

COMPARISON OF SEVERAL SYSTEM IDENTIFICATION METHODS FOR FLEXIBLE STRUCTURES

J.-S. LEW

*Center of Excellence in Information Systems, Tennessee State University, Nashville,
Tennessee 37203-3401, U.S.A.*

J.-N. JUANG

*Spacecraft Dynamics Branch, M.S. 297, NASA Langley Research Center, Hampton,
Virginia 23681-0001, U.S.A.*

AND

R. W. LONGMAN

*Department of Mechanical Engineering, Columbia University, New York,
New York 10027, U.S.A.*

(Received 25 April 1991, and in final form 18 May 1992)

In the past few years various methods of identifying structural dynamics models from modal testing data have appeared. This paper presents a comparison of four of these algorithms: the Eigensystem Realization Algorithm (ERA), the modified version ERA/DC, where DC indicates that it makes use of data correlations, the Q-Markov cover algorithm, and an algorithm due to Moonen, DeMoor, Vandenberghe and Vandewalle (MDVV). The comparison is made using a five-mode computer model of the 20 m Mini-Mast truss structure at the NASA Langley Research Center, and various noise levels are superimposed to produce simulated data. The results show that, for the example considered, ERA/DC generally gives the best results; that ERA/DC is always at least as good as ERA, which is shown to be a special case of ERA/DC; that Q-Markov requires the use of significantly more data than ERA/DC to produce comparable results; and that in some situations Q-Markov cannot produce comparable results. The MDVV algorithm can avoid pre-processing of data to obtain Markov parameters, but requires significantly larger matrix singular value decomposition for large data sets.

1. INTRODUCTION

In recent years, several time domain system identification techniques from the control theory community have become useful in structural modal testing, and still others have appeared which might offer advantages for modal testing. The purpose of this paper is to perform a comparison of four of these closely related identification methods on a typical structural dynamics identification problem. The aim is to make an initial determination of what the advantages and disadvantages of each method are, and under various different conditions, determine which method is best to use.

In recent years, the Eigensystem Realization Algorithm [1, 2] has seen considerable use on flexible spacecraft identification. This is one of the four closely related algorithms compared. An extension of this algorithm using data correlation (ERA/DC) is also studied (reference [3]). The third approach is based on Q-Markov cover theory, originally developed for model reduction [4, 5]. It was generalized to identification using both white

noise input tests and impulse input tests in reference [6], and then finally reformulated for tests using either type of input in reference [7]. Although obtained from a different starting point, the resulting mathematical algorithm bears a family resemblance to ERA/DC. The fourth algorithm studied is that of reference [8], which allows direct use of general input tests. This algorithm will be referred to as MDVV, using the first letter of each author's last name. A mathematical description of each algorithm is given in Appendix A.

2. THE TEST EXAMPLE

One must use a mathematical or computer model rather than a real physical structural test article to compare the performance of the identification algorithms, since one must know what the correct identification result is. Here we pick a model obtained by finite element analysis of the Mini-Mast [9] structure illustrated in Figure 1. This model is sufficiently complex to be a reasonable test of identification performance. The model for testing considers the first two bending modes as the lowest frequency modes with essentially identical frequencies, the first torsional mode, and the second two bending modes, again with essentially identical frequencies. These frequencies, and the associated damping expressed as the negative of the real part of the eigenvalues, are given in Table 1. The mathematical model is given in Appendix B.

The mathematical model considers two inputs and two outputs. The outputs are from non-orthogonal displacement sensors at the top of the structure, and the inputs are two torque wheels for the x and y axes, as shown in Figure 1.

3. THE APPROACH USED IN THE COMPARISON

It is not a simple matter to determine a set of tests that constitute a reasonable comparison of the different identification approaches. Each approach has a somewhat different set of parameters that the user can adjust in order to obtain good performance.

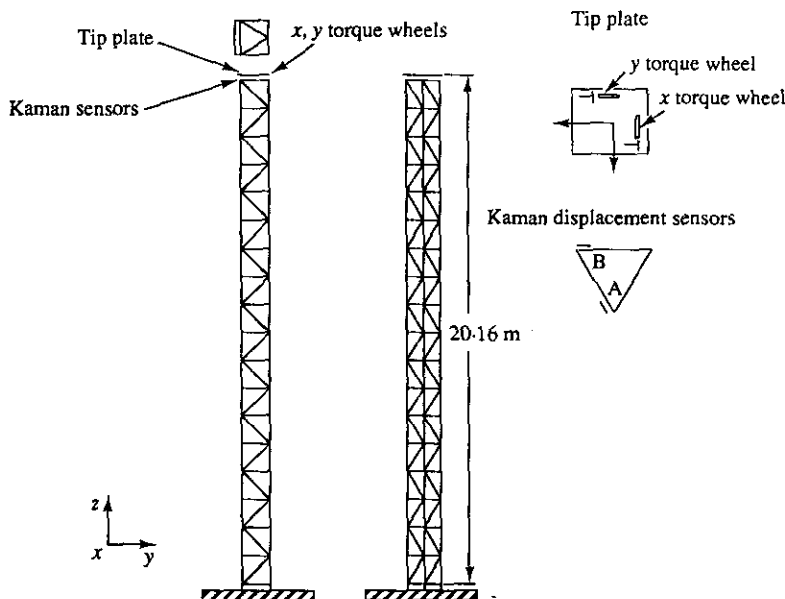


Figure 1. Mini-Mast test article showing the x and y torque wheel locations, the A and B displacement sensor locations and their axes on the bay 18 tip plate.

TABLE I
Modal information

Mode no.	Damping factor	Frequency	
		rad/s	Hz
1	0.09059	5.032	0.8008
2	0.09066	5.036	0.8014
3	0.32907	27.420	4.3640
4	0.38352	38.351	6.1038
5	0.38683	38.682	6.1565

It would be of interest to optimize these parameters so that in each case one compares the best results that each algorithm can produce. However, usually it is a matter of experience and intuition to decide how to make these parameter choices. In the case of ERA, methods based on perturbation theory have recently become available which help the ERA user know the accuracy of his identification results, and hence help him to adjust these parameters in order to improve the accuracy [10–12]. Similar results have been generated for ERA/DC and the Q-Markov cover approaches, and will be reported in the literature. No attempt is made here to optimize these parameter choices. The choices made are as follows.

In Q-Markov, the measurement matrix used for singular value decomposition is necessarily square, but in ERA and ERA/DC the measurement matrix can be rectangular, and in fact certain results from other sources [12] suggest that rectangular data matrices for ERA give better results. Nevertheless, for ease of comparison, we restrict ERA and ERA/DC to square data matrices and recognize that their identification accuracy may be degraded. Also, in ERA/DC there is considerable freedom in the choice of the correlation matrix $\mathcal{U}(q)$, but here we restrict ourselves to the special case in which $\mathcal{U}(q) = H(q)H^T(0)$ (see Appendix A). During the comparison, the size of the measurement matrix used for singular value decomposition for each algorithm is kept the same (at 120×120), with the idea that the singular value decomposition in each algorithm is the most time consuming step. However, this makes the comparison of ERA with ERA/DC and with Q-Markov somewhat biased against ERA. Both ERA/DC and Q-Markov involve a parameter, s and d respectively, which determines the upper limit on a possibly large summation of products that does not appear in ERA. Hence, with the same computation, ERA could have used a larger data matrix which could produce improved identification results. Some computer runs were made to study this. A summary of each of the algorithms considered is given in Appendix A. Note that equal values of the parameters s and d in ERA/DC and Q-Markov imply essentially equivalent amounts of computation in setting up the respective data matrices.

