Annual Reviews in Control 42 (2016) 126–142

Contents lists available at ScienceDirect

Annual Reviews in Control

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

Review

Advanced autonomous model-based operation of industrial process systems (*Autoprofit*): Technological developments and future perspectives

Leyla Özkan[a,*], Xavier Bombois[b], Jobert H.A. Ludlage[a], Cristian Rojas[c], Håkan Hjalmarsson[c], Per Erik Modén[d], Michael Lundh[d], Ton C.P.M. Backx[a], Paul M.J. Van den Hof[a]

[a] Department of Electrical Engineering, Eindhoven University of Technology, 5612 AJ, Eindhoven, The Netherlands
[b] Laboratoire Ampère UMR CNRS 5005, Ecole Centrale de Lyon, 69134 Ecully Cedex, France
[c] School of Electrical Engineering, Royal Institute of Technology, SE-100 44, Stockholm, Sweden
[d] ABB AB, Corporate Research, SE 72178 Västerås, Sweden

ARTICLE INFO

Article history:
Received 19 September 2016
Accepted 20 September 2016
Available online 4 November 2016

Keywords:
Model based operation support system
Autonomous maintenance
Performance diagnosis
Hypothesis testing
System identification
Experiment design
Randomized algorithms
(Auto)Tuning
Controller matching
Extremum seeking

ABSTRACT

Model-based operation support technology such as Model Predictive Control (MPC) is a proven and accepted technology for multivariable and constrained large scale control problems in process industry. Despite the growing number of successful implementations, the low level of operational efficiency of MPC is an existing problem, specifically the lack of advanced maintenance technology. To this end, within the EU FP 7 program, a project (Autoprofit [1]) has been executed to advance the level of autonomy and automated maintenance of MPC technology.

Taking linear model-based technology as a starting point, in the project a philosophy has been developed for autonomous performance monitoring, diagnosis, experiment design, model adaptation and controller re-tuning, that is driven by economic criteria in each step, working towards an operation support system in which effective maintenance and adaptation of MPC controllers becomes feasible.

In this development, challenging research questions have been addressed in the areas of on-line performance monitoring and diagnosis, least costly experiment design, automated adaptation of models, and auto-tuning, and new fundamental techniques have been developed. Although a full fledge and industrially proven (semi-)automated system is not yet realised, parts of the on-line system have been implemented and validated on real life cases provided by the industrial partners, showing that the formulated objectives are within reach.

1. Introduction

Model-based operation support systems such as Model Predictive Control (MPC), Real Time Optimisation (RTO) and Model-Based Process Monitoring have become, without doubt, standard tools in process industry, especially, in refineries and petrochemicals. They provide advanced solutions in order to achieve appropriate process control and optimisation of economic performance, while ensuring operational conditions and product specifications to be fulfilled, now and in the near future. In these capital intensive industrial sectors, relatively small improvements in throughput of the plant result in major financial benefits. However, despite these substantial benefits, the costs of model-based operation support technology are considered to be high while at the same time their full potentials can not be exploited easily yet. The high costs are typically caused by high costs of implementation and maintaining the desired performance level. To this end, plant modelling plays a dominant role as the most expensive and time-consuming part of the implementation campaign.

A further drawback of the current technology is its limited capability to effectively deal with varying operating conditions and - in particular for MPC - nonlinear plant behaviour, thereby limiting not only the operating range but also the performance and economical benefit of most currently implemented controllers. These aspects restrict the overall life-time performance of this technology and hinder its further widespread implementation in several other industrial sectors. If model-based operation support systems are to become essential tools to meet the current and future challenges for industrial process operations then *the life time cost*

* Corresponding author.
E-mail address: l.ozkan@tue.nl (L. Özkan).

[1] Upon its completion, the project was rated 'Excellent' by the EU.

http://dx.doi.org/10.1016/j.arcontrol.2016.09.015
1367-5788/

of these systems need to be drastically reduced and the life time performance needs to be significantly improved over varying operating conditions.

Within this scope, firstly, the accuracy of models is of paramount importance in establishing high performance of model-based operation support systems. Often more than 50% of the time is spent on the modelling phase and during commissioning again 60% of the time is spent on models (Darby & Nikolaou, 2012). In particular plant testing, as a way to identify or validate plant models, is very expensive, requiring around the clock presence of specialised engineers, while at the same time production quality may have to be compromised temporarily because of the plant excitations. Secondly, critical to the lifetime performance, is performance maintenance which is not directly addressed in a structural way. Lack of maintenance is frequently the major reason for the limited benefits of these model-based operation support systems. 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.* In particular for MPC applications, varying operating conditions as well as model deterioration over time often lead to switching off the automated MPC and returning to manual operation. Additionally, model maintenance is expensive and requires skilled process engineers.

These observations lead to the conclusion that if we can make performance maintenance and modelling campaigns cheaper, then the application of model based operation support technology will become more profitable not just for processes currently controlled by MPC but also for an extended range of processes. Hence, the *Autoprofit* project has been set up with the objective of automated *just-in-time just-enough maintenance* of models and their model-based applications. This requires continuous performance monitoring and diagnosis, and automated procedures for recalibrating the model and/or retuning the model-based application (RTO, MPC, soft-sensing) if, based on economic considerations, there is evidence from the diagnosis that these actions can establish the desired performance of the model-based application. In this respect it is important to distinguish between a required model calibration and a required (re)tuning of the model-based application. Note indeed that for MPC controllers, one cannot simply tune an MPC controller to compensate for a poor model (Darby & Nikolaou, 2012). However, in the sense of achieving maximum performance, tuning can improve the actual performance obtained from the model-based system if the model accuracy degrades.

There have been several studies and developments for sustaining the operational efficiency of model based technology (Chan, Dozal, Cheng, Kephart, & Ydstie, 2014; Ji et al., 2013; Jimoh, 2013; Marafioti, 2010; Zhu & Padwarthan, 2012). The majority of these studies is in the area of model maintenance since the performance of any model based controller depends intrinsically on the accuracy of the plant model over the nominal operating range of the plant (Darby & Nikalaou, 2014). Naturally, a closely related topic is productivity preserving closed loop experiment design (input signal design) to adjust the model on the basis of the data. To this end, some researchers have adopted the dual control approach (Heirung, Ydstie, & Foss, 2013; Marafioti, 2010) in which the input design requirement for model updating is built in the MPC formulation. It takes the form of an additional constraint which ensures that the minimal level of excitation is achieved while the process constraints are respected. Some of these ideas of closed loop identification have been put in practice by Kalafatis, Patel, Harmse, Zheng, and Craig (2006) and Zhu, Patwardhan, Wagner, and Zhao (2013).

In the context of autonomous maintenance, performance monitoring and performance diagnosis play an essential role. Closed loop performance monitoring is frequently based on Minimum Variance Control (MVC) (Harris, 1989; Harris, Seppala, & Desborough, 1999). However, such a performance benchmark is not realistic. A controller that shows poor performance compared to a minimum variance control system is not necessarily a poor controller (Huang & Shah, 1998). This observation has resulted in other performance indices and approaches that are closer to the actual achievable and specified closed-loop behaviour (Xu, Huang, & Akande, 2007). The model in model-based systems opens other possibilities for performance monitoring (McNaab & Qin, 2003; 2005). The ability to deal with constraints and pushing for the economic performance, which is unique for MPC, have only been considered recently in performance analysis (Agarwal, Huang, & Tamayo, 2007a; 2007b; Modén & Lundh, 2013; Xu et al., 2007). In comparison to performance monitoring, little attention has been paid to performance diagnosis (Patwardhan & Shah, 2002; Schafer & Cinar, 2004). Performance degradation can be caused either by a change in external circumstances (like changed disturbance characteristics, failure of actuators or sensors) or a change in the dynamic behaviour between the relevant inputs and outputs of the process (changing specifications, changes made in parts of the process unit (revamps or rebuilds)), which results in a less accurate model. In order to make an accurate diagnosis, it is generally required to make an experiment on the plant. The theory of hypothesis testing can then be used in order to accurately determine the cause of the performance drop (Basseville, 1997; 1998; Gustafsson & Graebe, 1998). The main challenge lies in an unexplored aspect of the problem: the design of an appropriate detection experiment.

The performance of a model-based system depends on both model accuracy and tuning of the system. Finding a good controller tuning which results in the desired closed loop behaviour is a challenge. For MPC systems the tuning problem is equivalent to a selection of the specification priorities and of the weighting matrices in the steady state as well as the dynamic optimisation criterion in such a way that the resulting closed loop behaviour is a good trade-off between the different conflicting requirements given the specific characteristics of the process and the model accuracy. The tuning of MPC has attracted attention of both practitioners and researchers (Garriga & Soroush, 2010; Lee & Yu, 1994; Maurath, Seborg, & Mellichamp, 1988; Shridhar & Cooper, 1998). Recently, there have been several developments in which the tuning parameters are computed such that MPC matches an arbitrary LTI controller (Di Cairano & Bemporad, 2010; Hartley & Maciejowski, 2011; 2013). These studies fall into the topic of controller matching and most of them consider the infinite horizon unconstrained case. On the other hand, MPC is a constrained optimising finite time controller and hence is time variant and nonlinear.

An integrated approach for realising a highly autonomous model-based control performance maintenance system is sketched in Fig. 1.

- The performance of the model-based control system will be continuously monitored by computing a performance measure evaluating the closed-loop performance at the present time. As long as this performance measure remains below a certain threshold, the algorithm remains in monitoring mode. If the performance measure exceeds that threshold (i.e. a performance drop is observed), the diagnosis algorithm is triggered.
- If the performance measure exceeds that threshold (i.e. a performance drop is observed), the diagnosis algorithm, which is based on a “cheap” identification, is triggered. By “cheap”, we mean an identification whose economical costs are much smaller than the one of a full re-identification. Using the identified model, the diagnosis algorithm allows to distinguish between the case of a control-relevant plant change and the case of a change of disturbance dynamics. If the performance drop is due to a change of disturbance dynamics, the controller is retuned on the basis of the same model. For this purpose, the

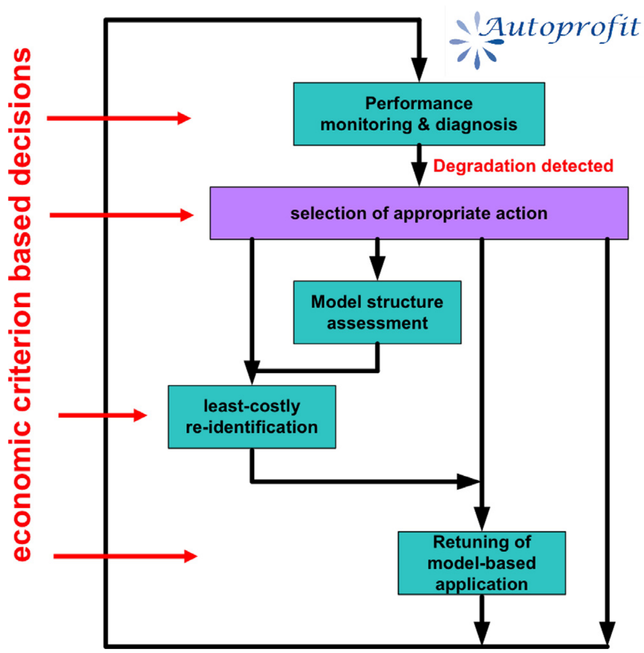

Fig. 1. Autoprofit decision tree.

information on the current disturbance dynamics that can be deduced from the identification experiment used for the diagnosis can be exploited. If the performance drop is due to a plant change, a full re-identification can be decided for. The data collected for the diagnosis experiment can also be used to reduce the cost of this re-identification, which is an advantage of an identification-based diagnosis algorithm. This step has to indicate the expected gain in performance from a model calibration and/or an application of retuning.

- An economic decision then has to be made whether a plant test has to be performed, balancing the cost of plant testing and the expected gain in model accuracy and subsequent model-based application performance.
- If a full re-identification is decided for, this full re-identification will be performed in an automated way and in the least-costly way i.e. using an excitation signal that causes the least experimental cost (perturbation) on the real system while guaranteeing that the identified model is sufficiently accurate to be able to restore the initial control performance by re-tuning the controller using that model.
- Based on the calibrated model the model-based application is retuned and activated.
- If a plant test is not desired the model-based application will be retuned on the basis of the existing model, and subsequently activated.

Critical to the success of this procedure is the introduction of economic criteria in designing test signals in plant testing, but also in deciding whether the results of the performance monitoring and diagnosis warrant a new test campaign. Indeed, any maintenance should only be decided when it is economically profitable to do so. The diagnosis step is, in this aspect, particularly important to guarantee a cost-effective maintenance. The methodology that will be needed to restore the closed-loop performance is highly dependent on the cause of the performance drop. Indeed, while an accurate and thus costly re-identification of the plant dynamics seems the only solution to restore the initial closed-loop performance when the cause of the performance drop is a control-relevant plant change, it may by far not be the optimal solution when the cause of the performance drop is a change of disturbance characteristics. In the latter case, controller retuning will generally be sufficient. Hence, the economic costs of a new campaign need to be balanced with the expected increase of accuracy of the model and of expected increase of performance of the model-based application.

Having defined the procedure for the autonomous operation of model-based support systems, the technological developments within the project naturally have been clustered in different modules which are explained briefly in the following.

Performance monitoring. A classical way to evaluate the performance of a closed-loop system is to estimate the present input/output variances and to verify whether the performance requirements are still satisfied (see Huang & Shah, 1999). Even though other monitoring criterion has also been considered in the Autoprofit project (Modén & Lundh, 2013), we will consider this classical approach in this paper.

Performance diagnosis. The diagnosis algorithm is started every time a performance drop is detected with the performance monitoring algorithm and it must be able to accurately distinguish between a (closed-loop) performance drop due to a plant change (hypothesis $\mathcal{H}_1$) or due to a disturbance change (hypothesis $\mathcal{H}_0$). The closed-loop performance diagnosis step can thus be considered as a hypothesis test.

Note that we do not consider here the case of performance drops which are due to actuator or sensor failures (we assume that such hypothesis have been excluded using available techniques). We will also make no assumption on the possible plant or disturbance changes that may occur. In this respect, our work is different from the work presented e.g. in Olaru, De Doná, and Seron (2008) and Seron and De Doná (2010). Note finally that we are not interested in detecting *any* plant change, but only those plant changes leading to a decrease of closed-loop performance. The diagnosis problem under consideration is thus different from the one considered e.g. in Basseville (1998). As mentioned in Mesbah and Bombois (2012), our work can nevertheless be related to the work in Gustafsson and Graebe (1998) and Tyler and Morari (1996).

