Faculty of Engineering and Information Sciences - Papers: Part A

Faculty of Engineering and Information Sciences

2001

Wiener model identification and predictive control for dual composition control of a distillation column

H H J Bloemen
Delft University of Technology

C T. Chou
University of Wollongong, ctchou@uow.edu.au

T J J van den Boom
Delft University of Technology

V Verdult
University of Twente

M Verhaegen
University of Twente

See next page for additional authors

Wiener model identification and predictive control for dual composition control of a distillation column

Abstract

The benefits of using the Wiener model based identification and control methodology presented in this paper, compared to linear techniques, are demonstrated for dual composition control of a moderate-high purity distillation column simulation model. An identification experiment design is presented which enables one to identify both the low and high gain directions of the distillation column, properties which are important for control and hard to identify in a conventional identification experiment setup as is demonstrated in the paper. Data from the proposed experiment design is used for indirect closed-loop identification of both a linear and a Wiener model, which shows the ability of the Wiener model to approximate the nonlinearity of the distillation column much closer than the linear model can. The identified Wiener model is used in a MPC algorithm in which the nonlinearity of the Wiener model is transformed into a polytopic description. In this way a convex optimisation problem is retained while the effect of the nonlinearity on the input-output behaviour of the plant is still taken into account. The performance of the proposed Wiener MPC is compared with linear MPC based on the identified linear models, and with a Wiener MPC in which the nonlinearity of the Wiener model is removed from the control problem via an inversion, a popular way to handle Wiener models in a MPC framework. The simulations demonstrate that the proposed Wiener MPC outperforms the other MPC algorithms.

Keywords

distillation, composition, dual, column, control, wiener, predictive, identification, model

Disciplines

Engineering | Science and Technology Studies

Authors

H H J Bloemen, C T. Chou, T J J van den Boom, V Verdult, M Verhaegen, and T C. Backx

Wiener model identification and predictive control for dual composition control of a distillation column

H.H.J. Bloemen [a],*, C.T. Chou [b], T.J.J. van den Boom [a], V. Verdult [c], M. Verhaegen [c], T.C. Backx [d]

[a]*Delft University of Technology, Department of ITS, PO Box 5031, 2600 GA Delft, The Netherlands*
[b]*University of Wollongong, TITR Institute, Wollongong NSW 2522, Australia*
[c]*University of Twente, Department of Applied Physics, PO Box 217, 7500 AE Enschede, The Netherlands*
[d]*IPCOS Technology, Bosscheweg 145a, 5282 WV Boxtel, The Netherlands*

Abstract

The benefits of using the Wiener model based identification and control methodology presented in this paper, compared to linear techniques, are demonstrated for dual composition control of a moderate–high purity distillation column simulation model. An identification experiment design is presented which enables one to identify both the low and high gain directions of the distillation column, properties which are important for control and hard to identify in a conventional identification experiment setup as is demonstrated in the paper. Data from the proposed experiment design is used for indirect closed-loop identification of both a linear and a Wiener model, which shows the ability of the Wiener model to approximate the nonlinearity of the distillation column much closer than the linear model can. The identified Wiener model is used in a MPC algorithm in which the nonlinearity of the Wiener model is transformed into a polytopic description. In this way a convex optimisation problem is retained while the effect of the nonlinearity on the input–output behaviour of the plant is still taken into account. The performance of the proposed Wiener MPC is compared with linear MPC based on the identified linear models, and with a Wiener MPC in which the nonlinearity of the Wiener model is removed from the control problem via an inversion, a popular way to handle Wiener models in a MPC framework. The simulations demonstrate that the proposed Wiener MPC outperforms the other MPC algorithms. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Wiener model; Closed-loop identification; Predictive control; Distillation column

1. Introduction

Linear model based control techniques have shown to be useful in controlling chemical processes in a limited operating region, because then the behaviour of the process can be approximated by a linear model. However, when the operating region is extended, the nonlinearity of the process can not be ignored and the performance of linear control techniques degrades because of the inability of the linear model to accurately approximate the real process. In order to improve the performance, nonlinear models are needed that can cope with the observed nonlinearity. For modelling of the nonlinear process, different model structures can be used, for example white box models, or black box models like neural networks, fuzzy models and block oriented models. The choice for a model structure is partly motivated by the purpose of the model. In this paper the model is used within a model-based predictive control (MPC) framework. White box models are usually of a large dimension and, therefore, are not attractive for use in MPC. Among the nonlinear black box models the block oriented models are related very closely to linear models, because these models consist of a series connection of a linear dynamic element and a static nonlinear element. The linear dynamic block can either be preceded by a static input nonlinearity or followed by a static output nonlinearity in which case these models are referred to as Hammerstein and Wiener models respectively. A schematic representation of a

---

* Corresponding author. Tel.: +31-15-2782087; fax: +31-15-2786679.

E-mail address: h.bloemen@its.tudelft.nl (H.H.J. Bloemen).

Nomenclature

a_i	coefficient
A	state space matrix
b_i	coefficient
B	state space matrix
C	state space matrix
C_{lin}	state space matrix of the linearised output equation
d	measurable disturbance
F	state feedback gain
Fc	feed composition
Fl	feed rate
G	the plant
h	static (output) nonlinearity
$H(v)$	nonlinear matrix in the output nonlinearity
$\mathcal{H}$	vertex of Ω
H_s	switching horizon
I	identity matrix
J	performance index
K	the controller
L	reflux rate
L^{obs}	(extended) Kalman filter gain
m	input dimension
M	state space matrix
n	state dimension
N	number of data points
N_{FIR}	number of FIR coefficients
p	output dimension
P	end-point state weighting matrix
Q_v	weighting matrix for v
Q_y	weighting matrix for y
r	set-point
$R(\omega)$	Fourier transform of r
R_u	weighting matrix for Δu
u	input
$U(\omega)$	Fourier transform of u
v	output of the LTI block of the Wiener model
V	boil-up rate
x	state
x_e	extended state
y	output
y_m	model-output
$\hat{y}_m$	predicted output in FIRMPC
$y_{m,\mathrm{lin}}$	linearised model-output
α_i	principal input direction
β_i	principal output direction
Δu	increment of the input
ϵ	output error
θ	FIR parameter
σ_i	principal gain
Φ_u	spectral density of u
Φ_y	spectral density of y
Ω	polytope
$\underline{}$	used to express deviation variables
ref	used to express reference value

Wiener model is given in Fig. 1. Although these block oriented models represent only a small subclass of all nonlinear models it has been proven in [8] that any time invariant system with fading memory can be approximated to arbitrary accuracy by a finite dimensional Wiener system. An identification algorithm for Wiener models can be found in [31].

In this paper a distillation column simulation model is used as a benchmark to demonstrate the benefits of a Wiener model based identification and control methodology compared to a methodology based on linear models. The objective is to perform dual composition control of a distillation column operated in the *LV*-configuration, which means the manipulated variables or inputs (u) are the reflux (L) and boil-up (V) rate and the controlled variables or outputs (y) are the top and bottom compositions. Besides the two manipulated variables also two disturbances enter the column, the feed rate (Fl) and the feed composition (Fc). See Fig. 2 for a schematic representation and Section 2 for a more detailed description of the distillation column. When the column is operated over a relatively wide operating region instead of at a single setpoint, it reveals a significant nonlinear behaviour, see for example [15,28]. Part of this nonlinearity originates from the fact that the

products can not exceed 100% purity, which acts as a kind of saturation. Therefore this nonlinearity will be more profound for operation at high purities. A common approach to make the plant less nonlinear is the use of logarithmic compositions, as proposed by [23]. A theoretical justification for this is given in [27]. This motivates the choice for a Wiener model to approximate the behaviour of the distillation column.

Apart from the nonlinearity the distillation column is also characterised by an ill-conditioned nature [24,25], which means that the gain is highly depending on the direction of the input vector. Roughly speaking, the high gain direction corresponds to the direction which causes one product to become purer and the other impurer simultaneously. This corresponds to a change in the external flows, i.e. an increase (or decrease) in L with constant V or vice versa. The low gain direction corresponds to the direction in which both products become either purer or impurer. This corresponds to a change in the internal flows, i.e. a simultaneous increase (or decrease) in L and V. Because the objective is to perform dual composition control it is clear that the low gain direction is important for control. However, obtaining accurate information about this direction is difficult as it is hard to change the internal flows without changing the external flows [27].
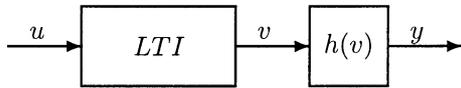
Fig. 1. Schematic representation of a Wiener model, u is the input to the linear time-invariant dynamic block (LTI) with output v, which is mapped to y by the static output nonlinearity $h(v)$.
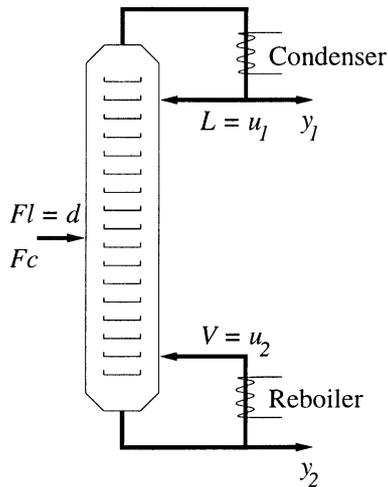

Fig. 2. A distillation column in LV-configuration.

In Section 3 a motivation is given why the sensitivity to disturbances, the ill-conditioned nature and the multiple input multiple output (MIMO) nature of the distillation column requires MIMO closed-loop experiments for identification of a model which is suitable for control purposes. A system identification algorithm for Wiener models is presented that is able to handle these closed-loop data. This is a tailor-made approach for closed-loop identification of Wiener models presented in [11]. The benefits of doing a closed-loop MIMO identification experiment setup are shown in Section 4 by means of a comparison with a conventional setup for identification of a finite impulse response (FIR) model. Also a comparison between indirect closed-loop identification for linear and Wiener models is presented.

The purpose of the identified models is to use them within a MPC framework. In MPC the input is calculated by on-line optimisation of a performance index based on predictions by the model, subject to possible constraints. In case of a linear model, a quadratic performance index and linear constraints, the optimisation problem is a quadratic program which can be solved on-line for many applications. For more information concerning linear MPC, the reader is referred to [12,13,16]. When the model is non-linear, the optimisation problem becomes non-convex in general because the model acts as an equality constraint. Solving a non-convex optimisation problem on-line is not attractive because it is computationally expensive and may suffer from the existence of several local minima

[18]. See [1,19,21] for a review on nonlinear MPC and possibilities to handle the non-convex optimisation problem. The advantage of using a Wiener model, instead of other nonlinear black box models, in a MPC scheme is the possibility to exploit the special structure of the Wiener model such that a convex optimisation problem is retained in the MPC algorithm, see [6,7,22] and Section 5 for an elaborate explanation.

In Section 6 it is demonstrated how the quality of the different identified models affects the performance of the MPC. Moreover the difference in performance of the presented Wiener MPC algorithm and the algorithm of [22] is shown.

Finally the paper is concluded with a discussion in Section 7. This paper combines and extends the work presented in [5,10]. Summarising, the main contribution of the paper is to demonstrate the benefits of the presented Wiener model closed-loop identification and predictive control methodology for nonlinear ill-conditioned processes, compared to existing (linear) techniques, using a distillation column simulation model as a benchmark process. For the identification part this involves a comparison between direct and indirect closed-loop identification of a linear model and indirect closed loop identification of a Wiener model. For the control part this involves a comparison of the performance of the MPC algorithms based on the identified linear models, and of two different approaches to handle the Wiener model within a predictive control framework.

2. Distillation column benchmark

The benchmark problem of this paper is a white box simulation model of a distillation column, presented as Column A in [25], operated in LV-configuration, see Fig. 2. The choice for this configuration is motivated by the fact that the LV-configuration is the one most commonly used in industrial practice [26] and because it is not very sensitive to the level control tuning. The LV-configuration is very sensitive to the feed rate disturbance (Fl), but the effect of the feed composition disturbance (Fc) is reduced in this configuration [25]. The feed rate disturbance is assumed to be measurable, therefore, its influence can be incorporated in the Wiener model as a non-manipulable input (d), which enables to do feed-forward control in the MPC to compensate for disturbances in the feed flow rate. The feed composition is generally not measurable on-line. Therefore, its influence is not incorporated in the Wiener model and it thus acts as an unknown process disturbance. In all simulations in this paper (identification experiments and control experiments) both disturbances are generated by low-pass filtering of white noise signals, superimposed on their nominal values. The maximum amplitudes of these two disturbances are adjusted to 10% of the nominal feed

flow rate and feed composition respectively. Such disturbance levels are realistic in an industrial environment.

