Computational performance of aggregated distillation models

Andreas Linhart, Sigurd Skogestad *

Department of Chemical Engineering, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

A B S T R A C T

Compartmental and aggregated modeling is used to derive low-order (reduced) dynamic models from detailed models of staged processes. In this study, the aggregated modeling method of [Lévine, J., & Rouchon, P. (1991). Quality control of binary distillation columns via nonlinear aggregated models. Automatica, 27, 463] is revised with the objective of deriving computationally efficient models for real-time control and optimization applications. A simple implementation of the original method not requiring the specification of compartments is presented. The resulting DAE models are converted into ODE models by pre-solving and substituting the algebraic equations resulting from the reduction procedure, which is the key step to increase simulation speed. To study this, the performances of several full and reduced distillation models, with and without base-layer controllers, are compared using different numerical integrators. It is found that while the reduced DAE models are computationally not advantageous, the reduced ODE models decrease the simulation time by a factor of 5–10.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

With the establishment of dynamic real-time optimization and model predictive control as state-of-the-art methods to efficiently operate industrial processes, reduced models with low computational complexity are in the focus of current research (Allgöwer & Zheng, 2000; Marquardt, 2001; van den Berg, 2005). In particular, reduced nonlinear physically based models are of high interest for the prediction of the system behavior over a wide range of operating conditions. Many model reduction techniques have been developed for nonlinear systems (Marquardt, 2001; van den Berg, 2005), most of which produce models of lower order. This, however, does not guarantee that the reduced models show a computationally better performance than the original models they were derived from (van den Berg, 2005). This is because a reduced model is most likely less accurate than the original full model, and because the numerical complexity of the full model is often retained in the equations of the reduced model.

For nonlinear model reduction of distillation columns, several model reduction and simplification methods have been developed in the past (Benallou, Seborg, & Mellichamp, 1986; Cho & Joseph, 1983; Khowinij, Bian, Henson, Belanger, & Megan, 2004; Khowinij,

Henson, Belanger, & Megan, 2005; Kienle, 2000; Kumar & Daoutidis, 2003; Lévine & Rouchon, 1991; Marquardt, 1990; Skogestad, 1997). Among these are the method of compartmental modeling (Benallou et al., 1986) and later the improved variant of aggregated modeling by Lévine and Rouchon (1991). The latter method is used for deriving the reduced models investigated in this study. It is based on partitioning a distillation column into "compartments" consisting of "steady-state" trays and dynamic "aggregation" trays, and using a singular perturbation argument (Kokotovic, Khalil, & O'Reilly, 1986) to derive a reduced-order model. Among its advantages is the perfect steady-state agreement with the original model, a simple derivation, and good control of the reduced model complexity.

Originally, these methods were intended for nonlinear controller design, for which a low-order model is necessary. More recently, they have been used to reduce the simulation time in real-time applications (Bian, Khowinij, Henson, Belanger, & Megan, 2005; Khowinij et al., 2004, 2005). However, it is shown in this study that while only transforming the original system into a reduced system in differential-algebraic equation (DAE) form does not improve the simulation speed of the reduced model, a subsequent elimination of the algebraic equations is necessary to obtain a reduced model in ordinary differential equation (ODE) form, which shows a significantly improved computational performance compared to the original model. On a more fundamental level, it is shown that the notion of compartments is not necessary in the derivation of the reduction method. This greatly simplifies the derivation and makes

* Corresponding author. Tel.: +47 73594154.
  E-mail address: skoge@ntnu.no (S. Skogestad).

the extension of the method to more complex systems straightforward.

The paper is organized as follows: In Section 2, the full model for a binary distillation column, and the derivation of reduced models from this using aggregated modeling is described. Important implementation details and properties of the models are given. In Section 3, the framework for testing the computational performance of the models is explained, discussing the input signal, the model parameters and the numerical solvers used for simulating the models. The results of the simulations are given in Section 4. Section 5 discusses the results of the simulations, and details and possible extensions of the reduction method. A summary and conclusions are given in Section 6.

2. Models

2.1. System and modeling assumptions

The system investigated is a binary distillation column with 72 trays plus reboiler and condenser. Two variants of this system are studied:

1. An "uncontrolled" column with level controllers for condenser and reboiler, but with no temperature or composition control. The reflux L and the boil-up V remain as degrees of freedom ("LV-configuration").
2. A "controlled" column with an additional composition controller in the lower column part that manipulates the boil-up rate V.

The controlled column is shown schematically in Fig. 1. In the following, the uncontrolled system is used to explain the system equations and the reduction procedure. Later, in Section 2.4, the inclusion of the composition controller is explained.

All assumptions made in this simplified distillation model are discussed in detail by Skogestad (1997). The major modeling assumptions are: Ideal trays, which means that liquid and vapor are in equilibrium at each tray; ideal mixture, which means that the vapor composition y can be expressed as a function of the liquid composition x assuming the constant relative volatility

$$y = k(x) = \frac{\alpha x}{1 + (\alpha - 1)x},\qquad(1)$$

where α is the relative volatility; constant molar flows, which means that the energy balance is simplified; constant molar holdup on each tray and negligible mass in the vapor phase.

The assumption of constant molar flows may not be good if the model is to be used for control purposes (Skogestad, 1997), but the focus here is on longer time scales.

The column has one feed flow F at tray number n_F. z_F denotes the concentration of the first (light) component in the feed. A liquid flow L (or $L + F$ for trays below the feed tray) and a vapor flow V enter and leave each tray. The condenser and reboiler levels are assumed to be controlled using the distillate flow D and bottom flow B, respectively. For simplicity, perfect level control is assumed, such that $D = V - L$ and $B = L + F - V$. Note that the assumption of perfect level control is not important with the LV-configuration. The concentrations in these flows determine the purity of the distillation products and are therefore the most important output variables in the process. The feed flow rate F and the feed concentration z_F can be seen as disturbance variables, and the flows L and V are manipulated variables for control.

2.2. Full uncontrolled model

The full model consists of one component material balance for each tray and the condenser and reboiler. For ease of notation, the condenser and reboiler are written as tray 1 and N:

$$H_1\dot{x}_1 = Vy_2 - Vx_1,\qquad(2)$$

$$H_i\dot{x}_i = Lx_{i-1} + Vy_{i+1} - Lx_i - Vy_i,\quad i = 2,\dots,n_F - 1\qquad(3)$$

$$H_{n_F}\dot{x}_{n_F} = Lx_{i-1} + Vy_{i+1} - (L+F)x_i - Vy_i + Fz_F,\qquad(4)$$

$$H_i\dot{x}_i = (L+F)x_{i-1} + Vy_{i+1} - (L+F)x_i - Vy_i,$$

$$i = n_F + 1,\dots,N-1\qquad(5)$$

$$H_N\dot{x}_N = (L+F)x_{N-1} - (L+F-V)x_N - Vy_N,\qquad(6)$$

where H_i is the total liquid molar holdup, x_i and $y_i = k(x_i)$ are the concentrations of the first component in the liquid and vapor phase, respectively, of tray i, N is the number of trays including the condenser and reboiler, n_F is the index of the feed tray, and V, L, F, z_F are as described above.
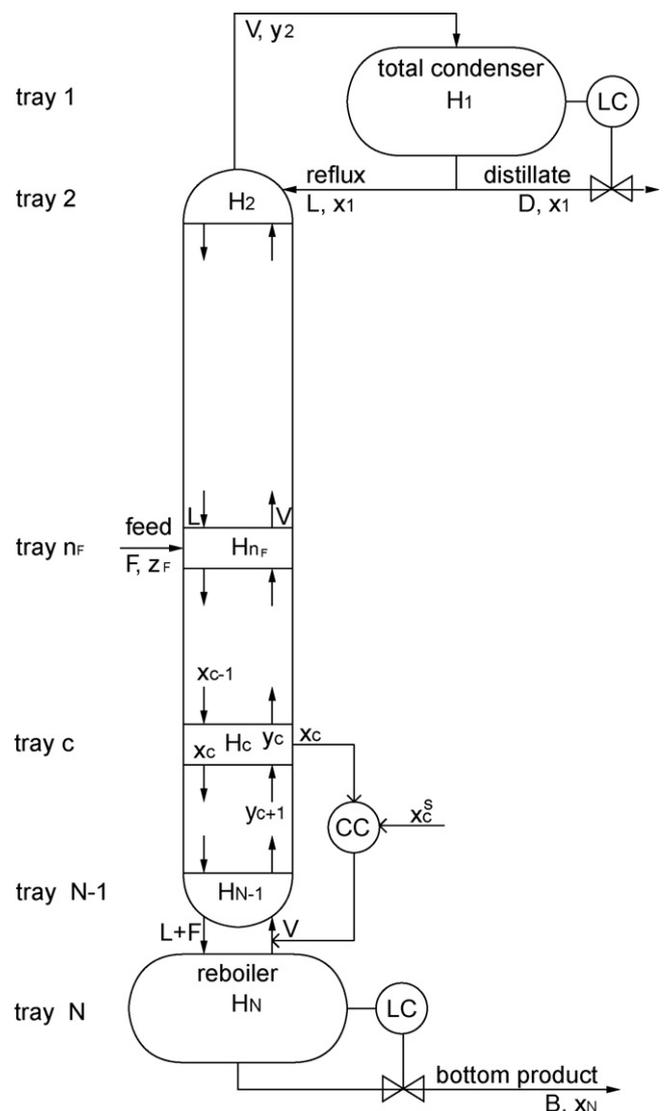

Fig. 1. Schematic diagram of a binary distillation column with a composition controller in the lower column section.

Table 1
Parameters of the full model.

Parameter	Value
N	74
n_F	36
H_1	20 mol
H_N	20 mol
$H_i, \quad i = 2, \ldots, N-1$	1 mol
F	0.04 mol/s
α	1.33
Input	**Nominal value**
z_F	0.45
L	0.12 mol/s
V	0.14 mol/s

The parameters for the system used in this study are given in Table 1. For the input variables, nominal values are given. The parameters are scaled average values of a real industrial distillation column, such that the dynamic behavior and the time constants are similar. SI units (mol, mol/s, s, etc.) are used for all variables.

