ERROR DETECTION AND CONTROL IN GREY-BOX IDENTIFICATION OF DISTRIBUTED PARAMETER PROCESSES

Yi Liu and Elling W. Jacobsen

Automatic Control
Dept. of Signals, Sensors and Systems
Royal Institute of Technology
SE-100 44 Stockholm, Sweden
Email: yi.liu@s3.kth.se, jacobsen@s3.kth.se

Abstract: A systematic approach to grey-box identification of distributed parameter processes is proposed. The principle idea is to integrate a novel moving finite element method for model reduction, with the toolbox MoCaVa for grey-box modelling of lumped parameter systems. A method is proposed for separating the model reduction error from discrepancies between model structure and measurement data, based on estimating the model reduction error in combination with simple hypothesis testing. This method can be used for validating the PDE structure as well as the order of the reduced model. To keep the model order at a minimum, the reduced model is augmented with a mesh controller, which uses error feedback to determine the size and location of the finite elements.

1. INTRODUCTION

Grey-box identification concerns calibration and validation of dynamic process models with partially known but uncertain structure, and in the presence of disturbances from the environment. A number of grey-box identification methods and toolboxes have been developed for modelling lumped parameter systems governed by ordinary differential equations (ODE). However, they do in principle not cover distributed parameter system, which are nevertheless very common in the process industries. In order to extend the methods to distributed parameter systems, some form of model reduction, i.e., approximating PDEs with a set of ODEs, is required.

In principle, model reduction can be performed a priori using any discretization method, e.g., finite differences or finite elements, on any given discretization mesh. This is the typical approach, e.g., in software for parameter estimation (gPROMS, 2004). However, as model reduction will introduce a model error, it is in general difficult to a priori determine an appropriate discretization method and model order that will provide an acceptable accuracy. Moreover, it will in general be impossible to decide whether discrepancies between the model structure and the available data are due to the original model structure or the model reduction. Also, by employing an a priori fixed order discretization, valuable information about the original model structure is lost. Since a key-idea in grey-box identification is to exploit information about the underlying model structure, the knowledge that a model structure indeed involves PDEs should therefore somehow be conserved.

Our aim is to propose a systematic method for grey-box identification of distributed parameter processes. The ultimate goal is to integrate a flexible method for model reduction, based on moving finite elements (Liu and Jacobsen, 2001), with a method for grey-box identification as pro-

posed in (Bohlin, 1991) and implemented in the software MoCaVa (Bohlin, 2001). The principle idea is to use the tools available in MoCaVa for calibration and validation also of the model reduction. See Figure 1. In this paper, we consider two specific problems. 1) Separating model reduction error from model-data discrepancies, by combining model reduction error estimation with simple hypothesis testing; 2) Model order selection, i.e., determining the required order of the discretized PDE model for identification purposes. The proposed method is demonstrated on a chromatography process used in bio-separation.

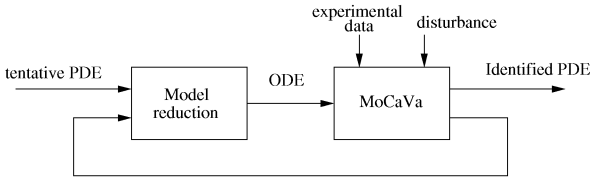

Figure 1. Integration of model reduction with MoCaVa

2. PROBLEM FORMULATION

Model reduction introduces a model error, and an important issue in grey-box identification of PDEs is to therefore distinguish between error sources, as illustrated in Figure 2. There are discrepancies between the tentative PDE model and measurement data (e_1), discretization errors from model reduction of the tentative PDE model (e_2), and discrepancies between the discretized model and the experimental data (e_3). Identification of the PDE models aims at making e_1 small, while the size of e_3 can be obtained when applying identification methods on the discretized model. In principle, a significant error e_3 can be caused by a poor model reduction rather than a poor structure of the tentative PDE model. Contrary, the ODE model may fit the experimental data well although e_1 is large. This may happen if also e_2 is large, and will thus lead to a falsely identified PDE model. Therefore, to ensure a small model error e_1, it is essential to establish whether the error e_3, as estimated by available identification methods, is due to e_1 or e_2.

In general, the problems discussed above cannot be solved by a priori discretization. Moreover, the required accuracy of the model reduction depends not only on the model structure, but also on the experimental data as well as the intended use of the model. For identification purpose, it is usually crucial to use a low order ODE model as the model order directly influences the complexity of the succeeding identification procedure.