The white box simulation model consists of 39 theoretical stages plus a re-boiler and condenser. In the development of this simulation model, the following assumptions are made: binary component, constant pressure, negligible vapour holdup, total condenser, equimolar flow, vapour–liquid equilibrium on all stages with constant relative volatility.

In this paper it is assumed that the measurements of the top and bottom compositions are directly available for feedback in the MPC. In practice this will imply that these measurements are provided by a software sensor, for example on the basis of temperature measurements [29] which are directly available. The reason for this is that the direct measurement of the product compositions introduces very large dead times because these measurements are carried out by gas chromatographs which are characterised by dead times of about 30 min [14]. Incorporation of such a software sensor into the model based predictive controller is beyond the scope of this paper.

Throughout the paper the outputs are expressed in impurities, where the impurity corresponds to the mole fraction of the main component in the product stream, i.e. the light component in the top product and the heavy component in the bottom product.

3. Model identification

At first it is argued in this section why the distillation column benchmark, and a nonlinear ill-conditioned plant in general, requires a MIMO closed-loop identification experiment in order to obtain relevant information concerning the plant dynamics. Next the identification algorithms are presented for the proposed indirect closed-loop identification algorithm to identify a Wiener model, Section 3.3, and a linear model, Section 3.4, and for the conventional algorithm to identify a Finite Impulse Response (FIR) model, Section 3.5. In Table 1 an overview of the presented identification algorithms is given.

3.1. The need for multivariable experiments

For the identification of linear models, the principle of superposition applies and one can choose between performing either a number of single input multiple output (SIMO) experiments or a single MIMO experiment. In fact, the former choice is almost the standard practice in the process industry. In this subsection it is demonstrated that such a choice no longer exists for the identification of multivariable nonlinear models by means of a simple example. Consider the identification of the nonlinear static map h defined as

$$h : \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \mapsto \begin{bmatrix} a_1 u_1 + a_2 u_2 + a_{12} u_1 u_2 \\ b_1 u_1 + b_2 u_2 + b_{12} u_1 u_2 \end{bmatrix} \quad (1)$$

where a_i and b_i are unknown coefficients to be identified. Assume two SIMO experiments are performed similar to a linear identification experiment. In the first experiment u_1 is excited while u_2 is kept zero, in the second u_2 is excited while u_1 is kept zero. The data from the first experiment allow to estimate both a_1 and b_1 while the second experiment gives estimates of both a_2 and b_2. However, no information regarding the two nonlinear terms can be obtained. Generally speaking, SIMO experiments do not provide any information on the nonlinear cross terms of a multivariable nonlinear map, therefore identification of multivariable nonlinear systems almost always requires MIMO experiments.

Besides the above observation, which holds for the identification of nonlinear systems in general, the aim to identify a model for the purpose to perform dual composition control of a distillation column in LV-configuration, also calls for MIMO experiments. Doing SIMO experiments in this case results in either L or V being excited while the other remains constant. This corresponds to a change in the external flow, thus information regarding the high gain direction of the column is obtained. For dual composition control the low gain direction is very important, which means the internal flows should be excited, which requires a simultaneous increase (or decrease) in L and V, i.e. MIMO experiments.

3.2. The need for closed-loop experiments

Apart from the nonlinearity and directionality of the distillation column, which require MIMO identification experiments as motivated in the previous subsection, the column also suffers from a large difference in gain in both directions, i.e. the column is ill-conditioned. As a result, a MIMO identification experiment in open-loop provides data that almost only reflect information

Table 1
Overview of the different identification methods that are used in the comparison

Identification experiment	Model	Identification algorithm	Results
Closed-loop to excite low gain direction	Wiener model	Section 3.3	Section 4.1
	Linear state space model	Section 3.4	Section 4.1
Closed-loop to prevent drifting	FIR model	Section 3.5	Section 4.2

regarding the high gain direction if no extra action is undertaken. To demonstrate this, a simulation is performed in which the inputs u were provided by two mutually independent pseudo-random binary signals (PRBSs), which is common practice for MIMO identification experiments. No disturbances were used in this simulation. The results are plotted in Fig. 3. From this figure it is clear that almost only information concerning the high gain direction of the plant is generated, corresponding to one product becoming purer and the other impurer, which is reflected by a small band of data points in the output space.

To explain this and to argue for the need of closed-loop identification, an analysis is made for the linear case (for simplicity, the idea behind these arguments can be generalised easily to the case of Wiener models because of their closely related structure), without disturbances. Let G denote the discrete time multivariable transfer function from the inputs (the reflux and boil-up) to the outputs (the top and bottom compositions). The frequency response $G(e^{j\omega})$ has the following singular value decomposition (SVD) (see [2] for example):

$$G(e^{j\omega}) = \begin{bmatrix} \beta_1(e^{j\omega}) & \beta_2(e^{j\omega}) \end{bmatrix} \begin{bmatrix} \sigma_1(\omega) & 0 \\ 0 & \sigma_2(\omega) \end{bmatrix} \begin{bmatrix} \alpha_1^*(e^{j\omega}) \\ \alpha_2^*(e^{j\omega}) \end{bmatrix}$$

(2)

where * denotes complex conjugate transpose. The vectors α_i and β_i are the principal input and output directions respectively. The singular values σ_i are the principal gains of the plant. Due to the ill-conditioned nature $\sigma_1(\omega) >> \sigma_2(\omega)$ over a large frequency range. Let $\Phi_u(\omega)$ and $\Phi_y(\omega)$ denote the spectral densities of the input u and output y respectively, then:

$$\Phi_y(\omega) = G(e^{j\omega})\Phi_u(\omega)G^*(e^{j\omega})$$

(3)

If the inputs are two mutually independent PRBSs of unit variance, then $\Phi_u(\omega)$ is the identity matrix and $\Phi_y(\omega)$ can be written as:

$$\Phi_y(\omega) = \begin{bmatrix} \beta_1(e^{j\omega}) & \beta_2(e^{j\omega}) \end{bmatrix} \begin{bmatrix} \sigma_1(\omega) & 0 \\ 0 & \sigma_2(\omega) \end{bmatrix} \begin{bmatrix} \alpha_1^*(e^{j\omega}) \\ \alpha_2^*(e^{j\omega}) \end{bmatrix} \cdot I \cdot$$

$$\begin{bmatrix} \alpha_1(e^{j\omega}) & \alpha_2(e^{j\omega}) \end{bmatrix} \begin{bmatrix} \sigma_1(\omega) & 0 \\ 0 & \sigma_2(\omega) \end{bmatrix} \begin{bmatrix} \beta_1^*(e^{j\omega}) \\ \beta_2^*(e^{j\omega}) \end{bmatrix}$$

$$= \begin{bmatrix} \beta_1(e^{j\omega}) & \beta_2(e^{j\omega}) \end{bmatrix} \begin{bmatrix} \sigma_1^2(\omega) & 0 \\ 0 & \sigma_2^2(\omega) \end{bmatrix} \begin{bmatrix} \beta_1^*(e^{j\omega}) \\ \beta_2^*(e^{j\omega}) \end{bmatrix}$$

(4)

Because $\sigma_1 >> \sigma_2$, the output of the distillation column mainly consists of a component in the β_1 direction, which is the high gain direction observed in Fig. 3. Note that the points in the lower right plot of Fig. 3 do not appear as a straight line (β_1 is a vector), but as a curved line due to the nonlinearity of the distillation column. From an identification point of view, this means that such an identification experiment provides a model with a good estimate in the high gain direction but a poor estimate in the low gain direction. However, for dual composition control it is clear that a good estimate of the low gain direction is desired as well. This implies that an identification experiment should be designed in which this low gain direction is excited sufficiently. Above derivation indicates that a solution to this problem is to use an input signal with a strong component in the weak input gain direction α_2.

In order to improve excitation of the low gain direction, Koung and MacGregor [20] suggest using this a
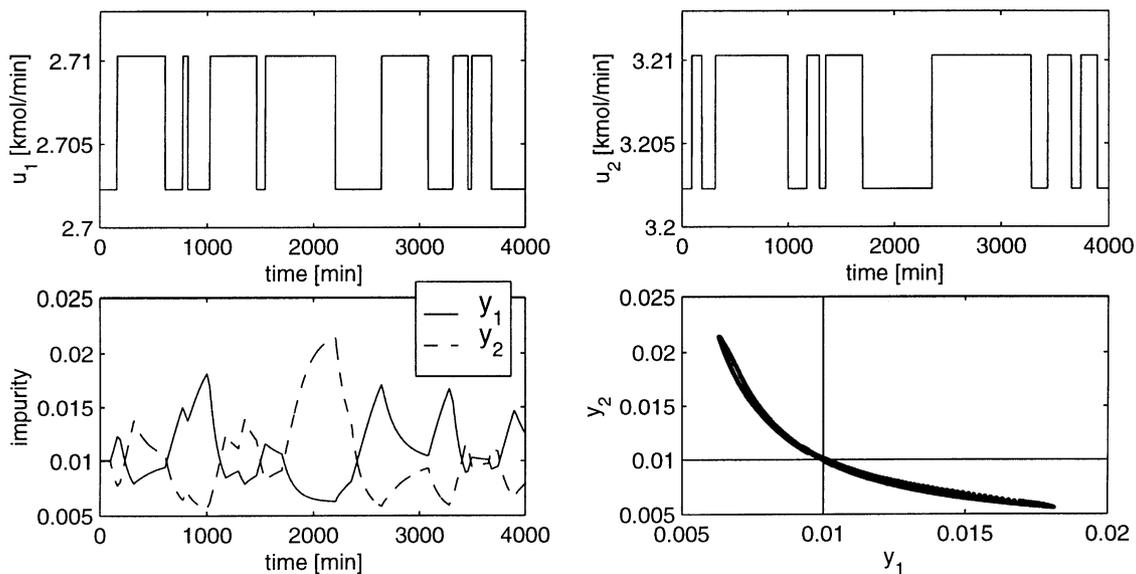

Fig. 3. Open-loop response of the distillation column (lower left plot) to manipulable inputs (upper two plots) that are provided by two mutually independent PRBSs. Lower right plot is the covered output space.

priori knowledge of the plant to design the input signal. Due to the ill-conditioned nature this requires large input amplitudes in the low gain direction in order to excite sufficiently the outputs corresponding to this direction. In [20] good results are obtained by designing open-loop identification experiments based on prior knowledge of the directionality. However, these simulations do not incorporate disturbances like feed rate and feed composition disturbances. Due to these disturbances it is almost impossible, in practice, to excite sufficiently the low gain direction, by means of changes in the internal flows, without changing the external flows [27]. A small disturbance causes a change in the external flows, which then dominates the response of the outputs due to the high gain in this direction. A solution for this problem, also indicated in [20], is to use closed-loop identification experiments since these enable to generate inputs with the correct correlations and relative magnitudes, without the need of accurate a priori knowledge concerning the directionality and ill-conditioned nature. A more detailed analysis and discussion concerning the benefits of closed-loop identification for ill-conditioned plants can be found in [2,3]. In [3] only one loop (bottom product to boil-up rate) of the distillation column is closed. This setup is not appropriate for the case study in this paper, since the uncontrolled output will drift too far from the desired operating region due to the disturbances acting on the column (not used in [3]). This is demonstrated in the next paragraph.

