Comparison of Various Model Order Reduction Techniques based on Moment Matching Methods

Seema Das*, P. Patnaik**, Dr. R.Jha***
**Dronacharya College of Engineering/ECE Department, Gurgoan,India*
***Global Institute of Technology and Management/ ECE Department, Gurgoan, India*
**** IET Bhaddal Technical Campus/ Director, Ropar, India*

Abstract— **Model Order reduction has been extensively used in study of dynamic behavior of many Engineering and industrial large scale systems as well as Electrical and Electronic systems including microelectronic systems such as integrated circuits. The motivation for appropriate MOR is to obtain an accurate model of smaller order which can be easily simulated and implemented in hard ware with ease saving effort, cost and time. In this paper an analysis of some MOR techniques based on moment matching and Padé approximation and related to Krylov sub spaces more particularly Asymptotic Waveform Evaluation(AWE), Padé Via Lanczos (PVL), Matrix- Padé Via Lanczos (MPVL), Symmetrical PVL (SymPVL) has been carried out. These techniques are applicable to matrix equations resulting when Finite Element Method (FEM) is used to model Electromagnetic wave propagation and radiation problems. AWE methods are straight forward to understand and implement but numerically not stable. Lanczos and PVL methods solve the Eigen value problems, numerically more stable and superior due to their computational efficiency but they can loose passivity for RLC systems. This paper also discusses the ARNOLDI algorithm and improved Arnoldi method like Passive Reduced order Interconnect Macro modelling Algorithm (PRIMA) and SVD Laguerre which are stable as well as retain the passivity.**

Index Terms— Model Order Reduction (MOR), Moment matching, Padé approximation, Finite Element Method (FEM), Krylov sub spaces, Asymptotic Waveform Evaluation(AWE), Arnoldi Algorithm, Lanczos process, Padé Via Lanczos (PVL), Matrix- Padé Via Lanczos (MPVL), Symmetrical PVL (SymPVL), Passive Reduced order Interconnect Macro modelling Algorithm (PRIMA)

I. INTRODUCTION

The basic idea of Model order reduction is to replace the original large scale system model with a much smaller one, yet still retain the original behavior with high accuracy. (PDE) Partial differential equations describe the physical behavior of the system under consideration. Depending on the dimension of the original PDE and desired spatial accuracy, the number of variables can extend from hundreds to several millions [1]. Every system can be generalized into the following time linear invariant linear differential algebraic equation:

$$Gx(t) + C\frac{d}{dt}x(t) = BU(t) \quad (1.1)$$

and output of system is as follows:

$$Y(t) = A^T x(t) \quad (1.2)$$

Where G, C € R^{nXn} , C is non singular B € R^{nXm} , A € R^{nXp}

Taking the Laplace transformation of equation Eq.(1.1) we have :

$$(G + sC)X(s) = BU(s) \quad (1.3)$$

The transfer function of the above system is:

$$H(s) = A^T(G + sC)^{-1}B \quad (1.4)$$

This transfer function is a function of s and can be expanded into a moment expansion around s=0:

$$H(s) = M_0 + sM_1 + s^2M_2 \dots .... \quad (1.5)$$

The matrices $M_i, i = 0,1,2 \dots$ are called moments of the transfer function. M_0 is the DC solution of the system for s=0, so for zero frequency.M_1 is the Elmore delay which represents the signal at the input port to reach the output port.

The transfer function in frequency domain is the Laplace transform of the impulse response.

$$H(s) = \int_0^\infty h(t)e^{-st}dt \quad (1.6)$$

If we expand e^{-st} in a Taylor polynomial we get

$$H(s) = \int_0^\infty h(t)\left(1 - st + \frac{1}{2}s^2t^2 + \cdots \dots\right)dt = \int_0^\infty h(t)dt - s\int_0^\infty th(t)dt + \frac{1}{2}s^2\int_0^\infty t^2h(t)dt \dots \quad (1.7)$$

Elmore delay is the first moment of the moment expansion. At least first two moments must be preserved in a reduction set up. The transfer function can also be expanded around another point s_0€ R for AC analysis.

$$H(s) = M_0 + (s - s_0)M_1 + (s - s_0)^2M_2 \dots .... \quad (1.8)$$

The goal of Krylov subspace MOR [2] which is based on moments is to find a projection based approximation of the original transfer function. In other words, the objective is to calculate a reduced order system with transfer function H(s) where moment expansion is given.