2.3. Reduced models

The reduced models investigated in this study are derived via the aggregated modeling method of Lévine and Rouchon (1991). The method is based on partitioning the column into N_c "compartments", where each compartment comprises a number of consecutive trays. By assigning the total holdup of all trays of a compartment to a single "aggregation" tray within the compartment, a time-scale separation caused by the difference between the large compartment holdup and the small tray holdups is made explicit in the model. This time-scale separation can be used for reducing the dynamic order of the model by applying quasi-steady-state assumptions to the trays within each compartment.

In their derivation of the method, Lévine and Rouchon (1991) perform one uncommented simplification step, which results from a mathematically inaccurate treatment of a term when applying the quasi-steady-state assumption. In Remark 3 in their paper, they note that the reduction procedure can be described much simpler. This simplification step has the major consequence for the resulting reduced model that only the compartment holdups, but not their boundaries appear in the model equations. Due to this simplification, the reduction procedure is not a true singular perturbation method (Kokotovic et al., 1986). However, the structure of the reduced model, which retains the tridiagonal Jacobian structure of the original model, is numerically favorable. Extensions to more complex models, as will be discussed in Section 5, are also much more straightforward. In the following, the derivation of the method is therefore described in some detail. First, the original derivation of Lévine and Rouchon is explained for one compartment. Then, the alternative, much simpler derivation is explained, and the relationship to the original derivation is shown. A "true" singular perturbation model, resulting from the derivation without the simplification step, is then compared to the reduced model of Lévine and Rouchon.

2.3.1. Original derivation of reduced compartment model by Lévine and Rouchon

The following derivation considers a certain compartment j, which is formed from a number m of consecutive trays (in this case not including a feed). For simplicity of notation, the numbering of the trays starts with 1 and ends with m. It is assumed here for simplicity that the holdups H of all individual trays are equal. Then, $H = (1/m)\bar{H}$, where $\bar{H}$ is the total compartment holdup. The

original model is

$$\frac{1}{m}\bar{H}\dot{x}_1 = Lx_0 + Vk(x_2) - Lx_1 - Vk(x_1), \tag{7}$$

$$\frac{1}{m}\bar{H}\dot{x}_2 = Lx_1 + Vk(x_3) - Lx_2 - Vk(x_2), \tag{8}$$

$$\ldots$$

$$\frac{1}{m}\bar{H}\dot{x}_m = Lx_{m-1} + Vk(x_{m+1}) - Lx_m - Vk(x_m), \tag{9}$$

with the same notation as in Eqs. (1)–(6).

A coordinate transformation is applied by introducing the average concentration of the compartment $\bar{x}_s$, where s is the index of the "aggregation tray" of the compartment, and leaving the remaining $m-1$ variables unchanged:

$$\bar{x}_s = \frac{1}{m}\sum_{i=1}^{m} x_i, \tag{10}$$

$$\bar{x}_i = x_i, \quad i = 1, \ldots, m, \quad i \neq s. \tag{11}$$

The transformed system now reads

$$\frac{1}{m}\bar{H}\dot{\bar{x}}_1 = Lx_0 + Vk(\bar{x}_2) - L\bar{x}_1 - Vk(\bar{x}_1), \tag{12}$$

$$\frac{1}{m}\bar{H}\dot{\bar{x}}_2 = L\bar{x}_1 + Vk(\bar{x}_3) - L\bar{x}_2 - Vk(\bar{x}_2), \tag{13}$$

$$\ldots$$

$$\frac{1}{m}\bar{H}\dot{\bar{x}}_{s-1} = L\bar{x}_{s-2} + Vk\left(m\bar{x}_s - \sum_{i \neq s}\bar{x}_i\right) - L\bar{x}_{s-1} - Vk(\bar{x}_{s-1}), \tag{14}$$

$$\bar{H}\dot{\bar{x}}_s = Lx_0 + Vk(x_{m+1}) - L\bar{x}_m - Vk(\bar{x}_1), \tag{15}$$

$$\frac{1}{m}\bar{H}\dot{\bar{x}}_{s+1} = L\left(m\bar{x}_s - \sum_{i \neq s}\bar{x}_i\right) + Vk(\bar{x}_{s+2}) - L\bar{x}_{s+1} - Vk(\bar{x}_{s+1}), \tag{16}$$

$$\ldots$$

$$\frac{1}{m}\bar{H}\dot{\bar{x}}_m = L\bar{x}_{m-1} + Vk(x_{m+1}) - L\bar{x}_m - Vk(\bar{x}_m). \tag{17}$$

This system is now in the standard form of singular perturbation (Kokotovic et al., 1986) with $\varepsilon := (1/m)$ as a small parameter, and a reduced model can be derived by making the quasi-steady-state assumption $\varepsilon \to 0$. In this way, all trays except for the aggregation tray are converted into "steady-state" trays, and are described by algebraic equations.

The term $m\bar{x}_s - \sum_{i \neq s}\bar{x}_i$ that appears in Eqs. (14) and (16) is replaced by $\bar{x}_s$ as a result of the quasi-steady-state assumption $\varepsilon \to 0$. Lévine and Rouchon (1991) use a slightly more complex formulation that admits nonuniform tray holdups. The derived reduced system then reads

$$0 = Lx_0 + Vk(\bar{x}_2) - L\bar{x}_1 - Vk(\bar{x}_1), \tag{18}$$

$$\ldots$$

$$0 = L\bar{x}_{s-2} + Vk(\bar{x}_s) - L\bar{x}_{s-1} - Vk(\bar{x}_{s-1}), \tag{19}$$

$$\bar{H}\dot{\bar{x}}_s = Lx_0 + Vk(x_{m+1}) - L\bar{x}_m - Vk(\bar{x}_1), \tag{20}$$

$$0 = L\bar{x}_s + Vk(\bar{x}_{s+2}) - L\bar{x}_{s+1} - Vk(\bar{x}_{s+1}), \tag{21}$$

$$\ldots$$

$$0 = L\bar{x}_{m-1} + Vk(x_{m+1}) - L\bar{x}_m - Vk(\bar{x}_m). \tag{22}$$

2.3.2. Simplified derivation

Lévine and Rouchon observe that the derivation of a reduced compartment model can be done much simpler in less mathematical form: The trays $i \neq s$ have no holdup ($H_i \to 0$), and the aggregation tray s has the compartment holdup ($H_s = \sum_{i=1}^{m} H_i$).

This can be seen from Eqs. (18) to (22) by subtracting all remaining equations from Eq. (20). Then, Eq. (20) can be replaced by equation

$$\bar{H}\dot{\bar{x}}_s = L\bar{x}_{s-1} + Vk(\bar{x}_{s+1}) - L\bar{x}_s - Vk(\bar{x}_s). \tag{23}$$

This is the same as shifting the compartment boundaries to the aggregation trays, as illustrated in Fig. 2.

The whole model reduction procedure can therefore be simplified to multiplying the left-hand sides of the aggregation tray equations with a constant $\gg 1$, and setting the remaining left-hand sides to 0. As it will be explained later, the scalar multiplying the left-hand side of the aggregation tray equation can be chosen arbitrarily without changing the steady-state of the model. Because of this and the fact that the compartment boundaries do not appear in the model anymore, the notion of "compartments" is not needed in the derivation of the method. It is sufficient to select some aggregation trays and assign them a large holdup.

Fig. 3. Responses of the top concentration of the full model, the model of Lévine and Rouchon (dashed line), and the "true" singular perturbation model (dash-dotted line) with five compartments to a step change in z_F from 0.45 to 0.55.

This is somewhat surprising, since the original derivation starts with the transformation (10) and (11), where the compartment boundaries are clearly specified (here trays 1 and m). The reason that these compartment boundaries dissappear during the derivation is the treatment of the term $m\bar{x}_s - \sum_{i \neq s} \bar{x}_i$ in Eqs. (14) and (16) while performing the quasi-steady-state approximation $\varepsilon \to 0$. Mathematically, since $m\bar{x}_s - \sum_{i \neq s} \bar{x}_i = x_s$, this term is invariant under the operation $\varepsilon \to 0$. If this term is left unchanged, the resulting model differs from the above model in Eqs. (19) and (21), which then read

$$0 = L\bar{x}_{s-2} + Vk\left(m\bar{x}_s - \sum_{i \neq s} \bar{x}_i\right) - L\bar{x}_{s-1} - Vk(\bar{x}_{s-1}) \tag{24}$$

and

$$0 = L\left(m\bar{x}_s - \sum_{i \neq s} \bar{x}_i\right) + Vk(\bar{x}_{s+2}) - L\bar{x}_{s+1} - Vk(\bar{x}_{s+1}). \tag{25}$$

In this model, the compartment boundaries appear explicitly in the summation terms. This model variant is a "true" singular perturbation model in a sense that first the model is transformed into the standard form of singular perturbation, and then the quasi-steady-state assumption is applied. In the derivation of Lévine and Rouchon, an additional step simplifying the transformed model is taken. This, however, has major advantages: first, the derivation of the reduced model is greatly simplified, as described above. Secondly, the resulting reduced model retains the tridiagonal structure of the Jacobian of the original model. This is advantageous from a numerical point of view, and allows for the elimination of the algebraic equations from the model as will be described in Section 2.3.4.