A second motivation for performing closed-loop identification experiments is the sensitivity to disturbances, especially to the feed rate disturbance in case of the *LV*-configuration [25]. To illustrate this, the distillation column is simulated using the same inputs as in Fig. 3, but now in the presence of both disturbances. The results are plotted in Fig. 4. Comparing Figs. 3 and 4 shows that the disturbances have a large impact on the response of the column, causing the outputs to drift far away from the desired operating region (a region around 0.01 impurity, see Fig. 3). Especially the trend of the feed rate disturbance is clearly recognisable in both output responses. The large influence of the disturbances on the output compared to the influence of both inputs, visualised in Figs. 3 and 4, is caused by a large difference in amplitude of the signals. The disturbances deviate up to 10% from their nominal values (1 kmol/min for the feed rate and 0.5 mol fraction of the light component for the feed composition) whereas the inputs change only up to 0.15%. However, since these disturbance magnitudes are common in practice, one will have to deal with them. Because the objective is to control the distillation column in an operating region up to 0.015 impurity, the drift caused by the disturbances should somehow be compensated for, as the data generated in the identification experiment should be within the desired operating region. Note again that the data in the output space in Fig. 4 mainly reflects information regarding the high gain direction of the distillation column. Due to the sensitivity to disturbances, a purely open-loop experiment is rare in practice, because the operator will interfere when necessary in order to keep the column operating around a certain set-point.

To analyse why a closed-loop identification experiment helps to generate signals that have a significant component in the low gain direction, let *r* denote the set-point reference for the top and bottom compositions and let *K* denote the feedback controller. The Fourier transform of the input *u* ($U(\omega)$) is given by:

$$U(\omega) = \left(I + K\left(e^{j\omega}\right)G\left(e^{j\omega}\right)\right)^{-1}K\left(e^{j\omega}\right)R(\omega). \qquad (5)$$

where $R(\omega)$ is the Fourier transform of *r*. At those frequencies where the gain of the controller *K* is sufficiently large, the above relation can be approximated by:

$$U(\omega) \approx G^{-1}\left(e^{j\omega}\right)R(\omega) \qquad (6)$$

$$= \left[\alpha_1\left(e^{j\omega}\right) \; \alpha_2\left(e^{j\omega}\right)\right]\begin{bmatrix} \frac{1}{\sigma_1(\omega)} & 0 \\ 0 & \frac{1}{\sigma_2(\omega)} \end{bmatrix}\begin{bmatrix} \beta_1^*\left(e^{j\omega}\right) \\ \beta_2^*\left(e^{j\omega}\right) \end{bmatrix}R(\omega) \qquad (7)$$

Thus, because $\sigma_1(\omega) >> \sigma_2(\omega)$, the input is likely to contain a large part of the low gain component α_2 at those frequencies where above relation holds. To meet the requirement on the controller *K* for above analysis two single-loop PI controllers can be used, because any PI controller has large gain at low frequencies. Furthermore, PI controllers are also useful in avoiding drifting of the outputs from the desired operating region, caused by the disturbances. The conditions, that the product compositions are not allowed to vary too much during the experiment and that larger variations on the reflux and boil-up rates are allowed, are ideal for performing closed-loop experiments.

In the above discussion, the importance of getting an accurate estimate of the low gain direction is emphasised. This does not imply that the high gain direction can be ignored altogether. As stated in [2] and demonstrated in [3] the design of high performance multivariable controllers requires a precise estimation of both the directions and the associated gains. In open-loop experiments the high gain direction overwhelms the response of the outputs resulting in hardly any information concerning the low gain direction. The purpose of doing closed-loop identification is to shift this balance such that both directions are sufficiently excited in order to enable the identification of both directions and associated gains. This is possible because by performing closed-loop identification the outputs can be driven in the directions determined by the reference rather than by the direction of the dominant plant gain.

3.3. The indirect closed-loop Wiener model identification algorithm

The structure of the Wiener model that is to be identified is:

$$\begin{cases} x(k+1) &= Ax(k) + Bu(k) + Md(k) \\ v(k) &= Cx(k) \\ y_m(k) &= h(v(k)) \end{cases} \quad (8)$$

where $u(k) \in \mathbb{R}^m$ represents the manipulable inputs L and V (i.e. $m = 2$), $d(k) \in \mathbb{R}$ represents the feed rate, $x(k) \in \mathbb{R}^n$ is the state of the linear dynamic block, $v(k) \in \mathbb{R}^p$ is the output of the linear dynamic block which is the input to the nonlinear static block, and $y_m \in \mathbb{R}^p$ represents the output of the model (i.e. the top and bottom compositions, so $p = 2$). (A,B,C,M) are state-space matrices of conformal dimensions, h is a static nonlinear mapping.

Because the identification experiment is carried out in closed-loop, an identification algorithm should be used that can cope with closed-loop data. The basic steps of the indirect closed-loop identification algorithm used, are presented in this subsection. For more information see [11]. The closed-loop Wiener model identification problem can be stated as follows: Given data sequences $\{r(k), d(k), u(k), y(k)\}_{k=1,...,N}$ and the controller K, the identification problem is to determine a state space realization (A,B,C,M) of the linear dynamic block, up to a similarity transform and an output scaling, and a parametric estimation of the static output nonlinearity of the Wiener model. The identification configuration is shown in Fig. 5. The top loop in Fig. 5 represents the closed-loop with the distillation column simulation model (G)
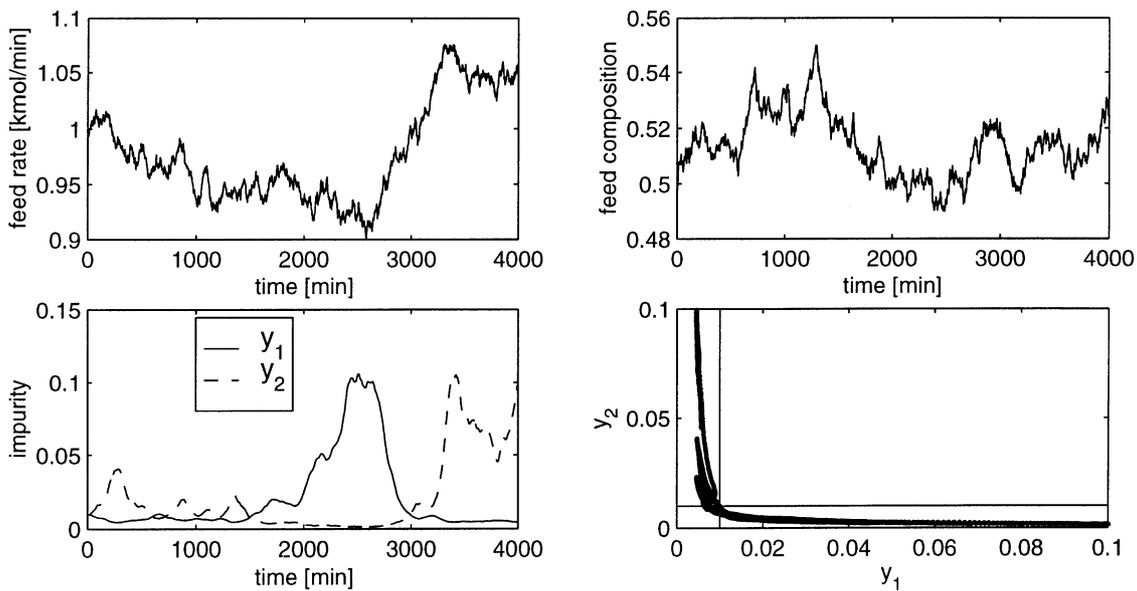

Fig. 4. Open-loop response of the distillation column (lower left plot) to manipulable inputs that are provided by two mutually independent PRBSs (see Fig. 3) in the presence of both disturbances (upper plots). Lower right plot is the covered output space.
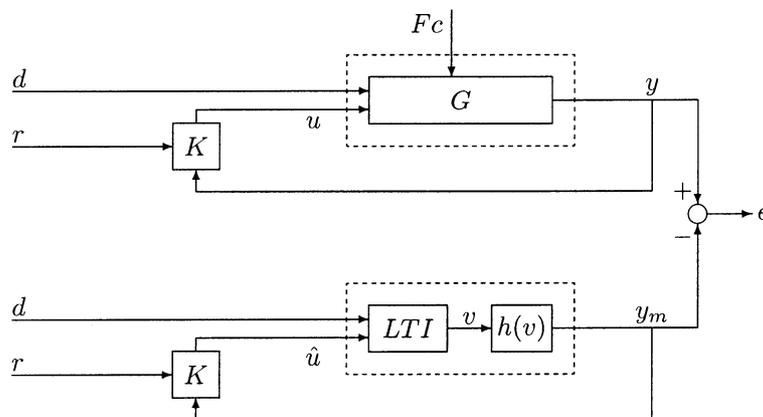

Fig. 5. Indirect closed-loop identification approach for identifying a Wiener model for the distillation column, $\hat{u}$ is used to indicate that it does not equal u.

and feed composition disturbance (*Fc*) entering the loop. The bottom loop differs from the top one in two aspects: first, the true plant is replaced by a parametrised Wiener model with linear dynamic block *LTI* and static nonlinearity *h*(*v*); second, there is no feed composition disturbance. The key idea is to adjust the parameterised model by minimising the summed and squared output error $\sum_{k=1}^{N} \epsilon(k)^2$. The algorithm basically consists of two steps:

1. Compute an initial estimate of the Wiener model. The initial estimate is obtained via a two-stage method.

    (a) In the first stage, a model is identified from the external inputs (*r* and *d*) to the plant input *u* in order to reconstruct the disturbance free part of the plant input. A Hammerstein model is identified in this stage. A motivation for this is given by the reasoning that if the controller tracks the reference perfectly, it sort of inverts the plant of which is assumed that it can be modelled by a Wiener model. The inverse then takes the form of a Hammerstein model, see [11] for more details. A non-iterative identification algorithm of Hammerstein models is given in [30].

    (b) In the second stage, a Wiener model is identified from the reconstructed input and the measured disturbance to the measured output using open-loop methods, in this case a subspace algorithm for Wiener models [31] is used. However, since the plant is nonlinear, the reconstructed input is generally not Gaussian distributed, which is assumed in the algorithm of [31]. Violation of this assumption still seems to give a model of reasonable quality [11] which can be further optimised in the next step.

2. The estimate of the previous step is used to initialise a nonlinear optimisation problem in which $\sum_{k=1}^{N} \epsilon(k)^2$ is minimised, given data sequences $\{r(k), d(k), y(k)\}_{k=1,\ldots,N}$ and the controller *K*, i.e. the model is part of a closed-loop, see Fig. 5.

Due to model mismatch and noise the output error ϵ will not be zero. For the on-line implementation of the model-based predictive controller, Section 5, the observed output error is used to update the state, by means of an extended Kalman filter, Eq. (11). Due to the nonlinearity of the model the observer gain L^{obs} in Eq. (11) is not a constant, but dependent on the linearisation at the current output, Eq. (12).

3.4. The indirect closed-loop linear model identification algorithm

Indirect closed-loop identification of a linear state space model is performed similar to that of the Wiener model, described in Section 3.3, using a linear model in every step. In this case the output nonlinearity *h* in Eq. (8) is replaced by $y_m = v$.

3.5. The direct closed-loop FIR identification algorithm

The conventional approach often used in practice, is to identify finite impulse response (FIR) models by performing open-loop SIMO identification experiments. Due to the disturbances acting on the column, the outputs drift far away from the desired operating region, see Fig. 4 for example. To prevent this an operator usually supervises the process and if necessary will take action by adjusting the inputs, thus acting as some kind of feedback controller. In fact this actually means that a closed-loop identification experiment is conducted. After collecting the input-output data the FIR model is identified as if the data were generated in open-loop, so called direct closed-loop identification. The structure of the FIR model is:

$$y_m(k) = \left(\sum_{i=1}^{2} \sum_{j=0}^{N_{\text{FIR}}-1} \theta_{i,j} u_i(k-j) \right) + \sum_{j=0}^{N_{\text{FIR}}-1} \theta_{3,j} d(k-j) \quad (9)$$

where N_{FIR} is the number of FIR coefficients. Given data sequences $\{u(k), d(k), y(k)\}_{k=1,\ldots,N}$, the parameters ($\theta$) are identified by minimising $\sum_{k=1}^{N} \epsilon(k)^2$, where $\epsilon(k) = y - y_m$. This is a linear least squares optimisation problem.

4. The identification experiment design and results

Because the difference between indirect closed-loop identification of a linear state space model and a Wiener model only involves a difference in the model structure, firstly the identification experiment design and results for these algorithms (Sections 3.3 and 3.4) are presented in Section 4.1. Secondly the identification experiment design and results for the FIR model identification are presented in Section 4.2.