Let us take different methods which use moments to reduce ODE. Lot of research is going on to enhance the methods and in this paper we have discussed a few methods in Krylov space which uses moments.

II. AWE (ASYMPTOTIC WAVEFORM EVALAUATION)

The AWE is based on approximating the Laplace domain transfer function of a linear network by a reduced order model.

The transfer function can be approximated by Pade's approximation. A Pade approximation is ratio of two polynomials P(s)/Q(s) where deg(Q(s) ≥ deg (P(s) [3] AWE calculates a Pade's approximation of finite degree so the degree of P(s) and Q(s) is finite. It computes the first two q moments of the transfer function H(s) and then find the Pade approximation whose 2q moments match with 2q moments with original transfer function.

Let us take the transfer function as considered in Eq.(1.3)

$$(G + sC)X(s) = BU(s)$$

Expand X(s) around some expansion point $s_0 \in C$

$$(G + s_oC + (s - s_o)C)(X_o + (s - s_o)X_1 + (s - s_o)^2X_2 \dots\dots\dots) = BU(s) \quad (2.1)$$

where $X_i = 0,1,2\dots$ are called the moments. Let us take U(s) =1 and take the terms according to powers of $(s - s_o)$

$$(G + s_oC)X_0 = B \quad (2.2)$$

$$CX_o + (G + s_oC)X_1 = 0$$

The following moments can be derived

$$(G + s_oC)X_i = -CX_{i-1} \text{ for } i > 0 \quad (2.3)$$

This leads to Krylov subspace:

$$K_q((G + s_oC)^{-1}B, (G + s_o C)^{-1}C)$$

$$=[(G + s_0C)^{-1}B, (G + s_0C)^{-1}C(G + s_0C)^{-1}B, \dots] \quad (2.4)$$

The transfer function can be written as :

$$H(s) = \sum_{k=0}^{n} m_k(s - s_o)^k \quad (2.5)$$

or simply as

$$H(s) = \sum_{k=0}^{n} m_k s^k \quad (2.6)$$

The transfer function in Pade's form can be written as:

$$H(s) = \frac{P(s)}{Q(s)} = \frac{1+a_1s+b_2s^2+\cdots a_m s^m}{1+b_1s+b_2s^2+\cdots b_n s^n} \quad (2.7)$$

Where transfer function is in moments form as

$$H(s) = 1 + m_1s + m_2s^2 + m_3s^3 + \cdots \quad (2.8)$$

Where P(s) are the zeroes of the transfer function H(s) and zeroes of Q(s) are equal to the poles of H(s). We assume P(s) to have order p, so P(s) can be written as

$$P(s) = \sum_{k=0}^{p} a_k s^k \quad (2.9)$$

and in terms of transfer function it can be taken as:

$$P(s) = H(s)Q(s) \quad (2.10)$$

$$\sum_{k=0}^{p} a_k s^k = \left(\sum_{k=0}^{n} m_k s^k\right)\left(\sum_{k=0}^{p+1} b_k s^k\right) \quad (2.11)$$

The equations can be solved for different powers of s setting $b_0 = 1$ this leads to the following matrix equation:

$$\begin{bmatrix} m_0 & m_1 \cdots & m_p \\ m_1 & \ddots & \vdots \\ m_p & m_{p+1} \cdots & m_{2p} \end{bmatrix} \begin{bmatrix} b_{p+1} \\ b_p \\ \vdots \\ b_1 \end{bmatrix} = \begin{bmatrix} m_{p+1} \\ m_{p+2} \\ \vdots \\ m_{2p+1} \end{bmatrix} \quad (2.12)$$

The AWE [4] can be used to determine the time domain or frequency domain response of the linear network over predetermined range of excitation frequencies. This method creates problem for matrices sizes larger than eight. Therefore more than eight poles cannot be approximated by AWE.

III. PADE'S VIA LANCZOS (PVL)

PVL was introduced by Gallivan, Grimme and Vandooren in 1994 and by Feldman and Freund in 1995 to overcome the drawbacks of AWE. PVL provides a numerically stable algorithm that computes the Pade approximation of a linear circuit[3]. PVL can be used to generate an arbitrary number of poles and zeros and in fact even all of them. This algorithm requires the same amount of computations as AWE as it generates more poles and can be increased in accuracy.