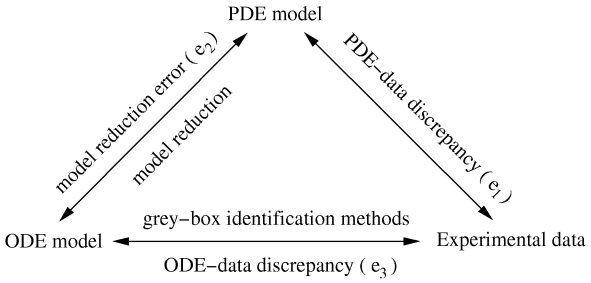

Figure 2. Error sources in grey-box identification of PDEs

2.1 Example: parameter estimation of a convection-diffusion process

To illustrate the potential problems with a priori discretization, consider a simple process involving convection and diffusion governed by

$$v_t = -v_x + 0.1 v_{xx}, \quad x \in [0,1], \ t > 0 \qquad (1)$$

where x represents space and t time, with initial and boundary conditions given by

$$v(x,0) = 0, \ \ v(t)|_{x=0} = \sin(bt), \ \ v_x(t)|_{x=1} = 0$$

Treat $u = v(0,t)$ as the input and $y = v(1,t)$ as the output of this system. For simplicity, we assume no process disturbances nor measurement noises. The problem is to identify the system based on measured discrete-time outputs $y(t_k)$. Two tentative PDE models are considered, one in correct form, one neglecting the diffusion term:

$$\mathcal{M}_\mathcal{A}: \quad v_t = -a v_x + \varepsilon v_{xx} \qquad (2)$$

$$\mathcal{M}_\mathcal{B}: \quad v_t = -a v_x \qquad (3)$$

First, model reduction is applied to the tentative models. For stability reasons, an upwinding scheme is employed for discretizing the convection term v_x while central difference is used for the diffusion term v_{xx}, both on a fixed uniform mesh. Then, a weighted least square method is employed, i.e., parameter estimates are determined as the minimizer of some cost function V depending on the prediction error, $\hat{\theta} = \arg\min_\theta V(\theta)$. To directly relate the prediction error to the goodness of the fit, we employ a relative error. The cost function is defined as variance of the prediction error, weighted by variance of the simulated outputs (with output mean removed).

$$V(\theta) = \frac{\frac{1}{N_s}\sum_{k=1}^{N_s}[y(t_k) - \bar{y}(t_k,\theta)]^2}{var(\bar{y}) - [E(\bar{y})]^2} \qquad (4)$$

where $\bar{y}(t_k,\theta)$ are simulated outputs of the discretized PDE with the parameter set θ at time t_k, N_s is the number of measured outputs. $var(\bar{y})$ and $E(\bar{y})$ are variance and mean of $\bar{y}$, respectively. For

Test set	Model	Input	N	$\hat{a}$	$\hat{\epsilon}$	$D(\hat{a})$	$D(\hat{\epsilon})$	V_{min}	$V_{real}(\hat{a},\hat{\epsilon})$
(a)	$\mathcal{M}_\mathcal{A}$	$sin(t)$	5	1.0908	0.0162	1.2423e-4	5.9197e-5	1.7840e-6	0.0017
(b)	$\mathcal{M}_\mathcal{A}$	$sin(t)$	20	1.0064	0.0815	4.9710e-5	1.1088e-5	3.0927e-7	1.5668e-4
(c)	$\mathcal{M}_\mathcal{B}$	$sin(t)$	5	1.1237	-	8.7238e-4	-	8.8801e-5	0.0024
(d)	$\mathcal{M}_\mathcal{B}$	$sin(t)$	20	1.1316	-	0.0050	-	0.0030	0.0024
(e)	$\mathcal{M}_\mathcal{A}$	$sin(15t)$	20	1.0245	0.0748	0.0011	1.9025e-4	0.0014	1.6480e-4

Table 1. Estimation results for convection-diffusion process. True parameters: a=1, ε=0.1; N: order of the reduced model; $\hat{a}$ and $\hat{\epsilon}$: parameter estimates; D: standard deviation; V_{min}: cost function with the estimates set, obtained from the discretized model; V_{real}: real cost function with tentative PDE.

the tentative model $\mathcal{M}_\mathcal{A}$, $\theta = [a\ \varepsilon]^T$, while for the tentative model $\mathcal{M}_\mathcal{B}$, $\theta = a$.

