ELSEVIER

IFAC

PUBLICATIONS

ORDER REDUCTION OF DAE MODELS

John D. Hedengren and Thomas F. Edgar

Department of Chemical Engineering
University of Texas at Austin
Austin, TX 78712

Abstract: This report outlines a technique for the order reduction of differential algebraic equations (DAEs). The order reduction is accomplished in two steps. First, algebraic states are partitioned into successive implicit sets of variables and equations by reconstructing the sparsity pattern into lower triangular block form. Second, proper orthogonal decomposition (POD) is used to reduce the number of differential states. As a test case for the theory, a dynamic binary distillation column model is analyzed with the generalized approach. The index 1 DAE model of 52 differential and 178 algebraic states is reduced to an ODE model of 26 differential states. Copyright © 2005 IFAC

Keywords: Order reduction, Model reduction, Nonlinear systems, Function approximation, Distillation columns, Dynamic models

1. INTRODUCTION

Differential algebraic equation (DAE) models of real processes can consist of hundreds of differential equations and thousands of algebraic equations. Order reduction of these models can give insight into the model structure, analysis of dynamic performance, and enable storage and retrieval of optimal control trajectories. Another reason for reduction of the order is to efficiently solve the DAE model in simulation and control applications. However, this last reason is not always valid. Order reduction destroys sparsity in favour of a reduced number of equations and variables. If the sparsity of the original system is exploited in the calculations then there may be little or no numerical advantage of solving a reduced order model.

Model reduction approaches such as balanced covariance matrices (BCM) and proper orthogonal decomposition (POD) have been developed to optimally reduce the number of differential states for

nonlinear models. Although effective for differential state reduction, these model reduction approaches do not reduce the number of algebraic equations. Because the algebraic equations often greatly outnumber the differential states, significant order reduction of the overall model is not achieved by POD and BCM. Physical significance of the variables is lost through model reduction. In many applications it is often desirable or required for a reduced model to retain physical significance of the original variables.

Additional approaches have been suggested for DAE model reduction, but they generally suffer from poor scaling to large scale problems or extensive model configuration (Carpanzano, 2000). The proposed technique in this work has the advantage of good scaling to large scale problems. An added advantage is that the physical significance of the algebraic equations is retained. A major focus of this work is also in making the model reduction approach

adaptive, meaning that the DAE model is reduced without a set of prior training simulations.

2. PREVIOUS WORK

2.1 Reduction of ODEs

Nonlinear ODE models can be reduced by projection methods, proper orthogonal decomposition (POD), balanced covariance matrices (BCM), perturbation methods, and model simplification (Marquardt, 2001). Perturbation methods are useful for models where there is a large separation of time scales allowing the fast dynamics to be eliminated (Vora and Daoutidis, 2001). Even though there are many types of model reduction techniques, few are optimal in some sense. Two optimal nonlinear model reduction approaches are balanced empirical gramians (Hahn and Edgar, 2002) and POD (Zhang, Henson, and Kevrekidis, 2003). Balanced empirical gramians were later found to be a special case of BCM (Hahn and Edgar, 2003). The two step process, in POD and BCM, first computes a similarity transform from step or impulse simulations of the original model. Next, a Galerkin projection constructs the reduced states from a linear combination of the original states. The reduced set of states from balanced covariance matrices is optimal in capturing input to output dynamics of the original nonlinear system. POD is optimal in capturing input to state dynamics. During the Galerkin projection step, the physical meaning of the variables is lost but can be recovered by an inverse transform.

2.2 Reduction of DAEs

Because algebraic equations often greatly outnumber the differential equations, reduction of the differential equations often does little to reduce the overall order (and also computational time) of a DAE model. Some attempts have been made to reduce the algebraic equations in a DAE model. These approaches search for an optimal precedence order and partitioning of the algebraic equations and variables. Obtaining a precedence order and partitioning can be done using a manual directed graph (digraph) as a graphical approach, using matrix methods to produce a block diagonal lower matrix (Tarjan, 1972), or through tearing (Borutsky and Cellier, 1996). These techniques attempt to maximize the number of algebraic variables that can be solved explicitly. However, this problem is NP complete, meaning that all possible combinations of variables must be attempted to find a maximum set of explicit equations (Carpanzano, 2000). Another technique for order reduction is through relaxation of the algebraic states (Otter, et. al., 1996). Relaxation directly generates a Gaussian elimination scheme when the algebraic equations are linear or made linear. The drawback to this technique is that the linearized algebraic equations will only be locally

accurate. Efficient algebraic reduction is a major obstacle to overall model reduction of DAEs.