4.1. Indirect closed-loop identification of a Wiener and a linear model

A closed-loop identification experiment is performed in which the top and bottom products of the distillation column are controlled by two single-loop PI controllers, see the left plot of Fig. 6. The purpose is to use the identified model for control of the column in an impurity range of about 0–0.015 for both products. The sampling time is 2 min, 10,000 samples are collected in this experiment. The PI-controllers have been tuned such that the closed-loop response time is roughly the same

as the open-loop response time. In the centre of the operating region, at $y = [0.01\ 0.01]^T$, this is about 800 min. The reference signal r (the set-points for the outputs) is a sum of two components. The first component is a slowly varying component (relative to the second component) which roughly determines a corner within the operating region $(y = [0.01\ 0.01]^T,\ y = [0.005\ 0.01]^T,\ y = [0.01\ 0.005]^T$ and $y = [0.005\ 0.005]^T)$ and is held constant for a quarter of the experimental time. The second component provides local excitation around one of the corners and is meant to provide a uniform excitation in the output space. It consists of a multi-level pseudo random signal. The multi-levels are chosen from a uniform distribution over $(-0.005,\ 0.005)$. The spectrum of this signal is determined by the switching time; the signal used has a minimum switching time of 150 min, maximum switching time of 700 min and average switching time of 300 min. Fig. 7 shows the reference signals. The disturbances are generated as described in Section 2. In Fig. 8 the collected data points are plotted in the output space. From this figure it is clear that the data is distributed more

evenly over the desired operating region compared to Figs. 3 and 4.

A Wiener model is identified using the identification algorithm described in Section 3.3. The identification data are divided into four sections of 2500 samples such that each section corresponds to data gathered around one corner of the operating region. The first 2000 samples of each section are used for estimation, the last 500 samples are used for cross-validation. The model order is selected using cross validation. The model order for the LTI part is chosen to be 3. The static nonlinearity consists of two univariate polynomials of order 3 such that the first input of the static nonlinearity affects the first output via the first nonlinear function and the same holds for the second nonlinear function. Following the discussion in Section 3.1, note that the use of two univariate polynomials to parameterise the static nonlinearity does not mean that cross terms in u do not appear since there is a MIMO LTI block preceeding the nonlinearity. Fig. 9 shows the predicted closed-loop output (with the PI controllers in the loop) of the identified Wiener
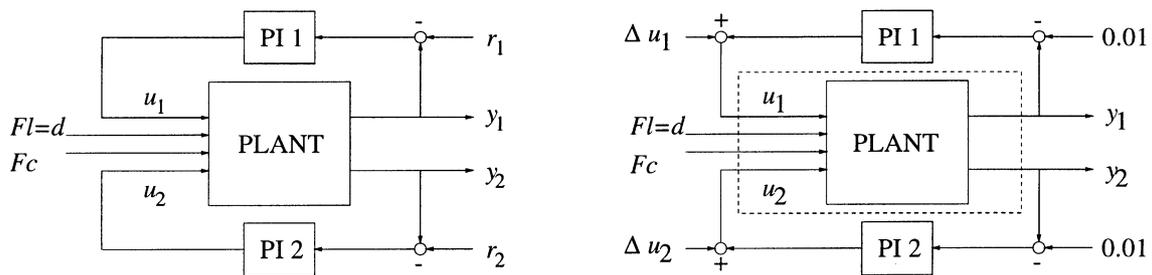

Fig. 6. Setup of the closed-loop identification experiments to identify a Wiener model and linear state space model (left) and to identify a FIR model in the conventional way (right). See the text for the meaning of the symbols.
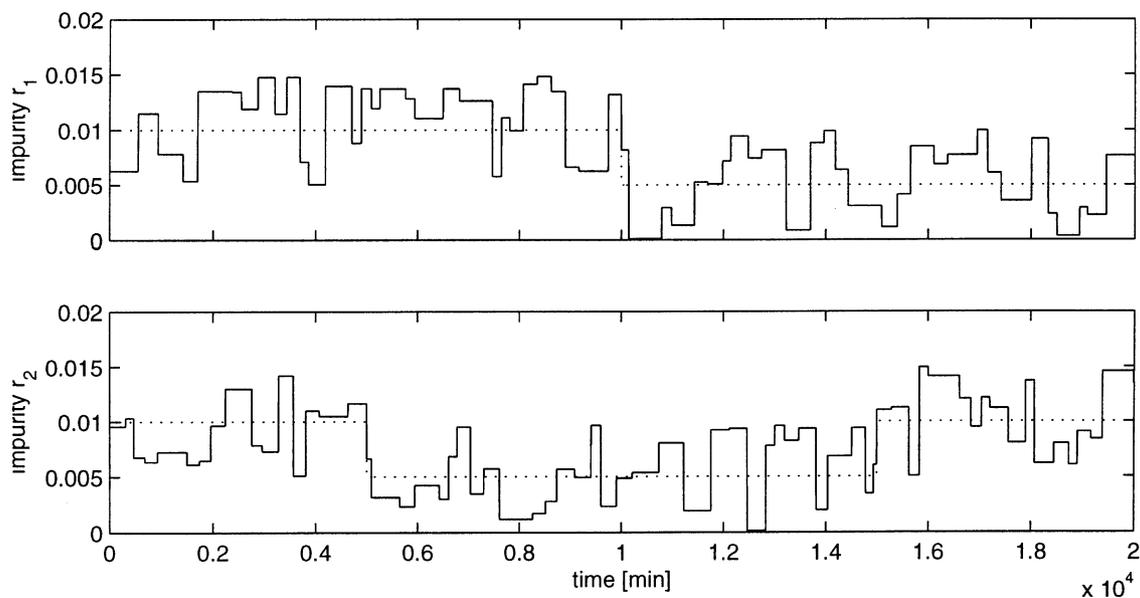

Fig. 7. Set-point reference signals used in the closed-loop identification experiment. The dotted lines show the slow components of the reference signals.
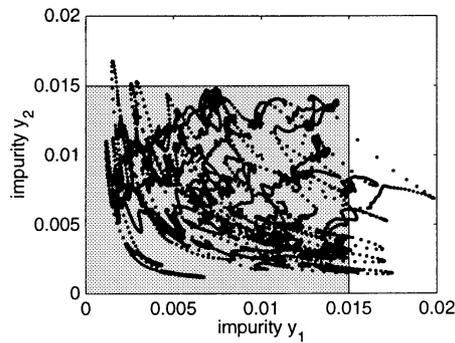
Fig. 8. The covered output space in the closed-loop identification experiment to identify the Wiener model and the linear state space model. The shaded area is the desired operating region.

model and the identification data. The regions between the narrow strips contain data used for cross-validation while the rest of the data is used for estimation. The difference between the measured output and predicted output is hardly distinguishable which indicates that the Wiener model is able to describe accurately the behaviour of the distillation column in the closed-loop setting. The input–output relation of the output nonlinearities of the identified Wiener model are plotted in Fig. 10. The saturating effect of the outputs y_m towards low impurities is clearly visible. Note that the nonlinearities are clipped at the two ends as they are only valid within the region where the identification experiment is conducted.

To demonstrate how the approximation of the distillation column by the Wiener model is improved compared to a linear model, a linear model was identified using the algorithm described in Section 3.4. The order of the model is selected using cross validation and is chosen to be 3. Fig. 11 shows the predicted closed-loop output (with the PI controllers in the loop) of the identified

linear model and the identification data. In this figure a significant difference between the measured output and predicted output is observed. Comparing Figs. 9 and 11 reveals that the Wiener model is a better approximation for the distillation column simulation model, than the linear model is. The worse fit of the linear model can partly be explained by its inability to model the saturating effect of the outputs towards low impurities, see Fig. 10. In the linear model some "average" gain is identified instead of this nonlinear relation. Therefore it underestimates the outputs towards the extremes in impurities. For high impurities the real gain is larger, see Fig. 10 in the region 2000–3000 min for example. For low impurities the real gain is lower, see Fig. 11 in the region around 8000 min for example (for identification the mean values of the data are subtracted, so the origin of the identified models corresponds to the mean of the identification data, i.e. about $y = [0.01\ 0.01]^T$. The concept of the above mentioned gains is related to a deviation from this origin).

4.2. Direct closed-loop identification of a FIR model

To compare the Wiener model identification setup with a conventional approach a FIR model has been identified using the algorithm presented in Section 3.5. Note that this comparison does not only include the difference between a Wiener model and a linear model, but also the difference in the setup of the identification experiment which is important for identifying the low gain direction as motivated in Section 3.

To mimic an operator, which should prevent the column from drifting from the desired operating region, an identification experiment is conducted in which the distillation column is controlled at the centre of the desired
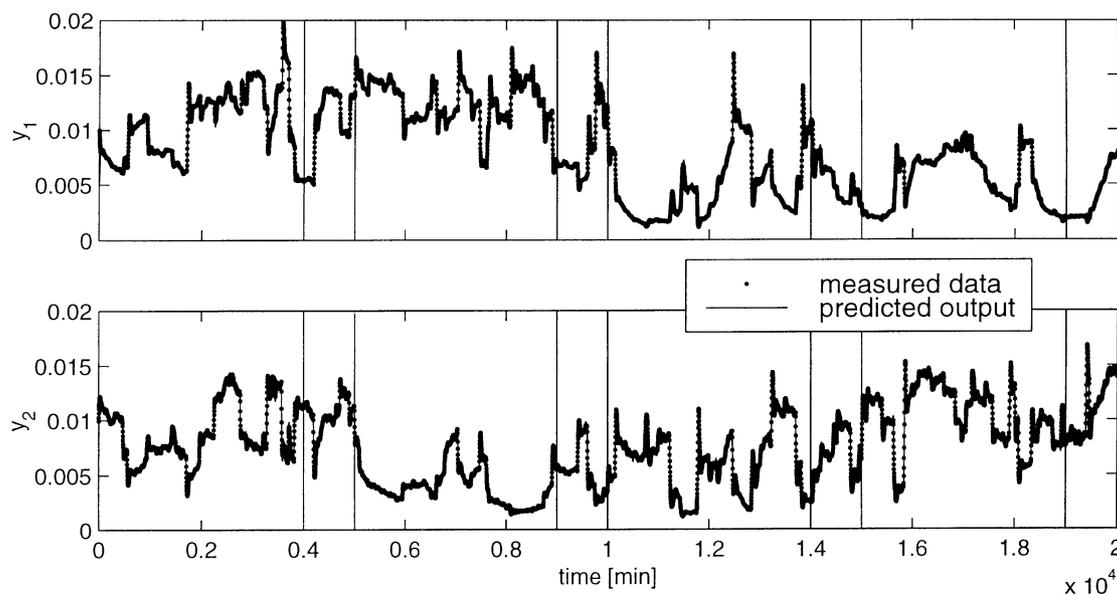

Fig. 9. Identification data and predicted output of the Wiener model in closed-loop with the PI controllers.

operating region, $y = [0.01\ 0.01]^T$, by two PI controllers. In order to excite the plant, the outputs of the PI controllers are corrupted by adding pseudo random binary signals (Δu_1 and Δu_2, see the right scheme of Fig. 6) with a minimum switching time of 50 min. Firstly the $y_1 - u_1$ loop is corrupted while Δu_2 is kept zero, secondly Δu_1 is kept zero and the $y_2 - u_2$ loop is corrupted. To cover the entire operating region and to avoid the use of large signals Δu_1 and Δu_2 the two PI controllers are tuned to give a slow response. This should mimic an operator whose objective is not to keep the outputs tight at the set-point (the objective is to excite the plant in the identification experiment), but who only tries to prevent the column from drifting too far from the desired operating region. An identification experiment of 42,000 min was performed. This relatively large time compared to the identification experiment for identifying the Wiener model is motivated by the large settling time of the distillation column in open-loop which is approximately