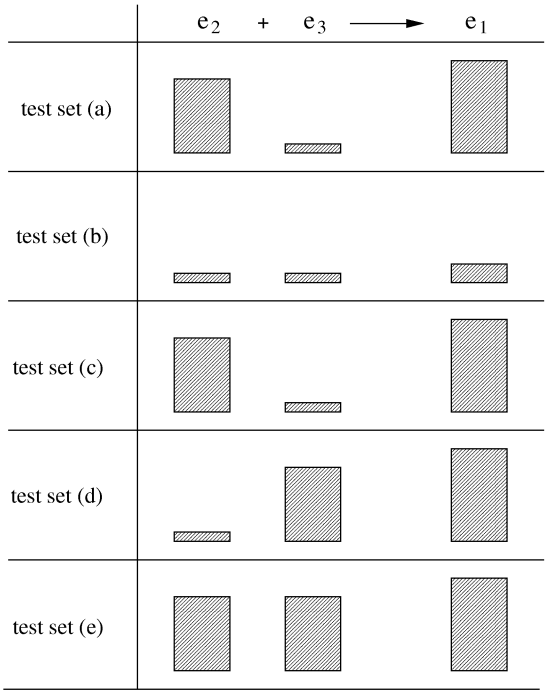

Figure 3. Error sources in different test sets for the illustrating example. The heights of the blocks represent size of the corresponding error.

In all test sets designed below, the input is a sinusoidal wave applied for 5 periods, and the output is recorded with 200 sampling points. The reason for using a simple sinusoid is to enable a simple illustration of the impact of the frequency content in the data. Tests (a) and (b) estimate parameters based on the model structure $\mathcal{M}_\mathcal{A}$, with different model orders. Tests (c) and (d) are similar, but for the erroneous model structure $\mathcal{M}_\mathcal{B}$. Test (e) is similar to test (b), except the input frequency is increased. A graphical illustration of the size of different error sources for all test sets are presented in Figure 3. The results are summarized in Table 1, from which three essential observations can be obtained. First, when the tentative model has a correct structure, significant biases in the estimates can be caused by a poor model reduction, which can then be improved by increasing the model order, as illustrated by (a) and (b). Note that in (a), although the minimized cost function V_{min} appears small, the *real* cost function, obtained from substituting the estimates into the tentative PDE model and discretizing it on a fine grid, is indeed large. This shows that the obtained parameter estimates involve significant biases even though the data-fitting seems fine. Second, (c) and (d) reveal that when a wrong model structure is used, e_3 may appear small, but it is due to a large e_1 counteracted by a large e_2. In such cases, increasing the model order decreases e_2 but increases e_3. The reason, in this particular example, is that the model reduction introduces an artificial diffusion which plays the role of the diffusion in the true system. Note, however, that this is not possible to detect by examining the standard deviations and the minimum cost function. Third, the required accuracy of the model reduction also depends on the frequency content of the data, as shown by (b) and (e). Due to the space limit, we do not include detailed discussion here but only point out that, for the purpose of identification, the discretized model should primarily approximate the tentative PDE model in the frequency range where the data are concentrated.

2.2 Separation of model reduction errors from model-data discrepancies

From the above, it is clear that in order to properly fit and validate PDE models it is essential to be able to have some knowledge about the model reduction error. If the model reduction error e_2 is guaranteed small, methods for evaluating estimation results applied on ODE systems can be more or less directly applied to PDE systems. In principle, e_2 decays to zero as the model order increases to infinity. However, higher order of the ODE model increases the complexity of the identification. Hence, it is critical to choose a suitable model order for accuracy and efficiency purposes. Below, two approaches to separate e_2 from model-data discrepancies, based on simple hypothesis postulation and testing, are considered.

The first approach starts by discretizing the tentative PDE model using the maximum model order N_{max} that one can practically accept, followed

by parameter estimation based on this high order model. Assuming a negligible model reduction error e_2, a minimum cost function V_{min} large than a pre-defined threshold $(V_{min})_{tol}$ indicates the need to improve the model structure. Otherwise, reduce the model order gradually and repeat the parameter estimation, so as to make $V_{min}(N)$ approach the given threshold. The resulting model order is then considered as *"optimal"*, in the sense it provides both acceptable accuracy and efficiency. The main drawback of this approach is that it typically will require excessive computational loads. Further, it is difficult to determine whether the starting maximum model order yields a negligible e_2.