3. REDUCTION OF ALGEBRAIC EQUATIONS

DAEs consist of differential equations and algebraic equations. In the general form, the DAE problem is as follows

$$f_{DAE}(\dot{z}, z, t) = 0 \qquad (1)$$

where z is a vector of variables and t is a scalar. The DAE is nonlinear when the vector f is a nonlinear function of the derivative of z, z, or t. In order for the problem to be a DAE, at least one of the coefficients for the derivative of z must be zero. The DAE can be grouped into differential equations (f_{ODE}) and algebraic equations (f_{AE}). The variables are also divided into differential variables (x) and algebraic variables (y).

$$f_{ODE}(\dot{x}, x, y, t) = 0 \qquad (2)$$
$$f_{AE}(x, y, t) = 0$$

Typically, the DAE equation residuals are time invariant and t can be dropped from the general equation form. However, it is included in subsequent derivations for the sake of generality. Because the implicit solution of a large number of algebraic equations is computationally expensive, variables that can be solved explicitly are removed from the set y (Hangos and Cameron, 2001). However, some algebraic variables cannot be explicitly solved or solved independently of other variables. The following techniques can be used to automatically identify independent sets of variables and equations.

3.1 Pairing Variables to Equations

Variables and equations are paired by rearranging the sparsity matrix to a maximum transversal. All variables are paired to equations when problem is completely specified (no degrees of freedom) and the maximum transversal is a zero-free diagonal (Duff, 1981). The sparsity matrix (J), also known as the incidence matrix, is generated by identifying the variables that are contained in each equation. The differential states are specified for an initial value problem (IVP). The remaining variables are the differential state derivatives and the algebraic variables.

$$J_{ij} = \begin{cases} 1 & \text{if } y_j \text{ or } \dot{x}_j \text{ appears in } f_{DAE\,i} \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

Computing a zero-free diagonal involves changing the equation or variable orders. The order of the row and column are then matched to give a variable/equation pairing.

3.2 Lower Triangular Block Form

The method proposed in this work differs from previous work by analyzing a dependency matrix M_D instead of the incidence matrix J (Duff and Reid, 1978). It will be shown that M_D can reveal more information about variable dependencies. The dependency matrix M_D is derived by first linearizing the DAE.

$$
\begin{aligned}
A\dot{x}' + Bx' + Cy' + \alpha t' &= 0 \\
Dx' + Ey' + \beta t' &= 0
\end{aligned}
\qquad (4)
$$

The prime indicates deviation from reference values. A, B, C, D, and E are coefficient matrices and α and β are coefficient vectors. The reference values are selected to give non-zero coefficients for the deviation variables. Because the selection of reference values is arbitrary, the coefficients can be arbitrarily selected to be 1 if the equation contains the variable and 0 otherwise. In this case the matrix E is equivalent to the incidence matrix J. Rearranging and combining the linear differential and algebraic equations results in the following matrix form:

$$
\begin{bmatrix} A & C \\ 0 & E \end{bmatrix} \begin{bmatrix} \dot{x}' \\ y' \end{bmatrix} = - \begin{bmatrix} B & \alpha \\ D & \beta \end{bmatrix} \begin{bmatrix} x' \\ t' \end{bmatrix}
\qquad (5)
$$

The dependency matrix M_D reveals the solution dependencies among the linearized equations.

$$
M_D = \begin{bmatrix} A & C \\ 0 & E \end{bmatrix}^{-1}
\qquad (6)
$$

The variable dependency information in M_D can be illustrated by a linear system of $Ax = b$. When A is invertible, the solution to x is $A^{-1}b$. Each element of the vector x is computed from the corresponding row of A_{ij}^{-1} and the vector b.

$$
x_i = \sum_j A_{ij}^{-1} b_j
\qquad (7)
$$

However, the solution to x_i is independent of b_j if $A_{ij}^{-1} = 0 \ \forall \ j \neq i$. If x_i is independent of b_j then it is also independent of equation j. The dependencies in the linear system also apply to the corresponding nonlinear system. Therefore, linearizing the DAE reveals the structure of the nonlinear system dependencies.