The test example considered is a multiple-input and multiple-output one. As a result, it is of interest to have a scalar measure of the accuracy of the identification results to aid the comparisons of the various approaches. We accomplish this by considering the maximum singular value of the transfer function matrix of the identification error as a function of frequency. In other words, we take the maximum singular value at each frequency, of the true transfer function matrix of the computer test example minus the corresponding transfer function matrix of the identified system model. The symbol $\bar{\sigma}$ is used to indicate maximum singular value. This scalar measure gives a simple way to express the combined error in natural frequencies, damping and mode shapes of the identified

model. Note that this error criterion is used to design the model uncertainty weighting matrix for H_∞ control design [13].

4. RESULTS

4.1. NOISE FREE RESULTS

Identification results for ERA, ERA/DC and Q-Markov when the pulse response data is noise free are given in Figure 2. The measurement sampling interval of all the examples in this paper is 0.03 s. In Figure 2(a) is given $\bar{\sigma}$ of the true system transfer function matrix for comparison, and the $\bar{\sigma}$ of the error transfer function matrix for ERA, and for ERA/DC with the parameter s chosen as 200, corresponding to 260 data points. The smaller the values of $\bar{\sigma}$ of the error transfer function matrix curves, the better the identification results. It is shown in Figure 2(a) that both ERA and ERA/DC give essentially perfect identification, with the errors in the identification limited to the magnitude of numerical round-off. Figure 2(b) repeats Figure 2(a) but for the Q-Markov algorithm. The similarity of the true and the $d = 500$ curves in the second and third peaks means that these modes are not identified.

We observe that the Q-Markov algorithm requires the use of very large data sets (e.g., d significantly greater than 2000) before it can approach the accuracy of ERA (with s equal to 59) and ERA/DC (with s equal to 200) on noise-free data. The explanation of this phenomena is given later in section 4.5. The same observation can be seen in a different manner in Figure 2(c). Only the first ten singular values should be non-zero, but the remaining singular values for the Q-Markov algorithm do not reduce to the level of numerical round-off until d is very large.

4.2. INTERPRETATION OF PLOTS OF MAXIMUM SINGULAR VALUE OF THE ERROR TRANSFER FUNCTION MATRIX

We are using the maximum singular value of the model error transfer function matrix, as a function of frequency, as our scalar measure of the model error. It is of interest to interpret these plots in terms of errors of frequency and damping of the modes in the model. In Figure 3(a), we show the maximum singular value of the error transfer function matrix when the model and the true system are identical, except that the damping (i.e., the negative of the real part of the eigenvalues) of the first two modes is decreased by 0.1%, 1% and 10%. The same computation was repeated using frequency error instead of damping error. When the frequency errors are $10^{-3}\%$, $10^{-2}\%$ and $10^{-1}\%$, we obtain Figure 3(b), which looks nearly identical to that of Figure 3(a) but with slightly lower peaks. Hence, the $\bar{\sigma}(w)$ peak is approximately 100 times more sensitive to frequency than it is to damping error, and the signatures of such errors are very similar.

4.3. THE RELATIONSHIP BETWEEN ERA AND ERA/DC

In Appendix C.1 we prove that ERA is a special case of ERA/DC in the following sense. Use the same Hankel matrix of pulse response data $H(0)$ in both algorithms. Pick the special case of $\mathcal{U}(0) = R(0)$ in equation (A5) of Appendix A.2 for the singular value decomposition matrix in ERA/DC. Pick the same order k for the desired identified model from each algorithm. Then, the pulse response history of the k th order ERA model is identical to the pulse response of the k th order ERA/DC model, for all choices of k , independent of the true order of the system. Hence the identified models are identical to within a linear transformation of state variables.

ERA/DC with $s = 200$ is shown in Figure 2(a). If s is set to 59 in order to comply with the above assumptions in comparing ERA/DC to the ERA results, the $\bar{\sigma}$ error curve in

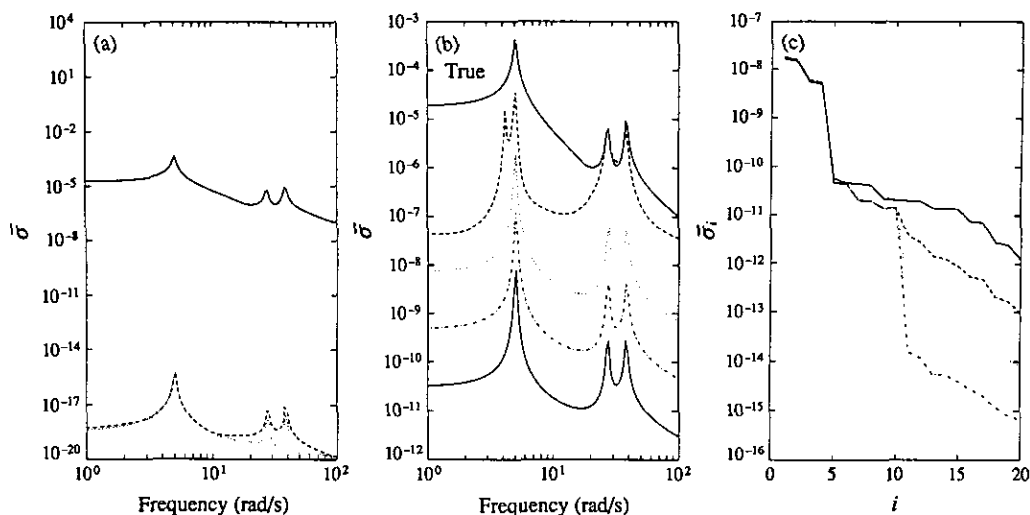


Figure 2. Noise free identification results: (a) Maximum singular value of true transfer function matrix vs. frequency, and of the model error transfer function matrices for ERA and ERA/DC ($s = 200$) vs. frequency: —, true; —, ERA/DC; ···, ERA. (b) Maximum singular value of true transfer function matrix vs. frequency, and of the Q-Markov model error transfer function matrices vs. frequency for various d : —, true; —, $d = 500$; ···, $d = 1000$; —·—, $d = 1500$; — — —, $d = 2000$. (c) Singular values of measurement matrix for Q-Markov identification for various d : —, $d = 500$; —·—, $d = 1000$; ···, $d = 1500$; — — —, $d = 2000$.

Figure 2(a) is similar to those shown, with differences again on the level of numerical round-off in the computations. This is consistent with the above theoretical result.

Note that unlike ERA, ERA/DC has the ability to increase the number of Markov parameters used in the identification, without increasing the dimension of the singular value decomposition matrix, simply by increasing the number of columns in $H(0)$. Since storage limitations constrain the size of the decomposition matrix, this extra freedom can be an important advantage of ERA/DC.

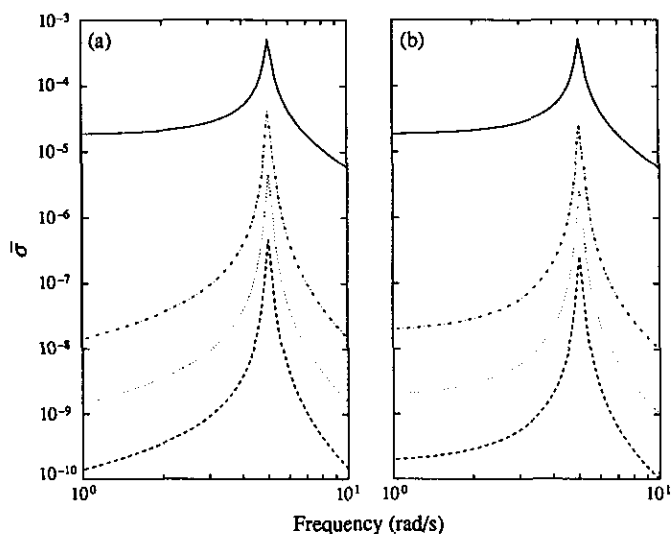


Figure 3. Maximum singular value of the model error transfer function matrix as a function of frequency: (a) for various damping errors in the first two modes (—, true; —, 0.1%; ···, 1%; —·—, 10%); (b) for various frequency errors in the first two modes (—, true; —, $10^{-3}\%$; ···, $10^{-2}\%$; —·—, $10^{-1}\%$).