Second, start with a low model order and apply parameter estimation on the discretized model. Then, increase the order N gradually, and perform estimation on each discretized N^{th} order model, so as to find an optimal N such that $V_{min}(N)$ is (slightly) below the given threshold. This approach seems simple but indeed involves some problems. Since e_2 may compensate for the model structure error e_1, V_{min} calculated from the reduced model may not always decrease as N increases. For example, $V_{min}(N)$ may first decrease then increase, for an increasing N. If this happens, it implies that counteraction of significant e_2 and e_1 produces a false minimum V_{min}, and hence the structure of the tentative PDE model may need improvement. To deal with the problem one may, after obtaining the corresponding estimate $\hat{\theta}$ for a given N, re-discretize the PDE model with a high model order N_{max}. Substituting the previously obtained estimate $\hat{\theta}$ into this model, one can compare the resulting loss function $V(\hat{\theta}, N_{max})$ with that obtained for the given N, $V(\hat{\theta}, N)$. If the difference between the two loss functions is large, one can assume that the model reduction error is significant with model order N. This approach is more computationally efficient than the above approach, but still demanding due to the need to consider a large number of parameter estimations as well as high order models for validation of the results.

To summarize, a suitable model order is essential for achieving both accuracy and efficiency in grey-box identification of PDE processes. Since, in principle, the model reduction error asymptotically goes to zero as the model order goes to infinity, it is possible to separate e_2 from the model-data discrepancies by simple hypothesis testing, as discussed above. However, these approaches are computationally inefficient and sometimes involve numerical problems with the use of high model orders. To reduce the complexity, we discuss in the following how the problem can be handled by estimating the model reduction error and combining it with hypothesis testing.

3. A SYSTEMATIC APPROACH TO GREY-BOX MODELLING OF PDES

To enable grey-box modelling of PDEs in a more systematic fashion, we propose to include estimation of the error e_2, introduced by model reduction. The estimated error $\hat{e}_2$ can be used in combination with simple hypothesis testing to separate e_2 from model-data discrepancies. In order to keep the reduced model order close to the minimum required, a moving finite element method named OCMFE (*Orthogonal collocation on moving finite elements*), proposed in (Liu and Jacobsen, 2001), is employed. Available grey-box modelling methods for ODE models can then be applied to the reduced model.

The underlying discretization method employed in OCMFE is *orthogonal collocation on finite elements* (OCFE), proposed in (Carey and Finlayson, 1975). In OCFE, the spatial domain is divided into elements, and the solution is approximated by low order polynomials within each element.

In general, it is not possible to obtain the true model reduction error. For finite elements methods, the equation residuals are often used as measures of the solution accuracy. With the use of OCFE, for instance, residuals are forced to zero on the collocation points, but are in general non-zero at non-collocation points. The integral size of the residuals can hence be used to provide some estimate of the model reduction error e_2. In particular, large residuals will in general imply significant model reduction errors. However, the opposite does not necessarily hold, that is, small residuals does in the general case not guarantee small model errors. In (Liu and Jacobsen, 2001), it is shown that the residuals for linear problems, discretized with OCFE, always will be identically zero while the solution may be arbitrarily in error. To get a more reliable error estimate we consider combining the residual with an error measure often used in polynomial interpolation, based on Nevilles interpolation algorithm (Press *et al.*, 1988). This error measure is based on removing one interpolation point and then performing a new interpolation based on the remaining points. The difference in predicted profile then serves as a measure of the solution error. We combine the residual based and interpolation based error estimates, and denote it Q. By comparing the size of Q to the solution profile, an estimate of the model reduction error is proposed as

$$\hat{e}_2 = \int_t \frac{\int_x Q dx}{\int_x u dx} dt \tag{5}$$

where the integration in time and space easily can be approximated by a simple quadrature.

Note that the error estimator as defined is data dependent.

Quantitatively, one can not draw any exact conclusion regarding the true error e_2 from the size of $\hat{e}_2$. However, it is reasonable to assume that a decrease in the size of the estimated error implies a decrease also in the true error. This makes the error estimation useful when combining it with simple hypothesis testing.