The matrix M_D can be converted to lower triangular block diagonal form with Tarjan's algorithm (Tarjan, 1972). Each block along the diagonal is a set of algebraic equations that require a simultaneous solution. The reduction of algebraic equations occurs by explicitly solving for independent groups of equations. Test cases with moderate sized DAE systems (<300 states) show that many of the

equations included in the implicit set can be transformed for explicit calculation. Once an algebraic variable is explicitly calculated, it can be removed from the model as a variable that the solver must calculate. Explicit approximations to implicit solutions can be attempted to further reduce the DAE order as Bosley (1994) did for batch distillation.

3.3 Application to Large Scale Problems

Another consideration relevant to large scale DAE systems is the computational time that is required for the analysis. In this paper, n and τ are the order of the matrix and the number of non-zeros, respectively. The maximum transversal algorithm has a worst case bound of $O(n\tau)$ although typical examples are more like $O(n) + O(\tau)$ (Duff, 1981). The lower block triangular algorithm also exhibits excellent scaling for large problems with an upper bound of $O(n) + O(\tau)$ (Duff and Reid, 1978).

4. ADAPTIVE REDUCTION OF DIFFERENTIAL EQUATIONS

Non-adaptive model reduction requires special training simulations that should adequately cover the state space of interest for an accurate reduced model. When the training data set does not cover the entire nonlinear region of interest, the singular values may be a poor indication of reduced model accuracy outside of the training domain. The most important tuning parameter, variable accuracy, is not directly controlled through the non-adaptive approach. Adaptive reduction adjusts the form and order of the model as new simulations become available. The order of the model is controlled to maintain a required variable accuracy. Adaptive reduction also eliminates the need for special training simulations. To adaptively reduce the order of the differential equations, a measure of the reduced model accuracy must be introduced. One possible solution is to directly solve the full model and reduced model at various checkpoints to determine the accuracy of the reduced model (Ryckelynck, 2004). Another option that avoids the periodic solution of the full model is to control the equation residuals. For linear systems at steady state, the equation residuals and variable residuals are exactly related. A linearized model is used to predict the variable residuals from the equation residuals.

4.1 Model reduction in the Semi-Explicit Equation Format

A Galerkin projection maps a full set of variables onto a reduced set of variables that make up the reduced model. For balanced covariance matrices, the Galerkin projection is a set of vectors that optimally captures the highest degree of input-output

dynamics. For POD, the Galerkin projection is a set of orthogonal vectors but captures the highest degree of input-state dynamics. Each successive vector is the direction that maximizes the amount of variance in the model states while maintaining orthogonality to the previous directions. For the semi-explicit ODE model form

$$\dot{x}(t) = f(x(t)) \tag{8}$$

the Galerkin projection ($\widetilde{P}$) is applied by defining a reduced set of variables

$$x(t) = \widetilde{P}^T \overline{x}(t) + r(t) \tag{9}$$

where $r(t)$ is a state residual to account for the inaccuracy of the reduced model. The reduced model exactly represents the original model when the residual is retained.

$$\dot{\overline{x}}(t) = \widetilde{P}f(\widetilde{P}^T \overline{x}(t) + r(t)) + \widetilde{P}\dot{r}(t) \tag{10}$$

At this point the residual and its derivative are typically ignored and some of the system dynamics are necessarily lost due to the reduced order of the model.