4.4. THE RELATIONSHIP BETWEEN ERA AND Q-MARKOV COVER ALGORITHMS

The relationship between ERA/DC and the Q-Markov cover algorithms is developed in Appendix C.2. ERA/DC uses a singular value decomposition of a matrix $\mathcal{U}(0)$ in equation (A6) to produce its realization, while Q-Markov uses singular value decomposition of the matrix D_q in equation (A11). Appendix C.2 shows that D_q can be produced from $\mathcal{U}(0)$ by setting certain Markov parameters to zero in $\mathcal{U}(0)$ (see equations (C8) and (C9) in Appendix C.2).

As the parameter d in Q-Markov increases, the number of Markov parameters set to zero becomes a smaller percentage of the total number of Markov parameters included in the identification. They also refer to later sample times, which allows for more decay of the pulse response.

The ERA/DC results for $s = 59$ and $s = 200$, which both have errors limited to the level of numerical round-off, suggest that ERA/DC gives accurate system realization from noise-free data for all r, s above the values needed to give row and column numbers of the Hankel matrix that are greater than the system order. Numerical simulations from the Q-Markov algorithm illustrate that it produces an accurate realization of the input-output relation only asymptotically, as the amount of data used tends to infinity, i.e., $d \rightarrow \infty$. This is consistent with the analytical results in Appendix C.2; once d is sufficiently large that the Markov parameters that are set to zero in the Q-Markov algorithm (equations (C8) and (C9)) become negligible, then Q-Markov produces results equivalent to ERA/DC with model errors reduced to the level of numerical round-off. This suggests that ERA/DC should give superior performance to Q-Markov cover for all finite data sets. The Q-Markov algorithm is not valid for finite data sets, whereas ERA/DC gives exact answers for finite noise free data sets.

4.5. A NOTE ON THE COMPUTATIONAL EFFORT OF EACH ALGORITHM

The ERA, ERA/DC, and the Q-Markov cover algorithms are compared by constraining the singular value decomposition in each case to be of the same dimension, with the idea that the effort involved in this decomposition dominates the computational effort of the algorithm. Such a decomposition is roughly a cubic function of the matrix dimension [14]. ERA/DC and Q-Markov both require an additional multiplication of matrices to obtain the matrix which is decomposed. Such a multiplication is generally cubic in the matrix dimension as well. However, the special structure of the data matrices allows one to generate a recursive procedure to simplify the computation. It produces a computation count that is quadratic in the matrix dimensions, so that the decomposition does in fact dominate.

It is difficult to compare computational effort of the different identification methods, because the details of the programming can be significant. Here, we cite the CPU times used in running the different algorithms. All computations were carried out using MATLAB on a Micro VAX 3200 workstation, running the high noise case discussed below. In Table 2 are given the results for $n = 120$ (the number of rows of H_q , or $H(d)$ in equation (C9) of Appendix C.2, or $H(0)$) in all cases, and for various values of s or d , where $s + 1$ and $d + 1$ are the number of columns of Markov parameters in $H(0)$ and $H(d)$ respectively. The actual number of columns is $2(s + 1)$ and $2(d + 1)$, because the system has two inputs.

Increasing s from 100 to 1500 in ERA/DC increased the CPU time by 24 s, which represents a 26% increase in run time to include roughly 15 times as much data. The Q-Markov algorithm took slightly longer than ERA/DC for corresponding values of d and s . Also, as d is increased from 100 to 1500, the CPU time increased by 40 s, representing a 41% increase in run time. Note, however, that a much larger value of d is required in

TABLE 2

*Comparison of CPU times (s) for ERA, ERA/DC and Q-Markov cover
($n = 120$ in all cases)*

ERA				
	$s = 59$ 90	$s = 79$ 120	$s = 99$ 150	
ERA/DC				
$s = 100$ 92	$s = 300$ 96	$s = 500$ 100	$s = 1000$ 110	$s = 1500$ 114
Q-Markov				
$d = 100$ 98	$d = 300$ 110	$d = 500$ 117	$d = 1000$ 128	$d = 1500$ 138

Q-Markov than of s in ERA/DC for comparable identification accuracy. Q-Markov with $d = 1500$ and ERA/DC with $s = 500$ give roughly comparable identification results, and Q-Markov takes 38% longer to run. We conclude that, at least for the programming implementations of the algorithm used, ERA/DC has a substantial advantage over Q-Markov in computer time needed.

The ERA algorithm requires the singular value decomposition of a rectangular Hankel matrix $H(0)$ rather than of a square matrix such as $H(0)H^T(0)$, one dimension of which is smaller. The increased dimension proportional to $(s + 1)$ causes memory problems which prevented use of $s = 300$ or above. In order to include additional data, one must convert to ERA/DC above this value. Since the CPU time in ERA is proportional to $(s + 1)$, we would expect a CPU time of 2252 s for $s = 1500$ if memory problems had not prevented such a computation. We conclude that ERA suffers from memory limitations, and that ERA/DC is to be preferred. Since ERA is a special case of ERA/DC in the prescribed sense given above, ERA/DC can always give as good results as ERA, and the extra parameters in ERA/DC can be optimized (e.g., by use of the methods of references [10–12]) to give better results.

4.6. LOW NOISE RESULTS

A low noise level is introduced in the data in Figure 4. The noise is modelled as additive, white, zero-mean Gaussian with standard deviation equal to 8×10^{-9} . Note that the signal initial root mean square value is 1×10^{-6} .

In Figures 4(b) and 4(d) it is again shown that the Q-Markov method needs a long data set, e.g., $d = 1500$, before all modes are clearly identified. It is shown in Figures 4(a) and 4(c) that the same level of identification error is achieved by ERA/DC with many fewer data points, e.g., for $s = 400$, a number too small for successful identification with Q-Markov. It is interesting to note in Figures 4(a) and 4(b) that the drop in the singular values, which is an indicator of the true system order, becomes worse (i.e., smaller) as either d or s is increased. ERA/DC is again better for corresponding values of d and s . This behavior can be explained by examining the time constants for decay of each mode. For the highest frequency mode, four time constants is 10.34 s, which corresponds to a decay of the amplitude of the mode to 0.018 of its original value. When $s = 800$, the last data entry in the ERA/DC Hankel matrix corresponds to 25.8 s, well after the highest frequency mode has disappeared into the noise level. Including more data points adds only noise and no signal as far as identification of the high frequency modes are considered. This added