Fig. 3 shows the response of both models to a step change in the feed concentration z_F. The "true" singular perturbation model shows an inverse response which is not present in the response of the full model. Except for this, the approximation quality of the "true" singular perturbation model is better than that of the model of Lévine and Rouchon. Both models yield the same steady-state values of the output variables. This is obvious from the facts that the right-hand side of the model of Lévine and Rouchon coincides with the right-hand side of the full model, and that the "true"
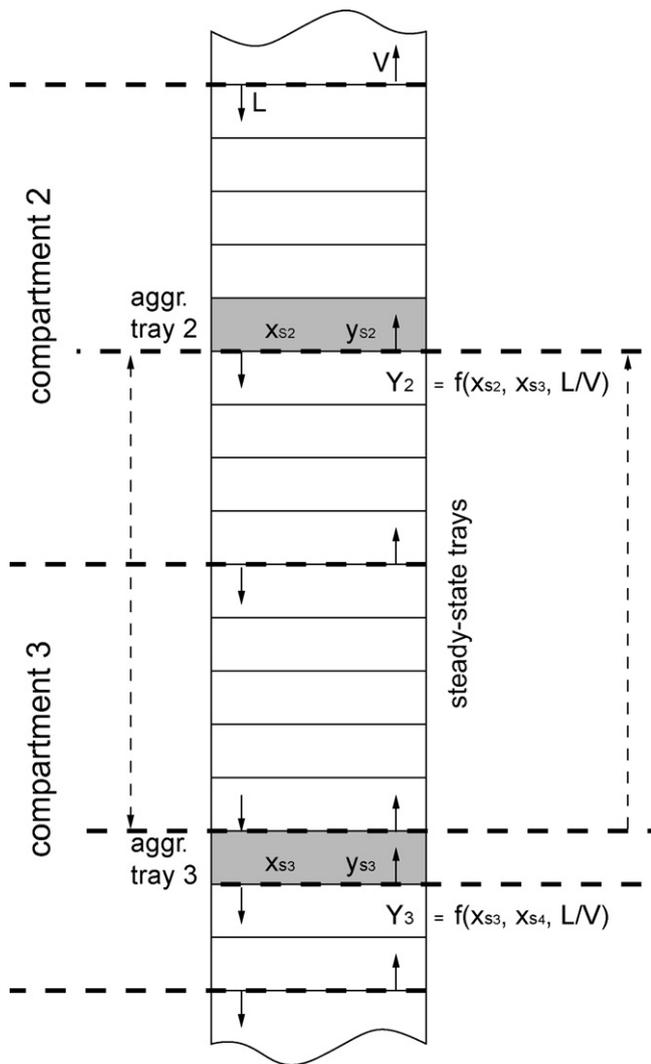

Fig. 2. Schematic diagram of two compartments. Since the trays between two aggregation trays are in steady-state, balance boundaries can be moved arbitrarily between the aggregation trays without changing the balances.
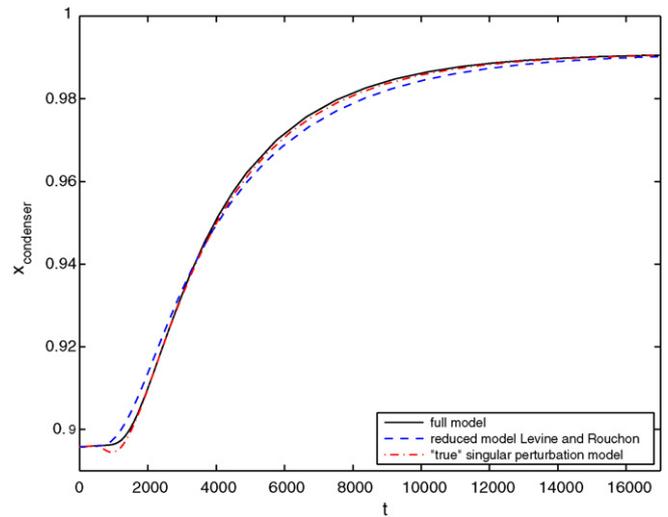
singular perturbation model results from a variable transformation that leaves the top and bottom variables unchanged.

Because of the favorable structure of the model of Lévine and Rouchon, and the wrong inverse response of the "true" singular perturbation model, only the reduced model as derived by Lévine and Rouchon will be investigated in the rest of this paper. Instead of the term "compartment", the term "aggregation tray" is used, since the method is basically "compartment-free", as explained above. Although the term "stage" is more general than "tray", the term "aggregation tray" is used throughout the paper to stay consistent with the terminology used by Lévine and Rouchon (1991).

2.3.3. Reduced model in DAE form

In this study, reduced models with between three and seven aggregation trays are used. The reboiler and condenser are always chosen as aggregation trays due to their large holdups. Applying the reduction method described above, the resulting DAE system for a reduced model with five aggregation trays reads

$$\bar{H}_1\dot{\bar{x}}_1 = Vk(\bar{x}_2) - V\bar{x}_1, \tag{26}$$

$$0 = L\bar{x}_{i-1} + Vk(\bar{x}_{i+1}) - L\bar{x}_i - Vk(\bar{x}_i), \quad i = 2, \ldots, s_2 - 1, \tag{27}$$

$$\bar{H}_2\dot{\bar{x}}_{s_2} = L\bar{x}_{s_2-1} + Vk(\bar{x}_{s_2+1}) - L\bar{x}_{s_2} - Vk(\bar{x}_{s_2}), \tag{28}$$

$$0 = L\bar{x}_{i-1} + Vk(\bar{x}_{i+1}) - L\bar{x}_i - Vk(\bar{x}_i), \quad i = s_2 + 1, \ldots, s_3 - 1, \tag{29}$$

$$\bar{H}_3\dot{\bar{x}}_{s_3} = L\bar{x}_{s_3-1} + Vk(\bar{x}_{s_3+1}) - (L+F)\bar{x}_{s_3} - Vk(\bar{x}_{s_3}) + Fz_F, \tag{30}$$

$$0 = (L+F)\bar{x}_{i-1} + Vk(\bar{x}_{i+1}) - (L+F)\bar{x}_i - Vk(\bar{x}_i),$$
$$i = s_3 + 1, \ldots, s_4 - 1, \tag{31}$$

$$\bar{H}_4\dot{\bar{x}}_{s_4} = (L+F)\bar{x}_{s_4-1} + Vk(\bar{x}_{s_4+1}) - (L+F)\bar{x}_{s_4} - Vk(\bar{x}_{s_4}), \tag{32}$$

$$0 = (L+F)\bar{x}_{i-1} + Vk(\bar{x}_{i+1}) - (L+F)\bar{x}_i - Vk(\bar{x}_i),$$
$$i = s_4 + 1, \ldots, N - 1, \tag{33}$$

$$\bar{H}_5\dot{\bar{x}}_N = (L+F)\bar{x}_{N-1} - (L+F-V)\bar{x}_N - Vy_N, \tag{34}$$

where $\bar{H}_j$ and s_j are the aggregated holdup and the index of aggregation tray j, respectively. The $\bar{x}$ notation is used to distinguish the variables in Eqs. (26)–(34), which are partially algebraic and partially dynamic, from the purely dynamic variables in the full model (2)–(6).

2.3.4. Reduced model in ODE form

The equations of the steady-state trays between two aggregation trays contain only the algebraic variables of the steady-state trays, the state variables of the two aggregation trays, and either the ratio L/V or $(L+F)/V$. By writing the mass balance around each aggregation tray and all "steady-state" trays above it, the reduced model can be formulated such that the only algebraic variables appearing in the dynamic equations are the vapor concentrations of the trays directly below each aggregation tray, i.e. $y_{s_j+1}, j = 1, \ldots, N_c$ (Fig. 2). By pre-solving the systems of algebraic equations for these variables as a function of the dynamic variables and the ratio $V/L, y_{s_j+1} = Y_j(x_{s_j}, x_{s_j+1}, V/L)$, and by substituting them into Eqs. (26)–(34), the DAE system can be cast in ODE form as

$$\bar{H}_1\dot{\bar{x}}_1 = VY_1(\bar{x}_1, \bar{x}_2, V/L) - V\bar{x}_1, \tag{35}$$

$$\bar{H}_2\dot{\bar{x}}_2 = L\bar{x}_1 + VY_2(\bar{x}_2, \bar{x}_3, V/L) - L\bar{x}_2 - VY_1(\bar{x}_1, \bar{x}_2, V/L), \tag{36}$$

$$\bar{H}_3\dot{\bar{x}}_3 = L\bar{x}_2 + VY_3(\bar{x}_3, \bar{x}_4, V/(L+F)) - (L+F)\bar{x}_3 - VY_2(\bar{x}_2, \bar{x}_3, V/L) + Fz_F, \tag{37}$$

$$\bar{H}_4\dot{\bar{x}}_4 = (L+F)\bar{x}_3 + VY_4(\bar{x}_4, \bar{x}_5, V/(L+F)) - (L+F)\bar{x}_4 - VY_3(\bar{x}_3, \bar{x}_4, V/(L+F)), \tag{38}$$

$$\bar{H}_5\dot{\bar{x}}_5 = (L+F)\bar{x}_4 - (L+F-V)\bar{x}_5 - VY_4(\bar{x}_4, \bar{x}_5, V/(L+F)). \tag{39}$$

The $\bar{x}$ notation is used to distinguish the variables in Eqs. (35)–(39) from the variables in the full model (2)–(6). The functions $Y_j, j = 1, \ldots, N_c$ can be calculated off-line and stored and later retrieved in a suitable way. In this study, look-up tables as described in Section 2.7 are used for this purpose.

2.4. Model with composition controller

To stabilize the composition profile, it is common in distillation control to use the boil-up or reflux rate to control the temperature at a certain location inside the column. The setpoint can serve as a manipulated variable for a higher control or optimization layer. In this study, composition is controlled instead of temperature, which is equivalent for a binary mixture.

If the boil-up rate is used, a proportional-integral control law for the concentration at the controlled tray c reads

$$V(t) = K_c(x_c(t) - x_c^s(t) + x_I(t)/T_I), \tag{40}$$

$$x_I(t) = \int_{-\infty}^t (x_c(\tau) - x_c^s(\tau))\,d\tau, \tag{41}$$

where K_c is the gain of the controller, x_c^s is the composition setpoint, x_I is the state of the integral part of the controller, and T_I is the integral time. This control law can be used directly in the full model Eqs. (2)–(6) and the reduced models Eqs. (26)–(34) and (35)–(39).

In this study, $K_c = 0.5$ and $T_I = 400$ s were chosen using the Skogestad IMC tuning rules for PI-controllers (Skogestad, 2003) to achieve a closed-loop response of the controlled concentration with a time-constant of 100 s, assuming an effective delay of 0 s for the boil-up.