As already mentioned, the advantage of using an identification approach as the basis for the diagnosis step is that the data collected for this diagnosis experiment can also be used for the (eventual) full re-identification step and therefore reduce the cost of this subsequent step. Besides the identification of a model, another important ingredient of our diagnosis approach is the introduction of the set $\mathcal{D}_{diag}$ containing all plant transfer functions resulting in a satisfactory closed-loop performance with the controller designed at commissioning under the original disturbance level (or, more precisely, delivering satisfactory performance for the set of disturbance spectra considered as realistic at commissioning). By definition of the set $\mathcal{D}_{diag}$, if the true plant transfer function is within $\mathcal{D}_{diag}$, an observed performance drop can only be explained by a change in the disturbance characteristics. Combining the identified plant transfer function $\hat{G}$ and the set $\mathcal{D}_{diag}$, we have introduced the following decision rule: if the identified model $\hat{G}$ is outside of $\mathcal{D}_{diag}$, we conclude that a plant change is the root cause of the performance drop and conversely we conclude that a change in the disturbance characteristics is the cause of the performance drop when $\hat{G} \in \mathcal{D}_{diag}$. Attention is also paid to the problem of checking the diagnosis accuracy a-posteriori.

Re-identification module. If a performance drop due to a plant change is detected by the diagnosis module, the current dynamics of the plant may have to be re-identified. This re-identification will be performed in the least-costly context. The topic of “least costly identification for control” has been introduced in Bombois, Scorletti, Gevers, Van den Hof, and Hildebrand (2006). This work have considered the problem of the design of the least disturbing excitation signal that can be applied to a real-life true system (operated e.g. in closed loop) while guaranteeing the model identified in this way is nevertheless sufficiently accurate for the design of a satisfactory controller for the true system. In Bombois et al. (2006), it is assumed that the controller that is present in the loop for the identification is linear (LTI). Moreover, it is also supposed that the performance of the controller that has to be designed with the identified model is expressed in the H_∞ framework. Using these assumptions, we have deduced an LMI constraint that has to be satisfied by the inverse P_θ^{-1} of the covariance matrix of the identified parameter vector to guarantee that this identified model is sufficiently accurate for the design of a satisfactory H_∞ controller. This LMI is thus a constraint on how accurate the identified model has to be for control design. Then, using the fact that, when the controller present in the loop for the identification is LTI, the inverse of the covariance matrix of the identified model is affine in the excitation spectrum, we could express the design of the least disturbing excitation signal for the identification of a model that is sufficiently accurate for control as an LMI optimisation problem where the excitation spectrum is the decision variable.

In the situation considered in Autoprofit, both the controller, that is present in the loop for the (re-)identification, and the controller that must be designed with the identified model are MPC controllers and thus are controllers that are not LTI. Thus, the framework of Bombois et al. (2006) has been extended to the case of MPC controllers. By estimating numerically the sensitivity to the modelling error of the performance of the MPC controller designed with an identified model, we have been able to approximate the constraint on the accuracy of the identified model for control design by an LMI on P_θ^{-1}. Due to the fact that the identification is performed in closed loop with an MPC controller in the loop, the inverse P_θ^{-1} of the covariance matrix is unfortunately not affine in the (power spectrum of the) excitation signal. Two different techniques have been established to be able to also consider MPC controllers in our closed-loop experiment design framework. The first solution consists of *fooling* the MPC controller so that the excitation signal affects the plant in an open-loop fashion (stealth MPC, Potters et al., 2014). The second solution consists of modifying the MPC algorithm in such a way that this algorithm ensures both the identification and control objectives (MPC-x, Larsson, Rojas, Bombois, & Hjalmarsson, 2015).

(Auto)Tuning. In the Autoprofit philosophy, the tuning of the MPC is required when a new calibrated model is developed or when it is economically more beneficial to retune the controller rather than executing a new identification campaign. In this project, by tuning of MPC, we mean the selection of the weighting matrices in the objective function. The closed loop performance of any system depends on the model accuracy and the tuning of the controller. In robust control, this dependency has been studied extensively using frequency domain techniques and it was shown that if the closed loop system bandwidth exceeds the bandwidth of the plant model mismatch, model uncertainty starts to be significant resulting in closed loop performance degradation. A similar observation, where the weighting matrices in the objective function define the closed loop bandwidth, is also shown for the case of MPC performance (Skogestad & Postlethwaite, 2005). The tuning methods developed in this project are based on this observation.

In the sequel, we provide a high level discussion on the different modules and we present technical developments in each module (Sections 2–4). We will suppose throughout that an MPC controller is present in the loop and that this controller has been designed at commissioning with an initial model of the plant. We will also assume that, at commissioning, the performance of this MPC control system has been satisfactory. One of the key activities in the project is the industrial validation of developed technologies. To this end, two cases have been provided by the industrial partners SASOL and BOLIDEN. In Section 5, we explain the SASOL case and present experimental results in the area of tuning of MPC. The paper concludes with Section 6 which provides an extensive discussion on the future needs of autonomous model based operation support systems.

2. Performance monitoring and performance diagnosis [2]

2.1. Preliminaries:MPC algorithm designed at commissioning

We consider that the true system has n outputs and m inputs ($y(t) \in \mathbf{R}^n$, $u(t) \in \mathbf{R}^m$) and that it can be described by

$$y(t) = \underbrace{G(z,\theta_0)}_{G_0(z)}u(t) + \underbrace{H(z,\theta_0)e(t)}_{v(t)}, \tag{1}$$

where $\theta_0 \in \mathbf{R}^k$ is the unknown true parameter vector; $e(t) \in \mathbf{R}^n$ is white noise with a covariance matrix Λ; $G(z, \theta_0)$ and $H(z, \theta_0)$ are stable discrete-time transfer matrices. $H(z, \theta_0)$ is further assumed to be monic and minimum-phase (Ljung, 1999).

A model (G_{mod}, H_{mod}) of (G_0, H_0) is available at commissioning. The model $G_{mod} = G(z,\theta_{mod})$ has been used to design an MPC algorithm for the system (1). This MPC algorithm will be denoted in the sequel by $C(G_{mod})$. An MPC controller computes the input $u(t)$ of the system (1) recursively by minimising a LQG-type criterion over a certain horizon (Maciejowski, 2002). More precisely, at time instant t, the input sequence $u(t), u(t+1), \dots u(t+M)$ (with M the control horizon) is determined as the one which minimises a criterion of the form

$$J_{MPC} = \sum_{j=1}^{P}((\hat{y}(t+j) - s(t+j))^T Q(\hat{y}(t+j) - s(t+j)) + \Delta u^T(t+j-1)R\Delta u(t+j-1)) \tag{2}$$

with $P \geq M$ the prediction horizon, Q, R user-chosen weightings, $\Delta u(t) = u(t) - u(t-1)$ and $\hat{y}(t)$ a prediction of the output $y(t)$ based on the past data and the available model G_{mod} of the plant:

$$\hat{y}(t) = G_{mod}u(t). \tag{3}$$

The prediction (3) is generally performed using a state-space model[3] of G_{mod}. Finally, $s(t)$ is the setpoint for the *predicted* output (see below).

Since $P \geq M$, we enforce $u(t+j) = u(t+M)$ for all $j > M$. The minimization is here performed under the presence of constraints on the process variables:

$$|\hat{y}_i(t)| < y_{cons,i} \qquad |u_l(t)| < u_{cons,l} \tag{4}$$

with $\hat{y}_i$ and u_l the entries of the vectors $\hat{y}$ and u. Even though the inputs $u(t), u(t+1), \dots u(t+M)$ are all computed at time t, only the input $u(t)$ is applied to the system (1) at time t and the whole procedure is repeated at time $t+1$ (receding horizon principle). In order to take disturbances and model mismatch into account, a kind of integrator is added to the MPC controller. This is done by defining the setpoint $s(t+j)$ for the *predicted* output as:

$$s(t+j) = y_{ref}(t+j) - c_t \quad (j = 1 \dots P)$$

with $y_{ref}(t+j)$ the reference for the *actual* output $y(t+j)$ and a correction term c_t which is at time t equal to:

$$c_t \triangleq y(t) - \hat{y}(t). \tag{5}$$

The correction term c_t thus corresponds to the error observed at instant t between the value $\hat{y}(t)$ predicted by the model and the

[2] The contents of this section have already appeared in Bombois, Potters, and Mesbah (2014) and Mesbah, Bombois, Forgione, Hjalmarsson, and Van den Hof (2015)

[3] In this way, it is indeed not necessary to use the full input history to predict the output at time t.

measured output $y(t)$. This correction is the only feedback term in the MPC algorithm (Maciejowski, 2002).

Remark. Note that the model H_{mod} can also be used in the MPC formulation, but it is not necessary.

2.2. Performance monitoring module

Let us introduce some notations. We denote by $(G, H, C(G_{mod}))$ the loop made up of a system $y(t) = Gu(t) + He(t)$ and the MPC algorithm $C(G_{mod})$ based on G_{mod}. In other words, a loop where the input u of the system is computed recursively with the MPC algorithm presented in the previous section. The performance of such a loop can be evaluated in different ways. Here, we will use a classical measure i.e. the variance of the tracking error. Inspired by (2), we define the performance of the loop $(G, H, C(G_{mod}))$ at time t as:

$$\mathcal{J}(t, G, H, C(G_{mod})) \;=\; (y(t) - y_{ref}(t))^T Q (y(t) - y_{ref}(t)) \tag{6}$$

A penalty for constraint violations and input variances can also be added, but will be omitted here for simplicity. This quantity will be generally averaged over a given time window and the performance of the loop will be deemed satisfactory if this average value remains below a carefully chosen threshold.

Remark. Note that, in the project, we have also introduced another performance measure for the cases where the control objective consists of pushing one important (money-making) output as close as possible to a constraint[4]. The performance of the loop $(G, H, C(G_{mod}))$ is then measured using an index which is equal to a linear combination of the number of observed constraint violations over the time window and of the average distance between the money-making variable and the constraint over the same time window (Modén, 2012; Potters et al., 2012).

The performance of the MPC algorithm $C(G_{mod})$ on the model (G_{mod}, H_{mod}) is of course satisfactory. Moreover, it is assumed that, at commissioning, this model is accurate enough for the performance of the loop $(G_0, H_0, C(G_{mod}))$ to be satisfactory at commissioning. Suppose now that the MPC controller has been operated for a while and we now observe that $\mathcal{J}(t, G_0, H_0, C(G_{mod}))$ (averaged over the chosen time window) remains structurally above the chosen threshold. The performance of the MPC control system has thus dropped and it has to be decided whether this performance drop is due to a change of the plant dynamics G_0 or due to a change of the disturbance dynamics $H_0(z)e(t)$. This diagnosis is very important since the performance drop will be often due to a disturbance change and a full re-identification of the plant G_0 will not be necessary. Only, in the eventuality that the performance drop is really due to a plant change, a full re-identification will be decided for (see Section 2.4).

2.3. Performance diagnosis module

2.3.1. Set $\mathcal{D}_{diag}$ with MPC controller

As mentioned in the introduction, an important ingredient for our diagnosis methodology is the determination of the set $\mathcal{D}_{diag}$. We show in this section how we can determine this set in the case of a system controlled by a MPC controller. For this purpose, let us consider a given transfer function $G(z, \theta) \neq G_{mod}$ and the following system:

$$y(t) = G(z, \theta)u(t) + \underbrace{H_{mod}(z)e(t)}_{v_{mod}(t)} \tag{7}$$

Note that this system has the same noise model as the model deduced at commissioning, but another plant model. Let us consider the loop $(G(\theta), H_{mod}, C(G_{mod}))$. Note that we can easily simulate this loop in any simulation environment. Let us denote the performance index computed during such a simulation by $\mathcal{J}(t, G(\theta), H_{mod}, C(G_{mod}))$.

The set $\mathcal{D}_{diag}$ can then be defined by comparing $\mathcal{J}(t, G(\theta), H_{mod}, C(G_{mod}))$ with the performance index $\mathcal{J}(t, G_{mod}, H_{mod}, C(G_{mod}))$ of the loop $(G_{mod}, H_{mod}, C(G_{mod}))$ i.e. the loop designed at commissioning. More precisely, we define $\mathcal{D}_{diag}$ as the set of plants $G(z, \theta)$ for which the loop $(G(\theta), H_{mod}, C(G_{mod}))$ leads to a performance index $\mathcal{J}(t, G(\theta), H_{mod}, C(G_{mod}))$ which is close to the performance index $\mathcal{J}(t, G_{mod}, H_{mod}, C(G_{mod}))$. Indeed, this set $\mathcal{D}_{diag}$ is the set of all plant transfer functions which leads to good performance under the original disturbance characteristics. Consequently, if $G_0 \in \mathcal{D}_{diag}$ and a performance drop is observed, the performance drop can only be due to a change in disturbance characteristics.

Let us formalize this mathematically. For this purpose, let us compute the performance indices $\mathcal{J}(t, G(\theta), H_{mod}, C(G_{mod}))$ and $\mathcal{J}(t, G_{mod}, H_{mod}, C(G_{mod}))$ over a long simulation (duration N_{sim}) of the two loops $(G(\theta), H_{mod}, C(G_{mod}))$ and $(G_{mod}, H_{mod}, C(G_{mod}))$ with the same realization[5] of $v_{mod}(t) = H_{mod}(z)e(t)$ and let us introduce the following notation:

$$V_{diag}(\theta) = \frac{1}{N_{sim}} \sum_{t=1}^{N_{sim}} (\mathcal{J}(t, G(\theta), H_{mod}, C(G_{mod})) - \mathcal{J}(t, G_{mod}, H_{mod}, C(G_{mod})))^2 \tag{8}$$

The set $\mathcal{D}_{diag}$ can then be defined as:

$$\mathcal{D}_{diag} = \{G(z, \theta) \mid V_{diag}(\theta) < \frac{1}{\gamma}\} \tag{9}$$

for some carefully chosen threshold γ. This is an implicit definition of the set $\mathcal{D}_{diag}$. It is very easy to verify whether a model $G(z, \theta)$ lies in $\mathcal{D}_{diag}$. However, from (9), it is not possible to deduce the contour and the size of $\mathcal{D}_{diag}$. But, it is possible to deduce an approximative expression of this set which is explicit. This observation is important for the sequel. We have indeed that $V_{diag}(\theta_{mod}) = 0$ and $V_{diag}(\theta)$ has its minimum in θ_{mod}. Consequently, $\mathcal{D}_{diag}$ can be approximated using a Taylor expansion by an ellipsoid in the parameter space:

$$\mathcal{D}_{diag} \approx \{G(z, \theta) \mid (\theta - \theta_{mod})^T \mathcal{H}_{diag} (\theta - \theta_{mod}) < \frac{2}{\gamma}\} \tag{10}$$

with $\mathcal{H}_{diag}$ the Hessian of $V_{diag}(\theta)$ evaluated in θ_{mod} (Hjalmarsson, 2009). The matrix $\mathcal{H}_{diag}$ can be determined numerically using a set of simulations around θ_{mod}.