The method we propose here consist of starting with parameter estimation based on a low-order discretized model, and estimate the resulting model reduction error $\hat{e}_2(N, \hat{\theta})$. Then, increase the model order slightly to N^+ and re-discretize the PDE with the previously obtained $\hat{\theta}$. In general, this should decrease the estimated error i.e., $\hat{e}_2(N^+, \hat{\theta}) < \hat{e}_2(N, \hat{\theta})$. If the reduction is significant, N should be increased and the above procedure repeated until the difference becomes negligible. When so, check if V_{min} is below the pre-defined threshold, and if the standard deviations of the parameter estimates are reasonably low. If both are satisfied, it indicates acceptable estimation results and a suitable model order. Otherwise, we may consider to improve the PDE model structure and/or re-design experiments.

In addition to use the above proposed error estimate for validation of the model reduction, we here also use it to improve the efficiency of the model reduction itself. This is done by augmenting the discretized model with a feedback control law using the element sizes and positions as manipulated variables with the aim of spatially equidistributing the instantaneous discretization error, i.e., the time derivative of $\hat{e}_2$. Equidistribution of error measures is common in numerical methods, and is based on the inherent assumption that it leads to a reduced total error. The method amounts to a novel *moving mesh method*, and is described in (Liu and Jacobsen, 2001).

4. APPLICATION ON A CHROMATOGRAPHY PROCESS

We here apply the systematic approach discussed above on a chromatography process used for bio-separation. Chromatography processes typically possess a sharp and moving composition front and hence represent a challenging problem for finite dimensional modelling. Consider a real process studied by (Borgquist, 1999), and described by

$$\frac{\partial C}{\partial t} = \frac{1}{B_o}\frac{\partial^2 C}{\partial z^2} - \frac{\partial C}{\partial z}, \quad B_o = 1039.5 \qquad (6)$$

with initial and boundary conditions

$$C(z,0) = 0, \quad \left.\frac{\partial C}{\partial z}\right|_{z=0} = B_o(C-1), \quad \left.\frac{\partial C}{\partial z}\right|_{z=1} = 0$$

where C is the adsorbant concentration, z space and t time. $z \in [0,1]$, $t \geq 0$. Consider C at the right boundary, $C(z=1,t)$, as the output of this system, and assume that the tentative model is in correct form. For simplicity we assume noise free measurement data. We consider estimating the parameter B_o using the weighted lest square method presented in (4) above.

Below, four tests are performed using different discretization methods and model orders. The estimation results are summarized in Table 2 while the simulated outputs are plotted in Figure 4. Assume that the threshold for measuring the goodness of data-fitting is chosen as $(V_{min})_{tol} = 10^{-4}$, indicating that a relative prediction error below 0.01% is acceptable. In test (i) and (ii) we discretize the tentative model using central difference, resulting in 30 and 100 ODEs, respectively. Clearly, the parameter estimate is improved by increasing the model order. Both the minimum cost function and the standard deviation of $\hat{B}_o$ are decreased in (ii), compared to (i). However, even with the model order as high as 100, V_{min} from the discretized model is still larger than the threshold $(V_{min})_{tol}$, which hence requires further increase of N. Moreover, it is not trivial to determine a suitable model order unless using involved hypothesis testing.

Test (iii) employs OCFE on 5 elements with a 6^{th}-order polynomial for each element, i.e., a total number of 29 ODEs. Thanks to this more efficient model reduction method, with a similar model order as in (i), the estimation results are apparently improved. Moreover, the proposed error estimation method can now be utilized. As can be seen, compared to the error estimate based on 29 ODEs, an error estimate based on a slightly higher order model is decreased about 30%, indicating that the model reduction needs to be improved.

In test (iv) we employ the moving mesh method OCMFE. The spatial discretization mesh again consists of 5 elements with 6^{th}-order polynomials on each element. As the mesh control contributes 4 ODEs, the total model order now becomes 33. In this case, the difference between the estimated model reduction errors based on the selected model order and a slightly higher model order is much smaller. Compared to (iii), the estimated error $\hat{e}_2$ is reduced by a factor 8, and better estimation results are obtained, as indicated by the decrease of standard deviation of $\hat{B}_o$ and V_{min}. An excellent output fitting is also obtained, and V_{min} is now below the specified threshold. Indeed, the estimated parameter is very close to its true value.