800 min. A FIR model of 60 lags (a usual length) then corresponds to a sample time of about 14 min for the tail of the FIR model to have reached steady state. Using 3000 data points then results in an identification experiment of $3000 \times 14 = 42,000$ min. Note that the number of data points is not too large when trying to identify $2 \times (2 + 1) \times 60 = 360$ impulse response coefficients (the model consists of 2 outputs, 2 inputs and one measured disturbance). After collecting the input–output data the FIR model is identified as if the data were generated in open–loop, so called direct closed-loop identification, concerning only the dashed box in the right scheme of Fig. 6. In Fig. 12 the collected data points are plotted in the output space. Although the experiment essentially is a closed-loop experiment it is not designed in order to provide sufficient excitation of the low gain direction. The only goal of the controller is to prevent drifting of the outputs from the desired operating region. Again, the result is that the data
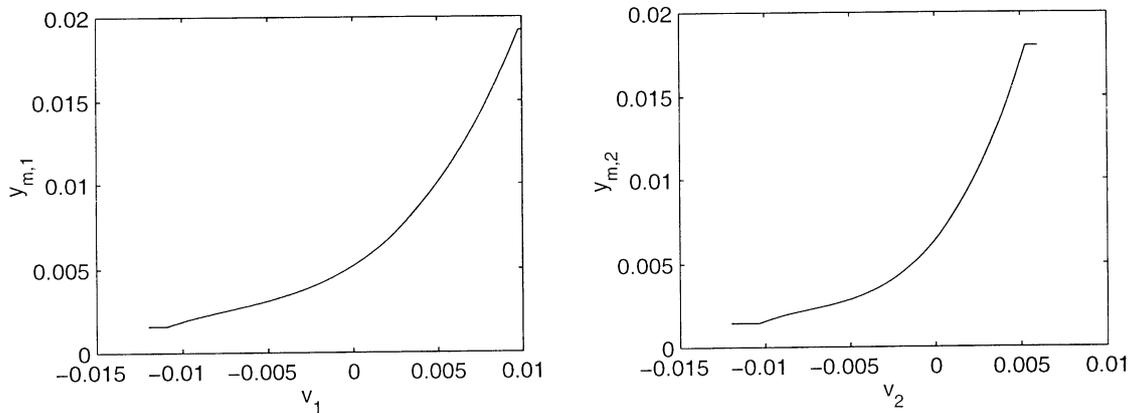

Fig. 10. Input–output relation of the identified output nonlinearities of the Wiener model.
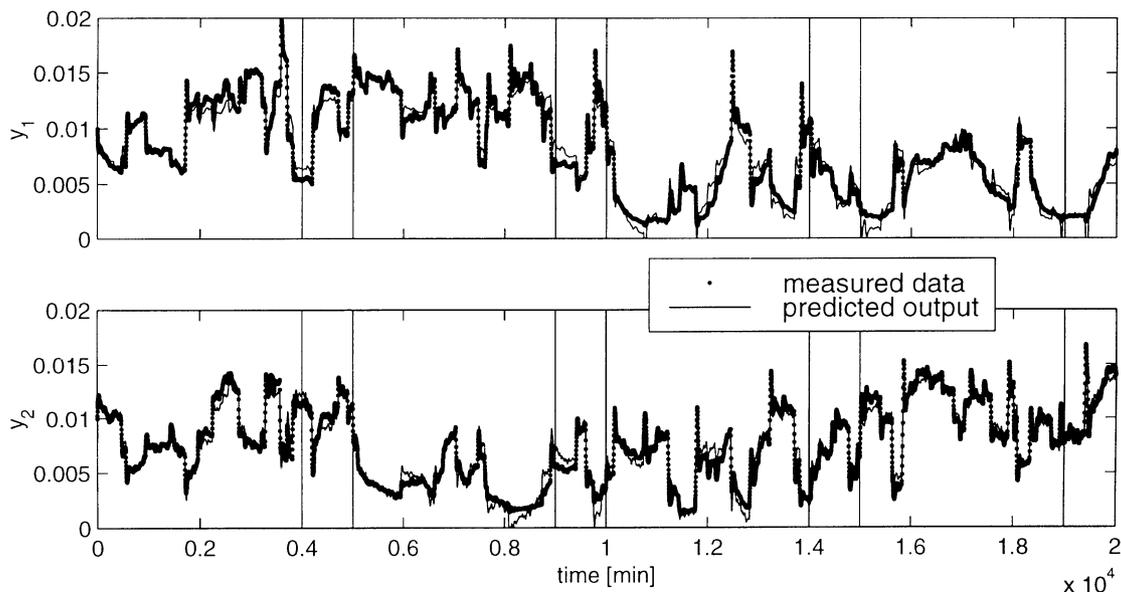

Fig. 11. Identification data and predicted output of the linear state space model in closed-loop with the PI controllers.
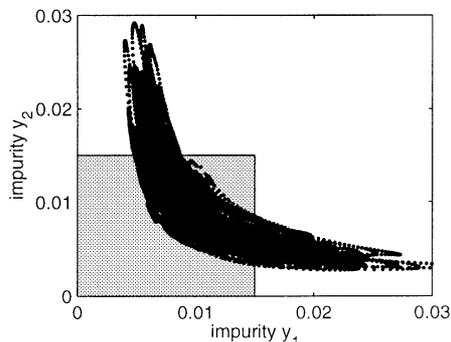
Fig. 12. The covered output space in the identification experiment used to identify the FIR model. The shaded area is the desired operating region.

Fig. 13. Detail of the identification data and the one-step ahead prediction of the raw FIR model.

mainly reflect the high gain direction of the distillation column, compare with Fig. 8, although the result is better than a purely open-loop experiment, Fig. 4. Excitation at a low impurity level of one output is very difficult, as it is always coupled with excitation of the other output at high impurity levels. A higher amplitude of the excitation signals, or a different tuning of the controller can not remedy this. The objective is to obtain a model that is valid at low impurities, however in this experiment setup impurity levels down to 0.002, as in the identification experiment used for identification of the Wiener model, can not be reached, especially in case of y_1.

Although in open-loop the settling time of the distillation column is approximately 800 min, the tail of the impulse response identified from the collected data dies out after about 70 samples with a sample time of 2 min. This deviation from a priori knowledge indicates the inability of the FIR identification approach to cope with the data which is generated in closed-loop. Another remarkable phenomenon in the identified responses is an extreme non-minimum phase behaviour for both manipulable inputs. A theoretical background for this effect can be found in [4]. Moreover, the linear FIR model can not approximate the non-linear behaviour of the distillation column accurately. Fig. 13 shows a detail of the identification data and the one-step ahead prediction of the identified FIR model in the case of the top impurity. The identified extreme non-minimum phase behaviour causes the peaks in the FIR simulation, which correspond to sample instants at which Δu changes sign. Because of the complete mismatch of the initial response, the identified "raw" FIR model appeared to be useless for application within a MPC algorithm. A FIR model of 101 elements is identified. To remedy the mismatch of the initial response the direct feed-through terms are set to zero. This impulse response is then filtered to obtain a reasonably smooth FIR. The response is truncated after 70 lags. In Fig. 14 the raw FIR and the modified FIR model which is used for the MPC are plotted.
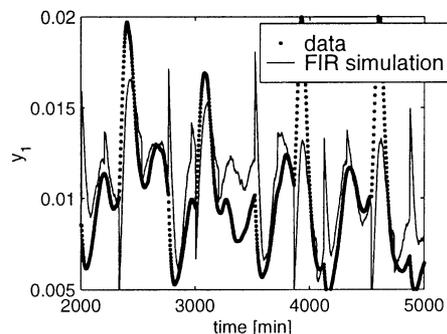
5. Model-based predictive control

The three different models obtained in the identification step, i.e. the Wiener model, the linear state space model and the FIR model, can all be used within a Model-based Predictive Control (MPC) scheme. In this section the basic steps of the algorithms are presented. For the Wiener model, two different algorithms are presented, which differ in the way they handle the nonlinearity of the Wiener model. In Section 5.1 the MPC algorithm is outlined which transforms the nonlinearity of the Wiener model into a polytopic description. The changes in the algorithm in case the nonlinearity is removed from the control problem via an inversion are presented in Section 5.2. The changes in the algorithm in case a linear state space model is used in stead of a Wiener model are presented in Section 5.3. The basic steps of the conventional MPC algorithm for FIR models are presented in Section 5.4. In Table 2 an overview of the presented MPC algorithms is given.

5.1. Wiener MPC algorithm based on polytopic descriptions

In order to achieve off-set free control integrating action is incorporated in the identified Wiener model by writing it in incremental-input–output form:

$$\begin{cases} x_e(k+1|k) = \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} x_e(k|k) + \begin{bmatrix} B \\ I \end{bmatrix} \Delta u(k|k) + \begin{bmatrix} M \\ 0 \end{bmatrix} d(k|k) \\ v(k|k) = [C\ 0] x_e(k|k) \\ y_{m,1}(k|k) = h_1(v_1(k|k)) \\ y_{m,2}(k|k) = h_2(v_2(k|k)) \end{cases}$$

(10)

with $\Delta u(k|k) = u(k|k) - u(k-1|k)$, and $x_e(k|k)^T = \left[x(k|k)^T u(k-1|k)^T \right]$ the extended state. I is the identity matrix. The argument $(k+1|k)$ reflects the prediction at time $k+1$, using information up to time k. By means of an extended Kalman filter (EKF) the states x_e are updated every sampling instant according to:
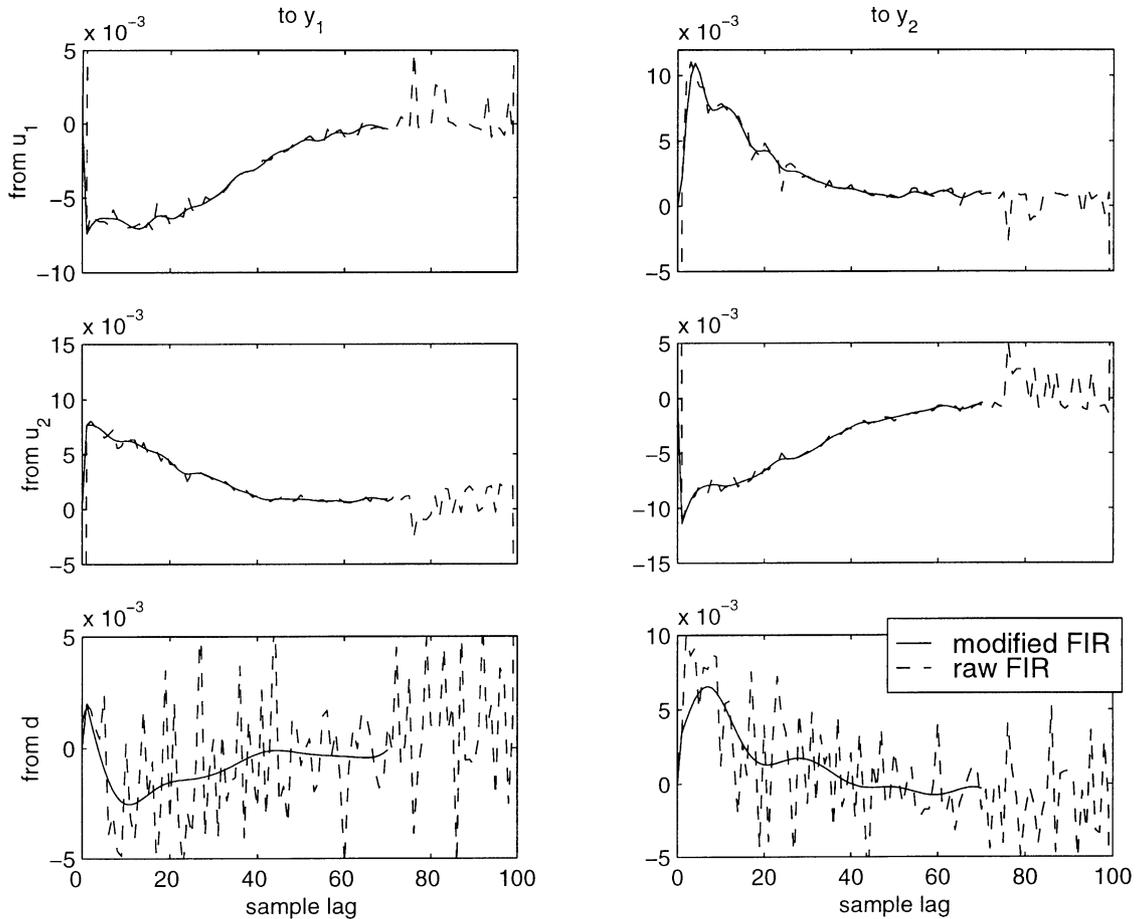
Fig. 14. The raw FIR model and the modified FIR model (direct feed through set to zero, filtered and truncated).