Remark. The parametrization $G(z, \theta)$ corresponds to the true system G_0 as it is at the moment of the diagnosis. If the order of G_0 is larger than the one at commissioning, some parameters of G_{mod} are put to zero.

2.3.2. Diagnosis experiment and decision rule

The set $\mathcal{D}_{diag}$ defined in the previous section will allow us to distinguish a performance drop due to a change in disturbance characteristics and one due to a plant change. First note that, at commissioning, G_0 lay in $\mathcal{D}_{diag}$ since the performance was deemed satisfactory at that moment. However, with time, the performance of the controller can drop due to changes in the characteristics of G_0 or H_0. Suppose that a performance drop is observed and that the plant transfer function G_0 is no longer in $\mathcal{D}_{diag}$ at that moment.

[4] Note that, instead of using a single money-making output, it is possible (as proposed in Potters, Lundh, Modén, & Bombois, 2012) to use several of these outputs in the definition of the performance measure

[5] A relatively long simulation is required to sufficiently attenuate the random effects linked with the realization of $v_{mod}(t)$.

Since G_0 is no longer in $\mathcal{D}_{diag}$, the MPC controller based on G_{mod} is no longer able to deliver sufficient performance even under the initial disturbance characteristics (i.e. $H_{mod}e$). Consequently, we can conclude that the performance drop is due to a control-relevant plant change. On the other hand, if a performance drop is observed and G_0 is still within $\mathcal{D}_{diag}$, we will conclude that the cause of the performance drop is a change in disturbance characteristics.

Consequently, the diagnosis problem under study consists in deciding between the following hypotheses when a performance drop is observed:

$$\begin{aligned} &\mathcal{H}_0 : G_0(z) \in \mathcal{D}_{diag} \\ &\mathcal{H}_1 : G_0(z) \notin \mathcal{D}_{diag}. \end{aligned} \tag{11}$$

To be able to discriminate between the two hypotheses stated in Eq. (11), we will perform a *diagnosis experiment* i.e. we will identify the unknown true system (1) in closed-loop operation with the existing MPC controller $C(G_{mod})$. For this purpose, an external signal $r_{diag}(t) \in \mathbf{R}^m$ can be added to the input of the system: $u(t) = u_{MPC}(t) + r_{diag}(t)$ where $u_{MPC}(t)$ is the input signal computed by the MPC controller. By applying the excitation signal $r_{diag}(t)$ for $(t = 0, \ldots, N-1)$ to the closed-loop system and measuring signals $\{u(t), y(t)\,|\,t = 1, \ldots, N\}_{diag}$, a model $\{G(z, \hat{\theta}_N^{diag}), H(z, \hat{\theta}_N^{diag})\}$ of the true system can be identified using prediction error identification. In this situation, the identified parameter vector is defined as:

$$\hat{\theta}_N^{diag} = arg\min_{\theta} V_{id}(\theta) \quad \text{with} \quad V_{id}(\theta) = \frac{1}{N}\sum_{t=1}^{N} \epsilon^T(t,\theta)\Lambda^{-1}\epsilon(t,\theta), \tag{12}$$

where $\epsilon(t,\theta) = H(z,\theta)^{-1}\big(y(t) - G(z,\theta)u(t)\big)$. If r_{diag} is sufficiently exciting and the model has been identified in a model structure containing the true system (1), the parameter vector $\hat{\theta}_N^{diag}$ identified through Eq. (12) will be asymptotically normally distributed around the true parameter vector θ_0. This suggests that $\hat{\theta}_N^{diag} \sim \mathcal{N}(\theta_0, P_{\theta,diag})$ with $P_{\theta,\,diag}$ being a strictly positive definite matrix defined by (13) and that can be estimated from $\hat{\theta}_N^{diag}$ and $\{u(t), y(t)\}_{diag}$ using (14) (Ljung, 1999):

$$P_{\theta,diag} = \frac{1}{N}\left(\bar{\mathbb{E}}\left[\left(\frac{\partial\epsilon(t,\theta)}{\partial\theta}|_{\theta_0}\right)\Lambda^{-1}\left(\frac{\partial\epsilon(t,\theta)}{\partial\theta}|_{\theta_0}\right)^T\right]\right)^{-1} \tag{13}$$

$$\approx \frac{1}{N}\left(\frac{1}{N}\sum_{t=1}^{N}\left[\left(\frac{\partial\epsilon(t,\theta)}{\partial\theta}|_{\hat{\theta}_N^{diag}}\right)\hat{\Lambda}^{-1}\left(\frac{\partial\epsilon(t,\theta)}{\partial\theta}|_{\hat{\theta}_N^{diag}}\right)^T\right]\right)^{-1} \tag{14}$$

Using this distribution, we can build a $\alpha\,\%$ confidence ellipsoid $U(\hat{\theta}_N^{diag}, \chi)$ for θ_0.

$$U(\hat{\theta}_N^{diag}, \chi) = \{\theta \in \mathbf{R}^k \mid (\theta - \hat{\theta}_N)^T P_{\theta,diag}^{-1}(\theta - \hat{\theta}_N) < \chi\} \tag{15}$$

The size χ of this $\alpha\,\%$ confidence ellipsoid is chosen in such a way that $Pr(\chi^2(k) < \chi) = \alpha$.

Once a model $G(z, \hat{\theta}_N^{diag})$ of the true system is identified, we can utilize this model to choose between hypotheses $\mathcal{H}_0$ or $\mathcal{H}_1$. The decision rule is the following one:

$$\begin{aligned} &G(z, \hat{\theta}_N^{diag}) \in \mathcal{D}_{diag} \Rightarrow \text{choose} \quad \mathcal{H}_0 \\ &G(z, \hat{\theta}_N^{diag}) \notin \mathcal{D}_{diag} \Rightarrow \text{choose} \quad \mathcal{H}_1. \end{aligned} \tag{16}$$

Eq. (16) indicates that $\mathcal{H}_0$ is chosen as the correct hypothesis when $G(z, \hat{\theta}_N^{diag})$ is inside $\mathcal{D}_{diag}$, whereas $\mathcal{H}_1$ is chosen as the correct hypothesis if $G(z, \hat{\theta}_N^{diag})$ lies outside $\mathcal{D}_{diag}$. Note that verifying that $G(z, \hat{\theta}_N^{diag}) \in \mathcal{D}_{diag}$ (resp. $G(z, \hat{\theta}_N) \notin \mathcal{D}_{diag}$) can be performed by evaluating $V_{diag}(\hat{\theta}_N)$ using a long simulation (see (8) and (9)).

2.3.3. Checking the diagnosis accuracy a-posteriori

Like any decision rule, the decision rule (16) may lead to erroneous decisions since $G(z, \hat{\theta}_N^{diag})$ is only an estimate of the true plant G_0. As an example, the null hypothesis $\mathcal{H}_0$ may indeed be chosen erroneously when $G(z, \hat{\theta}_N^{diag}) \in \mathcal{D}_{adm}$ has been generated by $G_0 \notin \mathcal{D}_{adm}$. This is in effect a wrong decision since the performance drop is not due to variations in disturbance characteristics, but due to changes in plant dynamics.

In hypothesis testing (Kay, 1998), the accuracy of the decision rule is determined by two probabilities: the probability $Pr_{\mathcal{H}_0}$ of deciding $\mathcal{H}_0$ when $\mathcal{H}_0$ is true, on the one hand, and the probability $Pr_{\mathcal{H}_1}$ of deciding $\mathcal{H}_1$ when $\mathcal{H}_1$ is true, on the other hand. The probability $Pr_{\mathcal{H}_1}$ is often called "detection rate" in the literature while $1 - Pr_{\mathcal{H}_0}$ is called "false alarm rate". It is obvious that both $Pr_{\mathcal{H}_0}$ and $Pr_{\mathcal{H}_1}$ must be high for the hypothesis test to be accurate. It is important to note that these two probabilities depend on the diagnosis experiment. Indeed, a diagnosis experiment leading to a model $G(z, \hat{\theta}_N^{diag})$ very close to G_0 will increase both probabilities. In Mesbah et al. (2015), techniques to optimally design the excitation signal $r_{diag}(t)$ used during the diagnosis experiment to guarantee pre-specified values have been discussed for both probabilities at the lowest cost. However, these techniques have only been tested and validated in an H_∞ framework and thus not in an MPC framework as in this paper.

If an arbitrary and cheap excitation signal r_{diag} is chosen, it is a good idea to assess the diagnosis accuracy based on the identified model $G(z, \hat{\theta}_N^{diag})$ and the $\alpha\,\%$ confidence ellipsoid given in (15). Let us for this purpose define the following uncertainty set in the transfer function space:

$$\mathcal{D}(\hat{\theta}_N^{diag}, \chi) = \{G(z,\theta) \mid \theta \in U(\hat{\theta}_N^{diag}, \chi)\} \tag{17}$$

Using the property of $U(\hat{\theta}_N^{diag}, \chi)$, we know that $G_0 = G(z, \theta_0)$ lies in the set $\mathcal{D}(\hat{\theta}_N^{diag}, \chi)$ in $\alpha\,\%$ of the identification experiments.

Suppose first that $G(z, \hat{\theta}_N^{diag}) \notin \mathcal{D}_{diag}$. We thus conclude that $\mathcal{H}_1$ is true. However, we will only be fully confident in this decision if the uncertainty set $\mathcal{D}(\hat{\theta}_N^{diag}, \chi)$ corresponding to a large α lies also entirely outside $\mathcal{D}_{diag}$ i.e.

$$\mathcal{D}(\hat{\theta}_N^{diag}, \chi) \subset \mathcal{C}\mathcal{D}_{diag} \tag{18}$$

with $\mathcal{C}\mathcal{D}_{diag}$ the complement of $\mathcal{D}_{diag}$. Indeed, if that is the case for a large α (and thus a large χ), it is very *unlikely* that $G_0 \in \mathcal{D}_{diag}$ and thus it is also very unlikely that we have made an erroneous diagnosis. Since $\mathcal{D}_{diag}$ can be described as an ellipsoid in the parameter space, verifying (18) for a given χ is straightforward. It is even possible to determine the largest value of χ for which (18) holds using LMI optimization (see Bombois et al., 2014). Let us denote, for further use, this largest value of χ by $\chi_{max,\mathcal{H}_\infty}$.

Suppose now that the identified model $G(z, \hat{\theta}_N^{diag})$ lies in $\mathcal{D}_{diag}$. We thus conclude that $\mathcal{H}_0$ is true. However, we will here also only be fully confident in this decision if the uncertainty set $\mathcal{D}(\hat{\theta}_N^{diag}, \chi)$ corresponding to a large α lies entirely inside $\mathcal{D}_{diag}$ i.e.

$$\mathcal{D}(\hat{\theta}_N^{diag}, \chi) \subset \mathcal{D}_{diag} \tag{19}$$

Indeed, if that is the case for a large α (and thus a large χ), it is very *unlikely* that $G_0 \notin \mathcal{D}_{diag}$ and thus it is also very unlikely that we have made an erroneous diagnosis. Since $\mathcal{D}_{diag}$ can be described as an ellipsoid in the parameter space, it is possible to determine the largest value of χ for which (19) holds (see Bombois et al., 2014).

Important remark. Due to the fact that, if $\mathcal{H}_1$ is chosen, a full re-identification will be generally decided for, choosing erroneously $\mathcal{H}_1$ is generally much more problematic. In addition, $\mathcal{H}_0$ is generally much more likely in practice than $\mathcal{H}_1$. Consequently, the value of α that we will require before being confident in our choice will be larger in the case $G(z, \hat{\theta}_N) \notin \mathcal{D}_{diag}$ than in the case $G(z, \hat{\theta}_N) \in \mathcal{D}_{diag}$. This suggests that, instead of (16), we could consider a decision rule that will opt for $\mathcal{H}_1$ only if the value of α corresponding to $\chi_{max,\mathcal{H}_1}$ is larger than a given threshold.

2.4. Re-identification

2.4.1. After performance diagnosis

When an observed performance drop is due to changes in disturbance characteristics ($\mathcal{H}_0$ is the correct hypothesis), it can be decided to let the controller intact (as disturbance changes are often temporary) or to restore the closed-loop performance by retuning the controller using the knowledge of the new disturbance characteristics $H(z, \hat{\theta}_N^{diag})$ identified along with $G(z, \hat{\theta}_N^{diag})$. Alternatively, when an observed performance drop is due to a control-relevant plant change ($\mathcal{H}_1$ is chosen), the MPC controller should be redesigned based on a model of the new plant dynamics G_0 to restore the closed-loop performance to its nominal level. As the model $G(z, \hat{\theta}_N^{diag})$ may not be sufficiently accurate for redesigning the controller, the diagnosis step is typically proceeded with another identification step when $\mathcal{H}_1$ is the true hypothesis. In this case, a new excitation signal $r_{id}(t)$ is applied to the closed-loop system to collect the data set $\{u(t), y(t)\,|\,t = 1, \cdots, N\}_{id}$.[6] In Section 3, we will show how we can design this excitation signal $r_{id}(t)$ in the least costly way while guaranteeing that the model identified in this way is sufficiently accurate for retuning of the MPC controller.

Since the proposed performance diagnosis method relies on system identification, a link is established between the diagnosis step and the plant re-identification step. This is done by identifying the new model not only based on the data $\{u(t), y(t)\}_{id}$, but also using the data $\{u(t), y(t)\}_{diag}$ obtained during diagnosis. This can be done by using $\hat{\theta}_N^{diag}$ and its covariance $P_{\theta,\,diag}$ in a regularization term. Hence, the parameter vector $\hat{\theta}_N^{id}$ of the re-identified plant model $\{G(z, \hat{\theta}_N^{id}), H(z, \hat{\theta}_N^{id})\}$ is determined by

$$\hat{\theta}_N^{id} = \arg\min_{\theta} \frac{1}{N}\left(\sum_{t=0}^{N-1} \epsilon^T(t,\theta)\Lambda^{-1}\epsilon(t,\theta) + \left\|\theta - \hat{\theta}_N^{diag}\right\|^2_{P^{-1}_{\theta,diag}}\right) \quad (20)$$

where $\epsilon(t, \theta)$ is computed using $\{u(t), y(t)\}_{id}$. Similarly as for $\hat{\theta}_N^{diag}$, the parameter vector $\hat{\theta}_N^{id}$ identified through Eq. (20) will be asymptotically normally distributed around the true parameter vector θ_0 i.e.

$$\hat{\theta}_N^{id} \sim \mathcal{N}(\theta_0, P_{\theta,tot}) \quad (21)$$

with $P_{\theta,\,tot}$ given by

$$P_{\theta,tot}^{-1} = P_{\theta,diag}^{-1} + P_{\theta,id}^{-1} \quad (22)$$

where $P_{\theta,\,diag}$ is the covariance matrix of $\hat{\theta}_N^{diag}$ and where $P_{\theta,\,id}$ is the covariance matrix that would have been obtained if only the data $\{u(t), y(t)\}_{id}$ would have been used; $P_{\theta,\,id}$ can be computed via the same formulas (13) and (14) as for $P_{\theta,\,diag}$. We see thus that, using both data sets to identify $\hat{\theta}_N^{id}$, enables us to increase the accuracy of $\hat{\theta}_N^{id}$ with respect to the situation where only $\{u(t), y(t)\}_{id}$ would have been used.