The reduced model can be further simplified when perfect composition control, that is

$$x_c = x_c^s, \tag{42}$$

where the setpoint is (piecewise) constant, $\dot{x}_c = \dot{x}_c^s = 0$, is assumed. Then, the equilibrium relation

$$0 = (L+F)x_{c-1} + Vk(x_{c+1}) - (L+F)x_c^s - Vk(x_c^s), \tag{43}$$

$$\Leftrightarrow V/(L+F) = \frac{(x_c^s - x_{c-1})}{k(x_{c+1}) - k(x_c^s)} \tag{44}$$

can be derived. Here, the liquid flow is $(L+F)$, because the controlled tray is assumed to be in in the lower column section below the feed.

Eqs. (42) and (44) can be used in the reduced DAE model (26)–(34) to eliminate one concentration and either the reflux rate L or boil-up rate V. If the controlled tray is chosen as an aggregation tray, its concentration variable will change from dynamic to algebraic, thereby reducing the dynamic model order by one. Compared to the corresponding PI-controlled reduced model, the dynamic order is reduced by two, since the dynamic I-part of the controller is not needed.

The model can be converted again into ODE form as described in Section 2.3.4 by solving the algebraic equations of the model off-line. For the part of the model containing the perfectly controlled tray, a separate table containing

$$Y_c = Y_c(x_k, x_N, x_c^s) \tag{45}$$

and

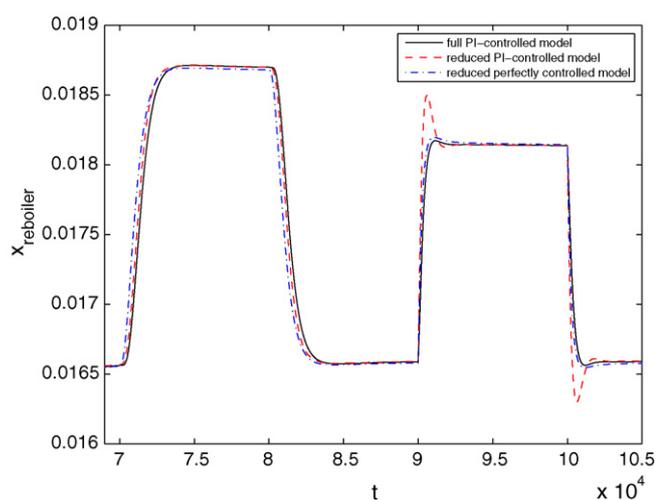$$V/(L+F) = f(x_k, x_N, x_c^s), \tag{46}$$

Fig. 4. Responses of the reboiler concentration of the full PI-controlled model (full line), a reduced PI-controlled model (dashed line) and a perfectly controlled model (dashed-dotted line), with seven and six aggregation trays, respectively, to a step change in z_F from 0.35 to 0.3 and back (left part of the plot), and in x_c^s from 0.105 to 0.1125 and back (right part of the plot). While the PI-controlled model shows a more accurate response to the change in z_F than the perfectly controlled model, it displays a too large overshoot as response to the change in x_c^s.

where k is the index of the aggregation tray above the controlled aggregation tray, is needed.

Fig. 4 shows the responses of the PI-controlled model and the perfectly controlled model, both with seven equally sized compartments, to step changes in z_F and in x_c^s. While the PI-controlled model is closer to the full model response for the step change in z_F, it shows a fast response with large overshoot and succeeding oscillations for the setpoint change in x_c^s. Note that the PI-controller was not retuned, leading to a much faster response of the reduced system because of the neglected dynamics.

2.5. Stiffness analysis

An approximate stiffness analysis is done by linearising the uncontrolled full and a five-aggregation tray model around a steady-state for a constant z_F, and by comparing the largest and smallest eigenvalue λ of the linearised systems. Fig. 5 shows the
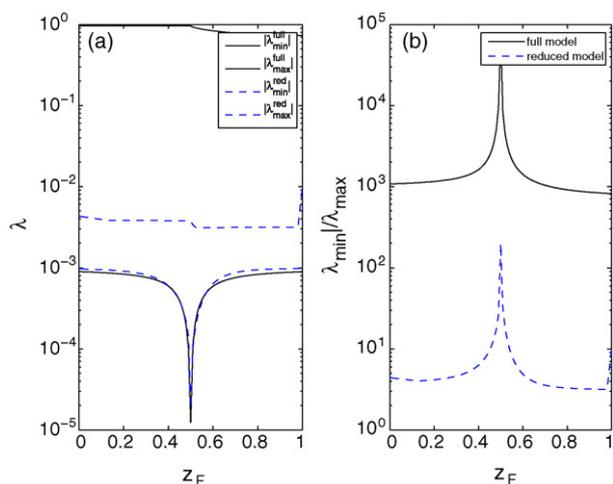

Fig. 5. Plots of the absolute values of the largest and smallest eigenvalues (plot a) of the uncontrolled full and a five-aggregation tray reduced model linearised around a constant z_F, and of their ratios (plot b).

largest and smallest eigenvalues (plot a), and their ratio (plot b), of the full and the reduced model for the full range $0 \leq z_F \leq 1$.

It can be seen that the full model is quite stiff with a maximal eigenvalue ratio (condition number) at $z_F = 0.5$ of 7.7×10^4. The reduced system does not contain the fast modes of the full system, which is indicated by the fact that the absolute values of the smallest eigenvalues of the full system are about 300 times larger than of the reduced system. The maximal eigenvalue ratio of the reduced system is therefore reduced to 198, which indicates that there is still some stiffness present in the system. This remaining stiffness is created by a sharp minimum for the smallest eigenvalue at $z_F = 0.5$, which is present in both full and reduced system. This phenomenon is described by Skogestad and Morari (1987). However, this slowest time-scale is actually not interesting for control, because a temperature or composition controller will cause the system to move to steady-state much faster. Therefore, for the numerical performance only the fastest eigenvalue, which dictates the step size of the integration, is of interest.

2.6. Choice of reduced model parameters

To the knowledge of the authors, no systematic procedure to determine compartment parameters that are optimal for a certain application is available. Khowinij et al. (2005) report that the accuracy of the compartmental models is not strongly affected by the location of the aggregation trays. They place the aggregation trays at the middle of the respective compartments. In order to investigate the influence of the reduced model parameter choice on the accuracy of the reduced model, a straightforward choice and an optimized choice of parameters are compared in this study.

It is worth noticing that although requiring that the sum of the aggregation tray holdups is equal to the sum of the tray plus condenser and reboiler holdups is physically meaningful, it is not necessary for obtaining dynamically functional models. Any choice of the aggregation tray holdups (that leads to a stable model) will yield a model that approaches the same steady-state as the full model. Thus the holdups can be used as degrees of freedom to optimize the dynamic behavior of the reduced model. In this study, however, models are restricted to the "physical" choice of aggregation tray holdups such that the reduced column model has the same holdup as the full column model.

For the study of the uncontrolled model, two versions of models with five and seven aggregation trays are compared: a straightforward version with equally sized and spaced aggregation trays, and an optimized version, where the parameters are determined by an optimization procedure to minimize the reduced model error.

In all models, the feed tray was chosen to be an aggregation tray. In this way, the input appears only in the dynamic part of the model, but not in the algebraic equations.

With this choice, the three-aggregation tray model is completely determined, because there is no freedom of choice for the aggregation tray parameters: condenser, reboiler, and the feed tray are the aggregation trays. For the study of the controlled model, the location of the controlled tray was fixed to be 62, and the last aggregation tray before the reboiler was positioned there as well, allowing a straightforward implementation of the controller. The parameter values for the five- and seven-aggregation tray models used in this study are given in Table 2.

2.7. Implementation details of reduced models

The steady-state functions $Y_j(x_{s_j}, x_{s_{j+1}}, V/L)$ were calculated offline and stored in a look-up table with $30 \times 40 \times 200$ entries, and the functions $Y_c(x_{s_l}, x_{reboiler}, x_{i_c}^s)$ and $V/(L+F) = f(x_{s_l}, x_{reboiler}, x_{i_c}^s)$ in tables with $60 \times 20 \times 200$ entries. From the tables, function

Table 2
Positions and holdups of the aggregation trays of the reduced models.

| Aggregation tray index j	1	2	3	4	5	6	7
Uncontrolled models							
Three aggregation trays							
s_j	1	36	74				
$\bar{H}$	20	72	20				
Optimized							
s_j	1	14	36	60	74		
$\bar{H}$	20	21	28	23	20		
s_j	1	8	20	36	53	67	74
$\bar{H}$	20	10	15	19	18	10	20
Equally sized							
s_j	1	13	36	62	74		
$\bar{H}$	20	24	24	24	20		
s_j	1	8	22	36	53	67	74
$\bar{H}$	20	14	14	16	14	14	20
Controlled models							
Optimized							
s_j	1	13	36	62	74		
$\bar{H}$	20	21	45	10	20		
Equally sized							
s_j	1	13	36	62	74		
$\bar{H}$	20	24	24	24	20		
s_j	1	8	22	36	53	62	74
$\bar{H}$	20	14	14	16	14	14	20

Fig. 6. Step input signals used for simulating the uncontrolled models and output trajectories of the full model to this input. The stepwise changing curves are the feed concentration z_F, the reflux L and the boil-up rage V. V is always 0.02 higher that L. The curves on top and bottom are the concentrations of condenser and reboiler, respectively.

values are obtained by applying three-dimensional interpolation. In this study, multi-linear and multi-cubic interpolation was used (Press, Teukolsky, Vetterling, & Flannery, 2007). Investigations on the table interpolation accuracy showed deviations in the fourth digit (less than 1%) for linear interpolation and in the sixth digit (less than 0.01%) for cubic interpolation. The table formats are the outcome of an optimization procedure, where the minimum average interpolation error was identified among possible entry number combinations all yielding the same table size. The obtained optimal combinations reflect the fact that the function shows stronger curvature in some dimensions than in others. For example, the function $Y_j(x_{s_j}, x_{s_{j+1}}, V/L)$ is almost linear in the first variable, because of the dominant influence of the liquid concentration from above on the function value.

3. Framework for evaluation of reduced model performance

The computational performance of the different models was assessed by multiple simulation runs, where the models were simulated with different solvers at varying simulation tolerances. The simulation time for a given model and solver depends on the simulation tolerance, which also influences the simulation error compared to the exact solution (Section 3.3.2). The performance of the models can therefore only be compared by considering both simulation time and simulation accuracy.