Table 2
Overview of the different MPC methods that are used in the comparison

Model	MPC methodology	Description	Results
Wiener	PWMPC: use a polytopic description to capture the effect of the nonlinearity	Section 5.1	Section 6.1
	IWMPC: remove the nonlinearity from the control problem via an inversion	Section 5.2	Section 6.1
Linear	LMPC based on the linear state space model	Section 5.3	Section 6.2
	FIRMPC based on the FIR model	Section 5.4	Section 6.3

$$x_e(k|k) = x_e(k|k-1) + L^{obs}(k)(y(k) - y_m(k|k-1)) \tag{11}$$

where $L^{obs}(k)$ is the Kalman gain for the Wiener model, linearised at the current linear output $v^* = v(k|k-1)$, turning the nonlinear output mapping into (the bar is used to express deviation variables):

$$\bar{y}_{m,\text{lin}}(k|k-1) = \frac{\partial h(v)}{\partial v}\Big|_{v=v^*}\bar{v}(k|k-1) = C_{\text{lin}}\bar{x}_e(k|k-1) \tag{12}$$

where C_{lin} describes the linear mapping from $\bar{x}_e$ to the linearised model outputs $\bar{y}_{m,\text{lin}}$. The EKF provides the

feedback path of the controller to compensate for the unmeasured disturbance (feed composition) and model-mismatch.

The design goal of the MPC is to minimise the following performance index:

$$J(k) = \sum_{i=1}^{\infty} y_m(k+i+1|k)^T Q_y y_m(k+i+1|k)$$
$$+ \Delta u(k+i|k)^T R_u \Delta u(k+i|k) \tag{13}$$

Note that the minimisation of $J(k)$ which starts at time k, yields $\Delta u(k+1|k)$ as the first input signal, which will be implemented at time $k+1$. In a real time experiment

this delay of 1 sample is necessary to be able to solve the optimisation problem.

Assume $(\Delta u_{\mathrm{ref}}, d_{\mathrm{ref}}, x_{e,\mathrm{ref}}, v_{\mathrm{ref}}, y_{m,\mathrm{ref}}) = (0, 0, 0, 0, 0)$, when this is not the case the system can be shifted such that the origin of the shifted system corresponds to the desired reference signal. Because future disturbances are unknown it is assumed that $d(k + i|k) = d(k|k)$, $\forall i \geqslant 0$, because the variations in the feed rate are relatively small compared to the sample time. These future disturbances will affect the setpoints of u_{ref} and x_{ref} corresponding to a desired setpoint $y_{m,\mathrm{ref}}$. As a result these future disturbances essentially affect $x_{e,\mathrm{ref}}$. After shifting the system, the shifted disturbance equals zero and drops out of the predictions of the shifted system. Thus feed-forward compensation of the feed rate disturbance is accomplished by adjustment of $x_{e,\mathrm{ref}}$.

Directly using the Wiener model in above performance index will in general lead to a non-convex optimisation problem due to the output nonlinearity. Solving a non-convex optimisation problem on line is not attractive because it is computationally expensive and may suffer from the existence of multiple local minima [18].

However, by exploiting the structure of the Wiener model, the need for solving a non-convex optimisation problem can be avoided. In this paper a technique is used in which the nonlinear output mapping is transformed into a polytopic description, which enables to use robust linear MPC techniques [6,7]. In Section 6 the performance of this technique is compared to a methodology in which the nonlinearity is removed from the control problem via an inversion [22]. The Wiener MPC algorithm presented in this section is a short and tailor-made description of the one presented in [6,7], for more information the reader is referred to these references.

Because the origin corresponds to the desired setpoint (possibly after shifting the system), and the identified output nonlinearity consists of two third order univariate polynomials, the output nonlinearity can be represented as:

$$y_m = h(v) = H(v) \cdot v = \begin{bmatrix} H_{1,1}(v_1) & 0 \\ 0 & H_{2,2}(v_2) \end{bmatrix} \cdot v \qquad (14)$$

where $H_{1,1}$ and $H_{2,2}$ are second order polynomials. By restricting the operating region for v, which is motivated by the fact that the model is only valid within the region of the identification data, lower and upper bounds on $H_{1,1}$ and $H_{2,2}$ can be calculated. All the possible combinations of these extreme values then generate 2^2 vertices $\mathcal{H}_1, \ldots, \mathcal{H}_{2^2}$, which span a convex hull (Co) Ω in which the nonlinear matrix $H(v)$ is contained, i.e.:

$$H(v) \in \Omega = Co\{\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4\} \qquad (15)$$

The constraints for the operating region for v can be implemented easily in the MPC algorithm.

What results from the above procedure is a linear uncertain model which is used in a robust linear MPC algorithm. To avoid an infinite number of degrees of freedom due to the infinite prediction and control horizon, Eq. (13), the inputs in the tail of this horizon, beyond the so called switching horizon H_s, are parameterised as a state feedback law:

$$\Delta u(k + i|k) = F x_e(k + i|k), \text{ for } i \geqslant H_s + 1 \qquad (16)$$

where F is a variable in the optimisation problem. The inputs until H_s are degrees of freedom as well, and in this respect H_s is related to the control horizon used in other MPC algorithms, see [19] for example. In the simulations presented in Section 6, H_s is set to 5. Now the performance index, Eq. (13), can be split into two parts [6,7]:

$$J(k) = J_1(k) + J_2(k) \qquad (17)$$

$$\begin{aligned} J_1(k) =& \sum_{i=1}^{H_s} y_m(k + i + 1|k)^T Q_y y_m(k + i + 1|k) \\ &+ \Delta u(k + i|k)^T R_u \Delta u(k + i|k) \end{aligned}$$

$$J_2(k) = x_e(k + H_s + 1|k)^T P x_e(k + H_s + 1|k)$$

The minimisation of $J(k)$ is transformed into a minimisation of a linear objective function, subject to matrix inequalities which are affine in the nonlinearity $H(v)$. Through the use of the polytopic description, Eq. (15), these nonlinear matrix inequalities are satisfied when they hold for every vertex of the polytopic description, leading to linear matrix inequalities (LMIs). Thus the minimisation problem results in an LMI-based optimisation problem, which is convex, see [6,7] for details.

Basically the nonlinearity is handled as uncertainty in this optimisation problem. Calculating the uncertainty based on the entire operating region for v may result in a very conservative controller, depending on the shape of the nonlinearity. By means of an iterative approach the conservatism is reduced by artificially narrowing the operating region, and thus the uncertainty. At first an input sequence is calculated by minimising the performance index based on the polytopic description calculated for the entire valid operating region. Then "artificial" constraints are introduced around the predicted trajectory for v. This narrows the operating region for v. Now a new polytope for $H(v)$ can be calculated, that is less conservative since it is based on a smaller operating region. Then a new input sequence is calculated by minimising the performance index based on the new polytope, which results in a less conservative control action, as the uncertainty is reduced. This improves the performance of this control algorithm. Because the "artificial" constraints on v lie around a trajectory, the

polytope for $H(v)$ changes along the predictions. This refinement of the calculated control sequence can be done repeatedly. It can be seen as an automatic gain scheduling procedure. In the simulations, see next section, two extra iterations are used. Because of model mismatch every new sampling instant the MPC starts with a polytope based on the entire operating region for v. For more details, see [6,7].

By transforming the nonlinearity into a polytopic description the effect of the nonlinearity on the output is still taken into account. This is the main advantage of this approach compared to the approach of [22] for example, in which the nonlinearity is essentially removed from the control problem by an inversion. This is demonstrated in Section 6.

In short, the controller does the following; at time k, apply input $\Delta u(k|k-1)$ and measure the output of the white box distillation column simulation model and the feed rate disturbance, then update the states in the observer and calculate the input to be applied at time $k + 1$, wait until time $k + 1$ and repeat the procedure.

The MPC algorithm described in this subsection will be referred to as PWMPC: it uses a Polytopic description of the output nonlinearity of the Wiener model within the MPC algorithm.

5.2. Wiener MPC algorithm based on an inversion of the nonlinearity

The basic concept of this Wiener MPC algorithm is presented in [22] and consists of inverting the output nonlinearity, thus removing it from the control problem. What remains is the linear block of the Wiener model. For this block a linear MPC can be used which minimises the following performance index:

$$J(k) = \sum_{i=1}^{\infty} v(k + i + 1|k)^T Q_v v(k + i + 1|k) + \Delta u(k + i|k)^T R_u \Delta u(k + i|k) \tag{18}$$

For comparison with the PWMPC algorithm the weighting matrix Q_v is chosen as:

$$Q_v = \left(\left. \frac{\partial h(v)}{\partial v} \right|_{v = h^{-1}([0.01\ 0.01]^T)} \right)^T \cdot Q_y \cdot \left. \frac{\partial h(v)}{\partial v} \right|_{v = h^{-1}([0.01\ 0.01]^T)} \tag{19}$$

thus at the point $y = [0.01\ 0.01]^T$ impurity both performance indices, (13) and (18), are equivalent. The extended Kalman filter, Eq. (11), is used to update the states. The linear MPC algorithm for the LTI block of the Wiener model essentially is the same as the algorithm described in Section 5.1. In the simulations presented in Section 6, H_s is set to 5. Because this MPC only considers the LTI

block of the Wiener model, there is no polytopic description and therefore there is no need for refining the control action in an iterative way.

The MPC algorithm described in this subsection will be referred to as IWMPC: it uses an Inversion of the output nonlinearity of the Wiener model in order to remove the nonlinearity from the optimisation problem within the MPC algorithm.

5.3. Linear MPC algorithm for the linear state space model

The linear MPC algorithm for the identified linear state space model essentially is the same as that for the Wiener model in Section 5.1. Since the output mapping is linear the observer now is a regular Kalman filter, thus $L^{obs}(k)$ in Eq. (11) is a constant. In the simulations presented in Section 6, H_s is set to 5. Because the output mapping is linear, there is no polytopic description, so there is no need for refining the control action in an iterative way.

The MPC algorithm described in this subsection will be referred to as LMPC: it uses the Linear state space model, identified by an indirect closed-loop identification algorithm, within the MPC algorithm.

5.4. Linear MPC algorithm for the FIR model

For the FIR model, the MPC algorithm described in Section 2.3 of [9] is used. Guidelines for tuning of the prediction and control horizon indicate that the product of prediction horizon and sample time should be larger than the time for the FIR to reach 95% of the steady state, and that the product of control horizon and sample time should be larger than the time for the FIR to reach 60% of the steady-state [17]. A prediction horizon and control horizon of 1.5 respectively 0.5 times the length of the impulse response were used, which satisfy the guidelines.

Unmeasured disturbances and model mismatch cause an error between the measured output of the nonlinear simulation model ($y(k)$) and the predicted output of the FIR model ($y_m(k)$). To compensate for this error, it is added to the output of the FIR model resulting in the following predictor within the MPC:

$$\hat{y}_m(k + i|k) = y_m(k + i) + (y(k) - y_m(k)) \tag{20}$$

where $\hat{y}_m(k + i|k)$ are the predicted values of the output at time $k + i$ using measurements up to time k. Again for the measurable disturbance is assumed that $d(k + i|k) = d(k|k)$, $\forall i \geqslant 0$. It is compensated for in a feed-forward fashion because the influence of d on the outputs is modelled in the FIR model.

The MPC algorithm described in this section will be referred to as FIRMPC: it uses a FIR model, identified as described in Section 3.5, within the MPC algorithm.

6. Control results and discussion

The extended Kalman filter for the PWMPC and IWMPC algorithms and the Kalman filter for the LMPC algorithm are tuned by adjusting the assumed noise covariances such that a smooth estimate of the real output is obtained.

The matrices that weight the deviation of the outputs from the setpoints (Q_y) and the control effort (R_u) are kept the same for all MPC algorithms. Note that the IWMPC algorithm of Section 5.2 does not use the physical output y_m in the performance index. To obtain a good comparison, the weighting of the deviation of the linear output v from its setpoint, used in this algorithm, is given by Eq. (19). This means that in the centre of the operating region, at $y = [0.01 \ 0.01]^T$, both weightings are equivalent.