The model $G(z, \hat{\theta}_N^{id})$ will be used to replace G_{mod} in the expression of the MPC controller. The obtained controller will thus be $C(G(z, \hat{\theta}_N^{id}))$. In the sequel, we will show how we can optimally design the excitation signal r_{id} in such a way that the accuracy of the identified model is sufficient for the new MPC controller $C(G(z, \hat{\theta}_N^{id}))$ to deliver satisfactory performance (i.e. a performance level similar to the one at commissioning).

[6] For simplicity, it is assumed that the measurement sets $\{u(t), y(t)\}_{diag}$ and $\{u(t), y(t)\}_{id}$ have the same length N.

2.5. Constraint on the matrix $P_{\theta,\,tot}$

The accuracy of the identified model $G(z, \hat{\theta}_N^{id})$ can be measured by the covariance matrix $P_{\theta,\,tot}$ of $\hat{\theta}_N^{id}$. In order to know how *large* this matrix has to be for satisfactory control design, let us first define the performance objective for the new MPC controller $C(G(z, \hat{\theta}_N^{id}))$. For this purpose, let us suppose that the controller has been designed using the model $G(z, \theta)$ for an arbitrary value of θ. This leads to the controller $C(G(z, \theta))$. Using the performance index $\mathcal{J}$ used for performance monitoring (6), the MPC controller $C(G(z, \theta))$ will be deemed to deliver satisfactory performance if $\mathcal{J}(t, G_0, H_0, C(G(z,\theta)))$ is sufficiently close to the performance we would obtain if the MPC controller could have been designed with the actual $G_0 = G(z, \theta_0)$ i.e. $\mathcal{J}(t, G_0, H_0, C(G_0))$. In order to compare the performance of these two loops, let us for the moment suppose that we could simulate these two loops for a very long time N_{sim} and let us introduce the following quantity:

$$V_{perf}(\theta) = \frac{1}{N_{sim}}\sum_{t=1}^{N_{sim}}(\mathcal{J}(t, G_0, H_0, C(G(z,\theta))) - \mathcal{J}(t, G_0, H_0, C(G_0)))^2. \quad (23)$$

Using this metric V_{perf}, the set $\mathcal{D}_{perf}$ of admissible models $G(z, \theta)$ for control design is given by:

$$\mathcal{D}_{perf} = \{\ G(z,\theta)\ |\ V_{perf}(\theta) < \frac{1}{\gamma}\ \} \quad (24)$$

for some carefully chosen threshold[7] γ. As for the set $\mathcal{D}_{diag}$ used for diagnosis, (24) is an implicit definition of the set $\mathcal{D}_{perf}$. However, it is here also possible to deduce an approximative expression of this set which is explicit. We have indeed that $V_{perf}(\theta_0) = 0$ and $V_{perf}(\theta)$ has its minimum in θ_0. Consequently, like $\mathcal{D}_{diag}$, $\mathcal{D}_{perf}$ can be approximated using a Taylor expansion by an ellipsoid in the parameter space:

$$\mathcal{D}_{perf} \approx \{\ G(z,\theta)\ |\ (\theta-\theta_0)^T\mathcal{H}_{perf}(\theta-\theta_0) < \frac{2}{\gamma}\ \} \quad (25)$$

with $\mathcal{H}_{perf}$ the Hessian of $V_{perf}(\theta)$ evaluated in θ_0 (Hjalmarsson, 2009). The matrix $\mathcal{H}_{perf}$ can be determined numerically using a set of simulations around θ_0, see Ebadat et al. (2014).

Using the reasoning above, the identified parameter vector $\hat{\theta}_N^{id}$ after the re-identification step will be deemed satisfactory for MPC control design if $G(z, \hat{\theta}_N^{id}) \in \mathcal{D}_{perf}$ or equivalently if:

$$(\hat{\theta}_N^{id}-\theta_0)^T\mathcal{H}_{perf}(\hat{\theta}_N^{id}-\theta_0) < \frac{2}{\gamma}. \quad (26)$$

Using now the distribution (21) of $\hat{\theta}_N^{id}$, we see that, with probability α, $\hat{\theta}_N^{id} \in \{\ \theta \in \mathbf{R}^k\ |\ (\theta-\theta_0)^T P_{\theta,tot}^{-1}(\theta-\theta_0) < \chi\ \}$ when χ is chosen in such a way that $Pr(\chi^2(k) < \chi) = \alpha$. We can thus conclude with a probability α that (26) will be satisfied if the following

$$P_{\theta,tot}^{-1} > \frac{\gamma\chi}{2}\ \mathcal{H}_{perf}. \quad (27)$$

Important remark. Since θ_0 is unknown, we will replace it by an initial estimate to compute $\mathcal{H}_{ref}$. Since $\mathcal{H}_1$ has been chosen by the performance diagnosis module, the parameter vector θ_{mod} of the commissioning model G_{mod} is no longer a good estimate of θ_0. The best estimate of θ_0 at this moment is $\hat{\theta}_N^{diag}$ i.e. the parameter vector of the model identified via the diagnosis experiment.

Using (22), the inequality (27) can be rewritten as:

$$P_{\theta,id}^{-1} > \mathcal{R}_{perf} \quad (28)$$

with $\mathcal{R}_{perf} \triangleq \frac{\gamma\chi}{2}\ \mathcal{H}_{perf} - P_{\theta,diag}^{-1}$

[7] γ can be e.g. chosen equal to the one in (9)

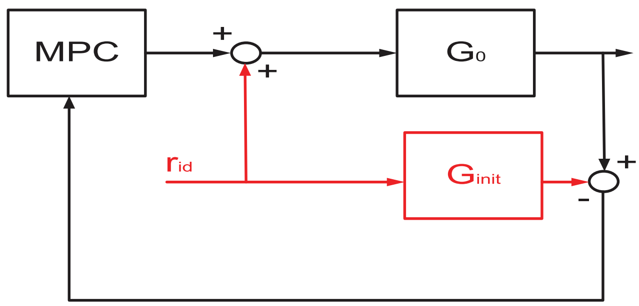

Fig. 2. Stealth excitation r_{id} in the presence of a nonlinear controller C_{id}.

where $P_{\theta,\,diag}$ is the covariance matrix of $\hat{\theta}_N^{diag}$, covariance matrix that can be easily estimated after the diagnosis experiment and where $P_{\theta,\,id}$ is the covariance matrix that can be tuned by the choice of the excitation signal r_{id} that we want to design optimally. The loop on which this excitation signal is applied is the loop (G_0, H_0, $C(G_{mod})$) with the MPC controller. As mentioned in the introduction, due to the presence of this implicit and nonlinear controller $C(G_{mod})$ in the loop, $P_{\theta,id}^{-1}$ is not affine in the power spectrum of the to-be-designed r_{id}. Consequently, the optimal design of r_{id} is not as straightforward as in Bombois et al. (2006) which considers a setting where (28) constitutes an LMI in the decision variables. One could of course linearize the MPC controller (Maciejowski, 2002) and consider this linearization in the design. In the following, we will also present two other approaches, namely stealth excitation and MPC-x, to circumvent that problem.

3. Optimal design of the least costly excitation[8]

3.1. Stealth excitation

The first solution to the problem of having an MPC controller in the loop instead of an LTI controller is the so-called *stealth excitation* (Potters et al., 2014). The idea behind this concept is to ensure that the excitation $r_{id}(t)$ affects the true system in an open-loop fashion. The stealth excitation procedure is illustrated in Fig. 2. Our objective is thus to have the following relations:

$$y(t) = G_0(z)r_{id}(t) + y_e(t)$$
$$u(t) = r_{id}(t) + u_e(t) \tag{29}$$

where $y_e(t)$ and $u_e(t)$ correspond to the output and input signals that would be generated by the closed loop if there was no excitation $r_{id}(t)$. Such relations can be obtained by modifying the feedback term: instead of feeding back $y(t)$ to the controller, we feed back $y(t) - G_0(z)r_{id}(t)$. We ensure in this way that the MPC controller $C(G_{mod})$ does not "see" $r_{id}(t)$ via the feedback term. In fact, this modification is achieved by modifying the correction c_t of the MPC controller given in (5) since this correction is the only feedback term of the MPC controller.

Since the system is operated in open loop with respect to $r_{id}(t)$, the prediction error can be written as follows using (29):

$$\epsilon(t,\theta) = H^{-1}(z,\theta)(G_0(z) - G(z,\theta))r_{id}(t) + \underbrace{H^{-1}(z,\theta)(y_e(t) - G(z,\theta)u_e(t))}_{\epsilon_e(t,\theta)}$$

where $\epsilon_e(t,\,\theta)$ is the prediction error that would be obtained if $r_{id}(t) = 0$. Since ϵ_e is independent of $r_{id}(t)$, using (13), the inverse of the covariance matrix can be expressed as:

$$P_{\theta,id}^{-1} = M_r \;\; + \;\; M_e$$

with M_r a matrix which is affine in the spectrum $\Phi_{r_{id}}(\omega)$ (see Barenthin, Bombois, Hjalmarsson, & Scorletti, 2008) and a matrix M_e corresponding to the term $\epsilon_e(t,\,\theta)$. The matrix M_e could be evaluated via simulation (when $r_{id} = 0$) or neglected for experiment design since $M_e \geq 0$.

It is important to note that the modified feedback term (i.e. $y(t) - G_0(z)r_{id}(t)$) is a function of the unknown true system G_0. Consequently, to realize the stealth excitation, we replace G_0 by an initial estimate $G_{init}(z)$ (chosen e.g. equal to $G(z,\hat{\theta}_N^{diag})$) in this modified feedback term.

3.2. Optimal experiment design problem

In this subsection, we will formulate the optimal experiment design problem that will lead to the least costly (re-)identification experiment to identify a model $G(z,\hat{\theta}_N^{id})$ that is sufficiently accurate for control. We will suppose in the sequel that the length N of the re-identification experiment is fixed. The only decision variable is thus the power spectrum $\Phi_{r_{id}}(\omega)$ of r_{id}. To obtain a sufficiently accurate model for control, the inverse of the covariance matrix $P_{\theta,\,id}$ should satisfy the constraint (28). To formalize our optimal experiment design problem, we also need to define what we mean by the cost $\mathcal{J}_{r_{id}}$ of the identification experiment. This cost can be e.g. measured by the power $\mathcal{P}_r$ of the excitation signal $r_{id}(t)$. However, we can also consider an alternative definition which seems closer to the actual cost of a closed-loop experiment i.e. a linear combination of the power of the perturbation induced by r_{id} on the input and output of the system (Bombois et al., 2006). In the stealth framework, these perturbations are equal to $u_r = r_{id}$ and $y_r = G_0(z)r_{id}$, respectively (see (29)). The cost is thus defined as:

$$\mathcal{J}_{r_{id}} = \beta_y \mathcal{P}_{y_r} + \beta_u \mathcal{P}_{u_r} \tag{30}$$

where the scalars β_u and β_y can be freely chosen. It is important to note that this cost $\mathcal{J}_{r_{id}}$ is linear in the power spectrum $\Phi_{r_{id}}(\omega)$ of r_{id}.

Now, we are ready to formulate the convex optimization problem that will lead to the spectrum of the optimal excitation r_{id}:

$$\min_{\Phi_{r_{id}}(\omega)} \quad \mathcal{J}_{r_{id}}$$

subject to the constraint that (31)

$$P_{\theta,id}^{-1} > \mathcal{R}_{perf}$$

Since both $\mathcal{J}_{r_{id}}$ and $P_{\theta,id}^{-1}$ are affine in the power spectrum $\Phi_{r_{id}}(\omega)$ of r_{id}, this optimization problem can be solved easily. For this purpose, a linear parametrization of $\Phi_{r_{id}}(\omega)$ has to be chosen (see e.g. Bombois et al., 2006; Jansson & Hjalmarsson, 2005).

Remark. When applying an external excitation, the constraints (4) can be violated. As far as the output constraints are concerned, it is a good idea to first move the set-points for the outputs further away from the constraints than when $r_{id} = 0$. In order not to violate the input constraints, we should limit the amplitude of $r_{id}(t)$ by playing with the length N of the identification. However, research has still to be performed to determine on how to deal optimally with the MPC constraints in the stealth framework. For this purpose, the work in Larsson, Hägg, and Hjalmarsson (2016) could be an important inspiration.

[8] The technical contents of this section have appeared in Potters et al. (2014), Larsson, Rojas, Bombois, and Hjalmarsson (2015) and Larsson, Ebadat, Rojas, Bombois, and Hjalmarsson (2016).

3.3. Optimal design of the least costly excitation - the MPC-x approach

Stealth excitation consists of fooling the MPC controller $C(G_{mod})$ in such a way that we can apply the classical least costly experiment design approach of Bombois et al. (2006) with the spectrum $\Phi_{r_{id}}(\omega)$ of the excitation signal r_{id} as decision variable. This signal r_{id} is an external signal whose spectrum can be chosen by the user. An alternative is to integrate the excitation constraint directly in the MPC. This is done in MPC-x: Model Predictive Control with excitation. For this purpose, we use the fact that an MPC controller is in essence an optimization algorithm subject to constraints (see Section 2.1). The MPC algorithm is modified by adding the additional constraint that (28) has to be satisfied after the identification duration N. This approach remains a least costly approach since the modified MPC controller will still minimize a cost function aiming at reducing the variance of the input and the output of the system, and it will at the same time guarantee that the input-output data are sufficiently informative to obtain a model with sufficiently small uncertainty for control. From a technical point of view, integrating (28) as an extra constraint to the MPC algorithm is quite a challenge (e.g. the design variable is here the actual input sequence $u(t)$ and no longer a spectrum).