A. Single Input Single Output PVL

Let us take transfer function of first order differential equations with single input and single output [6-8] :

$$C\dot{x} = -Gx + bu \quad (3.1)$$

$$y = I^H x + du \quad (3.2)$$

I^H is a nx1 output vector that selects output of interest from the matrix system where bu and du represents the excitations from independent sources. This transfer function is applied to complex system where s ∈ C with the condition $(G + s_0C)$ is non singular. Consider a problem in which the equations describing the system can be cast into the Laplace form

$$sCX = -GX + bU \quad (3.3)$$

$$Y = I^H X \quad (3.4)$$

$$H(s) = I^H(G + sC)^{-1}b \quad (3.5)$$

Using change of variables $s = s_0 + \sigma C$

$$D = -(G + s_0C)^{-1}C, \quad (3.6)$$

$$r = (G + s_0C)^{-1}b$$

Where s_0 is the point of expansion in the complex s plane.

$$H(s_0 + \sigma) = I^H(I - \sigma D)^{-1}r \quad (3.7)$$

The PVL algorithm uses the iterative Lanczos[4] process to reduce the matrix D to a triagonal matrix T_q where $q \leq n$

Assuming that the matrix D is diagonalizable, we obtain

$$H(s_0 + \sigma) = I^H V(I - \sigma\tau)^{-1}V^{-1}r \quad (3.8)$$

Where $D = V\tau V^{-1}$ and $\tau = diag(\lambda)$ where λ_1, λ_2 ……are the eigen values and D is a diagonal matrix and V contains the corresponding eigenvectors as columns.

$$I^H V = f^T, V^{-1}r = g \quad (3.9)$$

From eq. 3.8 we have :

$$H(s_0 + \sigma) = \sum_{i=1}^{N} \frac{f_i g_i}{1 - \sigma\lambda_i} \quad (3.10)$$

f_i and g_i are the components of the vectors f and g.

In PVL every iteration leads to the preservation of two extra moments. This makes PVL a very efficient and powerful algorithm. The disadvantage with this method is that it does not always preserve stability.

B. Multi Input Multi Output PVL(MPVL)

The original PVL algorithm is applicable to SISO system

Freund reported a matrix-Pade via Lanczos (PVL) IN 1998 for a multi input multi output (MIMO) system. Let us take the transfer function with MIMO system as in Eq.(1.1) the transfer function is as follows[9]:

$$H(s) = A^H(G + sC)^{-1}B \quad (3.11)$$

Let us expand around s_0 we have :

$$s = s_0 + \sigma C,$$

$$M = -(G + s_0C)^{-1}C, \quad (3.12)$$

$$R = (G + s_0C)^{-1}B \quad (3.13)$$

Where M and R are known as Block Krylov subspace.

The new transfer function is as follows:

$$H(s_0 + \sigma) = A^H(I - \sigma M)^{-1}R \quad (3.13)$$

This leads to Krylov subspace as:

$$K_q(R, MR, M^2R \dots \dots \dots M^{j-1}R) \quad (3.14)$$

MPVL is applied to solve models resulting from applying the finite element method (FEM) to model electromagnetic wave propagation problems in the frequency domain. Like PVL the disadvantage of this method is also that it does not always guarantee stability.

Another method presented by Freud and Fieldman [10] is SymPVL Symmetric PVL [5] which is the efficient version of PVL for the case of symmetric matrices. This method cures the stability problem observed for. The main idea of these methods is to make use of the fact that the matrix is symmetric so that it can be decomposed using [11] a choleskey decomposition. This then automatically leads to stability of the associated approximation methods. Rodney D Slone et al[17] reviewed various explicit and implicit computer based algorithms viz AWE (explicit), Lanczos (implicit) , PVL and MPVL (implicit) and concluded that they are computationally efficient. Though they have some individual advantages and disadvantages, MPVL is better suited for radiation problems.

IV. ARNOLDI METHOD

Arnoldi method can also be used as the basis for MOR method like PVL. Similar to PVL, one can define an expansion point s_0 and work with the shift and input transfer function.