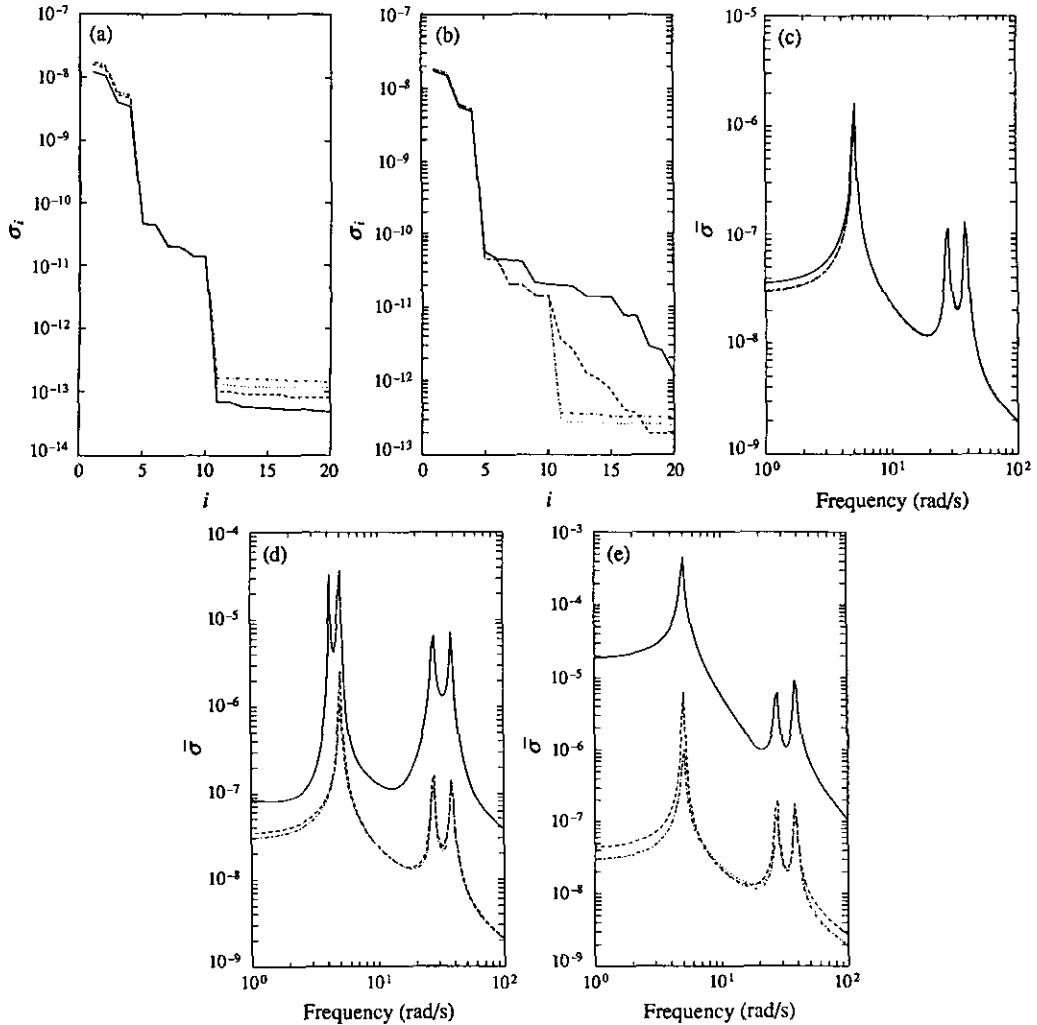


Figure 4. Low noise identification results: (a) Singular values of measurement matrix for ERA/DC identification for various s : —, $s = 200$; —, $s = 400$; ···, $s = 600$; —, $s = 800$. (b) Singular values of measurement matrix for Q-Markov identification for various d : —, $d = 500$; —, $d = 1000$; ···, $d = 1500$; —, $d = 2000$. (c) Maximum singular value of the ERA/DC model error vs. frequency for various s : —, $s = 200$; —, $s = 400$; ···, $s = 600$; —, $s = 800$. (d) Maximum singular value of the Q-Markov model error vs. frequency for various d : —, $d = 500$; —, $d = 1000$; ···, $d = 1500$; —, $d = 2000$. (e) Comparison of maximum singular value of true transfer function matrix and of the model error transfer function matrices for ERA, ERA/DC ($s = 800$), and Q-Markov ($d = 2000$): —, true; —, ERA; ···, Q-Markov; —, ERA/DC.

noise is reflected in higher singular values associated with noise, i.e., singular values 11 and above. On the other hand, four time constants of the lowest frequency mode is 44.16 s, and adding data above $s = 800$, i.e., after 25.8 s, still helps the identification of this mode, as is reflected in the increase of the first singular value with s . Obviously, there are different optimal values for s for different modes, generally with smaller values of s for higher frequency modes.

In Figure 4(e) it is shown that both ERA/DC and Q-Markov, using large data histories associated with $s = 800$ and $d = 2000$, somewhat outperform ERA with its much shorter 119 point data history.

4.7. HIGH NOISE RESULTS

Figure 5 repeats Figure 4 but with the noise level increased by a factor of 10. In Figure 5(b) it is shown that the Q-Markov algorithm fails to exhibit the drop in singular values at the true system order of 10, for all choices of d . On the other hand, ERA/DC exhibits a drop at the true system order for lower values of s . One can interpret this in terms of the damping in the system. The Q-Markov algorithm appears to inherently need more data points than ERA/DC, but as more data points are taken, the decay of the true system modes causes the extra data, for times after significant decay has occurred, to have poor signal to noise ratio. This increases the level of the singular values associated with noise, and decreases the step change in the singular values at the true system order.