Suppose that we are the beginning of the re-identification experiment. Let us denote this instant $t = 1$. Without extra identification constraint, the MPC algorithm has to determine the optimal value of $u(t)$, $u(t+1)$, .. $u(t+M)$ in such a way that the criterion (2) is minimized under the constraints (4). Recall that $u(t+j) = u(t+M)$ for all $M < j < P$ in (2) and that the prediction horizon P is much smaller than the duration N of the identification experiment. Since (28) has to be satisfied at the end of the identification experiment, the constraint that we will add to the MPC algorithm at $t = 1$ is the following:

$$\mathcal{I}_1^P > \frac{P}{N}\,\mathcal{R}_{perf} \tag{32}$$

where $\mathcal{I}_1^P$ is the part of the inverse of $P_{\theta,tot}^{-1}$ that will be realized with the data $\{u(t),\ y(t),\ t = 1\ldots P\}$. According to (14), this part is given by:

$$\mathcal{I}_1^P = \frac{N}{P}\sum_{t=1}^{P}\left[\left(\frac{\partial\epsilon(t,\theta)}{\partial\theta}|_{\hat{\theta}_N^{diag}}\right)\hat{\Lambda}^{-1}\left(\frac{\partial\epsilon(t,\theta)}{\partial\theta}|_{\hat{\theta}_N^{diag}}\right)^T\right] \tag{33}$$

We evaluate $\frac{\partial\epsilon(t,\theta)}{\partial\theta}$ in $\hat{\theta}_N^{diag}$ since it is the best estimate of θ_0 (or of $\hat{\theta}_N^{id}$) that we have at this moment. The gradient $\frac{\partial\epsilon(t,\theta)}{\partial\theta}$ of the prediction error at time $t = t^*$ is in general both a function of the input and the output until $t = t^*$. If the input and output are known for $t < 1$ and the input signal for $t = 1\ldots P$ is the decision variable of our optimization problem, the output signal $y(t)$ for $t = 1\ldots P$ is unfortunately unknown due to the fact that this output is not only a function of the input, but also of the unknown noise realization $H_0e(t)$. This is a classical problem in the literature on dual control (see Pronzato, 2008; Tse & Bar-Shalom, 1973). However, if we suppose that $H_0 = 1$ (output error true system), the gradient $\frac{\partial\epsilon(t,\theta)}{\partial\theta}$ becomes only a function of the input signal.

Note that the actual re-identification will pertain to both the identification of the plant model $G(z,\theta)$ and noise model $H(z,\theta)$ to avoid a biased estimate (see (20)). However, for the optimal experiment design, we will suppose that the noise model H_0 is equal to one and consider the covariance matrix corresponding to this modified true system for the optimal experiment design.

Under this approximation, the added identification constraint (32) has the same decision variable as the one in the original MPC algorithm (see Section 2.1). In Larsson, Ebadat, Rojas, Bombois, and Hjalmarsson (2016), it is furthermore shown that, using the convex relaxation presented in, e.g. Manchester (2010), the MPC algorithm with the extra constraint (32) can be approximated as an LMI optimization problem.

From the input sequence delivered by this LMI optimization problem, only the first element i.e. $u(t)$ will be applied owing to the receding horizon mechanism and the procedure will have to be restarted. Let us suppose that we have now arrived at an arbitrary time t $(t < N)$. The original MPC algorithm will have to be re-run with the additional *identification* constraint:

$$\mathcal{I}_1^{t+P} > \frac{t+P}{N}\,\mathcal{R}_{perf} \tag{34}$$

where $\mathcal{I}_1^{t+P}$ is the sum of the part $\mathcal{I}_1^{t-1}$ of the inverse of $P_{\theta,tot}^{-1}$ that has been realized with the data until the present time t and the one $\mathcal{I}_t^{t+P}$ that will be realized over this prediction horizon. In this sum, the first part is known and the second part is a function of the decision variable, similarly as in (33).

Remark. Unlike the stealth approach, the MPC-x approach deals with the constraints (4) in an elegant way since these constraints have become constraints of our combined *optimal identification experiment design-MPC control* problem. On the other hand, the approximation of considering $H_0 = 1$ for the design of the informative signal u can lead to a situation where the identified model will not be as accurate as desired.

4. (Auto)Tuning of model based control systems[9]

The tuning of MPC involves selecting the parameters in the cost function, the disturbance model and the state observer if the state-space formulation is used. In Garriga and Soroush (2010) it was reported that a majority of the studies on MPC tuning fix the prediction horizon at a value that covers the main dynamics of the open-loop system, and select the control horizon based on computational capacity. Garriga and Soroush (2008), Trierweiler and Farina (2003) and Lee and Yu (1994) address the tuning problem by analysing the performance specifications of MPC such as closed-loop poles, robustness and sensitivity functions. In Maurath et al. (1988) and Shridhar and Cooper (1998) the tuning parameters are found by considering the conditioning of the control law. In current practice of MPC, the tuning problem is handled in a heuristic manner.

The review of tuning and auto-tuning methods shows that, generally, the input weight is tuned to find a good trade-off between nominal performance and robustness. In process industry, this optimum is not always obtained due to various reasons. For example, the tuning settings at commissioning cannot adapt to changes in plant dynamics or disturbance characteristics. With this observation in mind, in this project we aimed to develop a (auto)tuning method which finds the optimal trade-off between nominal performance and robustness.

The objective of MPC in process industry is to reduce the variance of the key performance (money making) variable and then to push it towards the system constraints so that the system operates closely to its economically optimal condition. Therefore, the variance of the key variable is a good indicator of the performance of the closed-loop system. In addition, the closed-loop performance, the tuning of controllers and the model accuracy are interrelated. This relationship has been extensively studied and presented in robust control theory (Skogestad & Postlethwaite, 2005) using frequency-domain techniques. It was shown that the performance of the closed-loop system becomes sensitive to the model uncertainty at a certain bandwidth of frequency. Increasing the

[9] The technical contents of this section have appeared in Tran (2015), Tran et al. (2015) and Tran, Octaviano, Özkan, and Backx (2014)

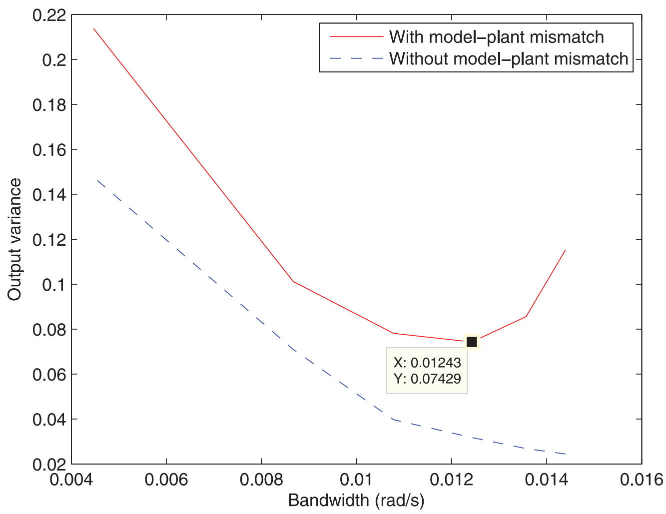

Fig. 3. Relation between closed-loop bandwidth and output variance in the quadruple tank system.

closed-loop bandwidth further results in closed-loop performance deterioration.

A similar analysis can also be done for MPC by representing the cost function as functions of sensitivity, complementary sensitivity functions and weighting matrices. In robust control theory, the sensitivity and complementary sensitivity functions are tuned to adjust the closed-loop bandwidth. In the case of MPC, the closed-loop bandwidth is determined by the weighting matrices on the controlled variables (CV·s) and manipulated variables (MV's). The same impact of uncertainty as in the performance of robust controllers is observed in the closed loop performance of MPC. The closed loop performance (i.e. decrease in the output variance) is improved as the bandwidth is increased until the plant-model mismatch becomes significant. Further increasing the bandwidth results in performance degradation (increase in the output variance). The analysis is illustrated using the model of quadruple tank example of Johansson (1997) and Åkesson (2006) in Fig. 3. More detailed explanation of this analysis can be found in Tran (2015).

The observation in Fig. 3 forms the basis of the work in this research. To this end, our research efforts are directed towards

- Developing an auto-tuning method that enables the closed loop system to achieve the optimal closed loop bandwidth. The auto-tuning method must also be able to steer the system to a new optimal closed loop bandwidth if changes in the plant dynamics or disturbance characteristics occurs.
- Developing a systematic approach for the calculation of the weights in the objective function such that the resulting closed loop bandwidth matches the desired optimal bandwidth which balances the performance and robustness.

These objectives are expressed in a two layer approach schematically shown in Fig. 4

In the top layer, we seek to find the optimal closed bandwidth that corresponds to the minimum point depicted in Fig. 3. This search can be done manually or can be automated. For the latter, we have exploited extremum seeking. Once the closed loop optimal bandwidth is identified, it is passed to the bottom layer. This information is then used for the calculation of the weighting matrices such that the optimal closed loop is achieved. To this end, we have investigated optimisation based or reversed engineering type of methods. In all these developments, we assume that the constraints are inactive.

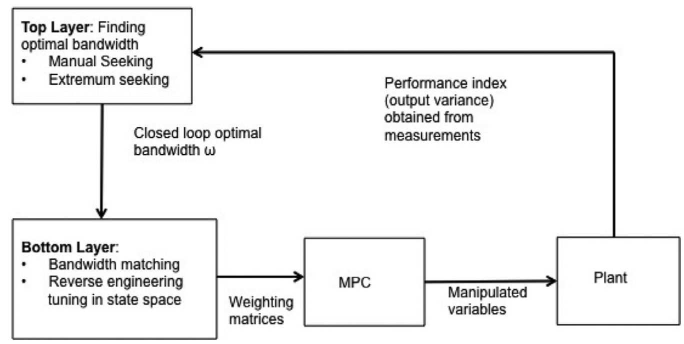

Fig. 4. Two-layer tuning procedure.

The use of the tuning approach explained above is not just limited to the commissioning phase of MPC or the adjustment of the control strategy when it has been shown that retuning would improve the closed loop performance. The tuning algorithm can also be used to determine the optimal weights for the controller $C(G(z, \theta))$ that will be used (see Section 2.5) to design the least costly identification experiment whenever a performance drop due to a plant change is observed.

This section is organised as follows. Section 4.1 describes the methods developed in the top layer, specifically, how the extremum seeking method is used in automating the calculation of the optimal closed loop bandwidth. This section is followed by technological developments in the bottom layer which aims at finding the corresponding objective function (weighting matrices) that will achieve the desired closed loop bandwidth.

4.1. Determining the optimal closed loop bandwidth

4.1.1. Manual seeking

In this approach, the optimal closed-loop bandwidth is found by manually changing the bandwidth and monitoring the output variance. Only the weighting matrices are tuned in the procedure while the horizons are selected based on common engineering rules. In addition, based on the model of the MPC, an initial range for the closed-loop bandwidth can also be determined. This range is then divided into a number of segments.

1. Initialization: Determine a certain initial bandwidth based on disturbance characteristics and available information on modelling errors. The information on disturbance and noise can be obtained by investigating the behaviour of the open-loop plant at steady state, while an initial idea on the frequency range in which the model is correct can be found from identification tests. In this method, the initial bandwidth is chosen close to the bandwidth of the open-loop plant and sent to the bottom layer where the weighting matrices are computed and implemented. The output variance for this initial bandwidth is then calculated.
2. Seeking: Increase the closed-loop bandwidth to the next value in the bandwidth range. Find the corresponding MPC tuning parameters with the bottom layer and implement them. Compute the output variance.
3. Monitoring: Monitor the output variance. Compare it with the previous tuning. If the variance decreases, keep raising the bandwidth. If it increases, the previous tuning is assumed to be optimal.

If the output variance still decreases at the end of the initial frequency range, the closed-loop bandwidth will be increased

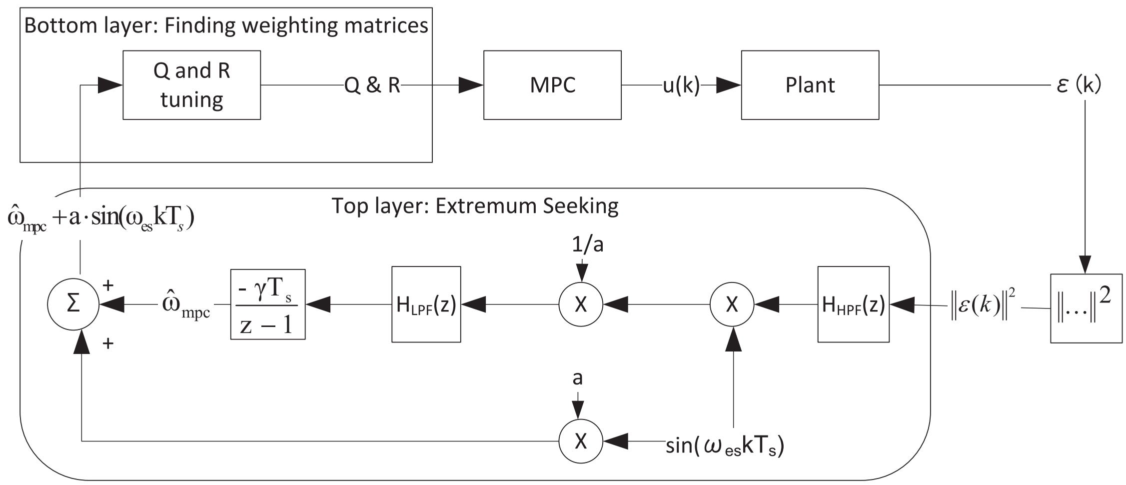

Fig. 5. Overview of the 2-layer auto-tuning method using extremum seeking.

until the output variance increases. This implementation is used to ensure that a robust starting point is applied.

4.1.2. Extremum Seeking and its application to MPC tuning

Extremum seeking is a form of non-model-based adaptive optimisation. It belongs to the class of gradient-based optimisation techniques and deals with systems that have an unknown input/output relation, but are known to have an extremum. The basic idea is to perturb the input of the extremum seeking system with a so-called dither signal and multiply the extremum seeking system output (i.e. the output variance of the plant) with the same dither signal. The result of this multiplication gives an estimate of the system's gradient and this gradient estimate is used to move the input parameter (i.e. the closed-loop bandwidth) to its extremum, which minimises (or maximises) the output of the extremum seeking system.

The method has existed since the 1920s, was extensively investigated in the 1950s and 1960s and has gained renewed interest since the beginning of the new millennium (Stanković & Stipanović, 2010; Tan, Moase, Manzie, Nešić, & Mareels, 2010). It has been used successfully for PID tuning (Killingsworth & Krstić, 2006), maximum power point tracking of a photo-voltaic system (Levya, Artillan, Cabal, Estibals, & Alonso, 2011), tuning the parameters of mobile sensors (Stanković & Stipanović, 2010) and many other applications (Tan et al., 2010).

The ability to find the optimum of an unknown mapping of extremum seeking makes it suitable for finding the optimal closed-loop bandwidth in MPC, under the assumption that the function from the closed-loop bandwidth to the output variance is convex. Not only is the relation between the closed-loop bandwidth and the output variance difficult to obtain analytically in the presence of disturbances and model-plant mismatch, the optimum of this relation is subject to changes due to variation in the plant dynamics and disturbance characteristics. These changes can be gradual or drastic. Hence, a method that continuously perturbs the closed-loop bandwidth to bring it to the optimum is also capable of steering it to the new optimum if that optimum varies over time. Extremum seeking is therefore a suitable method to use in the top layer of the two-layer auto-tuning method. When the extremum seeking method is applied to the auto-tuning approach, the prediction horizon is fixed so as to cover the main dynamics of the system, and the control horizon is chosen according to the computational capacity of the system. A schematic overview of the extremum seeking algorithm in the two-layer auto-tuning method is given in Fig. 5.