Where

$$\tilde{G} = (G - s_0C)^{-1}C \quad (4.1)$$

$$\tilde{B} = (s_0C - G)^{-1}B \quad (4.2)$$

In the Arnoldi process, Krylov space is generated which is associated with the matrices $\tilde{G}$ and $\tilde{B}$ as follows:

$$K_q\left(\tilde{G}, \tilde{G}\,\tilde{B}, \tilde{G}^2\tilde{B} \dots \dots \dots \tilde{G}^q\tilde{B}\right) \quad (4.3)$$

If the size of the Krylov space q , is smaller than the size of the system n, a reduction can be performed by projecting the system matrices onto the Krylov space, in the following way where V is the orthonormal basis of the ktylov space[12]

$$\tilde{G} = V^TGV \quad (4.4)$$

$$\tilde{C} = V^TCV \quad (4.5)$$

$$\widetilde{A^T} = A^TV \quad (4.6)$$

The transfer function of the reduced system approximates the transfer function of the original system well within a certain frequency range. The main difference with PVL is that only one Krylov space is generated namely with the block Arnoldi process [13] and the projections are performed with orthogonal vectors. Block Arnoldi methods are well known in the context of solving linear systems. The accuracy of the block Arnoldi approximations gradually increases as the order is increased since more moments of the original matrix will be matched.

V. PASSIVE REDUCED ORDER INTERCONNECT MACRO MODELLING ALGORITHM (PRIMA)

PRIMA [12] is an improved Arnoldi method which is described below. Fundamental difference is that the projection of the system matrices is done explicitly, in contrast to PVL and Arnoldi, where the tridiagonal or Hessenberg matrix is used directly.

The Krylov space is generated

$$K_q = ((G + s_0C)^{-1}B, (G + s_0C)^{-1}C(G + s_0C)^{-1}B, \dots \dots) \quad (5.1)$$

Instead of taking the block Hessenberg matrix H as approximation of $(G + s_0C)^{-1}C$, the system matrices G and C are explicitly projected onto the basisV:

$$G_q = V^TGV \quad (5.2)$$

$$C_q = V^TCV \quad (5.3)$$

Although the method is expensive, the explicit projection onto the Krylov space as done here has strong advantages. It makes PRIMA more accurate than the Arnoldi method. It ensures preservation of stability and passivity. Explicit projection [14] comes down to pre-multiplying the equation of the system by V^T and replacing the state space vector x by V $\bar{x}$.

$$V^T(G + sC)\mathrm{V}\,\bar{x} = V^TBu. \quad (5.5)$$

$$\bar{y} = A^TV\bar{x} \quad (5.6)$$

Since $x = V\bar{x}$ the approximation of the original state space can be found if $\bar{x}$ is available. The projection onto the smallest possible Krylov space, consisting of only one block preserves the first moment of the system. Sequentially, a moment more is preserved every time a large Krylov space is used for projection, whereas in PVL with every iteration only two moments are preserved. Joao M S Silva et al [18] reviewed the projection base methods viz PVL and PRIMA and observed that though they are simple, efficient and accurate, they have the short comings of computational complexity, high cost and lack of general strategy for error control and are dependent on

the structure of original models viz number of ports, power grid etc.

VI. SVD – LAGUERRE METHOD

In this method the transfer function is [16] not shift and inverted as done in PRIMA and PVL, but the method is based on Laguerre expansion of transfer function.

For the expansion scaled Laguerre functions are used, defined as:

$$\varphi_n^t(t) = \sqrt{2\alpha}e^{-\alpha t} l_n(2\alpha t) \quad (6.1)$$

Where α is a positive scaling parameters and $l_n(t)$ is the laguerre polynomial.

$$l_n(t) = \frac{e^t d^n}{n! dt^n}(e^{-t}t^n) \quad (6.2)$$

In it was shown that the Laplace transformation of $\varphi_n^{\alpha}(t)$ is :

$$\varphi_n^t(s) = \frac{\sqrt{2\alpha}}{s+\alpha}\left(\frac{s-\alpha}{s+\alpha}\right)^n \quad (6.3)$$

The Laguerre expansion looks like this:

$$H(s) = (G+sC)^{-1}\mathrm{B} = \frac{\sqrt{2\alpha}}{s+\alpha} A^T \sum_{n=0}^{\infty}((G+\alpha C)^{-1}(G-\alpha C))^n (G+\alpha C)^{-1} B \left(\frac{s-\alpha}{s+\alpha}\right)^n \quad (6.4)$$

The expansion gives raise to generally Krylov space starting with the matrix $(G+\alpha C)^{-1}\mathrm{B}$ and using a generating matrix $(G+\alpha C)^{-1}(G-\alpha C)^{-1}$ and then Krylov space leads to

$$K_q = [\,(G+\alpha C)^{-1}\mathrm{B}, (G+\alpha C)^{-1}, (G-\alpha C)^{-1}] \quad (6.5)$$

The number of linear systems equations to be solved is equal to that in PRIMA, so the method is comparable in computation demand. The orthogonalisation of the columns can be done by a SVD after all columns have been computed. Analogously to the PRIMA method, the system matrices are explicitly projected onto the generated Krylov subspace. Consequently stability and passivity are preserved.