4.2 Estimate of the Variable Error from the Equation Residuals

Ideally, one would like to adjust the order of the reduced model to control the variable errors directly. Barring simultaneous solution of the reduced and full order model, the variable error cannot be directly calculated for nonlinear systems. The next best option is to estimate the variable residual ($r(t)$) from the equation residual ($R(t)$).

$$\dot{\overline{x}}(t) = \widetilde{P}f(\widetilde{P}^T \overline{x}(t)) + R(t) \tag{11}$$

When the system is linear, the equation residuals are related to the variable residuals by the state matrix (A).

$$\dot{x}(t) = Ax(t) \tag{12}$$

with

$$x(t) = \widetilde{P}^T \overline{x}(t) + r(t) \tag{13a}$$
$$\dot{x}(t) = \widetilde{P}^T \dot{\overline{x}}(t) + \dot{r}(t) \tag{13b}$$

the linear reduced model becomes

$$\widetilde{P}^T \dot{\overline{x}}(t) = A\left(\widetilde{P}^T \overline{x}(t)\right) + Ar(t) - \dot{r}(t) \tag{14}$$

The equation residual and variable residual are related to each other by the state matrix and the variable residual derivative.

$$R(t) = Ar(t) - \dot{r}(t) \tag{15}$$

By assuming that the variable residual is locally constant, the variable residual derivative term can be ignored and an estimate of the variable residual can be obtained.

$$\hat{r}(t) = A^{-1}\big(R(t)\big) \tag{16}$$

By linearizing the nonlinear model, an estimate of the variable residuals can be obtained from the equation residuals. The predictive capability of this relation for nonlinear models depends on the severity of nonlinearity and closeness to the point of linearization.

4.3 Model Reduction in the Open Equation Format

A semi-explicit ODE model is a restricted form of the more general open equation format.

$$f(\dot{x}, x) = 0 \tag{17}$$

Applying the Galerkin projection to the open equation format changes the solution procedure. By reducing the number of variables and maintaining the same number of equations, extra degrees of freedom arise. Physically, this is the result of giving up some of the least important dynamic degrees of freedom. The reduced order model is solved by minimizing the residuals instead of finding equation roots.

$$R = f(\widetilde{P}^T \dot{\overline{x}}, \widetilde{P}^T \overline{x}) \tag{18}$$

The minimized equation residuals will generally be small for a good reduced order model. As the order of the reduced model is decreased, the equation residuals will generally increase. The differential equation model reduction approach is made adaptive by increasing or decreasing the number of states in the reduced model to meet the required variable tolerances.

5 APPLICATION STUDY

A binary distillation column model is employed to show a practical application of the DAE to ODE model conversion. The model is composed of the MESH (Material balances, Equilibrium stages, Summation of fractional quantities, and Heat balances) equations, and has 23 stages, a condenser, sump, and reboiler. The DAE model has 52 differential equations and 178 algebraic equations. The independent variables are:

$$\begin{bmatrix} \dot{x} \\ y \end{bmatrix} = \begin{bmatrix} \dot{x}_A & \dot{h} & | & y_A & x_L & T & \dot{n}_V \text{ or } \dot{n}_L & h_V & h_L & P_A^{sat} & P_B^{sat} \end{bmatrix}^T$$

$$\tag{19}$$

During the linearization step, the reference values are selected to give non-zero coefficients for the deviation variables. Because the reference values are arbitrary, the non-zero coefficients are shown by X if the equation contains the variable and 0 otherwise. The matrix $\mathbf{M}$ is a 230 x 230 matrix that has variables lumped together for display.

$$\mathbf{M}\begin{bmatrix}\dot{\mathbf{x}}\\\mathbf{y}\end{bmatrix} \rightarrow \left[\begin{array}{cc|cccccccc}X & 0 & X & 0 & 0 & X & 0 & 0 & 0 & 0\\0 & X & 0 & 0 & 0 & X & X & X & 0 & 0\\\hline 0 & 0 & X & 0 & 0 & 0 & 0 & 0 & X & 0\\0 & 0 & 0 & X & 0 & 0 & X & X & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X & X\\0 & 0 & 0 & X & 0 & X & 0 & 0 & 0 & 0\\0 & 0 & X & 0 & X & 0 & X & 0 & 0 & 0\\0 & 0 & 0 & 0 & X & 0 & 0 & X & 0 & 0\\0 & 0 & 0 & 0 & X & 0 & 0 & 0 & X & 0\\0 & 0 & 0 & 0 & X & 0 & 0 & 0 & 0 & X\end{array}\right]\begin{bmatrix}\dot{\mathbf{x}}_A\\\dot{\mathbf{h}}\\\mathbf{y}_A\\\mathbf{x}_L\\\mathbf{T}\\\dot{\mathbf{n}}_V \text{ or } \dot{\mathbf{n}}_L\\\mathbf{h}_V\\\mathbf{h}_L\\\mathbf{P}_A^{sat}\\\mathbf{P}_B^{sat}\end{bmatrix} \quad (20)$$