The weighting on deviations from the setpoints is equal for both outputs, and 1000 times larger than the weighting on the control effort. These weightings have been selected such that reasonably fast control is achieved, without too large overshoots. Note that the magnitude of the output signals is much smaller than that of the input signals, which causes the higher weighting on the output signals.

The performance of the different controllers is compared in case of several setpoint changes within the operating region. For the PWMPC, IWMPC and LMPC algorithms the setpoints for the outputs are either 0.003, 0.01 and 0.015 impurity, covering the valid operating region, see Fig. 8. For comparison with the FIRMPC the setpoints are either 0.005, 0.01 and 0.015 impurity, because the FIR model is not valid down to low impurities, see Fig. 12. Although the identification data for the FIR model does not cover the points $y = [0.005 \ 0.005]^T$ and $y = [0.015 \ 0.015]^T$, see Fig. 12, the corresponding setpoints have been included as the goal is to perform dual setpoint control, including setpoints where both outputs are at low or high impurities. Every 200 min one of the setpoints for the outputs of the distillation column is changed. The disturbances (feed rate and feed composition) are the same in all simulations and are plotted in Fig. 15. Firstly the PWMPC algorithm is compared to the IWMPC algorithm, which demonstrates the difference between taking the nonlinearity into account via the polytopic description, and removing it via an inversion. Secondly a comparison to the LMPC algorithm is made to show the difference between a linear model and a Wiener model within the MPC scheme. Thirdly a comparison to the FIRMPC is made to show the effect which is due to the difference in the identification experiment setup.

6.1. Differences between the Wiener MPC algorithms

The simulations with both the PWMPC and IWMPC algorithms are displayed in Fig. 16. Because around $y = [0.01 \ 0.01]^T$ the output non-linearity is rather linear, see Fig. 10, both methods are expected to give the same results around this operating point, because the performance index that is used by the IWMPC algorithm, Eq. (18), is tuned, Eq. (19), to give the same performance at this point, when minimising the performance index in Eq. (13). This is visible in the first part of the simulation in Fig. 16. Both algorithms indeed give the same performance. However, when going to lower or higher impurities the results differ. Going in the direction of higher impurities, for y_1 at time step $t = 400$, $t = 1400$, and $t = 1600$, the gain of the output nonlinearity increases and therefore the weighting on v in the performance indices should increase. The PWMPC algorithm takes care of this increased gain, resulting in a faster setpoint change compared to the IWMPC algorithm, which does not take this increase in gain into account. The contrary happens when going in the opposite direction, i.e. lower impurity, at time step $t = 200$, $t = 1000$, $t = 1200$, which means that the gain of the output nonlinearity decreases, and, therefore, the weighting on v in the performance indices should decrease. Now the PWMPC algorithm, which takes care of this decreased gain, shows a slower response than the IWMPC algorithm which in these cases uses a relatively larger weighting on v. However, these faster setpoint changes are coupled with a quite oscillatory behaviour of the input and output signals, especially at $t = 200$, $t = 1000$, and $t = 1200$. Moreover, the weighting of the output that operates at the higher impurity level will be relatively too small again, causing an oscillation with large amplitude at $t = 200$ and $t = 1000$. These negative side effects of the faster setpoint-changes of the IWMPC algorithm when going towards lower impurity, ask for a detuning of this controller, leading to a slower response, also for setpoint changes towards higher impurities. The PWMPC algorithm does not suffer from these negative side-effects and is able to balance the input-output behaviour over the entire operating region because it accounts for the nonlinearity.

6.2. Influence of the model: Wiener versus linear

The simulation with the LMPC algorithm, based on the linear state space model identified using the indirect closed-loop identification algorithm, is displayed in Fig. 17. When going to low impurity levels, the linear model over-estimates the gain. As a result the required input level to reach a steady state at low impurity is underestimated. Based on this observation one would expect a slower convergence to low impurities compared to the Wiener MPC algorithms, which are expected to give a better estimate of the required input level. This is true for the IWMPC algorithm, but not true for the PWMPC algorithm, compare with Fig. 16. The reason why PWMPC shows a similar behaviour in this case, at

$t = 200$, $t = 1000$ and $t = 1200$, comes from the fact that the weighting on v (i.e. $H(v)^T Q_y H(v)$) in this algorithm decreases at lower impurity levels, which makes the convergence for the PWMPC algorithm slower towards low impurities. In the LMPC algorithm the weighting remains the same. Because these two effects counteract

it results in a similar response for these setpoint changes. Note that these effects are caused by a model-mismatch of the linear model, which does not take the saturating effect of the output into account. Because the observer is tuned to be quite fast, the model-mismatch of this linear model can be compensated for rapidly. A similar
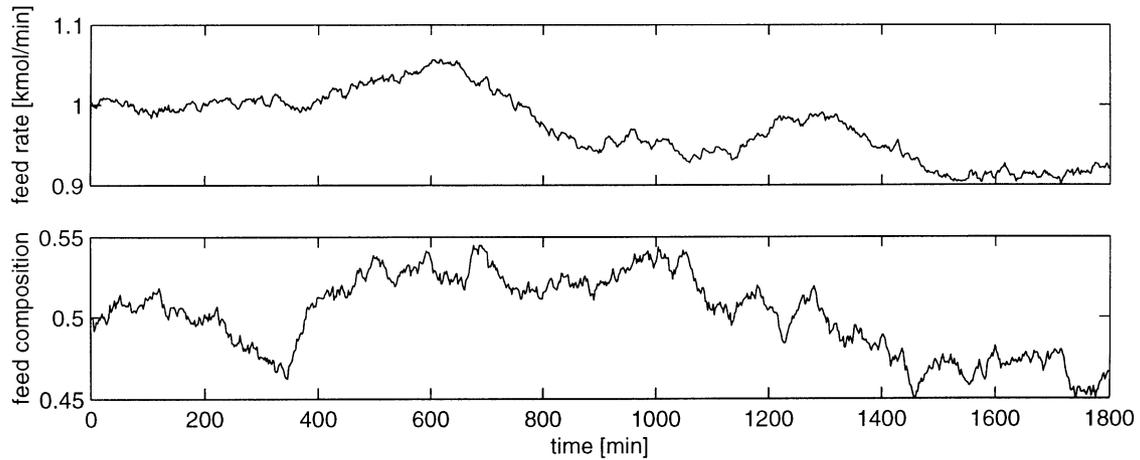

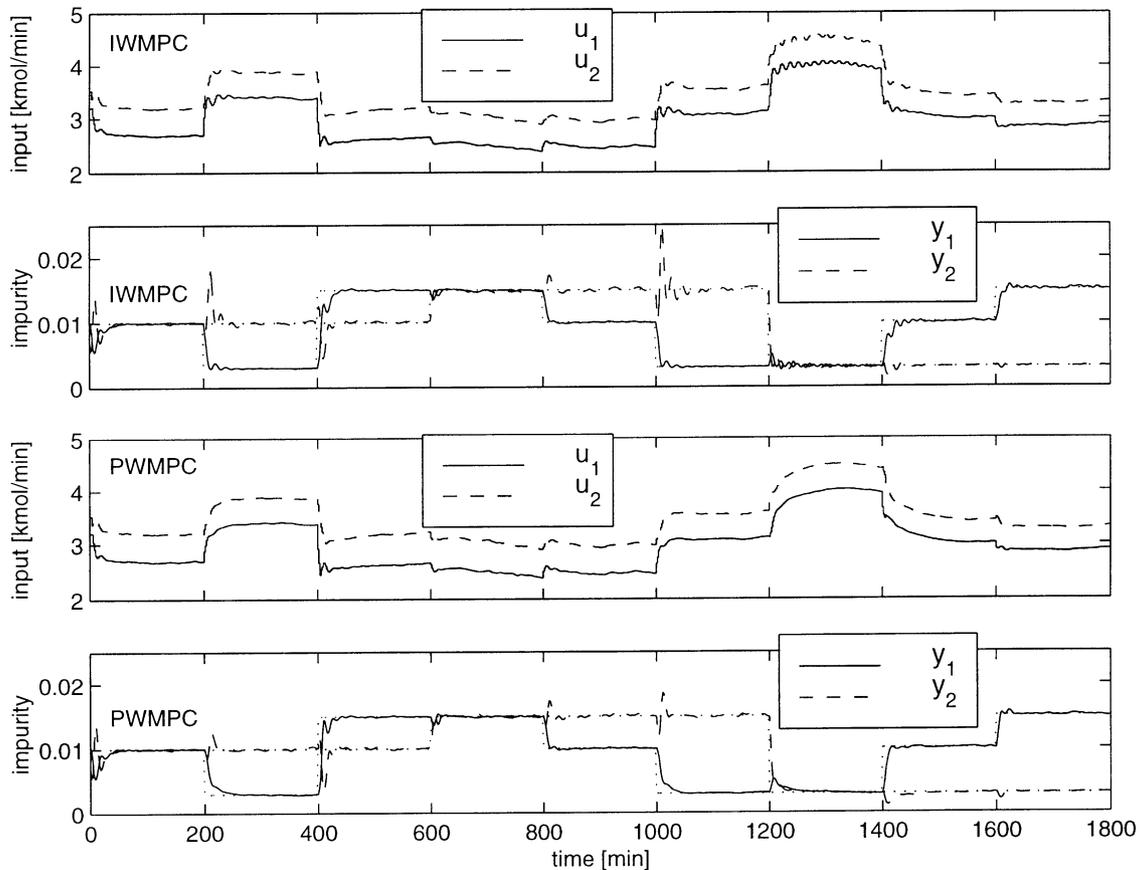Fig. 15. The disturbances used in the control experiments.

Fig. 16. Setpoint changes for the IWMPC algorithm (upper plots) and the PWMPC algorithm (lower plots). The dotted lines are the setpoints for the outputs.

effect can be observed when going towards higher impurity levels. Based on the same reasoning the response of this LMPC algorithm is expected to be faster than the IWMPC algorithm. This tendency can be observed at time $t = 400$, $t = 1400$ and $t = 1600$. However, the PWMPC algorithm shows a better response for these instants. This is due to the fact that the two effects caused by the model mismatch in the linear model do not completely counteract, resulting in overshoot at these instants. At $t = 600$ these oscillations even become so severe that detuning of the LMPC algorithm is required. Because of the model mismatch of the linear model, although it causes two effects which partly counteract, the performance of the LMPC algorithm is worse than the PWMPC algorithm. The model mismatch of the linear model can partly be compensated for by a fast observer tuning. When the observer has to be detuned in case of the presence of measurement noise for example, which is not considered in this paper, the performance of the LMPC algorithm will degrade further. As the control action calculated by the MPC heavily relies on the model, a better model favours a better control performance.