Test set	Model reduction	N	$\hat{B}_o$	$D(\hat{B}_o)$	$V_{min}(N)$	$\hat{e}_2(N, \hat{B}_o)$	$\hat{e}_2(N^+, \hat{B}_o)$
(i)	Central difference	30	600.0	1.5992	0.0101	-	-
(ii)	Central difference	100	1028.1	0.8535	4.3332e-4	-	-
(iii)	OCFE	29	896.2	0.0598	5.4650e-4	0.0126	0.0091
(iv)	OCMFE	33	1043.7	0.0387	9.5380e-5	0.0018	0.0016

Table 2. Parameter estimation of the chromatography process, based on a tentative model $C_t = C_z/B_o - C_{zz}$. True parameter $B_o = 1039.5$; N: order of the discretized model; $\hat{B}_o$: estimate of B_o; D: standard deviation; V_{min}: cost function obtained from the discretized model; $\hat{e}_2(N, \hat{B}_o)$: estimated model reduction error based on N^{th} order discretized model with the parameter estimate $\hat{B}_o$.

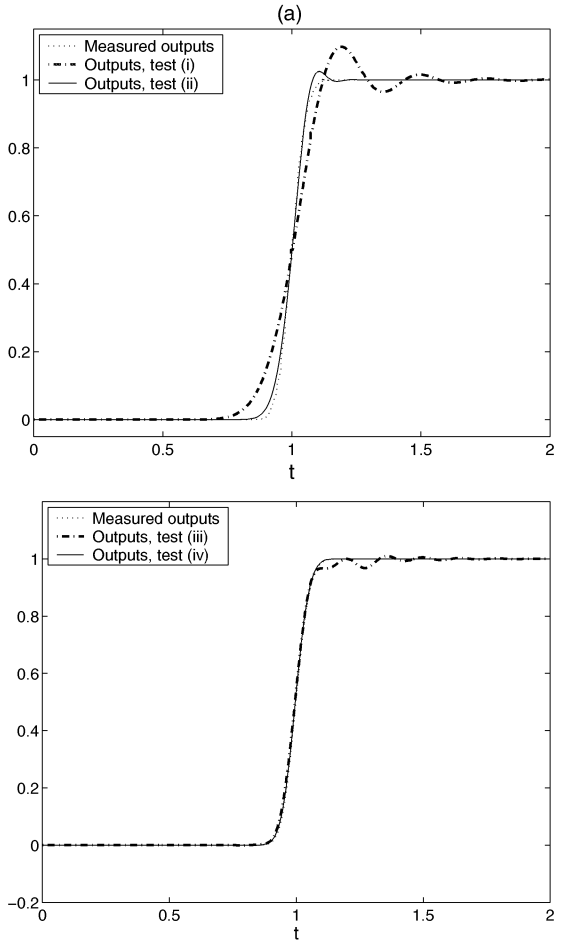

Figure 4. Data fitting for the chromatography process. (a) Measurements and modelled outputs in test (i) and (ii); (b) measurements and modelled outputs in test (iii) and (iv). Note that measured outputs and modelled outputs from test (iv) almost overlap.

5. CONCLUSIONS AND FUTURE WORK

We have considered the problem of grey-box identification of PDE systems in a systematic fashion. We showed that by estimating the model reduction error, the problem of separating this error from model-data discrepancies can be significantly simplified, which hence provides an effective complement to pure hypothesis postulation and testing. By employing the method of OCMFE for mesh control, a low order ODE model is achieved from the discretization of the original PDE model while preserving the model accuracy. This significantly reduces the complexity of the identification procedure. In the future, further problems involved in the integration of OCMFE with MoCaVa will be considered, e.g., the initialization and calibration of model order and control parameter in the model reduction, by utilizing statistical tools provided in MoCaVa.

REFERENCES

Bohlin, T. (1991). Interactive system identification: Prospects and pitfalls. *Springer-Verlag.*

Bohlin, T. (2001). A grey-box process identification tool: Theory and practice. *Technical report, IR-S3-REG-0103, Royal Institute of Technology, Sweden.*

Borgquist, P. (1999). Simulering av kromatografi processer. *Master thesis, TKKA-5005, Lund Institute of Technology, Sweden.*

Carey, G.F. and B.A. Finlayson (1975). Orthogonal collocation on finite elements.. *Chemical Engineering Science,* **30**, 587–596.

gPROMS (2004). *Process Systems Enterprise, London, UK.*

Liu, Y. and E.W. Jacobsen (2001). Effective model reduction for analysis of distributed parameter systems. *Proceedings EFChE ESCAPE11.* Kolding, Denmark.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling (1988). Numerical recipes in C: The art of scientific computing. *Cambridge university Press* p. Ch3.1.