In this study, only the open-loop performance of the various models was assessed. This means that the input signals (z_F, L and V for the uncontrolled column; or z_F, L and x_c^s for the controlled column) to the system were fixed at predetermined profiles during the simulations.

The input profile was selected with the aim of exposing the system to step changes in the disturbance and control variables over a wide range of operating conditions. The simulation accuracy was measured using the average two-norm of the output deviations from the exact solution.

3.1. Selection of input profile

The input profiles for z_F, L and V to test the performance of the uncontrolled models are shown in Fig. 6. They include changes
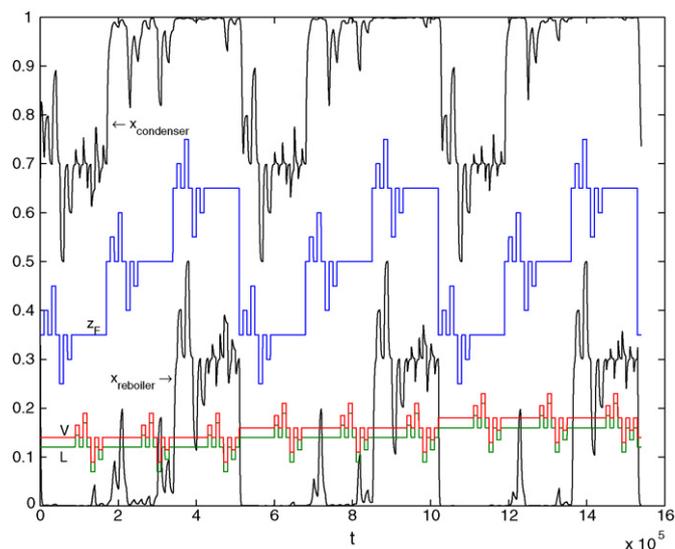
of two different magnitudes in the feed concentration z_F and the reflux rate L and vapor flow V. The two latter are changed simultaneously, such that their difference is constantly half of the feed flow F. This kind of simultaneous change is likely to appear in closed-loop optimal control of a high-purity column, because then the split

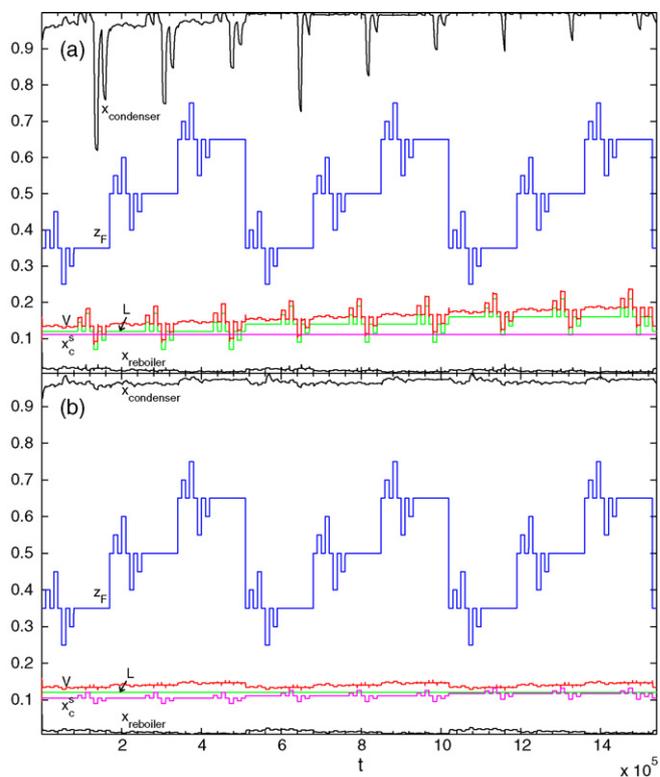**Fig. 7.** Step input signals used for simulating the *controlled* models and output trajectories of the full model to this input. Plot (a) shows the input signal where z_F and V is varied, and plot (b) shows the input signal where z_F and x_c^s is varied. The curves on top and bottom of both plots are the concentrations of condenser and reboiler, respectively. In both cases, V is manipulated by the composition controller.

is approximately constant (Shinskey, 1984). Nevertheless, we see from the simulations that the variations in product compositions (x_D, x_B) are large, because there is no compensation for changes in the feed composition (z_F). In the case of the controlled model, the same input signal was used, but the boil-up rate V was determined by the composition controller. The variation in product composition is here much less (Fig. 7). The time interval between changes in each variable was chosen to 10^4 s (about 3 h). To take into account fast and slowly varying input signals, both stepwise changes and slowly changing cubic B-spline signals (Schlegel, 2005) have been used. The cubic B-splines yield a continuously differentiable input signal, where the rate of change form one stationary value u_t to the next u_{t+1} is determined by a time constant δ. The total transition time is 2δ with a maximum slope of $(u_{t+1} - u_t)/\delta$. A value of 3000 s was chosen for δ.

The input profiles for z_F, L and x_c^s for the controlled column are shown in Fig. 7. Since the setpoint of the composition controller can be used as an additional manipulated variable for control, two sets of input profiles, one with step changes in z_F and L (plot a), and one with step changes in z_F and x_c^s (plot b) were used.

3.2. Selection of accuracy measure

As a measure for model accuracy, the condenser and reboiler concentrations, which are the most important variables for control, are evaluated by taking the average of the 2-norm of the deviations from the exact solution (see Section 3.3.2) over the whole simulation period:

$$\varepsilon = \frac{1}{t_{end}} \int_0^{t_{end}} \sqrt{\left(x_{condenser}^{exact} - x_{condenser}\right)^2 + \left(x_{reboiler}^{exact} - x_{reboiler}\right)^2} \, dt, \tag{47}$$

where $x_{condenser}$ and $x_{reboiler}$ are the condenser and reboiler concentrations, respectively, of either the full or the reduced model, and of the exact solution.

3.3. Simulation details

3.3.1. Solvers

In this study, the following four solvers have been used to simulate the full model and the different versions of the reduced models:

- *LIMEX*: This solver is an extrapolation scheme for the solution of linearly implicit differential-algebraic systems (Ehrig & Nowak, 2000; Hairer & Wanner, 2002). In this study, version 4.2A by Ehrig and Nowak (2000) for systems with dense or banded Jacobian was used. With this solver, all full and reduced uncontrolled model versions have been simulated. This solver is attractive for the use in real-time optimization, because of its good restarting qualities for problems involving frequent discontinuities (Brenan, Campell, & Petzold, 1996). It is used, for example, in the real-time optimization software DyOS (Schlegel, Stockmann, Binder, & Marquardt, 2005; Würth, Linhart, Preisig, & Marquardt, 2007).
- *RADAU*: This solver is an implicit Runge–Kutta scheme with coefficients based on Radau IIA quadrature formulas (Hairer & Wanner, 2002). The code used in this study is the classic fifth-order code RADAU5 by Hairer and Wanner (1996). It is suitable to solve stiff equations and was used to integrate the full model and the reduced models in ODE form.
- *DOPRI*: This solver is an embedded explicit Runge–Kutta scheme based on the Dormand–Prince pair for step control (Hairer, Nørsett, & Wanner, 2000). The code used in this study is the classic fifth-order code DOPRI5 by Hairer and Wanner (1996). It is a solver for non-stiff equations and was used to integrate

the reduced models in ODE form. This is possible because of the reduced stiffness of the reduced models. The same integration scheme is used in the MATLAB solver ode45 (Shampine & Reichelt, 1997). The higher-order code DOPRI853 has also been tested, but showed comparable performance to the fifth-order solver and was therefore not included into the detailed study.
- *DASPK*: This solver is an implicit multistep integrator based on the backward differentiation formula (Ascher & Petzold, 1998). Version 3.0 by Li and Petzold (2000) was used in this study. It was used to integrate the full and the reduced controlled models in this study. A similar integration scheme is used in the MATLAB solver ode15s (Shampine & Reichelt, 1997).

All solvers provide dense output, that is, interpolated solution values can be obtained at any desired point within the simulation period (Hairer et al., 2000). The LIMEX and DASPK codes were restarted after every change in the input signal. The RADAU and DOPRI codes were only started once at the beginning of a simulation, because restarting did not have any effect on their execution time.

The numerical properties of the controlled and the uncontrolled models are significantly different. The reason for this is a change of the Jacobian structure, which is caused by the composition controller. Fig. 8 shows the Jacobian structure of the full controlled model. The two vertical columns of entries on the right side of the plot reflect the influence of the state of the controlled tray (left column) and the I-state of the PI-controller (right column) on the equations of each tray via the vapor flow V. In contrast to this, the Jacobian of the uncontrolled model is purely tridiagonal because of the absence of the composition controller. This has important numerical consequences. The Jacobian of the uncontrolled models and the solution of the linear equations arising from the implicit integration schemes using this Jacobian can be done numerically efficient by exploiting the tridiagonally banded structure of the Jacobian. The Jacobian of the controlled models is, however, no longer banded. This implies that the calculation of a numerical Jacobian and the linear algebra is much slower. To avoid this loss in performance, the Jacobian for the controlled models was calculated analytically, and a specialized solver for the linear algebra which takes into account both banded structure and vertical lines in the Jacobian was used. To simulate the controlled systems, the
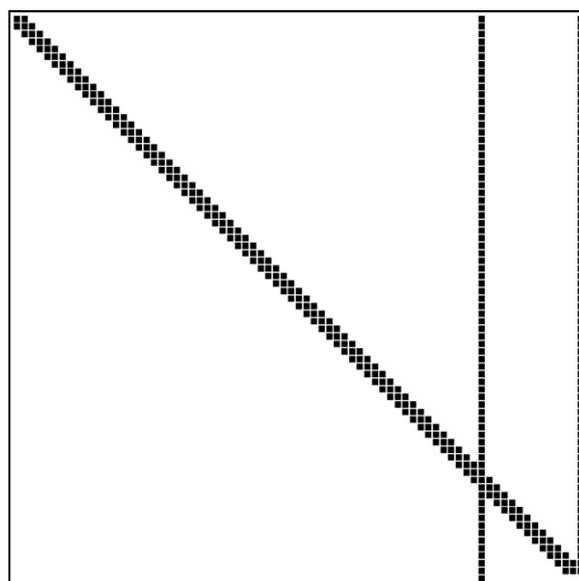

Fig. 8. Plot of the Jacobian structure of the full controlled model.

solver DASPK (Li & Petzold, 2000) was used, since this solver allows for the straightforward use of external linear algebra routines.