VII. CONCLUSION

In this paper a number of Model Order Reduction methods have been reviewed. AWE though simple but cannot be used for large matrices. Techniques based on the PRIMA and Arnoldi improve the numerical stability by finding a set of orthogonal Vectors with the same information as the moments. PRIMA preserves the stability and passivity of a system, whereas other methods are not successful while preserving the passivity of a system. A slight disadvantage of this method as compared to PVL is that only one moment per iteration is preserved. Just like PRIMA, in SVD Laguerre method stability and passivity is preserved.

.

REFERENCES

[1] A.C. Antoulas book. *Approximation of Large-Scale Dynamical Systems*. SIAM series on Advances in Design and Control, 2005.

[2] Boley DL. Krylov space methods on state-space control models. *Circuits Systems Signal Process* 1994, 13(6):733{758.

[3] E. Chiprout and M.S. Nakhla. *Asymptotic Waveform Evaluation and moment matching for interconnect analysis*. Kluwer Academic Publishers, 1994.

[4] L.T. Pillage and R.A. Rohrer. Asymptotic Waveform Evaluation for Timing Analysis. *IEEE Trans. Computer-Aided Design Int. Circ. and Syst.*, 9(4): 352-366, April 1990.

[5] Z. Bai, R.D. Slone, W.T. Smith, and Q. Ye. Error bound for reduced system model by Pad´e approximation via the Lanczos process. *IEEE Trans. CAD-IC and Syst*., 1999.

[6] P. Feldmann and R. Freund. Efficient linear circuit analysis by Pad´e approximation via the Lanczos process. *IEEE Trans. Computer-Aided Design*, 14:137-158, 1993.

[7] P. Feldmann and R.W. Freund. Reduced-order modeling of large linear subcircuits via a block Lanczos algorithm. *Proc. 32nd ACM/IEEE Design Automation Conf.*, June 1995.

[8] P. Feldmann and R. W. Freund. Efficient linear circuit analysis by Padéapproximation via the Lanczos process. In Proceedings of EURO-DAC '94 with EURO-VHDL '94, Grenoble, France, pages 170–175. IEEE Computer Society Press, 1994.

[9] R.W. Freund and P. Feldmann. Reduced-order modeling of large linear passive multiterminal circuits using Matrix-Pad´e approximation. *Proc. DATE Conf. 1998*, 530-537,1998.

[10] R.W. Freund and P. Feldmann. The SyMPVL algorithm and its application to interconnect simulation. *Proc. 1997 Int. Conf. Simulation of Semiconductor Processes and Devices*, 113-116, 1997.

[11] R. W. Freund, Reduced-Order Modeling Techniques Based on Krylov Subspaces and Their Use in Circuit Simulation. *Numerical Analysis Manuscript*, No. 98-3-02, Bell Laboratories, Feb. 1998.

[12] L. Knockaert and D. De Zutter. Passive Reduced Order Multiport Modeling: The Pad´e-Arnoldi-SVD Connection. *Int. J. Electron. Commun. (AEU)*, 53:254-260, 1999.

[13] Odabasioglu A, Celik M. PRIMA: Passive Reduced-Order Interconnect Macromodeling Algorithm. *IEEE* R. Coase, "The federal communications commission," *J. Law and Economics*, pp. 1–40, 1959.

[14] E.J. Grimme. *Krylov Projection Methods for model reduction*. PhD thesis, Univ. Illinois, Urbana-Champaign, 1997.

[15] W.H.A. Schilders and E.J.W. ter Maten. *Numerical methods in electromagnetics*. Handbook of Numerical Analysis, volume XIII, Elsevier, 2005.

[16] Y. Chen, V. Balakrishnan, C-K. Koh, and K. Roy. Model reduction in the time-domain using Laguerre polynomials and Krylov methods. *Proc. Design Automation and Test in Europe Conf.*, pages 931.935, March 2002.

[17] Rodney D. Slone, Jin-Fa Lee, Robert Lee. A Comparison of Some Model Order Reduction Techniques *Electromagnetics*, 22:275–289, 2002

[18] Joao M.S.Silva, Jeorge Fernandez Villena , Paulo Slores, L.Miguel Silveria outstanding issues in model order reduction scientific computing in electric engineering Vol.11 of mathematics in industry page 139-152, springer verlag may 2007