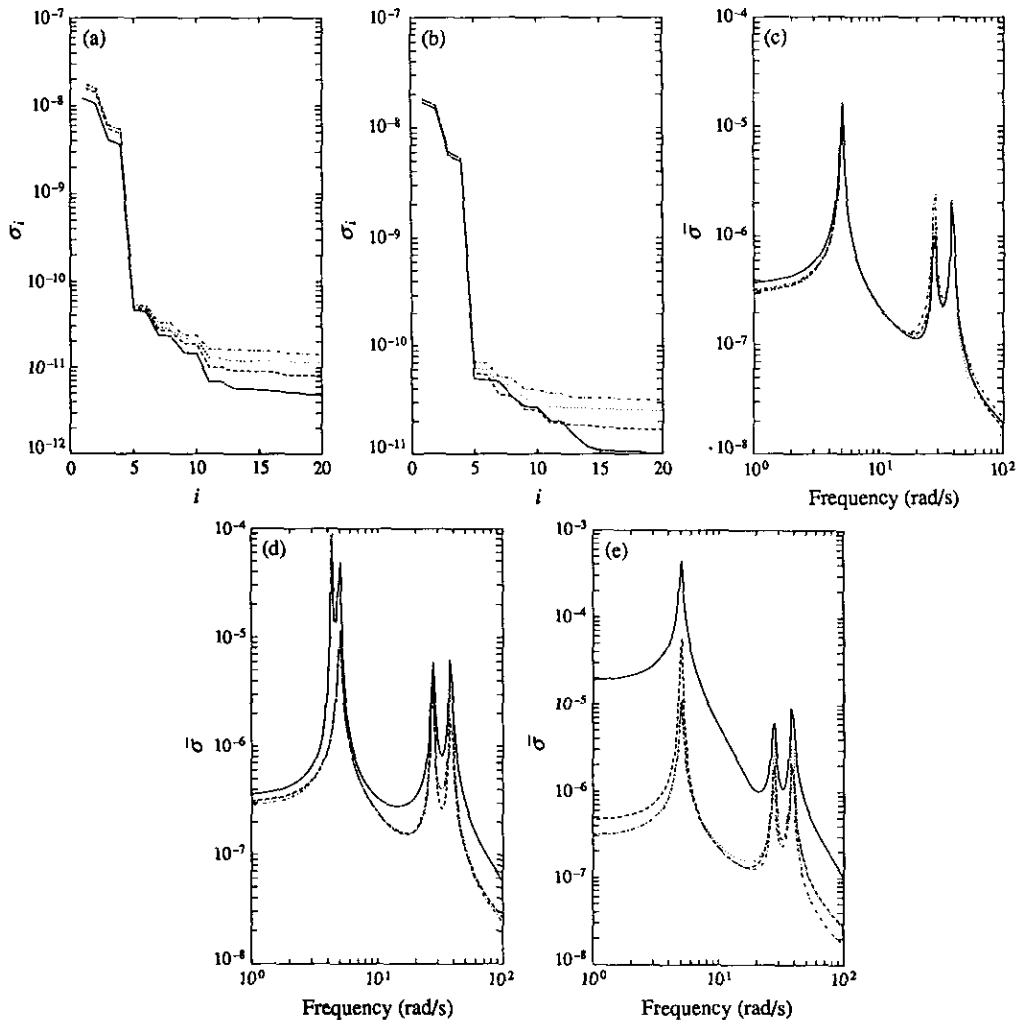


Figure 5. High noise identification results: (a) Singular values of measurement matrix for ERA/DC identification for various s : —, $s = 200$; —, $s = 400$; ···, $s = 600$; —, $s = 800$. (b) Singular values of measurement matrix for Q-Markov identification for various d : —, $d = 500$; —, $d = 1000$; ···, $d = 1500$; —, $d = 2000$. (c) Maximum singular value of the ERA/DC model error *vs.* frequency for various s : —, $s = 200$; —, $s = 400$; ···, $s = 600$; —, $s = 800$. (d) Maximum singular value of the Q-Markov model error *vs.* frequency for various d : —, $d = 500$; —, $d = 1000$; ···, $d = 1500$; —, $d = 2000$. (e) Comparison of maximum singular value of true transfer matrix and of the model error transfer function matrices for ERA, ERA/DC ($s = 400$), and Q-Markov ($d = 1500$): —, true; —, ERA; ···, Q-Markov; —, ERA/DC.

The comparison of the methods in Figure 5(e) shows that ERA/DC using $s = 400$, and Q-Markov using roughly four times as much data with $d = 1500$, produce comparable results.

4.8. HIGH DAMPING RESULTS

We consider the same high noise situation as in the previous section, but change the system by increasing the damping in each mode of the system. The new real parts of the eigenvalues are -0.6 , -0.6 , -1.5 , -1.8 and -1.8 , respectively. This corresponds to decreasing the time constants for the envelope of decay of each mode from 11.04, 11.03, 3.04, 2.61, 2.59 to 1.67, 1.67, 0.67, 0.56, 0.56.

Because of the fast decay, lower values of s are appropriate in ERA/DC so as not to include data after complete decay of the signal. ERA/DC succeeds in identifying all modes. Figure 6(a) is somewhat congested in the neighborhood of the second and third peaks, but one can observe that when s increases from 60 to 100, the model error for these peaks becomes worse while the model error for the first peak, corresponding to the first two modes that decay more slowly, remains essentially unchanged. Thus, $s = 100$ is beyond the optimal value for the high frequency modes.

It is shown in Figure 6(b) that the Q-Markov cover algorithm is essentially unable to identify, or gives very poor identification, of the three highest frequency modes. Unlike ERA/DC, the algorithm needs a d large enough that the Markov parameters that are set to zero either have sufficiently small values due to the decay of the signal or are insignificant in number relative to the total number of other Markov parameters. Neither situation occurs before d becomes so large that increasing d simply adds noise without adding signal. The best compromise appears to be $d = 200$, but this is still insufficient for good identification of these modes. In this high noise and fast decay situation, ERA/DC can successfully identify all modes, and Q-Markov cover cannot.

4.9. Q-MARKOV COVER ALGORITHM RESULTS USING RANDOM INPUT RESPONSE

The Q-Markov cover algorithm has the potential advantage that it can use random input response histories directly, rather than using data from such an input to obtain the Markov

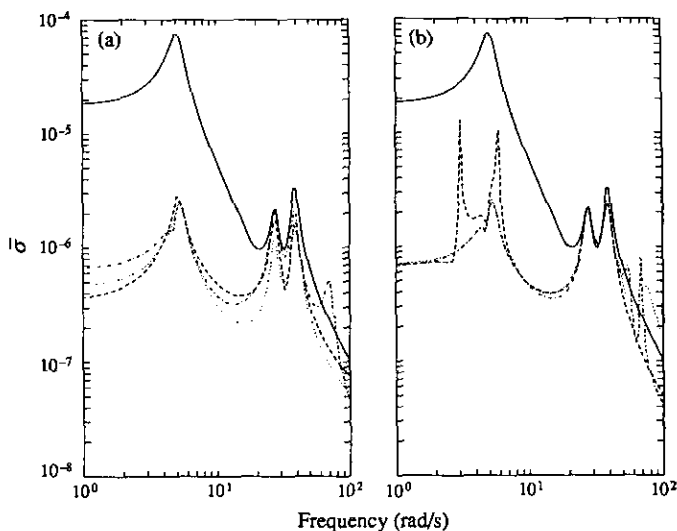


Figure 6. Maximum singular value of the model error transfer function matrix for fast decay system: (a) ERA/DC (—, true; ---, $s = 40$; ···, $s = 60$; -·-, $s = 100$); (b) Q-Markov cover (—, true; ---, $d = 100$; ···, $d = 200$; -·-, $d = 300$).

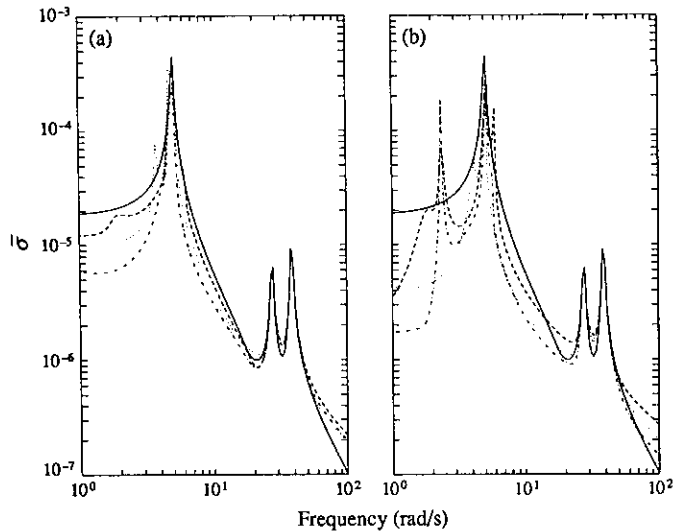


Figure 7. Maximum singular value of the model error transfer function matrix for Q-Markov cover using random input response data directly. (a) —, true; ---, $d = 500$; ···, $d = 1000$; -.-, $d = 1500$. (b) —, true; ---, $d = 2000$; ···, $d = 3000$; -.-, $d = 5000$.

parameters and then using these to perform the identification as described above. The results of numerical tests with noise free data are shown in Figure 7. The method identifies the first two modes with comparatively poor accuracy, and totally fails to identify the remaining three modes. Going above $d = 1500$ produces very modest improvement. Averaging a set of runs to obtain Markov parameters and covariance parameters can produce additional modest improvement. At least for the example considered, this potential advantage of Q-Markov does not appear to be realizable in practice.

4.10. MDVV RESULTS

In Figure 8, we turn to the MDVV algorithm, which is another algorithm that can use random input responses without the need to obtain Markov parameters first. In fact, it can use any sufficiently rich inputs. Like ERA, this algorithm uses a data matrix that need not be square. The originator of the algorithm suggests use of a rectangular matrix. Also, the matrix involved, which undergoes a singular value decomposition, contains not only the measurement data but also the input function. This means that one has the freedom to use general inputs as long as they are sufficiently rich, rather than obtaining impulse responses. This benefit is obtained, in our case, at the expense of doubling the size of the singular value decomposition matrix for the same amount of data.