6.3. Influence of the identification experiment setup

As motivated in the first part of Section 6, the lowest setpoint for the outputs is 0.005 impurity in these simulations. The results are plotted in Fig. 18. From this figure it is clear that the PWMPC algorithm is able to do faster setpoint changes than the FIRMPC. This is mainly caused by the fact that the low gain direction of the distillation column is modelled better in the Wiener model, compared to the FIR model. Observing the FIRMPC simulation in Fig. 18 one can see, for example, that when the impurity setpoint for y_1 decreases after $t = 200$ min, it causes an increase in the impurity of y_2. This corresponds to the high gain direction, i.e. one product becoming purer and the other impurer. On the contrary the PWMPC algorithm is very well able to keep y_2 at the setpoint when changing the setpoint for y_1. The better modelling of the low gain direction in the
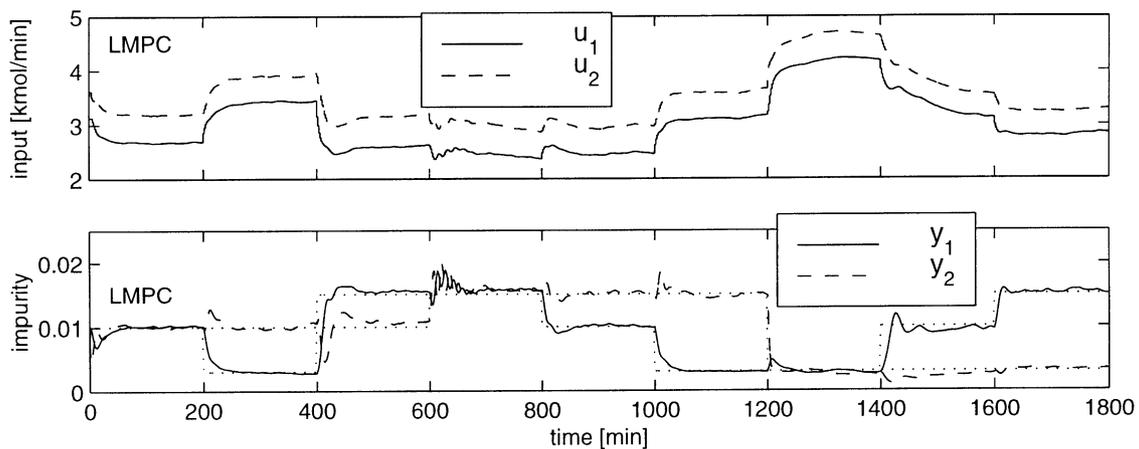

Fig. 17. Setpoint changes for the LMPC algorithm, based on the linear state space model identified in the indirect closed-loop identification algorithm. The dotted lines are the setpoints for the outputs.
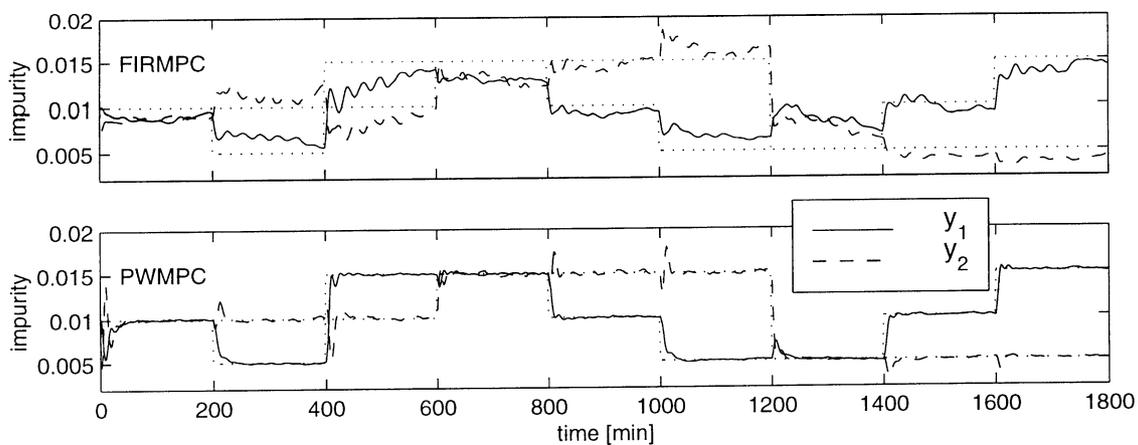

Fig. 18. Setpoint changes for the FIRMPC algorithm (upper plot) and PWMPC algorithm (lower plot). The dotted lines are the setpoints.

Wiener model is due to the different setup of the Wiener model identification experiment compared to the conventional setup used for FIR identification, see Section 4, which makes it easier to excite and identify this low gain direction. Note that also the LMPC algorithm, which uses the linear model identified by the indirect closed-loop identification algorithm, outperforms the FIRMPC algorithm, compare Figs. 17 and 18, because the low gain direction of the distillation column is modelled better in the linear state space model than in the FIR model. This is due to the identification experiment setup.

7. Conclusions

In this paper the benefits of a Wiener model based identification and control methodology, compared to one based on linear models, have been demonstrated for a nonlinear benchmark problem. The benchmark problem considered is a nonlinear white box distillation column simulation model. Besides the nonlinearity of the benchmark problem it is also characterised by an ill-conditioned nature. It has been motivated and demonstrated that such an ill-conditioned process requires closed-loop identification in order to be able to identify the low gain direction. A closed-loop identification experiment setup was presented which enables to excite the low gain direction. Based on these data the benefits of identifying a Wiener model instead of a linear model have been shown. In addition it has been demonstrated that both these models, that were obtained through an indirect closed-loop identification algorithm, performed better than a FIR model which was identified using a direct closed-loop identification algorithm, using data generated in a conventional identification experiment setup. This comparison demonstrates the benefits of the presented identification experiment setup. The poor quality of the FIR model stems from the facts that the identification data in the conventional setup do not contain enough information regarding the low gain direction of the plant, and that the data is used as if it is generated in open-loop, thus ignoring the controller which is needed in the identification experiment to avoid drifting from the desired operating region.

The identified models have been used within a MPC framework. Because the low gain direction of the FIR model is not correctly identified, due to the identification experiment setup, the FIRMPC gives an inferior performance as it uses too much the high gain direction of the plant, which is not suitable for dual composition control. The LMPC, that uses the linear state space model identified using the indirect closed-loop identification algorithm, offers a much better performance because the low gain direction is modelled better. However, the linear model can not capture the nonlinearity of the process, and this model-mismatch limits the performance of the LMPC. This model mismatch can be compensated for via the feedback path, the state observer. The Wiener model is able to approximate accurately the nonlinearity of the process. A MPC based on the Wiener model therefore can handle this nonlinearity in a feed forward fashion, which provides a better control performance. Two different MPC algorithms for Wiener models have been compared. In the optimisation problem of the IWMPC algorithm the nonlinearity is inverted and removed from the control problem, resulting in a linear MPC algorithm for the remaining linear block. The nonlinearity is used to specify the setpoint for the output of the linear dynamic block, therefore taking only the steady state effect of the nonlinearity into account. Transient effects such as an increased amplitude of an oscillation are not taken into account by the controller. The PWMPC does take these effects into account by transforming the nonlinearity into a polytopic description, resulting in a robust linear MPC algorithm for the remaining uncertain linear block. This offers the best performance and enables to balance the input-output behaviour over the entire operating region. This is demonstrated in the simulations.

In the paper no comparison is made with a linear model, identified using logarithmic compositions, a common approach to make the plant less nonlinear, proposed by [23]. Note that this in fact corresponds to a Wiener model in which the output nonlinearity is fixed a priori, and the resulting MPC corresponds to the IWMPC algorithm using a weighting on the logarithmic compositions. Note that this logarithmic transformation is application specific, and requires a priori knowledge about the static nonlinearity, in contrast to the presented Wiener model identification algorithm. If such a priori knowledge is available, it can be used to specify the structure of the nonlinearity in Eq. (8), e.g. incorporate exponential functions in $h(v)$. For the case study in this paper, incorporation of such knowledge was not necessary since the accuracy of the identified third order polynomials proved to be satisfactory, see Fig. 9.

References

[1] F. Allgöwer, T.A. Badgwell, J.S. Qin, J.B. Rawlings, S.J. Wright, Nonlinear predictive control and moving horizon estimation — an introductory overview, in: P.M. Frank (Ed.), Advances in Control, Highlights of ECC'99, Springer-Verlag, London, 1999, pp. 391–449.

[2] H.W. Andersen, M. Kümmel, Evaluating estimation of gain directionality. part 1: methodology, Journal of Process Control 2 (2) (1992) 59–66.

[3] H.W. Andersen, M. Kümmel, Evaluating estimation of gain directionality. part 2: a case study of binary distillation, Journal of Process Control 2 (2) (1992) 67–86.

[4] E.-W. Bai, Z. Ding, Zeros of sampled data systems represented by FIR models, Automatica 36 (1) (2000) 121–123.

[5] H.H.J. Bloemen, C.T. Chou, T.J.J. van den Boom, V. Verdult, M. Verhaegen, T.C. Backx, Wiener MPC for high purity dual composition control of a distillation column, in: Proceedings of Process Control and Instrumentation 2000, Glasgow, Scotland, July 2000, pp. 198–203.

[6] H.H.J. Bloemen, T.J.J. van den Boom, MPC for Wiener systems, in: Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, Arizona, USA, December 1999, pp. 4595–4600.

[7] H.H.J. Bloemen, T.J.J. van den Boom, H.B. Verbruggen, Model-based predictive control for Hammerstein-Wiener systems, International Journal of Control, in press.

[8] S. Boyd, L.O. Chua, Fading memory and the problem of approximating nonlinear operators with Volterra series, IEEE Transactions on circuits and systems CAS- 32 (11) (1985) 1150–1161.

[9] E.F. Camacho, C. Bordons, Model Predictive Control in the Process Industry, Springer-Verlag, London, 1995.

[10] C.T. Chou, H.H.J. Bloemen, V. Verdult, T.J.J. van den Boom, T. Backx, M. Verhaegen, Nonlinear identification of high purity distillation columns, in: Proceedings of SYSID 2000: Symposium on System Identification, Santa Barbara, California, June 2000.

[11] C.T. Chou, M. Verhaegen, An indirect approach to closed-loop identification of Wiener models, in: Proceedings of the American Control Conference, San Diego, California, 1999, pp. 3451–3455.

[12] D.W. Clarke, C. Mohtadi, P.S. Tuffs, Generalized predictive control — Part I. The basic algorithm, Automatica 23 (2) (1987) 137–148.

[13] D.W. Clarke, C. Mohtadi, P.S. Tuffs, Generalized predictive control — Part II. Extensions and interpretations, Automatica 23 (2) (1987) 149–160.

[14] G. Fieg, B. Landwehr, G. Wozny, About the possibility of a direct concentration control for distillation columns, Chemical Engineering and Technology 19 (1996) 299–307.

[15] C. Fuentes, W.L. Luyben, Control of high-purity distillation columns, Industrial and Engineering Chemistry, Process Design and Development 22 (3) (1983) 361–366.

[16] C.E. Garcia, D.M. Prett, M. Morari, Model predictive control: theory and practice — a survey, Automatica 25 (3) (1989) 335–348.

[17] A. Georgiou, C. Georgakis, W.L. Luyben, Nonlinear dynamic matrix control for high-purity distillation columns, AIChE Journal 34 (8) (1988) 1287–1298.

[18] M.A. Henson, A critical review of constraint compensation techniques for nonlinear process control, in: Proceedings of the European Control Conference, Brussels, Belgium, 1997.

[19] M.A. Henson, Nonlinear model predictive control: current status and future directions, Computers and Chemical Engineering 23 (1998) 187–202.

[20] C.W. Koung, J.F. MacGregor, Design of identification experiments for robust control. A geometric approach for bivariate processes, Industrial and Engineering Chemistry Research 32 (8) (1993) 1658–1666.

[21] R. Di Marco, D. Semino, A. Brambilla, From linear to nonlinear model predictive control: Comparison of different algorithms, Industrial and Engineering Chemistry Research 36 (5) (1997) 1708–1716.

[22] S.J. Norquay, A. Palazoglu, J.A. Romagnoli, Model predictive control based on Wiener models, Chemical Engineering Science 53 (1) (1998) 75–84.

[23] C.J. Ryskamp, Explicit versus implicit decoupling in distillation control, in: T.F. Edgar, D.E. Seborg (Eds.), Chemical Process Control II; Engineering Foundation Conference, United Engineering Trustees, New York, 1982, pp. 361–375.

[24] M.F. Sågfors, K.V. Waller, Multivariable control of ill-conditioned distillation columns utilizing process knowledge, Journal of Process Control 8 (3) (1998) 197–208.

[25] S. Skogestad, Dynamics and control of distillation columns: a tutorial introduction, Transactions of the Institution of Chemical Engineers, Part A, Chemical Engineering Research and Design 75 (1997) 539–562.

[26] S. Skogestad, M. Morari, LV-control of a high-purity distillation column, Chemical Engineering Science 43 (1) (1988) 33–48.

[27] S. Skogestad, M. Morari, Understanding the dynamic behavior of distillation columns, Industrial and Engineering Chemistry Research 27 (10) (1988) 1848–1862.

[28] G.R. Sriniwas, Y. Arkun, I.-L. Chien, B.A. Ogunnaike, Nonlinear identification and control of a high-purity distillation column: a case study, Journal of Process Control 5 (3) (1995) 149–162.

[29] M. Verhaegen, Identification of the temperature-product quality relationship in a multi-component distillation column, Chemical Engineering Communications 163 (1998) 111–132.

[30] M. Verhaegen, D. Westwick, Identifying MIMO Hammerstein systems in the context of subspace model identification methods, International Journal of Control 63 (2) (1996) 331–349.

[31] D. Westwick, M. Verhaegen, Identifying MIMO Wiener systems using subspace model identification methods, Signal Processing 52 (2) (1996) 235–258.