identification, we make a Monte Carlo simulation with 310 trials. We evaluate the benefit of the extra excitation added in MPC-X by running the same simulations without the excitation constraint active in the MPC. The parameters of MPC-X were chosen to be  $N_u = N_u = 8$ ,  $\alpha = 0.99$ .

Disturbance change The second scenario, a change in disturbance characteristics, is realised by changing the cutoff frequency of the second-order Butterworth low-pass filter to 0.1. The variance of the feed rate disturbance is increased from  $\operatorname{var}(\Delta F) = 64$  to  $\operatorname{var}(\Delta F) = 225$  and the feed composition disturbance from  $\operatorname{var}(\Delta zF) = 2.5 \times 10^{-3}$  to  $\operatorname{var}(\Delta zF) = 2.5 \times 10^{-3}$ .

# 5. IMPLEMENTATION RESULTS

The auto-tuning method is applied to the distillation column during the nominal operation to achieve the optimal bandwidth for the closed loop. To this end, we solve the optimization problem in (21). We consider the cut-off frequency  $\omega_{cs}(\sigma(\mathbf{Q}, \mathbf{R}))$  of the minimum singular value of the closed loop system. In the MPC,  $\mathbf{Q}$  is an identity matrix since the bottom and top compositions are of equal importance and  $\mathbf{R} = \rho I_2$ . In this case,  $\rho$  is the only optimization variable. The reasoning behind this choice is that increasing  $\rho$  will lead to a decrease in the maximum crossover frequency of the sensitivity function and reducing  $\rho$  will lead to an increase in that frequency. The bandwidth range the controller can attain is [0;0.1]rad/min (the maximum bandwidth of 0.1 rad/min is the cutoff frequency of the sensitivity function of the case  $\rho \approx 0$ ). Therefore, a low bandwidth of 0.01 rad/min is implemented as a starting point. The bandwidth is then increased by 0.01 rad/min after every 2000 minutes and the variance of the outputs is computed. The relation between the bandwidth of the controller and the output variance after running 10 simulations for 10 bandwidth points is given by the solid curve in Figure 3. The optimal bandwidth is 0.06 rad/min and the corresponding output variance is 0.3264.  ${\bf Q}$  and  ${\bf R}$  corresponding to the optimal bandwidth is then fixed as the tuning parameter values at the commissioning.



Figure 3. Relation between bandwidth and output variance of the distillation column at commissioning and after a change in the disturbance.

#### 5.1 Plant change

The plant change scenario results in a performance drop which activates the diagnostic component of the support strategy. Using the algorithm discussed in the previous section, we did a cheap experiment. The identified model has an order of m = 8. We, then, sample the 95-percentile confidence region and check whether the associated transfer functions lie in  $\mathcal{D}_{adm}$ . We conclude that, out of 2000 samples, 80.799% of the models lie outside  $\mathcal{D}_{adm}$ . Hence, we can conclude that  $\mathcal{H}_1$  is true. Hence, the performance drop was due to a plant change. This conclusion requires identification of a new model in closed loop.

Results The performance index for the four different phases and two simulation setups are illustrated in Figure 4. It is clear that adding the constraint results in degradation of performance during the re-identification phase. However, the new model results in acceptable performance in 94 % of the cases. We also note that the spread of the resulting performance measure is small. This should be compared to the case without the excitation constraint, where the re-identification (naturally) does not degrade performance but the resulting models are able to restore performance in 62 % of the cases. Furthermore, while the



Figure 4. Box plots of the performance measures of the four phases of the 310 Monte Carlo trials for the two cases where normal MPC or MPC-X is used. During the excitation phase (3) the cost for MPC-X is higher than for normal MPC where no extra excitation is added. However, in the new model phase (4), the number of models with acceptable performance is significantly higher for the MPC-X case. The acceptable level is below the solid line.



Figure 5. Top (red) and bottom (blue) compositions for the plant change scenario when MPC-X with extra excitation during phase 3 is used. The performance measures for the different phases are 0.08, 0.15, 0.20, 0.07, respectively.

median performance measure is reasonable, the spread is very large. Another observation is that when the model has been identified from normal closed loop data, in 13% of the cases the resulting closed loop is unstable, while this happened for only 2% of the models identified with MPC-X. One typical simulation is shown in Figure 5.

Discussion We do not reach the goal of 99 % re-identified models giving acceptable performance. There are a number of approximations made along the way which surely contribute to the discrepancy. It is also important to comment on the fact that because of the conflicting nature of the amplitude and excitation constraints of MPC-X, it is easy to run into a situation where the problem is infeasible. In these cases, MPC-X instead solves the original MPC problem, without extra excitation. In the Monte Carlo study here, this happened for at least one sample in 39 of the trials and for more than one sample in 7 of the trials.

#### 5.2 Disturbance characteristics change

The performance drop due the disturbance change is realized by changing its variance and frequency content. We, then, sample the 95-percentile confidence region and check whether the associated transfer functions lie in  $\mathcal{D}_{adm}$ . We conclude that, out of 2000 samples, 11.69% of the models lie outside  $\mathcal{D}_{adm}$ . Hence,  $\mathcal{H}_0$  is true, i.e., the performance drop is due to a disturbance change.

*Results* In this scenario, the auto-tuning method is re-implemented when the performance drop is due to a change in the disturbance. The relation between the bandwidth and the output variance is given by the dashed curve in Fig. 3. At the commissioned bandwidth of 0.06 rad/min, the variance goes up to 0.5066 due to the change in the disturbance. The auto-tuning method can bring the variance down to 0.4236 by increasing the bandwidth. The corresponding bandwidth 0.09 rad/min is the new optimal bandwidth of the current system.