In place of the 120×120 matrices used in all previous computations, the matrix size for H in these runs is chosen as 80×160 , i.e., a rectangular matrix with a roughly comparable computation time. The inputs were taken as random, producing displacement sensor outputs shown in Figure 8(a). The singular values of H are given in Figure 8(b). The first 40 are associated with the random input functions, and the next ten are associated with the system dynamics. Experience indicates that the input signals need to be very rich to obtain good results. Noise-free results are given, as well as results with three noise levels. The two higher noise levels of 5×10^{-7} and 5×10^{-8} correspond to the high and low noise level of the previous computations, and the third level is associated with a decrease by another factor of ten.

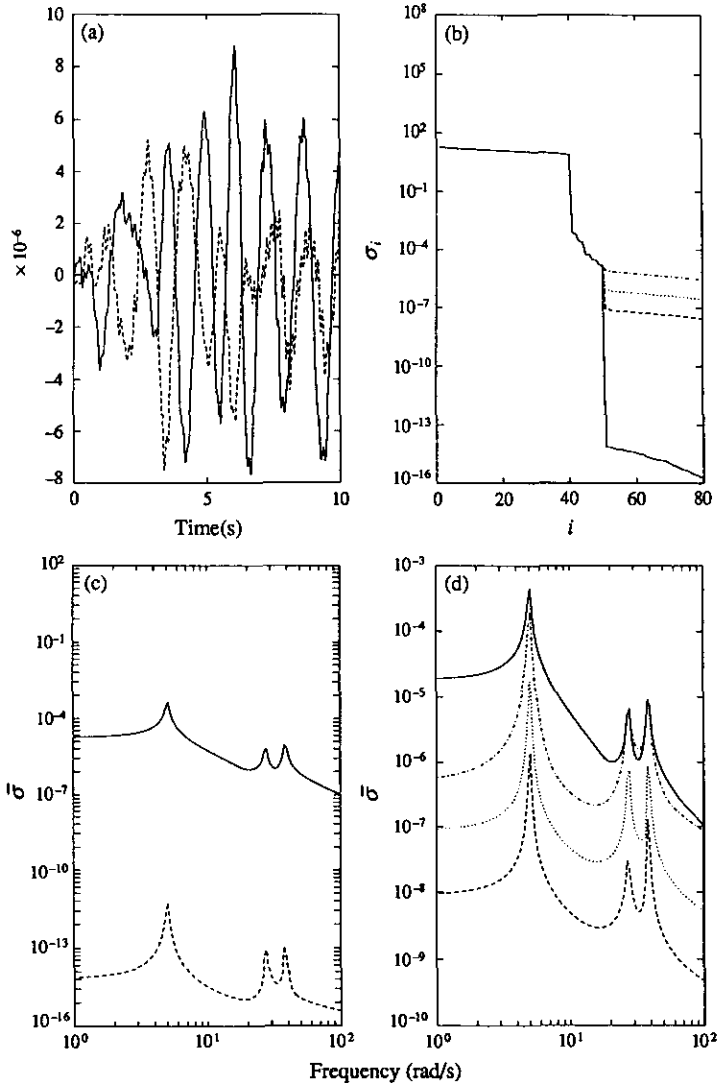


Figure 8. Identification results for the MMVV algorithm: (a) Time histories of the A and B simulated by 18 displacement sensor outputs for the chosen random input signals: —, A; ---, B. (b) Singular values of the MMVV input-output matrix with noise free data, and with three noise levels: —, no noise; ---, $5 \times 10^{-9}\%$; \cdots , $5 \times 10^{-8}\%$; -.-, $5 \times 10^{-7}\%$. (c) Maximum singular value of the true transfer function matrix and of the model error transfer function matrix (noise free) vs. frequency: —, true; ---, error. (d) Maximum singular value of the true transfer function matrix and of the model error transfer function matrices with various noise levels vs. frequency: —, true; ---, $5 \times 10^{-9}\%$; \cdots , $5 \times 10^{-8}\%$; -.-, $5 \times 10^{-7}\%$.

Like ERA and ERA/DC, the identification error is of the order of numerical round-off error in the no-noise case, Figure 8(c). In Figure 8(d) somewhat worse error behavior is indicated than for the other algorithms when noise is present.

5. CONCLUSIONS

Some of the conclusions suggested by the numerical test results are as follows:

1. Of the four algorithms tested, ERA/DC gives the best results.
2. ERA is shown to be a special case of ERA/DC in the sense that in the no-noise case and using the same Hankel matrix for each method, the models produced by the two

algorithms give identical input-output relations for all model orders chosen (i.e., for all choices of the number of non-zero singular values retained). The state space realizations for the models can be different, and there can be different numerical round-off properties but, otherwise, they both give the same input-output relationship.

3. Both algorithms produce models with identical input-output relations when each uses the same Hankel matrix, but ERA/DC has a computational advantage since it produces a smaller singular value decomposition for rectangular Hankel matrices. In practice, memory limitations associated with the singular value decomposition dictate the maximum size of the rectangular Hankel matrix in ERA, based on both the shorter and longer dimensions of the matrix. In ERA/DC, the singular value decomposition is of a square matrix, the dimension of which is the shorter dimension of the Hankel matrix. This allows ERA/DC to give significantly better results in practice.

4. The Q-Markov cover algorithm can be obtained from ERA/DC by setting a triangular set of Markov parameters to zero in the data matrix, and then using a different realization. As a result, Q-Markov is the only algorithm tested here that needs large sets of data to obtain accurate identification from noise free data. The setting of these Markov parameters to zero can be insignificant if they apply to times after all transients have decayed, or if enough Markov parameters are included that the set becomes insignificant. Hence, Q-Markov required much longer sets of data to be used (compare s and d values in the figures) in order to obtain results comparable to those of ERA/DC. Also, situations can arise in which the speed of decay of the transients prevents Q-Markov from identifying a mode that ERA/DC can successfully identify.

5. ERA and ERA/DC, as well as another form of Q-Markov, require computing pulse response histories from the data associated with the actual inputs, before the algorithms can be applied. Our example run with Q-Markov, using random input response data directly, gave significantly worse results than the other algorithms. This suggests that the potential advantage is not actually an advantage in practice.

6. The MDVV algorithm can directly use data from any sufficiently rich input signal, but care must be exercised to insure a well conditioned full rank input signal. This advantage is counterbalanced by the disadvantage of requiring a singular value decomposition of a larger matrix than all of the above algorithms when the same amount of data is used. In ERA/DC (and in Q-Markov) the dimension of the matrix used in singular value decomposition is independent of the amount of data used, whereas in MDVV the dimension increases as the number of data points increases, becoming prohibitively large for large data sets. Results appear to be inferior to the other methods.

We reiterate that the conclusions presented here are based on testing only one structural dynamics example, and that the comparisons are made without fully optimizing all the parameters of any of the methods. Nevertheless, most of the characteristics detailed above for each algorithm are general properties independent of the numerical examples considered. In practice, the variance and bias computation methods developed using perturbation methods (references [10-12]) can be used with experimental data to help the user of ERA, ERA/DC to optimize the parameters of these methods.

ACKNOWLEDGMENTS

The research reported here was performed while J.-S. Lew was a Research Associate of the Department of Mechanical Engineering at Old Dominion University under NASA Grant NAG-1-1117, and R. W. Longman was a National Research Council Senior Research Associate.