For all solvers, integration tolerances *AbsTol = RelTol* in the range 10^{-1} to 10^{-5} were used.

3.3.2. Calculation of "exact" solution

As "exact" solutions, simulations with very tight simulation tolerances (10^{-13}) were performed, using the RADAU code for the uncontrolled models and the DASPK code for the controlled models.

4. Results

4.1. Accuracy vs. simulation time

4.1.1. Uncontrolled models and general observations

Figs. 9 and 10 show the performance of the uncontrolled full and different optimized reduced models with linear table interpolation in combination with different solvers, simulated at varying tolerances, for the "slow" (cubic spline) and the "fast" (step) input signals, respectively. For better readability, the plots were split into three subplots, each of which shows the results for one solver. For the "fast" input signal, the following observations can be made:

- The reduction error limits the best achievable accuracy of the reduced models. As can be expected, these accuracies increase
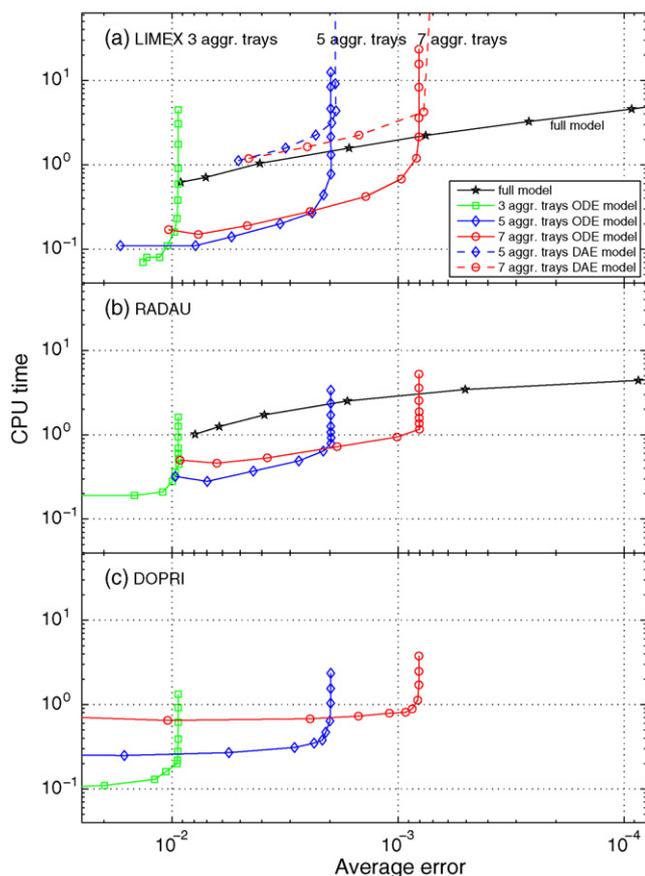
Fig. 9. Comparison of the performance of the uncontrolled models, simulated with three different solvers at different simulation tolerances, and with a stepwise input signal (Fig. 6). Part (a) shows simulations with the LIMEX solver, (b) with the RADAU solver and (c) with the DOPRI solver. Simulations of the same model with different simulation tolerances are connected with a line. Full lines are used for ODE models, and dashed lines for DAE models. In all simulations, linear table interpolation was used. Due to its high stiffness, the full model could not be simulated with the DOPRI-solver.
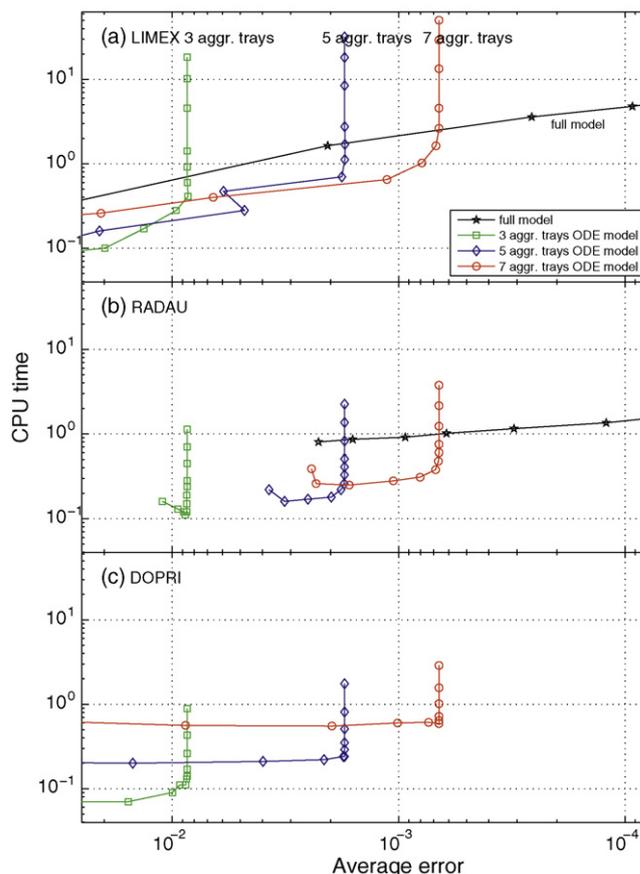
Fig. 10. Comparison of the performance of the uncontrolled models, simulated with various solvers and at various simulation tolerances, and with the "slow" input signal.

with increasing reduced model order. For crude simulation tolerances (low accuracy simulations), the accuracy of the reduced models is dominated by the simulation error. For tight simulation tolerances, the reduction error dominates and no further gain in accuracy can be achieved by increasing the simulation accuracy. Typically, for the LIMEX solver, tighter tolerances than 10^{-3} will mostly increase the simulation time, but not the accuracy.
- While the reduced models in DAE form do not show any improvement in computational speed, the reduced models in ODE form can be simulated much faster. This is explainable with the fact that a general-purpose DAE solver works basically like a stiff ODE solver (Hairer & Wanner, 2002), where the algebraic equations are treated similarly to fast dynamic equations. This means that there is no computational gain to be expected from converting dynamic equations into algebraic equations, although the use of specialized sparse DAE solvers and control algorithms might change this situation (Bian et al., 2005; Khowinij et al., 2005).
- For every model and solver, the accuracy curve intersects at a certain accuracy with the curve of the model of the next lower order. From this accuracy on downwards, it is computationally favorable to use the model of lower order. However, the seven-aggregation tray model is still performing well at low accuracies in some cases (LIMEX solver), such that it could be a more robust choice when there is doubt about the proper choice of simulation accuracy.
- Compared to the full model at the same accuracy, the highest achievable gain in simulation speed is between eight and 10.

For the "slow" input signal, similar observations can be made, but the performance of the individual solvers is different and the
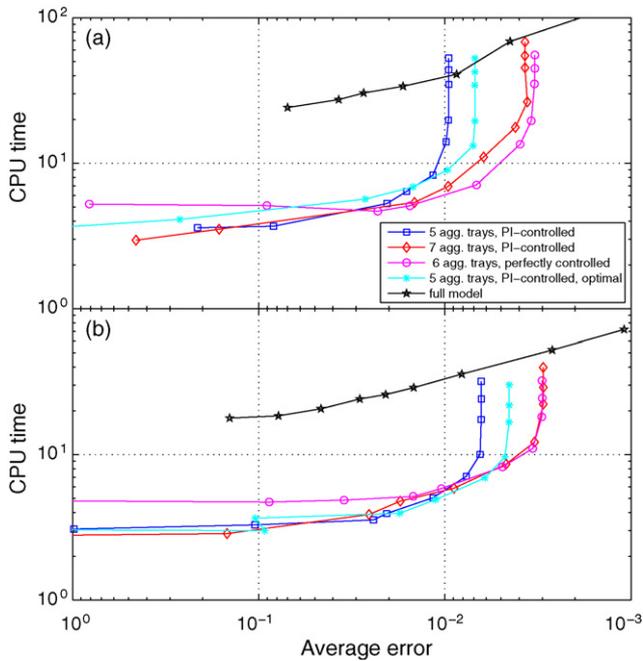
Fig. 11. Comparison of the performance of the full and four reduced controlled models, simulated with various simulation tolerances and input signals with step changes in z_F and L (a), and in z_F and x_c^s (b) (Fig. 7).

maximum gain in simulation speed is not as significant as in the stepwise case.

4.1.2. Controlled models

Fig. 11 shows the performances of four reduced models for stepwise input changes in z_F and L (plot a), and in z_F and x_c^s (plot b). The reduced models are models with five and seven equally sized aggregation trays, respectively, an optimized five-aggregation tray model, and a perfectly controlled model with six aggregation trays, where the reboiler and the controlled aggregation tray are merged together.

The following observations can be made:

- The general trends of the performance of the controlled model are the same as for the uncontrolled models.
- The optimal five-aggregation tray model has a lower reduction error than the equally sized model, but shows slightly worse performance at lower accuracies for changes in z_F and L.
- The perfectly controlled model shows slightly better performance than the seven-aggregation tray model at higher accuracies. The situation reverses at lower accuracies.
- At intermediate accuracies, the gain in computational speed is roughly of factor six.