*Discussion* The auto-tuning method always aims for the best balance between disturbance rejection and robustness starting at a low-bandwidth setting. This low-bandwidth point can be decided based on information available on the disturbance model and model uncertainty. Since MPC systems in practice are generally tuned conservatively, choosing a low bandwidth is consistent.

## 6. CONCLUSIONS

The automated support strategy and its components have been presented and demonstrated on a binary distillation column example. Once a performance drop is experienced, we are able to distinguish the cause of it with a high probability. The subsequent respective actions then can restore the nominal performance. More specifically, the MPC-X algorithm has been successfully implemented in MATLAB. Initial experiments on the case study show promising results. Further research is needed to understand the reason behind the numerical problems occurring when estimating the complete state-space model of the distillation column (including the C-matrix) and why the gain of using the optimal excitation signal compared to normal excitation signal is so small. The auto-tuning method succeeded in finding the tuning parameters resulting in the best balance between robustness and performance in commissioning and when a change in disturbance characteristics is observed. The optimization problem solved in the tuning algorithm has only considered  $\rho$  as the optimization variable but this could easily be extended to the complete set of diagonal elements in the weighing matrices. An additional research question is also considering system constraints in automated tuning of MPC systems.

#### REFERENCES

- Bauer, M. and Craig, I.K. (2008). Economic assessment of advanced process control: A survey and framework. *Journal of Process Control*, 18(1), 2 – 18.
- Bombois, X., Scorletti, G., Gevers, M., Van den Hof, P.M.J., and Hildebrand, R. (2006). Least costly identification experiment for control. *Automatica*, 42(10), 1651–1662.

- Boyd, S. and Vandenberghe, L. (2003). Convex Optimization. Cambridge University Press.
- Harrison, C. and Qin, S.J. (2009). Discriminating between disturbance and process model mismatch in model predictive control. *Journal of Process Control*, 19, 1610– 1616.
- Hjalmarsson, H. (2009). System identification of complex and structured systems. *European Journal of Control*, 15(4), 275–310.
- Johansson, K. (1997). Relay feedback and multivariable control. Phd thesis, Department of Automatic Control, Lund Institute of Technology, Sweden.
- Larsson, C.A., Annergren, M., Hjalmarsson, H., Rojas, C.R., Bombois, X., Mesbah, A., and Modén, P.E. (2013). Model predictive control with integrated experiment design for output error systems. In *Proceedings of the* 12th European Control Conference. Zürich, Switzerland.
- Lee, K., Xu, F., Huang, B., and Tamayo, E.C. (2008). Controller performance analysis technology for industry: Implementation and case studies. In 17th IFAC World Congress, 14912–144919.
- Ljung, L. (1999). Systems Identification: Theory for the User. Prentice Hall.
- Ljung, L. and Caines, P. (1979). Asymptotic normality of prediction error estimation for approximate system models. *Stochastics*, 3(1), 29–46.
- Lundh, M. and Modén, P.E. (2012). Simulation Environments for Benchmark Validation. Deliverable 6.1 of the AUTOPROFIT Project.
- Luo, Z.Q., Ma, W.K., Šo, A., Ye, Y., and Zhang, S. (2010). Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Processing Magazine*, 27(3), 20–34.
- Maciejowski, J. (2002). Predictive Control with Constraints. Pearson - Prentice Hall.
- Manchester, I.R. (2010). Input Design for System Identification via Convex Relaxation. ArXiv e-prints.
- Mesbah, A. and Bombois, X. (2011). *Performance Monitoring and Diagnosis*. Deliverable 4.1 of the AUTO-PROFIT Project.
- Mesbah, A. and Bombois, X. (2012). Closed-loop Performance Diagnosis: Least Costly Fault Detection and System Re-identification. Deliverable 4.2 (part 2) of the AUTOPROFIT Project.
- Modén, P. (2012a). Performance monitoring for MPC maintenance. ABB Technical Report.
- Modén, P. (2012b). *Performance monitoring for MPC maintenance*. Deliverable 4.2 (part 1) of the AUTO-PROFIT Project.
- Özkan, L., Meijs, J., and Backx, A. (2012). A frequency domain approach for mpc tuning. In *Proceedings of* the 11<sup>th</sup> International Symposium on Process Systems Engineering. Singapore.
- Potters, M., Lundh, M., Moden, P., and Bombois, X. (2012). Software for performance monitoring. Deliverable 4.3 of the AUTOPROFIT Project.
- Richalet, J., Rault, A., Testud, J.L., and Papon, J. (1976). Algorithmic control of industrial processes. In 4th IFAC symposium Identification and System Parameter Estimation, 1119–1167.
- Skogestad, S. (1997). Dynamics and control of distillation columns: A tutorial introduction. *Chemical Engineering Research and Design*, 75(6), 539–562.

- Skogestad, S. and Morari, M. (1987). Implications of large rga elements on control performance. *Ind.Eng.Chem.Res.*, 26, 2323–2330.
- Skogestad, S. and Morari, M. (1988). Robust control of ill-conditioned plants: High-purity distillation. *IEEE Transactions on automatic control*, 33(2), 1092–1105.
- Tran, Q.N., Özkan, L., and Backx, A. (2012). Mpc tuning based on impact of modelling uncertainty on closed-loop performance. In *Proceedings of the 2012 AICHE annual meeting*. Pittsburgh, PA.
- Tyler, M. and Morari, M. (1996). Performance monitoring of control systems using likelihood methods. *Automatica*, 32, 1145–1162.
- Zhu, Y. and Padwarthan, R. (2012). Towards a low cost and high performance mpc: The role of system identification. In *FOCAPO 2012*.