REFERENCES

1. J. N. JUANG and R. S. PAPPAS 1985 *Journal of Guidance, Control and Dynamics* **8**, 620–627. An eigensystem realization algorithm for modal parameter identification and model reduction.
2. R. S. PAPPAS and J. N. JUANG 1988 *Proceedings of the 6th International Modal Analysis Conference, Orlando, Florida*. Some experiences with the eigensystem realization algorithm.
3. J. N. JUANG, J. E. COOPER and J. R. WRIGHT 1988 *Control Theory and Advanced Technology* **4**, 5–14. An eigensystem realization algorithm using data correlations (ERA/DC) for modal parameter identification.
4. B. D. ANDERSON and R. E. SKELTON 1988 *IEEE Transactions for Circuits and Systems* **35**, 375–384. Generation of all Q-Markov covers.
5. D. A. WAGIE and R. E. SKELTON 1986 *IEEE Transactions for Automatic Control* **31**, 1114–1120. A projection approach to covariance equivalent realizations of discrete systems.
6. A. M. KING, U. B. DESAI and R. E. SKELTON 1988 *Automatica* **24**, 507–515. A generalized approach to Q-Markov covariance equivalent realization of discrete systems.
7. C. HSIEH, J. H. KIM, K. LIU, G. ZHU and R. E. SKELTON 1990 *Control Structural Interaction Guest Investigator Technical Report, NASA Langley Research Center*. Model identification and controller design for large flexible space structures—an experiment on NASA's Mini-Mast.
8. M. MOONEN, B. DEMOOR, L. VANDENBERGHE and J. VANDEWALLE 1989 *International Journal of Control* **49**, 219–232. On- and off-line identification of linear state-space models.
9. R. PAPPAS, B. MISERENTINO, J. BAILEY, K. ELLIOTT, S. PEREZ, P. COOPER and B. WILLIAMS 1989 *Mini-Mast CSI Testbed User's Guide*. NASA Langley Research Center.
10. R. W. LONGMAN and J. N. JUANG 1988 *Advances in the Astronautical Sciences* **65**, 581–601. A variance based confidence criterion for ERA identified modal parameters.
11. R. W. LONGMAN, M. BERGMANN and J. N. JUANG 1988 *Proceedings of the AIAA/AAS Astrodynamics Specialist Conference, Minneapolis, Minnesota*. Variance and bias confidence criteria for ERA modal parameter identification.
12. M. BERGMANN, R. W. LONGMAN and J.-N. JUANG 1989 *Proceedings of the IEEE Conference on Decision and Control, Tampa, Florida*. Variance and bias computation for enhanced system identification.
13. R. Y. CHIANG and M. G. SAFONOV 1988 *The Robust-control Toolbox for Use with Matlab*. South Natick: The MathWorks, Inc.
14. J. J. DONGARRA, C. B. MOLER, J. R. BUNCH and G. W. STEWART 1979 *LINPACK User's Guide*. Philadelphia: SIAM.

APPENDIX A: MATHEMATICAL DESCRIPTION OF THE IDENTIFICATION ALGORITHMS

A.1. ERA ALGORITHM

1. Form the Hankel matrix

$$H(j-1) = \begin{bmatrix} Y_j & Y_{j+1} & \cdots & Y_{j+s} \\ Y_{j+1} & Y_{j+2} & \cdots & Y_{j+s+1} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{j+r} & Y_{j+r+1} & \cdots & Y_{j+r+s} \end{bmatrix}, \quad (\text{A1})$$

where Y_i is the i th Markov parameter of the pulse response history.

2. Form the singular value decomposition of $H(0)$,

$$H(0) = U \Sigma V^T. \quad (\text{A2})$$

3. The system matrices of the ERA identified k th order discrete time model are given as

$$\begin{aligned} A_k &= \Sigma_k^{-1/2} U_k^T H(1) V_k \Sigma_k^{-1/2}, \\ B_k &= \Sigma_k^{1/2} V_k^T E_m, \quad C_k = E_l^T U_k \Sigma_k^{1/2}. \end{aligned} \quad (\text{A3})$$

The matrix Σ_k is the upper left-hand $k \times k$ partition of Σ containing the k largest singular values along the diagonal. Matrices U_k and V_k are obtained from U and V by retaining only the k columns, i.e., singular vectors, associated with the singular values retained in Σ_k . Matrix E_l is a matrix of appropriate dimension, containing l columns, that is all zero except that the top $l \times l$ partition is an identity matrix. E_m is defined analogously.

A.2 ERA/DC ALGORITHM

1. Form the Hankel matrix as in equation (A1).
2. Compute $R(q)$,

$$R(q) = H(q)H^T(0). \quad (\text{A4})$$

3. Form the block correlation matrix

$$\mathcal{U}(q) = \begin{bmatrix} R(q) & R(q+\gamma) & \cdots & R(q+\beta\gamma) \\ R(q+\gamma) & R(q+2\gamma) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ R(q+\alpha\gamma) & \cdots & \cdots & R(q+(\alpha+\beta)\gamma) \end{bmatrix}. \quad (\text{A5})$$

4. Form the singular value decomposition of $\mathcal{U}(0)$,

$$\mathcal{U}(0) = U\Sigma V^T. \quad (\text{A6})$$

5. The ERA/DC identified k th order discrete time model is

$$\begin{aligned} A_k &= \Sigma_k^{-1/2} U_k^T \mathcal{U}(1) V_k \Sigma_k^{-1/2}, \\ B_k &= \Sigma_k^{-1/2} V_k^T \begin{bmatrix} H(0) \\ H(\gamma) \\ \vdots \\ H(\alpha\gamma) \end{bmatrix} E_m, \\ C_k &= E_l^T U_k \Sigma_k^{1/2}. \end{aligned} \quad (\text{A7})$$

A.3. Q-MARKOV COVER ALGORITHM

1. Calculate X_i ,

$$X_i = Y_i(d, i) = \sum_{k=0}^d Y_{k+i} Y_k^T. \quad (\text{A8})$$

Here, Y_i is the pulse response at time i .

2. Form R_q , H_q and M_q

$$R_q = \begin{bmatrix} X_0 & X_1^T & \cdots & X_{q-1}^T \\ X_1 & X_0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ X_{q-1} & \cdots & \cdots & X_0 \end{bmatrix}, \quad H_q = \begin{bmatrix} Y_0 & 0 & \cdots & 0 \\ Y_1 & Y_0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ Y_{q-1} & \cdots & \cdots & Y_0 \end{bmatrix}, \quad M_q = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{q-1} \end{bmatrix}. \quad (\text{A9})$$

3. Compute D_q ,

$$D_q = R_q - H_q H_q^T. \quad (\text{A10})$$

4. Form the singular value decomposition of D_q :

$$D_q = U \Sigma V^T. \quad (\text{A11})$$

5. Form P_k and partition it:

$$P_k = U_k \Sigma_k^{1/2}, \quad P_k = \begin{bmatrix} O_0 \\ O_1 \\ \vdots \\ O_{q-1} \end{bmatrix}, \quad O_i \in R^{m \times k}. \quad (\text{A12})$$

6. Form P and \bar{P} :

$$P = \begin{bmatrix} O_0 \\ O_1 \\ \vdots \\ O_{q-2} \end{bmatrix}, \quad \bar{P} = \begin{bmatrix} O_1 \\ O_2 \\ \vdots \\ O_{q-1} \end{bmatrix}. \quad (\text{A13})$$

7. The Q-Markov identified k th order discrete model system matrices are

$$A_k = P^+ \bar{P}, \quad B_k = O_0, \quad C_k = P^+ M_q, \quad (\text{A14})$$

where $+$ indicates the Moore-Penrose pseudo-inverse.