Fig. 12 shows numerical statistics of the simulation runs for the full, the PI-controlled seven-aggregation tray model, and the perfectly controlled model with six aggregation trays. The numerical behavior of the PI-controlled model is very similar to that of the full model. The perfectly controlled model shows significantly lower numbers of steps for intermediate and tight tolerances, and accordingly low numbers for the other values as well. The number of residuum calculations is increased, since the Jacobian is calculated numerically. A possible explanation for this behavior is that the perfect model does not include the comparably fast dynamics of the composition controller in the bottom, allowing the integrator to take larger integration steps.
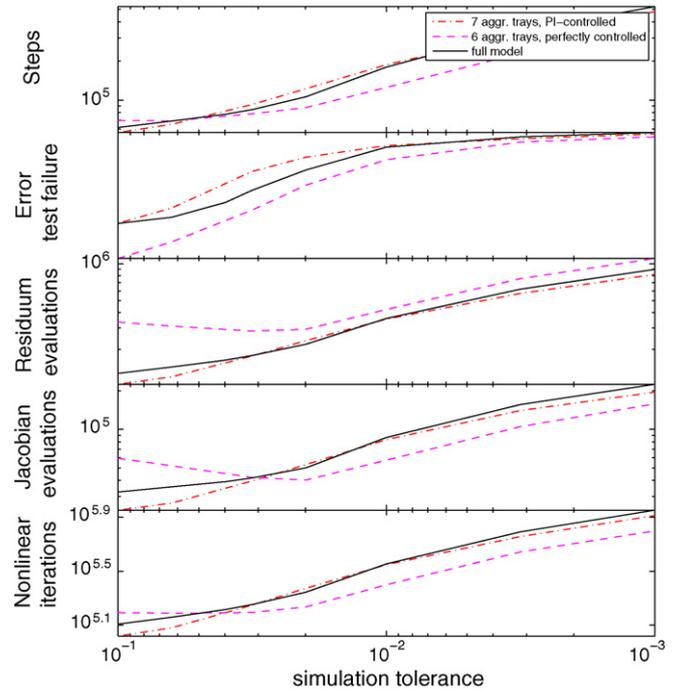

Fig. 12. Comparison of the numerical behavior of the full, the seven-aggregation tray PI-controlled reduced model, and the six-aggregation tray perfectly controlled model.

Fig. 13 shows the percentage of simulation time that is spent in the computationally most intensive parts of the integration for the full, the seven-aggregation tray model and the perfectly controlled model with six aggregation trays. For the full model, the functions for the solution of the linear equations dominate the simulation time. For the reduced model, residual and Jacobian evaluation time dominate. This is because of the computationally expensive look-up table interpolation that is used in the reduced models, and the reduction of the number of dynamic variables of the models. The perfectly controlled model is simulated with numerically calculated Jacobians, therefore the time of the Jacobian calculation is included in the residual calculation time. The remaining 40–50%
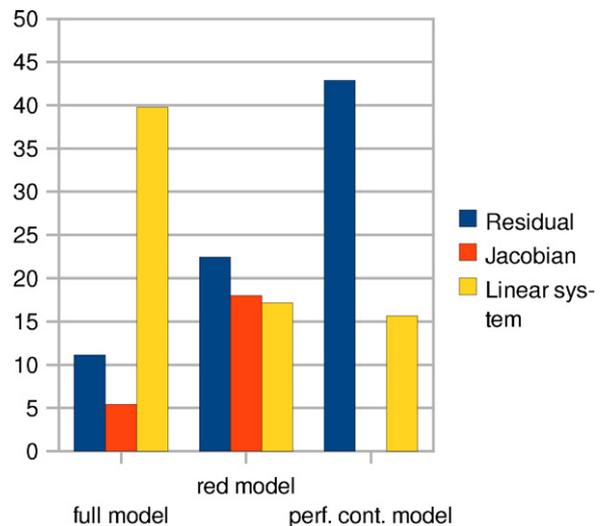

Fig. 13. Comparison of the percentage of simulation time of the computationally most intensive parts of the integration for the full, the seven-aggregation tray model and the perfectly controlled model with six aggregation trays.

Table 3
Reduction errors of the reduced models for input signals with step changes in z_F and L.

A	B	C	D	Reduction error	
Uncontrolled models					
3		Lin.		9.55	1e −3
3		Cub.		9.33	1e −3
5	Eqs.	Lin.		2.57	1e −3
5	Opt.	Lin.		2.00	1e −3
5	Eqs.	Cub.		2.45	1e −3
5	Opt.	Cub.		1.91	1e −3
7	Eqs.	Lin.		1.35	1e −3
7	Opt.	Lin.		8.13	1e −4
7	Eqs.	Cub.		1.26	1e −3
7	Opt.	Cub.		7.08	1e −4
Controlled models					
4	Eqs.	Lin.	pc	6.07	1e −3
4	Eqs.	Cub.	pc	5.34	1e −3
5	Eqs.	Lin.	PI	9.54	1e −3
5	Opt.	Lin.	PI	6.09	1e −3
5	Eqs.	Cub.	PI	8.09	1e −3
5	Opt.	Cub.	PI	5.23	1e −3
6	Eqs.	Lin.	pc	3.29	1e −3
6	Eqs.	Cub.	pc	2.44	1e −3
7	Eqs.	Lin.	PI	3.72	1e −3
7	Eqs.	Cub.	PI	2.64	1e −3

Column A gives the number of aggregation trays, column B gives the choice of position and size of the aggregation trays (equally sized or optimal), column C gives the type of interpolation used (linear or cubic), and column D gives the type of concentration control (PI or perfect control).

of the simulation time is spent in other parts of the solver. In the reduced models, the time spent for the linear look-up table interpolation amounts to about half of the time spent for the residual and Jacobian computations, which is about 20% of the total simulation time.

4.2. Choice of compartments

In Table 3, the reduction error of the reduced model variants are shown. The error of the optimized models is roughly between 50% and 80% of the models with equally sized aggregation trays. For higher-order reduced models, the ratio gets better in favor of the optimized variants, reflecting the increased number of degrees of freedom for optimization.

4.3. Solver performances

It has been observed before that the LIMEX solver performs well in simulations with low accuracy, while it is usually outperformed by other solvers at high accuracies (Schlegel, 2005). The same tendency can be observed in this study when comparing the full model being simulated with LIMEX and RADAU codes, respectively.

For the uncontrolled reduced models at low accuracies, in case of a stepwise input signal, the LIMEX solver shows a performance that is far superior to that of the RADAU solver. The DOPRI code shows comparable performance to the LIMEX code for high accuracies. While the LIMEX solver shows a gradual trade-off between accuracy and simulation time, the DOPRI solver rather abruptly looses performance when the simulation accuracy is not close to the maximal achievable accuracy. During simulations with crude tolerances, the DOPRI solver usually finds the problem to be stiff. This is probably the cause for the loss in performance.

In the "slow" input case, the performance of the RADAU solver is superior to the other solvers. A possible explanation is that the advantage of the LIMEX code when applied to discontinuous input signals cannot be exploited in this case. The DOPRI code again looses performance abruptly because of the problem becoming stiff for crude simulation tolerances.

The performance of the DASPK solver as the only solver used for integrating the controlled models cannot be compared to the other solvers. Still, the same gradual trade-off between accuracy and simulation time as with the LIMEX solver for the uncontrolled models can be observed.

5. Discussion

5.1. Reduction method

It was shown in Section 2.3.1 that due to a simplification step in the original derivation of Lévine and Rouchon (1991), the reduction method is a very simple procedure that can be applied without specifying any compartments, but only aggregation trays. Furthermore, it was shown that the algebraic equations resulting from the quasi-steady-state approximation can be eliminated from the model, resulting in significantly more compact reduced models that can be simulated several times faster than the original model.

The main bottleneck of the method lies in the second step, where the DAE system is converted into an ODE system. The off-line computed solutions of the algebraic equations have to be stored and retrieved as functions of the state variables of the aggregation trays encompassing the steady-state trays in a suitable way. In this study, this was done in a rather straightforward way by tabulating the relevant solutions and retrieve them by multi-dimensional interpolation of the table values. Since both table size and interpolation intensiveness increases exponentially with the number of independent variables, a high number of independent variables will lead to computationally expensive right-hand sides of the reduced models. In this study, the time spent for the linear look-up table interpolation is about 20% of the total simulation time. With every additional independent variable, this number will double, in case of linear interpolation, and increase four times, in case of cubic interpolation. The gain in computational performance will therefore depend on how computationally intensive the table interpolation is compared to the computational intensiveness of the residual and Jacobian calculation of the full model.

5.2. Relationship to singular perturbation methods

As described in Section 2.3.1, the model reduction method was originally derived as a singular perturbation method. However, since an additional modification of the equations is performed, the derivation is different from a standard singular perturbation approach, where the model is taken into the standard form of singular perturbation by some state transformation, and then a quasi-steady-state approximation is performed (Kokotovic et al., 1986). In Section 2.3.2, these "true" singular perturbation systems have been compared to the reduced models by Lévine and Rouchon (1991). Their reduced models still show a typical "slow model" behavior. This means that the model is getting the more accurate, the closer it is to steady-state. The initial response to a fast input change is typically faster than the response of the original model, because the fast dynamics of the original model are replaced by "instantaneous dynamics" by the quasi-steady-state approximations in the reduced model. The method of Lévine and Rouchon can therefore be described as a "singular perturbation-related" method. The physical interpretation of the method is also slightly different: Instead of exploiting the time-scale separation between slow compartments and fast trays, the method works by making the signal transport between the aggregation trays infinitely fast, and slowing the dynamics down again by increasing the time-constants of the aggregation trays.

5.3. Validity of investigated system

The distillation model used in this study is very simple. The constant molar holdup assumption is generally not a good assumption for the initial response to a change in reflux and for control purposes (Skogestad, 1997), but is acceptable in this study where only one control loop in the bottom is closed. The model is, however, well suited to demonstrate the application and performance of the model reduction method. It can be expected that the obtained results possess some predictive quality for more complex cases, since the dynamics and time-constants are similar.

5.4. Extension and generalization of reduction method

The simple derivation of reduced models presented in this paper makes it straightforward to extend the reduction method to more complex distillation models, and to generalize it to general one-dimensionally distributed systems.

In more realistic distillation models, each tray comprises a total mass balance (tray hydraulics), an energy balance, and possibly several component balances. These can be treated in the same way as the single component mass balance equations used in this study. The resulting algebraic equations are treated in the same way as well. However, the number of independent variables, which determine the solution of the algebraic equations, is larger than in the models in this study, yielding more complex functions that have to be approximated. Indeed, the reduction method has been successfully applied to a binary distillation column with energy balances and complex thermodynamic and hydraulic relationships. These results will be published separately.