4.2. Determining the weighting matrices

In the previous section, the top layer of the proposed (auto)tuning technology was presented. In this section, we explain the developments in the bottom layer.

4.2.1. Bandwidth matching-a practical approach

Once a desired closed-loop bandwidth ω_{mpc} is chosen, there exist several methods to find the corresponding tuning parameters **Q** and **R** of MPC. In Rowe and Maciejowski (2000b), Shah and Engell (2011) and Tran et al. (2014) the approach is to adapt controller matching in which MPC is matched to a desired H_∞ controller while Lee and Yu (1994) tunes the Kalman filter and disturbance model to achieve the desired closed-loop bandwidth.

It can be easily shown that the weighting factors of the input energy and output energy are correlated. Fixing the input weights and increasing the output weights, or fixing the output weights and reducing the input weights, both raise the closed-loop bandwidth. Hence, in this work, the input weight is selected as the free variable to adjust the closed-loop bandwidth. We choose the weighting matrices $\mathbf{Q} = \mathbf{I}$ and $\mathbf{R} = \rho\mathbf{I}$, where $\mathbf{I}$ is the identity matrix and ρ is a scalar. In that case, the closed-loop bandwidth is solely determined by the selection of the scalar ρ.

When the constraints are inactive, MPC can be considered to be a linear time-invariant controller. Therefore, it is possible to derive the sensitivity and complementary sensitivity functions of a system with MPC as a linear controller. In an $n \times n$ system, the sensitivity function has n cut-off frequencies corresponding to n directions of the system. To find the ρ that corresponds to the desired closed-loop bandwidth ω_{mpc}, the following optimisation problem is solved:

$$\min_{\rho} ||\omega_{mpc} - \omega_{cs}(\sigma_1(\rho))||^2 \tag{35}$$

where $\omega_{cs}(\sigma_1(\rho))$ is the crossover frequency of the maximum singular value of the sensitivity function. The crossover or cut-off frequency of the singular value plot is defined by the lowest frequency at which the singular value crosses 0 dB. In this optimisation problem ρ is the only optimisation variable. The rationale behind this choice is that increasing ρ will lead to a decrease in the maximum crossover frequency of the sensitivity function and reducing ρ will result in an increase in that frequency. Therefore, the optimisation problem becomes a convex one and can be easily solved by a search method. This is the tuning method that has been implemented on the experiments performed on the SASOL plant (Guidi, Larsson, Tran, and Backx, 2014, see Section 5).

This practical approach presented above considers only one cross over frequency and can be seen as a simple loop shaping using optimisation. If more elements in the frequency-domain sensitivity function (e.g. its slope) are considered, more complex weighting matrices will be required to obtain a feasible solution. Then, the optimisation problem is non-convex optimisation problem due to the non-linear mapping from the weighting matrices to the sensitivity function. In order to circumvent the non-convex optimisation problem, the well known controller matching problem is revisited.

4.2.2. Controller matching for general predictive control

A very active research topic in MPC tuning is controller matching, which aims to compute the tuning parameters such that MPC matches an arbitrary LTI controller, also referred to as the favourite controller. The motivation for this approach is to use the available degrees of freedom in the MPC cost function while the constraints are inactive. This is not straightforward (Hartley & Maciejowski, 2011). When MPC operates closely to the constraints and when active constraints occur frequently, the system will take advantage of the constraint handling ability of MPC. When MPC operates away from the constraints (e.g. at commissioning), the system can inherit the characteristics of an LTI controller, e.g. its robustness.

The matching of MPC with an LTI controller, when MPC is formulated based on the state-space models, has been investigated by several authors (Chmielewski & Manthanwar, 2004; Di Cairano & Bemporad, 2010; Hartley & Maciejowski, 2011; 2013; Rowe & Maciejowski, 2000b and Rowe & Maciejowski, 2000a). In such a formulation, the unconstrained solution of MPC can be written as a state feedback control law and the aim of the matching is to minimize the error between the state feedback gain of the favourite controller and that of MPC. The foundation of this approach is the inverse problem of linear optimal control, laid by Kalman (1964) and Anderson and Moore (1971). This problem aims at finding the weighting matrices of the linear quadratic regulator (LQR) in order to match a given linear feedback gain. The inverse optimality problem is extended to a more general cost function in Kreindler and Jameson (1972) with a cross-product term between the state and the control input. In Di Cairano and Bemporad (2010), a matching method based on formulating an optimization problem with linear matrix inequality (LMI) or bilinear matrix inequality (BMI) constraints is proposed. The cost function of the optimization problem is the error between the control action of the MPC and the favourite controller. The above matching methods based on the inverse problem of linear optimal control usually consider the case in which the MPC is equivalent to an LQR and the states of the system are available.

We extend the controller matching problem for MPC formulations based on transfer function models, also known as Generalized Predictive Control (GPC). The MPC based on transfer function models has been introduced in Clarke, Mohtadi, and Tuffs (1987a, 1987b) and further developed in Rossiter (2005). It is also adopted by several MPC providers in process industry (Qin & Badgwell, 2003). We consider a process of Controlled-AutoRegressive Integrated Moving Average(CARIMA) form with n_u inputs, n_d measurable disturbances and n_y outputs:

$$D(z)y(t) = N(z)u(t) + F(z)d(t) + \frac{T(z)}{\Delta(z)}\nu(t) \tag{36}$$

where $N(z)^{n_y \times n_u}$ and $D(z)^{n_y \times n_y}$ are the numerator and denominator matrices of the system respectively; $F(z)^{n_y \times n_d}$ is the numerator of the feed-forward model from d_k to y_k; $T(z) = l(z)I_{n_y}$ is a diagonal transfer matrix used to model the disturbance signal and usually considered as a design parameter (Clarke et al., 1987a; 1987b)); $\Delta(z)$ is the difference operator $\Delta(z) = 1 - z^{-1}$; $y(t) \in \mathbb{R}^{n_y}$, $u(t) \in \mathbb{R}^{n_u}$, $d(t) \in \mathbb{R}^{n_d}$ and $\nu(t) \in \mathbb{R}^{n_y}$ represent the output, input, measurable disturbance and a zero-mean random variable at time instant k respectively.

In the rest of this section, we consider a system where $n_u = n_y = n_d$ for the sake of simplicity, although the proposed method can easily be extended to the case of non-square systems. The output reference is assumed to be incorporated in the model and the aim of the controller is to steer the outputs of the system to zero. Since $T(z) = l(z)I_{n_y}$ and $n_u = n_y = n_d$, $T(z)$ can be treated as a scalar term in the computation.

The process output can be expressed as in the following

$$\underset{\rightarrow t}{y} = \tilde{H}\Delta U + \tilde{P}\Delta \underset{\leftarrow t-1}{\tilde{u}} + \tilde{Q}\underset{\leftarrow t}{\tilde{y}} + \tilde{R}\Delta \underset{\leftarrow}{\tilde{d}}_t \tag{37}$$

where $\tilde{H}$, $\tilde{P}$, $\tilde{Q}$, $\tilde{R}$ are obtained from D, N, F. $\underset{\rightarrow t}{y}$, $\underset{\leftarrow}{\tilde{y}}_t$, $\underset{\leftarrow t-1}{\tilde{u}}$, $\underset{\leftarrow}{\tilde{d}}_t$ are the vector of future outputs, vector of past outputs, vector of past incremental inputs and vector of past disturbances respectively. The model in (37) is used to compute the optimal input sequence by solving the following optimization problem:

$$\min_{\Delta U} J_t = \underset{\rightarrow t}{y}^{\top} \mathcal{Q} \underset{\rightarrow t}{y} + \Delta U^{\top}\mathcal{R}\Delta U \tag{38}$$

where $\mathcal{Q} \in \mathbb{R}^{H_p n_y \times H_p n_y}$ and $\mathcal{R} \in \mathbb{R}^{H_c n_u \times H_c n_u}$ are the weighting matrices penalizing the outputs and input increments. It is common that $\mathcal{Q}$ and $\mathcal{R}$ are chosen to be positive definite. The first element of the unconstrained solution to optimization problem (38) at time instant k is then given by:

$$\Delta u(t) = -\tilde{N}_t \underset{\leftarrow t}{\tilde{y}} - \breve{\tilde{D}}_t \Delta \underset{\leftarrow t-1}{\tilde{u}} - \tilde{F}_t \Delta \underset{\leftarrow}{\tilde{d}}_t \tag{39}$$

where

$$\begin{cases} \breve{\tilde{D}}_t = \Phi\left(\tilde{H}^{\top}\mathcal{Q}\tilde{H} + \mathcal{R}\right)^{-1}\tilde{H}^{\top}\mathcal{Q}\tilde{P} = \tilde{K}_{MPC}\tilde{P} \\ \tilde{N}_t = \Phi\left(\tilde{H}^{\top}\mathcal{Q}\tilde{H} + \mathcal{R}\right)^{-1}\tilde{H}^{\top}\mathcal{Q}\tilde{Q} = \tilde{K}_{MPC}\tilde{Q} \\ \tilde{F}_t = \Phi\left(\tilde{H}^{\top}\mathcal{Q}\tilde{H} + \mathcal{R}\right)^{-1}\tilde{H}^{\top}\mathcal{Q}\tilde{R} = \tilde{K}_{MPC}\tilde{R} \end{cases} \tag{40}$$

and $\Phi = \begin{bmatrix} I_{n_u} & 0 & \cdots & 0 \end{bmatrix}$. Let us define

$$\tilde{N}_t(z) = \tilde{N}_t\begin{bmatrix} I & z^{-1} & \cdots & z^{-n} \end{bmatrix}^{\top} \tag{41}$$

$$\breve{\tilde{D}}_t(z) = \breve{\tilde{D}}_t\begin{bmatrix} z^{-1} & z^{-2} & \cdots & z^{-n-1} \end{bmatrix}^{\top} \tag{42}$$

$$\tilde{F}_t(z) = \tilde{F}_t\begin{bmatrix} I & z^{-1} & z^{-2} & \cdots & z^{-n} \end{bmatrix}^{\top}. \tag{43}$$

The transfer matrix representation of the control law is given by:

$$\left(I + T^{-1}(z)\,\breve{\tilde{D}}_t(z)\right)\Delta u(t) = -T^{-1}(z)\tilde{N}_t(z)y(t) - T^{-1}(z)\tilde{F}_t(z)\Delta d(t) \tag{44}$$

$$\Rightarrow \left(T(z) + \breve{\tilde{D}}_t(z)\right)\Delta u(t) = -\tilde{N}_t(z)y(t) - \tilde{F}_t(z)\Delta d(t) \tag{45}$$

$$\Rightarrow \tilde{D}_t(z)\Delta u(t) = -\tilde{N}_t(z)y(t) - \tilde{F}_t(z)\Delta d(t). \tag{46}$$

Given a favourite output-feedback controller, we find the weighting matrices of GPC such that the GPC behaves like the favourite controller when constraints are inactive. The matching consists of two steps. Step 1 finds the linear feedback gain $\tilde{K}_{MPC}$ from the transfer matrix of the favourite controller and step 2 computes the weighting matrices in the cost function of GPC from the obtained feedback gain in step 1.

Assume that the proper favourite (but not necessarily strictly proper) controller is given by:

$$\begin{aligned} &\left(I + A_1 z^{-1} + A_2 z^{-2} + \cdots + A_{n-1} z^{-n+1}\right)u_k \\ &\quad = -\left(\tilde{B}_0 + \tilde{B}_1 z^{-1} + \cdots + \tilde{B}_{n-1} z^{-n+1}\right)y_k \end{aligned}$$

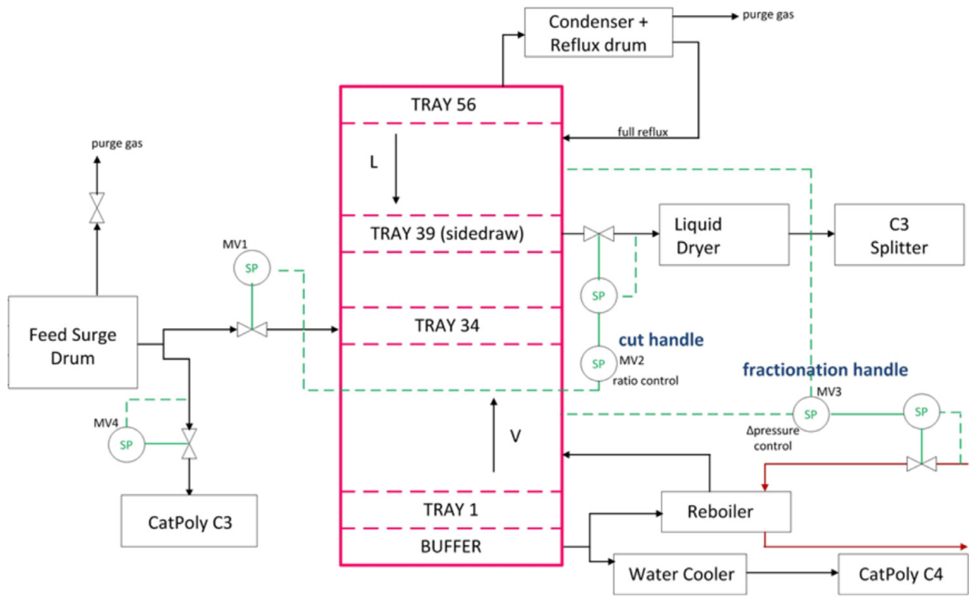

Fig. 6. FT-depropaniser.

$$- \left(C_0 + C_1 z^{-1} + \cdots + C_{n-1} z^{-n+1}\right) d_k. \tag{47}$$

Hence:

$$\begin{aligned}&\left(I + A_1 z^{-1} + \cdots + A_{n-1} z^{-n+1}\right) \Delta u_k \\ &= -\left(\tilde{B}_0 + \cdots + \tilde{B}_{n-1} z^{-n+1}\right)\left(1 - z^{-1}\right) y_k \\ &\quad - \left(C_0 + \cdots + C_{n-1} z^{-n+1}\right) \Delta d_k.\end{aligned} \tag{48}$$

$$\Rightarrow A(z) \Delta u_k = -B(z) y_k - C(z) \Delta d_k. \tag{49}$$

where $A(z) = I + A_1 z^{-1} + \cdots + A_{n-1} z^{-n+1}$, $B(z) = B_0 + B_1 z^{-1} + \cdots + B_n z^{-n}$ and $C(z) = C_0 + C_1 z^{-1} + \cdots + C_{n-1} z^{-n+1}$.