A.4. MDVV

1. Form H_1 , H_2 and H :

$$H_1 = \begin{bmatrix} u_k & u_{k+1} & \cdots & u_{k+j-1} \\ y_k & y_{k+1} & \cdots & y_{k+j-1} \\ u_{k+1} & u_{k+2} & \cdots & u_{k+j} \\ y_{k+1} & y_{k+2} & \cdots & y_{k+j} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ u_{k+i-1} & u_{k+i} & \cdots & u_{k+j+i-2} \\ y_{k+i-1} & y_{k+i} & \cdots & y_{k+j+i-2} \end{bmatrix},$$

$$H_2 = \begin{bmatrix} u_{k+i} & u_{k+i+1} & \cdots & u_{k+i+j-1} \\ y_{k+i} & y_{k+i+1} & \cdots & y_{k+i+j-1} \\ u_{k+i+1} & u_{k+i+2} & \cdots & u_{k+i+j} \\ y_{k+i+1} & y_{k+i+2} & \cdots & y_{k+i+j} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ u_{k+2i-1} & u_{k+2i} & \cdots & u_{k+j+2i-2} \\ y_{k+2i-1} & y_{k+2i} & \cdots & y_{k+j+2i-2} \end{bmatrix}, \quad H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix}, \quad (\text{A15})$$

where u_k and y_k denote the l -dimensional input vector and the m -dimensional output vector at time k .

2. Form the singular value decomposition of H :

$$H = U\Sigma V^T = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} S_{11} & 0 \\ 0 & 0 \end{pmatrix} V^T. \quad (\text{A16})$$

Here, the sizes of the matrices are:

$$\text{size}(U_{11}) = (mi + li) \times (2mi + n),$$

$$\text{size}(U_{12}) = (mi + li) \times (2li - n),$$

$$\text{size}(U_{21}) = (mi + li) \times (2mi + n),$$

$$\text{size}(U_{22}) = (mi + li) \times (2li - n),$$

$$\text{size}(S_{11}) = (2mi + n) \times (2mi + n).$$

3. Form the singular value decomposition of $U_{12}^T U_{11} S_{11}$:

$$U_{12}^T U_{11} S_{11} = [U_q \quad U_q^\perp] \begin{pmatrix} S_q & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_q^T \\ V_q^{\perp T} \end{pmatrix}, \quad (\text{A17})$$

where U_q is a $(2li - n) \times n$ matrix.

4. The identified n th order model is (using MATLAB notations)

$$\begin{pmatrix} U_q^T U_{12}^T U(m + l + 1:(i + 1)(m + l), :)S \\ U(mi + li + m + 1:(m + l)(i + 1), :)S \end{pmatrix} \\ = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} U_q^T U_{12}^T U(1:(m + l)i, :)S \\ U(mi + li + 1:mi + li + m, :)S \end{pmatrix}. \quad (\text{A18})$$

APPENDIX B: MINI-MAST MODEL

The model is

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad (\text{B1})$$

APPENDIX C: COMPARISON OF IDENTIFICATION ALGORITHMS

C.1. COMPARISON OF ERA AND ERA/DC

We can show that if one uses the same Hankel matrix $H(0)$ in ERA and in the special case of ERA/DC with $\mathcal{U}(0) = R(0)$, then the resulting realizations, for any chosen system order, have identical pulse response histories. In other words, ERA is a special case of ERA/DC in the sense that the realizations produced have the same Markov parameters, but they can have different choices for state variables.

The $(j+1)$ Markov parameter of the k th order model obtained by ERA in equation (A3) is

$$\begin{aligned} Y_{j+1} &= C_k A_k^j B_k \\ &= E_l^T U_k \Sigma_k^{1/2} (\Sigma_k^{-1/2} U_k^T H(1) V_k \Sigma_k^{-1/2}) \Sigma_k^{1/2} V_k^T E_m \\ &= E_l^T U_k \Sigma_k (\Sigma_k^{-1} U_k^T H(1) V_k) V_k^T E_m. \end{aligned} \quad (C1)$$

Construct the $(j+1)$ Markov parameter of the k th order model from ERA/DC using the $\mathcal{U}(0) = R(0) = H(0)H^T(0)$ special case. Use U , Σ and V for the singular value decomposition of $H(0)$ as in equation (C1), and use \bar{U} , $\bar{\Sigma}$ and \bar{V} for the singular value decomposition of equation (A6) for ERA/DC. Then, these decompositions are related by

$$\bar{U} = U, \quad \bar{\Sigma} = \Sigma \Sigma^T, \quad \bar{V} = U. \quad (C2)$$

Substituting these decompositions into equation (A7) produces the realization

$$\begin{aligned} A_k &= \Sigma_k^{-1} U_k^T \mathcal{U}(1) U_k \Sigma_k^{-1}, \\ B_k &= \Sigma_k^{-1} U_k^T H(0) E_m, \quad C_k = E_l^T U_k \Sigma_k, \end{aligned} \quad (C3)$$

the $(j+1)$ Markov parameter of which is

$$Y_{j+1} = E_l^T U_k \Sigma_k (\Sigma_k^{-1} U_k^T H(1) V \Sigma^T U^T U_k \Sigma_k^{-1}) \Sigma_k^{-1} U_k^T U \Sigma V^T E_m, \quad (C4)$$

and, after noting that

$$V \Sigma^T U^T U_k \Sigma_k^{-1} = V_k, \quad (C5)$$

one obtains equation (C1), which completes the proof.

C.2. COMPARISON OF $\mathcal{U}(0)$ IN ERA/DC AND D_q IN Q-MARKOV

ERA/DC creates its realization from a singular value decomposition of $\mathcal{U}(0)$. In this paper, we pick the special case $\mathcal{U}(0) = R(0) = H(0)H^T(0)$. The i, j th partition of $\mathcal{U}(0)$ can be written as

$$[\mathcal{U}(0)]_{ij} = \sum_{k=0}^s Y_{i+k} Y_{j+k}^T, \quad i, j = 1, 2, \dots, r+1. \quad (C6)$$

The Q-Markov algorithm obtains its realization from a singular value decomposition of the symmetric matrix D_q . The i, j th partition for $i \geq j$ can be written as

$$\begin{aligned} [D_q]_{ij} &= \sum_{k=0}^{d+1} Y_{i-j+k} Y_k^T - \sum_{k=0}^{j-1} Y_{i-1-k} Y_{j-1-k}^T \\ &= \sum_{k=0}^{d+1} Y_{i-j+k} Y_k^T - \sum_{k=0}^{j-1} Y_{i-j+k} Y_k^T \\ &= \sum_{k=j}^{d+1} Y_{i-j+k} Y_k^T \\ &= \sum_{k=0}^{d+1-j} Y_{i+k} Y_{j+k}^T, \quad i, j = 1, 2, \dots, q. \end{aligned} \quad (C7)$$

Comparison of these two formulas shows that if we pick $s = d$, and $q = r + 1$, then the first row partition of D_q , i.e., for $j = 1$, is identical to the first row partition of $\mathcal{U}(0)$. However, for larger j there are $j - 1$ terms missing in the summation for $[D_q]_{ij}$ that are present in $[\mathcal{U}(0)]_{ij}$.

An alternative way to view the difference is to write the following matrix expressions:

$$\begin{aligned}\mathcal{U}(0) &= H(0)H^T(0), \\ [D_q]_{ij} &= (H(0)H_d^T)_{ij}, \quad \text{for } i \geq j,\end{aligned}\tag{C8}$$

where

$$H_d = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_{d+1-q} & \cdots & Y_d & Y_{d+1} \\ Y_2 & Y_3 & \cdots & Y_{d+2-q} & \cdots & Y_{d+1} & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ Y_q & Y_{q+1} & \cdots & Y_{d+1} & 0 & \cdots & 0 \end{bmatrix}, \tag{C9}$$

and $H(0)$ is given in equation (A1). Thus, the $\mathcal{U}(0)$ differs from D_q by setting the lower right triangular set of Markov parameters in the $H(0)$ to zero in order to produce H_d .