The non-zero values of $\mathbf{M}^{-1}$ show the dependencies between the variables and equations. The non-zero values of $\mathbf{M}^{-1}$ in lower triangular block diagonal form are shown below with the corresponding variable order.

$$\mathbf{M}^{-1} = \left[\begin{array}{cccccccc|cc}X & X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0\\X & X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0\\X & X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0\\X & X & X & X & 0 & 0 & 0 & 0 & 0 & 0\\X & X & X & 0 & X & 0 & 0 & 0 & 0 & 0\\X & X & X & 0 & X & X & 0 & 0 & 0 & 0\\X & X & X & X & X & X & X & 0 & 0 & 0\\X & X & X & X & X & X & X & X & 0 & 0\\\hline X & X & X & X & X & X & X & X & X & 0\\X & X & X & X & X & X & X & X & 0 & X\end{array}\right] \quad (21)$$

$$\begin{bmatrix}\mathbf{y}\\\dot{\mathbf{x}}\end{bmatrix} \rightarrow \begin{bmatrix}\mathbf{T}\\\mathbf{P}_A^{sat}\\\mathbf{P}_B^{sat}\\\mathbf{h}_L\\\mathbf{y}_A\\\mathbf{h}_V\\\mathbf{x}_L\\\dot{\mathbf{n}}_V \text{ or } \dot{\mathbf{n}}_L\\\hline \dot{\mathbf{x}}_A\\\dot{\mathbf{h}}\end{bmatrix} \quad (22)$$

The first three rows of $\mathbf{M}^{-1}$ indicate that T, P_A^{sat}, and P_B^{sat} must be solved simultaneously because the corresponding equations form one block. The equations for P_A^{sat}, and P_B^{sat} can be explicitly substituted into the bubble point temperature equation.

$$P = x_A P_A^{sat}(T) + (1 - x_A)P_B^{sat}(T) \quad (23)$$

For the cyclohexane / heptane binary mixtures, an explicit temperature solution is approximated by a second order polynomial in composition.

$$T = c_1 + c_2 x_A + c_3 x_A^2 \quad (24)$$

The vector c was calculated with a least squares fit with data from the sets $x_A \in \begin{bmatrix}0 & 1\end{bmatrix}$ to be $c = \begin{bmatrix}385.42 & -21.57 & 3.736\end{bmatrix}^T$. The polynomial fit has a mean sample error of 0.012 K and a maximum sample error of 0.04 K.

The molar flow rates can form the last block along the diagonal. Because the molar flow rate equations are linear, they can be solved explicitly. After solving the flow rates the dependency matrix indicates that the differential equation variables can now be solved explicitly. If extraneous algebraic equations were present in the model, they could be identified at this point because the equations for $\dot{\mathbf{x}}_A$ and $\dot{\mathbf{h}}$ have no further dependencies. By explicitly solving all of the algebraic equations, the model is in an ODE form. In this form, nonlinear model reduction techniques can be applied to further reduce the number of differential states.

As an ODE, the distillation column model is available for further model reduction through balanced covariance matrices or proper orthogonal decomposition (POD). POD was chosen for this example and the number of differential states was reduced to 26. Fig. 1 shows the bottoms composition after a 5% increase in reboiler duty.
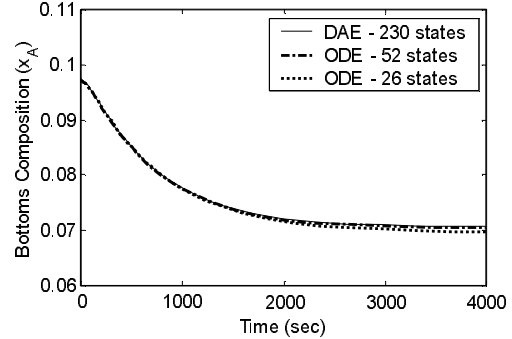

Fig. 1. Results of a 5% step change in reboiler duty. The 26 state reduced model closely approximates the 230 state DAE model.