Furthermore, it is possible to generalize the reduction method to general one-dimensionally spatially distributed systems, such as plug-flow reactors or heat exchangers. If the system consists of discrete units, such as the trays in a distillation column, or a model resulting from the discretization of a partial differential equation, the reduction method can be applied in exactly the same way. A detailed description of this method is in preparation.

5.5. Computational performance of reduced models

It has been found that for a given accuracy a gain in simulation speed of a factor five to ten can be achieved using the reduced models proposed in this study. The accuracy measure is the integrated deviation from a reference trajectory, when the model is exposed to step changes in the inputs, where the interval between two step changes is long enough to allow the system to come close to steady-state again. However, the performance of a reduced model will strongly depend on the application. In model predictive control, for example, the model will be exposed to step changes in the inputs at intervals of a few hundred seconds. Since the reduced models presented here are relatively slow models, their response in the initial period after an input change is possibly rather inaccurate (compare Fig. 3). However, the reduced models might be well suited for an application that emphasizes the long-term behavior of the system, such as in optimal trajectory planning. This is confirmed by a study by Würth et al. (2007), who found that reduced aggregated models yield significant savings in computation time for dynamic real-time optimization.

5.6. Tabulation of steady-state tray solutions

Multi-linear and multi-cubic interpolation has been used in this study to obtain continuous function values from the look-up tables of the steady-state tray solutions. Multi-linear interpolation is recommended because multi-cubic interpolation will be compu-

tationally expensive for a high number of independent variables as can be expected for models with more components. For a low number of independent variables, an increased accuracy (see Table 3) can instead be achieved by multi-linear interpolation with a higher table resolution.

5.7. Comparison with other model reduction methods for distillation columns

Besides tray aggregation, wave propagation theory (Hankins, 2007; Kienle, 2000; Marquardt, 1990) and orthogonal collocation methods (Cho & Joseph, 1983; Dalouti & Seferlis, 2006) have been proposed as alternative methods for deriving reduced distillation models. Wave propagation approximates the system dynamics by traveling wave profiles in response to input changes. Kienle (2000) reports very good dynamic and steady-state approximation qualities of the wave models, assuming constant molar flow rates, holdups and pressure in the system. Hankins (2007) presents extensions of the method to systems with energy balances and hydraulic effects. Khowinij et al. (2004) compare the performance of a wave model to that of a compartment model. They report that the approximation quality of the wave model is rather poor compared to that of the compartment model, which, however, might be due to the different assumptions incorporated in the two models.

Orthogonal collocation approximates the state variable profile as sums of orthogonal polynomials. For this, a discrete column model is first transformed into continuous partial differential equations. By enforcing the approximating profile to satisfy the differential equations at a number of collocation points, discrete dynamic equations are obtained. The approximation quality of collocation methods depend on the choice and number of collocation points. Dalaouti and Seferlis (2006) apply the approach to complex staged reactive separation processes. Their reduced models show good approximation quality, while a gain in simulation speed of several times is reported.

Of the three methods, tray aggregation is conceptually probably the simplest. However, the implementation requires some effort, as is described in this study. In terms of accuracy, each method has its strengths and weaknesses. The advantage of tray aggregation compared to the other methods is the perfect match of steady-states, making it interesting for applications where the long-term behavior is important. Wave propagation, on the other hand, seems to yield a good approximation of initial system responses to fast changing inputs. Collocation methods are probably the most widely used methods to obtain reduced distillation models. Since the method can be applied to complex systems without many restrictive assumptions, the accuracy is suitable for most applications. Of the three methods, wave propagation is computationally probably the least intensive. However, a detailed comparison of the three methods with careful implementations taking the application of the models into account has yet to be done in order to assess their strengths and weaknesses for that particular application.

6. Summary and conclusions

In this study, the tray aggregation method of Lévine and Rouchon (1991) has been used to derive reduced models of a binary distillation column model. It has been shown that the original derivation, which partitions the column into "compartments" of consecutive trays, can be simplified to a procedure which only requires the specification of aggregation trays and is independent of the choice of compartment boundaries. This gives the resulting reduced model a favorable numerical structure and makes the extension to more complex models straightforward. The original method yields a DAE model with the same number of variables as the original model.

Because of the banded Jacobian structure of the reduced model, the algebraic equations of the DAE models can be eliminated from the model by off-line solution and substitution of the tabulated solution values as functions of the dynamic variables into the dynamic equations, yielding low-order ODE models. The numerical performance of these models has been assessed by simulations using four different numerical integrators. While simulation of the DAE models showed no gain in simulation speed, the reduced models yielded an improvement of up to factor 10. Due to the reduced stiffness of the models, a non-stiff integrator can be used for integration. The inclusion of a controller in the bottom section of the column has been investigated. While the application of a PI-controller to the reduced model is easily possible, the controller can also be directly incorporated into the reduced model assuming perfect control. Both variants show good performance, decreasing simulation time about six times.

The computational gains that can be achieved by the reduced models form make these kind of models attractive for use in real-time optimization applications. In more complex models, the performance of the reduced models will depend on the complexity of the right hand sides as well as the number of independent variables for the tabulation of the steady-state tray solutions.

Acknowledgement

References

Allgöwer, F., & Zheng, A. (2000). Nonlinear model predictive control. *Progress in systems theory* (p. 26). Basel: Birkhäuser.

Ascher, U. M., & Petzold, L. R. (1998). *Computer methods for ordinary differential equations and differential-algebraic equations*. Philadelphia: SIAM.

Benallou, A., Seborg, D. E., & Mellichamp, D. A. (1986). Dynamic compartmental models for separation processes. *AIChE Journal, 32*, 1067.

Bian, S., Khowinij, S., Henson, M. A., Belanger, P., & Megan, L. (2005). Compartmental modeling of high purity air separation columns. *Computers & Chemical Engineering, 29*, 2096.

Brenan, K. E., Campell, S. L., & Petzold, L. R. (1996). *Numerical solution of initial-value problems in differential-algebraic equations*. Philadelphia: SIAM.

Cho, Y. S., & Joseph, B. (1983). Reduced-order steady-state and dynamic models for separation processes. Part I. Development of the model reduction procedure. *AIChE Journal, 29*, 261.

Dalaouti, N., & Seferlis, P. (2006). A unified modeling framework for the optimal design and dynamic simulation of staged reactive separation processes. *Computers & Chemical Engineering, 30*, 1264.

Ehrig R., & Nowak, U. (2000). *Konrad-Zuse-Zentrum für Informationstechnik Berlin*, http://www.zib.de.

Hairer, E., Nørsett, S. P., & Wanner, G. (2000). *Solving ordinary differential equations I—nonstiff problems*. Berlin: Springer.

Hairer, E., & Wanner, G. (2002). *Solving ordinary differential equations II—stiff and differential-algebraic problems*. Berlin: Springer.

Hairer, E., & Wanner, G. (1996). http://www.unige.ch/~hairer/software.html.

Hankins, N. P. (2007). A non-linear wave model with variable molar flows for dynamic behavior and disturbance propagation in distillation columns. *Chemical Engineering Research & Design, 85*, 65.

Khowinij, S., Henson, M. A., Belanger, P., & Megan, L. (2005). Dynamic compartmental modeling of nitrogen purification columns. *Separation and Purification Technology, 46*, 95.

Khowinij, S., Bian, S., Henson, M. A., Belanger, P., & Megan, L. (2004). Reduced order modeling of high purity distillation columns for nonlinear model predictive control. In *Proceedings of the American control conference* Boston, MA, (p. 4237).

Kienle, A. (2000). Low-order dynamic models for ideal multicomponent distillation processes using nonlinear wave propagation theory. *Chemical Engineering Science, 55*, 1817.

Kokotovic, P., Khalil, H. K., & O'Reilly, J. (1986). Singular perturbation methods in control: analysis and design. *SIAM classics in applied mathematics* (p. 25). London: SIAM

Kumar, A., & Daoutidis, P. (2003). Nonlinear model reduction and control for high-purity distillation columns. *Industrial and Chemistry Research, 42*, 4495.

Lévine, J., & Rouchon, P. (1991). Quality control of binary distillation columns via nonlinear aggregated models. *Automatica, 27*, 463.

Li, S., & Petzold, L. R. (2000). Software and algorithms for sensitivity analysis of large-scale differential algebraic systems. *Journal of Computational and Applied Mathematics, 125*, 131.

Marquardt, W. (1990). Traveling waves in chemical processes. *International Chemical Engineering, 30*, 585.

Marquardt, W. (2001). Nonlinear model reduction for optimization based control of transient chemical processes. In *Proceedings of the CPC VI* (p. 30).

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes: the art of scientific computing* (3rd ed.). Cambridge: Cambridge University Press.

Schlegel, M. (2005). Adaptive discretization methods for the efficient solution of dynamic optimization problems. In *Fortschritt-Berichte VDI*. Düsseldorf: VDI-Verlag., p. 829.

Schlegel, M., Stockmann, K., Binder, T., & Marquardt, W. (2005). Dynamic optimization using adaptive control vector parameterization. *Computers & Chemical Engineering, 29*, 1731.

Shampine, L. F., & Reichelt, M. W. (1997). The MATLAB ODE Suite. *SIAM Journal on Scientific Computing, 18*, 1.

Shinskey, F. G. (1984). *Distillation control for productivity and energy conservation* (2nd ed.). New York: McGraw-Hill.

Skogestad, S., & Morari, M. (1987). The dominant time constant for distillation columns. *Computers & Chemical Engineering, 11*, 607.

Skogestad, S. (1997). Dynamics and control of distillation columns: a critical survey. *Modeling, Identification and Control, 18*, 177.

Skogestad, S. (2003). Simple analytic rules for model reduction and PID controller tuning. *Journal of Process Control, 13*, 291.

van den Berg, J. (2005). *Model reduction for dynamic real-time optimization for chemical processes*. PhD Thesis, TU Delft.

Würth, L., Linhart, A., Preisig, H., & Marquardt, W. (2007). *Performance of reduced distillation models in dynamic real-time optimization*. Copenhagen: ECCE.