To investigate the matching problem, two subproblems are studied:

- Matching transfer matrices: Find the controller gain $\tilde{K}_{MPC}$ in (40) such that $\tilde{N}_k(z) = B(z)$, $\tilde{D}_k(z) = A(z)$ and $\tilde{F}_k(z) = C(z)$.
- Finding the tuning parameters: Construct the cost function (38) such that $\left(\tilde{H}^\top \mathcal{Q} \tilde{H} + \mathcal{R}\right)^{-1} \tilde{H}^\top \mathcal{Q} = \tilde{K}_{MPC}$.

In Step 1 of the problem formulation results we solve a set of linear equations to find the output feedback gain of GPC. To this end, the rank conditions of coefficient matrices are investigated. Once the rank conditions are fulfilled, an output feedback gain that guarantees the matching can always be found. In Step 2, a convex optimisation problem with LMI constraints similar to Di Cairano and Bemporad (2010) is used to find the tuning parameters which provide the computed output feedback gain. The degrees of freedom of the convex optimisation problem are increased by extending the objective function of the GPC with cross-product terms between the outputs and inputs. This approach does not require any loop-shifting technique to tackle the feed-through term from output to input in the controller and also allows a control horizon greater than 1. A more detailed explanation of this contribution can be found in Tran et al. (2015)

5. Industrial case study: SASOL FT-depropanizer

The FT-depropanizer is a distillation column which aims to separate C_3 (propane,propylene) and lighter components from C_4 (butane, ibutane, etc.) and heavier components. It is situated in the cold side of the SASOL Secunda Synthetic fuels refinery. The column has 56 trays, a total reflux in the top of the column and a side draw section above tray 38. The feed enters the column on tray 34. An illustration of the column is given in Fig. 6. The feed to the column comes from several units but is received directly from a feed surge drum. Variance in the feed quality, specifically C_4 content, is the most significant disturbance to the unit. All the streams leaving the unit are fed to other critical processing units hence the efficient control of this column is important for the general performance of the downstream processing. On top of the base layer control, an MPC is running with the flows from the feed surge drum to the column (MV1) and to the Cat Poly C3 header (MV4), the side draw flow (MV2 - cut handle), which is controlled as a ratio of the feed to the column, and the column delta pressure (MV3 - fractionation handle) as the manipulated variables. The MPC specifies the set-points (SP) of these four MVs with the individual base-layer PID controllers implementing the necessary adjustments on the final control elements. The control objectives for this column are

- Maximise the production rate by increasing the feed (when possible) and the side draw flow.
- Maintain the product specification which is reflected in the amount of C_4's in the side draw.
- Prevent flooding in the feed drum by manipulating feed flow to the Cat Poly C_3 header
- Prevent flooding in the column by imposing an upper bound on the differential pressure.
- Prevent flaring by imposing an upper bound on the column pressure.
- Prevent sudden and dramatic breakthrough of C_4's by imposing an upper limit on the bottom temperature.

One of the technical developments validated on this distillation has been tuning by manual seeking. As explained in Section 4.1.1, this approach seeks to find the optimal closed loop bandwidth by manually changing it and continuously monitoring the variance of the key performance variable. In this column, the key performance variable is the C_4 content in the side draw. Due to the limit on the amount of time for experiments in the plant, we were only able to investigate three closed loop bandwidths. The resulting C_4's content in the side draw for a closed loop bandwidth of 0.0005 rad/s,

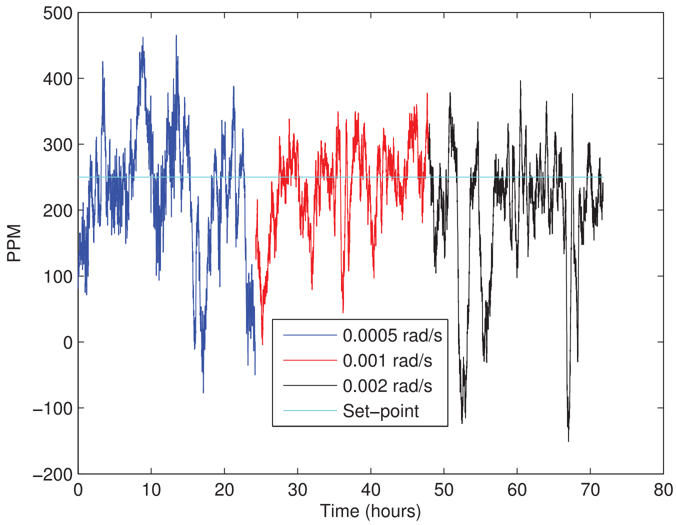

Fig. 7. C_4 content in the side draw at three different closed-loop bandwidths.

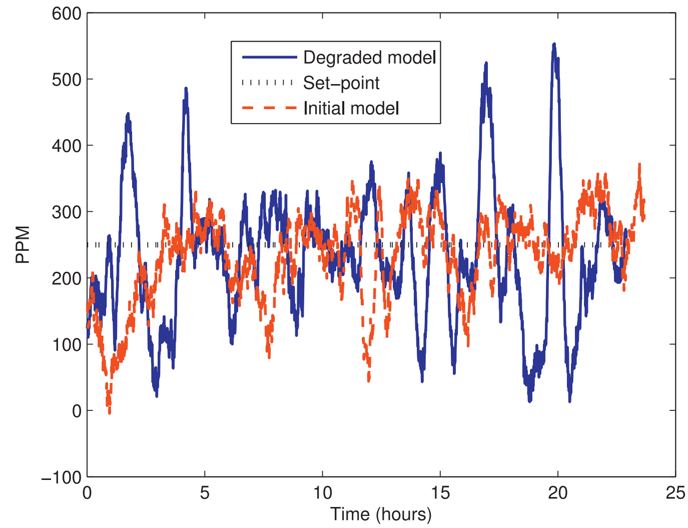

Fig. 9. C_4 content in the side draw at the optimal bandwidth of 0.001 rad/s.

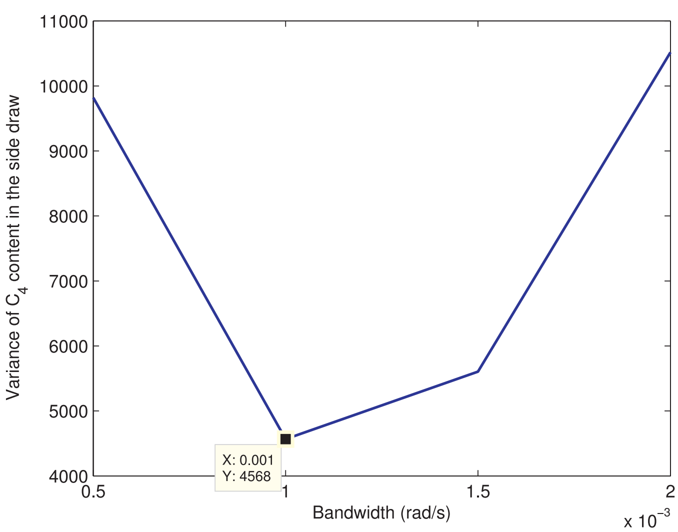

Fig. 8. Relation between variance of C_4 content and closed-loop bandwidth.

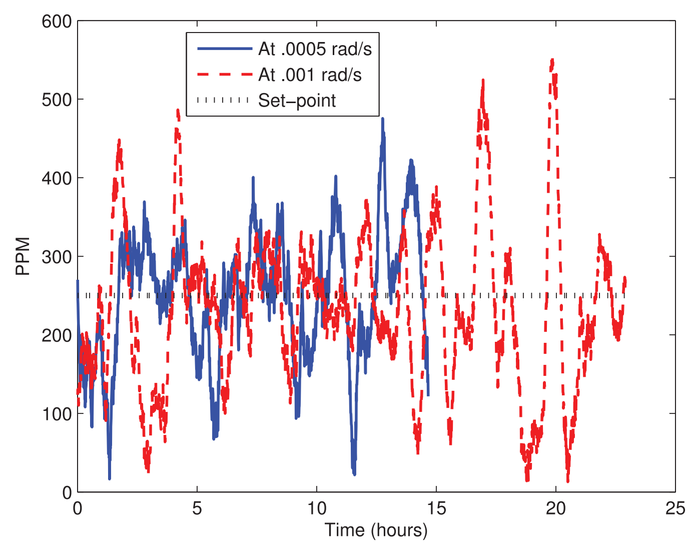

Fig. 10. C_4 content in the side draw at 0.0005 rad/s and 0.001 rad/s with model-plant mismatch.

0.001 rad/s and 0.0015 rad/s is presented in Fig. 7. The bandwidth and the closed performance relation is given in Fig. 8. It can be seen that 0.001 rad/s was the optimal closed-loop bandwidth which gave the minimum variance. Furthermore, the performance of the tuning approach was demonstrated when an model-plant mismatch was introduced by changing the model used in the MPC algorithm. In this scenario, the variance of the C_4 content in the side draw changes from 4568 to 1.007e4. The C_4 content in the two cases is presented in Fig. 9. It is obvious that the performance with the degraded model is worse than the commissioned performance.

As a result, the closed-loop bandwidth was reduced to de-tune the controller to obtain better performance. Fig. 10 shows that the that the closed-loop performance was improved when the bandwidth was reduced to 0.0005 rad/s.

With these tests, we had the chance to demonstrate some parts of the Autoprofit philosophy on a physical system. A more in depth description of tuning/retuning of this column can be found in Tran (2015). In addition, the dual control approach MPC-x was tested on the column and the results of these experiments can be found in Larsson et al. (2015).

6. Conclusions & future perspectives

In this paper, we present the technical developments and outcomes of the *Autoprofit* project that ran between 2010–2014. Within this limited time, we developed novel techniques in performance monitoring and diagnosis, least costly experiment design and MPC tuning which are fundamental for a high level automated model based support strategy. Some of these technical developments were implemented on one of the industrial case studies (Guidi et al., 2014).

At the end of the project, a workshop was also held and representatives of Advanced Process Control providers and users of this technology from process industries were invited. The discussions during this workshop highlighted the following points for future technical developments required to address the needs of the users.

- **Autopilot operation such as in the aircraft or automobile industry should be the objective**

  Autopilot type of operation requires control systems with fault tolerance such that the safety is ensured and an acceptable performance is attained within a limited time frame. In this project, we made a conscious decision of considering only changes in the input-output dynamic behaviour or disturbance characteristics as the sole reason for the performance degradations. Faults such as sensor/actuator malfunctioning were not considered. A first task in this direction would be including the fault detection/isolation technology combined by a fault tolerant control technology. This kind of technology is very important for safety of critical systems such as aircrafts, spacecrafts, power plants and chemical process industry (Goupil et al., 2014; Sojoudi, Lavaei, & Murray, 2011; Zhang & Jiang, 2008).

- **Reducing the engineering effort, especially in modelling, is essential**
  At the heart of this model based technology is the model of the process. Whether it is a black box model, hybrid model or mechanistic model, APC project costs and computational power are important factors to take into account. Project costs can be too high if modelling becomes too complex. The *Autoprofit* project has taken the steps in the right direction with the least costly experiment design and considering an economic criterion in decision making.
  The work presented here however, considers linear MPC and the technology can only handle model changes that occur gradually in time. The developed technologies are limited when it comes to processes operating over a nonlinear range continuously. Hence, a logical research direction would be extension of these technological developments so that the model based support system can cope with nonlinearities. In fact, within *Autoprofit*, first steps in this direction were already taken by investigating identification of linear parameter varying (LPV) models and experiment design (Bachnas, Toth, Mesbah, & Ludlage, 2014; Kauven, Piga, & Toth, 2013). In view of complementary extension to the nonlinear dynamic domain, the use of mechanistic models in model-based operation support technology is attractive. Despite their extensive offline use, the process knowledge within these mechanistic models is still not utilised extensively in the production/operation environment for real time online model based applications. On one hand, the development of such models is a costly activity for process operations alone and on the other hand these models are computationally very expensive and frequently not efficient enough for real time applications. Bringing physical structure in our data driven models, could be a good direction.
- **The controller has to communicate in English with operators**
  Human-automation interaction has been investigated quite extensively also in the case of MPC (Guerlain, Jamieson, Bullemer, & Blair, 2002). This is a challenging task due to the complexity of controller algorithms and complex interactions that exist in the underlying processes and the high amount of information needed to make a decision (Forbes, Patwardhan, Hamadah, & Gopalun, 2015). In normal operation, a common observation is that operators tend to stay away from the constraints which is in contradiction with the MPC philosophy that is to make processes to operate as close as possible to the constraints which is the most profitable condition (Backx & Van den Hof, 2013). Hence, the financial losses made during normal operation may be higher than the losses during the experiments. Operators prefer to go through modelling tests as quickly as possible. These are just signs of lack of confidence and trust in MPC technology by the operators. Therefore, it is utmost important that operators are informed what and why the controller and maintenance tools are doing whether that is tuning or closed loop identification, and the outcome of the possible decisions in terms of economical gains. In summary, automation action needs to be explained to the operator.
- **Plantwide-distributed maintenance**
  The Autoprofit model-based operation support philosophy has been developed for one unit operation in mind. The interaction of several (distributed) model based controllers or, even further, interaction between plantwide scheduling/optimization and MPC level have not been considered in the project. Plantwide and hierarchical control considerations have attracted considerable attention for quite sometime (Chu & You, 2015). The plantwide optimization may drive the process to new operating conditions that require adjustment of the control strategy of specific units or require the degraded accuracy of the model to be improved again. Hence, extension of the Autoprofit philosophy to enterprise level is a relevant and challenging step in the right direction.

Acknowledgements

The research leading to these results received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 257059, The *Autoprofit* project (www.fp7-autoprofit.eu).

We thank Ali Mesbah, Assistant Professor, UC-Berkeley; Quang Tran, APC Engineer, Smurfit Kappa; Christian Larsson, Development Engineer, Scania; Max Potters, Capital Management & Policies, Aegon, Mariette Annergren, KTH, and David Kauven, RWTH for their scientific contributions in making this project a success.

References

Åkesson, J. (2006). *MPCtools 1.0 - reference manual*. Lund, Sweden: Department of Automatic Control, Lund Institute of Technology.

Agarwal, N., Huang, B., & Tamayo, E. C. (2007a). Assessing model predictive control (mpc) performance 1 probabilistic approach for constraint analysis. *Industrial & Engineering Chemistry Research, 46*, 8101–8111.

Agarwal, N., Huang, B., & Tamayo, E. C. (2007b). Assessing model predictive control (mpc) performance 2 bayesian approach for constraint tuning. *Industrial & Engineering Chemistry Research, 46*, 8112–8119.

Anderson, B. D. O., & Moore, J. B. (1971). *Linear Optimal Control*. Prentice Hall.