The ODE model with 52 states approximates the 230 state DAE very well. The ODE model with 26 states also approximates the DAE model well but with a larger offset in the steady-state value of composition. ODE models with fewer than 20 states performed poorly, indicating that there are at least 20 dynamic degrees of freedom in the binary distillation column model.

The CPU time for the step simulations on a 2000 MHz Celeron Processor was 4.3 seconds for the DAE model and 3.5 seconds for the two ODE models. For the step tests, significant computational reduction was not achieved but may be more significant when

the reduced models are applied in optimization such as nonlinear MPC (NMPC).

6 CONCLUSIONS

An adaptive DAE model reduction technique is outlined in this report with special consideration for large scale models. The technique consists of two steps:

1. Algebraic states are partitioned into successive implicit sets of variables and equations by reconstructing the sparsity pattern into a lower triangular block form.
2. POD reduces the number of differential states.

POD is applied to the open equation format by minimizing the equation residuals instead of finding roots. Once a minimized solution is found, the equation residuals provide an estimate of the variable accuracy. The differential variable accuracy is controlled by increasing or decreasing the order of the reduced model. In this way, POD is made adaptive while dynamically constructing the similarity transform. The estimate of the variable accuracy can also be used to improve the reduced model accuracy.

In step tests the reduced order distillation model did not significantly lower the computational time. However, reduced models may be a big advantage in NMPC applications where the large size of the model can make the optimization computationally infeasible. Other motivations for model reduction include insight to the model structure, feasibility of storage and retrieval of NMPC solutions, and analysis of dynamic degrees of freedom.

REFERENCES

Borutsky, W. and F.E. Cellier (1996). Tearing in Bond Graphs with Dependent Storage Elements. *CESA'96, IMACS MultiConference on Computational Engineering in Systems Applications*, Lille, France, 1, pp. 1113-1119.

Bosley, J.R. (1994) An Experimental Investigation of Modeling, Control, and Optimization Techniques for Batch Distillation, *Dissertation*, The University of Texas at Austin.

Carpanzano, E. (2000) Order Reduction of General Nonlinear DAE Systems by Automatic Tearing. *Mathematical and Computer Modelling of Dynamical Systems*, 6, pp. 145-168.

Duff, I.S. and J.K. Reid (1978) An implementation of Tarjan's algorithm for the block triangularization of a matrix, *ACM Transactions on Mathematical Software*, 4, pp. 137-147.

Duff, I.S. (1981) On algorithms for obtaining a maximum transversal, *ACM Transactions on Mathematical Software*, 7, pp. 315-330.

Hahn, J. and T.F. Edgar (2002) An Improved Method for Nonlinear Model Reduction Using Balancing of Empirical Gramians. *Computers and Chemical Engineering*, 26, pp. 1379-1397.

Hahn, J., T.F. Edgar, and W. Marquardt (2003) Controllability and observability covariance matrices for the analysis and order reduction of stable nonlinear systems. *J. Proc. Cont.*, 13, pp. 115-127.

Hangos, K.M. and I.T. Cameron. (2001) *Process Modelling and Model Analysis*, Academic Press, San Diego, CA.

Marquardt, W. (2001) Nonlinear model reduction for optimization based control of transient chemical processes. *Proceedings of the CPC VI*, AIChE Symposium Series, 98, pp. 30-60.

Otter, M., H. Elmqvist, and F.E. Cellier. (1996) "Relaxing" – A Symbolic Sparse Matrix Method Exploiting the Model Structure in Generating Efficient Simulation Code. *CESA'96, IMACS MultiConference on Computational Engineering in Systems Applications*, Lille, France, 1, pp. 1-12.

Ryckelynck, D. (2004) A priori hyperreduction method: an adaptive approach. submitted to *J. Comp. Physics*.

Tarjan, R. (1972) Depth first search and linear graph algorithms. *SIAM Journal on Computing*, 1, pp. 146-160.

Vora, N., P. Daoutidis (2001) Nonlinear model reduction of chemical reaction systems. *AIChE Journal*, 40, pp. 2320-2332.

Zhang, Y., M.A. Henson, Y.G. Kevrekidis (2003) Nonlinear order reduction of discretized cell population models. *Proceedings of the ACC*, pp. 2383-2388.