Bachnas, R., Toth, R., Mesbah, A., & Ludlage, J. H. A. (2014). A review data-driven linear parameter varying modeling approaches: a high purity distillation column case study. *J. Process Control, 24*(4), 272–285.

Backx, A. C. P. M., & Van den Hof, P. M. J (2013). Towards advanced autonomous model-based operation of industrial process systems - where do we take it from here? *Technical report*. Delivrable 7.3 Autoprofit Project.

Barenthin, M., Bombois, X., Hjalmarsson, H., & Scorletti, G. (2008). Identification for control of multivariable systems: controller validation and experiment design via LMIs. *Automatica, 44*(12), 3070–3078.

Basseville, M. (1997). Information criteria for residual generation and fault detection and isolation. *Automatica, 33*(5), 783–803.

Basseville, M. (1998). On-board component fault detection and isolation using the statistical local approach. *Automatica, 34*, 1391–1415.

Bauer, M., & Craig, I. K. (2008). Economic assessment of advanced process control - a survey and framework. *Journal of Process Control, 18*(1), 2–18.

Bombois, X., Potters, M. G., & Mesbah, A. (2014). Closed-loop performance diagnosis for model predictive control systems. In *Proceedings of European control conference, Strasbourg*.

Bombois, X., Scorletti, G., Gevers, M., Van den Hof, P. M. J., & Hildebrand, R. (2006). Least costly identification experiment for control. *Automatica, 42*, 1651–1662.

Chan, K. H., Dozal-Mejorada, E. J., Cheng, X., Kephart, R., & Ydstie, B. E. (2014). Predictive control with adaptive model maintenance: Application to power plants. *Computers and Chemical Engineering, 70*(91–103).

Chmielewski, D. J., & Manthanwar, A. M. (2004). On the tuning of predictive controllers: Inverse optimality and the minimum variance covariance constrained control problem. *Industrial & Engineering Chemistry Research, 43*, 7807–7814.

Chu, Y., & You, F. (2015). Model based integration of control and operations: Overview, challenges, advances and opportunties. *Computers and Chemical Engineering, 83*, 2–20.

Clarke, D. W., Mohtadi, C., & Tuffs, P. S. (1987a). Generalized predictive control - part i. The basic algorithm. *Automatica, 23*(2), 137–148.

Clarke, D. W., Mohtadi, C., & Tuffs, P. S. (1987b). Generalized predictive control - part ii. Extensions and interpretations. *Automatica, 23*(2), 149–160.

Darby, M., & Nikalaou, M. (2014). Identification test signal design for multivariable model-based control: An industrial perspective. *Control Engineering Practice, 22*, 165–180.

Darby, M. L., & Nikolaou, M. (2012). MPC: Current practice and challenges. *Control Engineering Practice, 20*, 328–342.
Di Cairano, S., & Bemporad, A. (2010). Model predictive control tuning by controller matching. *IEEE Transactions on Automatic Control, 55*(1), 185–190.
Ebadat, A., Annergren, M., Larsson, C. A., Rojas, C. R., Wahlberg, B., Hjalmarsson, H., … Sjöberg, J. (2014). Application set approximation in optimal input design for model predictive control. *In proceedings of European control conference*.
Forbes, M. G., Patwardhan, R. S., Hamadah, H., & Gopalun, R. B. (2015). Model predictive control in industry: Challenges and opportunities. In *IFAC 9th international symposium on advanced control of chemical processes* (pp. 532–537).
Garriga, J. L., & Soroush, M. (2008). Model predictive controller tuning via eigenvalue placement. In *2008 American control conference* (pp. 429–434).
Garriga, J. L., & Soroush, M. (2010). Model predictive control tuning methods: A review. *Industrial & Engineering Chemistry Research, 49*(8), 3505–3515.
Goupil, P., Boada-Bauxell, J., Marcos, A., Cortet, E., Kerr, M., & Costa, H. (2014). Airbus efforts towards advanced real-time fault diagnosis and fault tolerant control. In *Proceedings of 19th IFAC world congress* (pp. 3471–3476).
Guerlain, S., Jamieson, G. A., Bullemer, P., & Blair, R. (2002). The mpc elucidator: A case study in the design for human-automation interaction. *IEEE Transactions on Sysrems, Man, and Cybernetics- Part A: Systems and Humans, 32*(1), 25–40.
Guidi, H., Larsson, C. A., Tran, Q. N., Özkan, L., & Backx, A. C. P. M. (2014). Autonomous maintenance of advanced process control: Application to an industrial depropanizer. *AICHE spring meeting*.
Gustafsson, F., & Graebe, S. F. (1998). Closed-loop performance monitoring in the presence of system changes and disturbances. *Automatica, 34*, 1311–1326.
Harris, T. J. (1989). Assesment of control loop performance. *Canadian Journal of Chemical Engineering, 67*(5), 856–861.
Harris, T. J., Seppala, C. T., & Desborough, L. D. (1999). Review of performance monitoring and assessment techniques for univariate and multivariate contrsystems. *Journal of Process Control, 9*(1), 1–17.
Hartley, E. N., & Maciejowski, J. M. (2011). Designing MPC controllers by reverse-engineering existing lti controllers. *Technical report*. Cambridge University Engineering Department, Cambridge, UK.
Hartley, E. N., & Maciejowski, J. M. (2013). Designing MPC controllers by reverse-engineering existing LTI controllers. *IEEE Transactions on Automatic Control*, (99), 2934–2939.
Heirung, T. A. N., Ydstie, B. E., & Foss, B. (2013). An adaptive model predicitive dual controller. *Adaptation and learning in control and signal processing, 11*, 62–67.
Hjalmarsson, H. (2009). System identification of complex and structured systems. In *Proceedings of the European control conference 2009, Budapest, Hungary* (pp. 3424–3452).
Huang, B., & Shah, S. (1999). *Performance assessment of control loops. Theory and applications*. Springer.
Huang, B., & Shah, S. L. (1998). Practical issues in multivariable feedback control performance assessment. *Journal of Process Control, 8*(5–6), 421–430.
Jansson, H., & Hjalmarsson, H. (2005). Input design via LMIs admitting frequency–wise model specifications in confidence regions. *IEEE Transactions on Automatic Control, 50*(10), 1534–1549.
Ji, G., Chen, L., Luo, Y., Lin, J., Hu, X., Zhang, X., … Zhu, Y. (2013). Design and implementation of an MPC system with model error detection. *Nirma university international conference on engineering (NUiCONE)*.
Jimoh, M. T. (2013). *A vision for MPC maintenance*. University of Glasgow Phd thesis.
Johansson, K. H. (1997). *Relay feedback and multivariable control*. Lund Institute of Technology, Sweden Phd thesis.
Kalafatis, K., Patel, K., Harmse, M., Zheng, Q., & Craig, M. (2006). Multivariable step testing for mpc projects reduces crude unit testing time. In *Hydrocarbon processing* (pp. 93–99).
Kalman, R. E. (1964). When is a linear control system optimal? *Journal of Basic Engineering, 86*(1), 51–60.
Kauven, D., Piga, D., & Toth, R. (2013). Autoprofit deliverable: Report on model refinement. *Technical report*.
Kay, S. M. (1998). *Fundamentals of statistical signal processing: Detection theory*. Prentice Hall.
Killingsworth, N. J., & Krstić, M. (2006). Pid tuning using extremum seeking: Online, model-free performance optimization. *IEEE Control Systems Magazine, 26*(1), 70–79.
Kreindler, E., & Jameson, A. (1972). Optimality of linear control systems. *IEEE Transactions on Automatic Control, 17*(3), 349–351.
Larsson, C. A., Ebadat, A., Rojas, C. R., Bombois, X., & Hjalmarsson, H. (2016). An application-oriented approach to dual control with excitation for closed-loop identification. *European Journal of Control, 29*, 1–16.
Larsson, C. A., Hägg, P., & Hjalmarsson, H. (2016). Generation of signals with specified seond order properties for constrained systems. *International Journal of Adaptive Control and Signal Processing, 30*, 456–472.
Larsson, C. A., Rojas, C. R., Bombois, X., & Hjalmarsson, H. (2015). Experimental evaluation of model predictive control with excitation (mpc-x) on an industrial depropanizer. *Journal of Process Control, 31*, 1–16.
Lee, J. H., & Yu, Z. H. (1994). Tuning of model predictive controllers for robust performance. *Computers & Chemical Engineering, 18*(1), 15–37.
Levya, R., Artillan, P., Cabal, C., Estibals, B., & Alonso, C. (2011). Dynamic performance of maximum power point tracking circuits using sinusoidal extremum seeking control for photovoltaic generation. *International Journal of Electronics, 98*(4), 529–542.
Ljung, L. (1999). *Systems identification: Theory for the user*. Prentice Hall.
Maciejowski, J. M. (2002). *Predictive control with constraints*. Pearson, Prentice Hall.
Manchester, I. (2010). *Input design for system identification via convex relaxation*. ArXiv e-prints
Marafioti, G. (2010). *Enhanced model predictive control: Dual control approach and state estimation issues*. Norwegian University of Science and Technology Phd thesis.
Maurath, P. R., Seborg, D. E., & Mellichamp, D. A. (1988). Predictive controller design by principal components analysis. *Industrial & Engineering Chemistry Research, 27*(7), 1204–1212.
McNaab, C. A., & Qin, S. J. (2003). Projection based mimo control performance monitoring i: Covariance monitoring in state space. *Journal of Process Control, 13*, 739–757.
McNaab, C. A., & Qin, S. J. (2005). Projection based mimo control performance monitoring ii: Measured disturbances and set-point changes. *Journal of Process Control, 15*, 89–102.
Mesbah, A., & Bombois, X. (2012). *Closed-loop performance diagnosis: Least costly fault detection and system re-identification*. Deliverable 4.2 (part 2) of the AUTOPROFIT Project
Mesbah, A., Bombois, X., Forgione, M., Hjalmarsson, H., & Van den Hof, P. M. J. (2015). Least costly closed-loop performance diagnosis and plant re-identification. *International Journal of Control, 88*(11), 2264–2276.
Modén, P. E. (2012). *Performance monitoring for MPC maintenance*. Deliverable 4.2 (part 1) of the AUTOPROFIT Project
Modén, P. E., & Lundh, M. (2013). Performance monitoring for mpc maintenance. In *European control conference* (pp. 3770–3775).
Olaru, S., De Doná, J. A., & Seron, M. M. (2008). Positive invariant sets for fault tolerant multisensor control schemes. In *Proceedings of the 17th IFAC world congress, Seoul, South Korea* (pp. 1224–1229).
Patwardhan, R. S., & Shah, S. L. (2002). Issues in performance diagnostics of model-based controllers. *Journal of Process Control, 12*, 413–427.
Potters, M. G., Bombois, X., Forgione, M., Modén, P., Lundh, M., Hjalmarsson, H., et al. (2014). Optimal experiment design in closed loop with unknown, nonlinear or implicit controllers using stealth identification. In *Proceedings of European control conference, Strasbourg*.
Potters, M. G., Lundh, M., Modén, P. E., & Bombois, X. (2012). *Software for performance monitoring*. Deliverable 4.3 of the AUTOPROFIT Project
Pronzato, L. (2008). Optimal experimental design and some related control problems. *Automatica, 44*, 303–325.
Qin, S. J., & Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice, 11*, 733–764.
Rossiter, J. A. (2005). Model-based predictive control: A practical approach. *Control series*. CRC Press.
Rowe, C., & Maciejowski, J. M. (2000a). Robust finite horizon mpc without terminal constraints. In *Proceedings of the 39st IEEE conference on decision and control, Sydney, Australia* (pp. 166–171).
Rowe, C., & Maciejowski, J. M. (2000b). Tuning MPC using H_∞ loop shaping. In *Proceedings of the American control conference, Chicago, IL*.
Schafer, J., & Cinar, A. (2004). Multivariable MPC system performance assessment, monitoring and diagnosis. *Journal of Process Control, 14*, 113–129.
Seron, M. M., & De Doná, J. A. (2010). Actuator fault tolerant multi-controller scheme using set separation based diagnosis. *International Journal of Control, 83*, 2328–2339.
Shah, G., & Engell, S. (2011). Tuning MPC for desired closed-loop performance for mimo systems. In *Proceedings of the 2011 American Control Conference, San Francisco, CA, USA*.
Shridhar, R., & Cooper, D. J. (1998). A tuning strategy for unconstrained multivariable model predictive control. *Industrial & Engineering Chemistry Research*, (37), 4003–4016.
Skogestad, S., & Postlethwaite, I. (2005). *Multivariable feedback control* (2nd ed.). Wiley.
Sojoudi, S., Lavaei, J., & Murray, R. M. (2011). Fault-tolerant controller design with applications in power systems and synthetic biology. In *Proceedings of 2011 American control conference* (pp. 4135–4142).
Stanković, M. S., & Stipanović, D. M. (2010). Extremum seeking under stochastic noise and applications to mobile sensors. *Automatica, 46*(8), 1242–1251.
Tan, Y., Moase, W. H., Manzie, C., Nešić, D., & Mareels, I. M. Y. (2010). Extremum seeking from 1922 to 2010. In *Proceedings of the 29th Chinese control conference, Beijing, China*.
Tran, Q. N. (2015). *Tuning model based controllers for autonomous maintenance*. Eindhoven University of Technology Phd thesis.
Tran, Q. N., Octaviano, R., Özkan, L., & Backx, A. C. P. M. (2014). Generalized predictive control tuning by controller matching. In *American control conference, Portland, OR, USA* (pp. 4889–4894).
Tran, Q. N., Özkan, L., & Backx, A. C. P. M. (2015). Generalized predictive control tuning by controller matching. *Journal of Process Control, 25*, 1–18.
Trierweiler, J. O., & Farina, L. A. (2003). Rpn tuning strategy for model predictive control. *Journal of Process Control*, (13), 591–598.
Tse, E., & Bar-Shalom, Y. (1973). An actively adaptive control for linear systems with random parameters via the dual control approach. *IEEE Transactions on Automatic Control, 18*, 109–117.
Tyler, M. L., & Morari, M. (1996). Performance monitoring of control systems using likelihood methods. *Automatica, 32*, 1145–1162.
Xu, F., Huang, B., & Akande, S. (2007). Performance assessment of model predictive control for variability and constraint tuning. *Industrial & Engineering Chemistry Research, 46*, 1208–1219.

Zhang, Y., & Jiang, J. (2008). Bibliographical review on reconfigurable fault tolerant control systems. *Annual Reviews in Control, 32*, 229–252.
Zhu, Y., & Padwarthan, R. (2012). Towards a low cost and high performance MPC: The role of identification. *FOCAPO*.
Zhu, Y., Patwardhan, R., Wagner, S. B., & Zhao, J. (2013). Towards a low cost and high performance MPC: The role of identification. *Computers & Chemical Engineering, 51*, 124–135.