Application of Least Squares MPE technique in the reduced order modeling of electrical circuits

Patricia Astrid and Arie Verhoeven

Abstract—Reduced order models are usually derived by performing the Galerkin projection procedure, where the original equations are projected onto the space spanned by a set of approximating basis functions. For Differential Algebraic Equations this projection scheme may yield an unsolvable reduced order model. This means that a model of an electrical circuit can become ill-posed if it is reduced by the Galerkin technique. As a remedy to the problem, in this paper the reformulation of the reduced order model problem in the least squares sense is suggested. The space where the original is projected is different to the space used in the Galerkin procedure. It is shown that the resulting reduced order model will be guaranteed to be well-posed when the problem of finding a reduced order model is cast into a least squares problem.

To accelerate the reduced order modeling computation, the Missing Point Estimation (MPE) technique which was successfully implemented in the PDE-models of heat transfer processes is also applied to the least-square reduced order model of the electrical circuit. The least-square based MPE model is derived by projecting a subset of the original equation onto the least-square space. The dynamics of the stiff DAE model can be approximated very closely by a reduced order model built from less than 28% of the original equations.

I. ELECTRICAL CIRCUIT MODELS

The dynamics of electrical circuits at time $t$ are generally given by the Differential Algebraic Equations (DAE)

$$
\frac{d}{dt} q(t, x) + j(t, x) = 0
$$

where $x \in \mathbb{R}^K$ represents the state of the electrical circuit. The voltages and currents usually serve as state variables for electrical circuits. The functions $q(t, x) \in \mathbb{R}^K$ and $j(t, x) \in \mathbb{R}^K$ can be nonlinear functions of the state. The state dependence originates from the dependence of the electrical components on voltages and electrical currents.

The circuit simulation models are used to perform various types of analysis such as the small signal analysis and the transient analysis. The transient analysis plays an important role in circuit design because then the real dynamics of the nonlinear model are captured. Usually model (1) is simulated for a time interval in the range of nanoseconds because of the typical operating frequencies of electrical circuits.

The advancement of large scale integrated circuit (chips) design drives the need of fast simulators which may provide fast analysis of the circuit designs and reduce the manufacturing cost and time. Typical dimensions of circuit models of type (1) lie between $10^5$ and $10^6$. Larger circuit models are seldomly used for a transient simulation because of the growing computational time. To reduce the required computational effort, the development of low order circuit models has been an active research area in circuit design.

Research on model reduction for linear time-invariant circuit models [6], [7] is extensive. Several popular reduction methods in circuit modeling include the Krylov subspace [6], [8] and the balanced truncation methods [13]. More recently, the proper orthogonal decomposition technique (POD) has also been applied to circuit models [12].

Due to the growing complexities in circuit designs, linear models are often inadequate to describe the original dynamics. The reduced order model of nonlinear circuit models are not favorable as the derivation requires complete evaluations of the nonlinear functions $q(t, x)$ and $j(t, x)$, which can be very expensive. To avoid the expensive calculations, Rewienski et.al [14] proposed the trajectory-piecewise linear approach to reduce nonlinear circuit models. The nonlinear dynamics are approximated by a set of linear models derived from linearizations of the nonlinear models at several operating points.

More recently, the Missing Point Estimation (MPE) approach was proposed in [1], [3], [4]. The idea was to build a nonlinear reduced order model by projecting a part of the original equations onto the subspace spanned by the chosen set of basis functions. The technique has been successfully implemented on partial differential equation (PDE) models. In this paper, the technique is going to be implemented to reduce a DAE model of an electrical circuit.

This paper is organized as follows. The convenient formulation of the DAE reduced order model using Galerkin projection is given in section II. Then the proper orthogonal decomposition (POD) method is described briefly in section III. Section IV describes the least squares formulation of the DAE reduced order model, while the MPE technique is presented in section V. The application of the MPE technique on an electrical circuit model is discussed in section VI.

II. REDUCED DAE MODELS BY GALERKIN PROJECTION

Let the state $x \in \mathbb{R}^K$ belong to a separable Hilbert space $X$, endowed with the Euclidian inner product space. Then for all $t$ the state $x$ can be expanded in an orthonormal basis...
\[ \Phi = (\varphi_1 \ldots \varphi_K) : \]

\[ x(t) = \sum_{i=1}^{n} a_i(t) \varphi_i. \quad (2) \]

The orthonormal basis is derived from various criteria based on the approximation quality of the original state \( x \) by its truncated expansion \( x_n \) as defined in (3).

\[ x(t) \approx x_n(t) = \sum_{i=1}^{n} a_i(t) \varphi_i. \quad (3) \]

The order of the truncated expansion, \( n \), is lower than the order of the original expansion, \( K \). Different reduction methods yield different bases.

The reduced order model is the model which describes the dynamics of the basis coefficients or the reduced state \( a = \{a_1, \ldots, a_n\} \). In many methods, the reduced order model is derived by replacing the original state \( x \) by its truncated expansion \( x_n \) and projecting the original equations onto the truncated basis

\[ \Phi_n = (\varphi_1 \ldots \varphi_n) \]

This projection scheme is known as the Galerkin projection scheme. The resulting reduced order model defines the evolutions of the reduced state

\[ a(t) = \{a_1(t), \ldots, a_n(t)\} \]

Suppose the original model is a DAE model defined in (1). Substituting the original state \( x \) by its truncated state \( x_n \) as defined in (3) yields

\[ \frac{d}{dt} q(t, \Phi_n a) + j(t, \Phi_n a) = 0 \quad (4) \]

Projection of (4) onto the truncated basis \( \Phi_n = (\varphi_1 \ldots \varphi_n) \) results in the reduced DAE model:

\[ \frac{d}{dt} \Phi_n^T q(t, \Phi_n a) + \Phi_n^T j(t, \Phi_n a) = 0 \quad (5) \]

The original \( K \)-th dimensional DAE model is reduced to \( n \)-th DAE reduced order model by means of the Galerkin projection. Unfortunately, the resulting reduced order model is not always solvable for any arbitrary truncation degree \( n \).

### III. Proper Orthogonal Decomposition

The proper orthogonal decomposition (POD) was also known as the Principal Component Analysis (PCA) and the Karhunen-Loève expansion.

The POD basis \( \Phi_n = (\varphi_1 \ldots \varphi_n) \) is an orthonormal basis and derived from the collected state evolutions (snapshots)

\[ X = (x(t_1) \ldots x(t_N)) \]

The POD method is particularly very popular for systems governed by nonlinear partial differential equations. Analytical solutions do not exist for such systems and the collected data may serve as the only adequate description of the system dynamics.

The POD basis is found by minimizing the time averaged approximation error given in (6)

\[ \text{av} (\| x(t_k) - x_n(t_k) \|_2) \quad (6) \]

The averaging operation \( \text{av}(\cdot) \) is defined as:

\[ \text{av}(f) = \frac{1}{N} \sum_{k=1}^{N} f(t_k) \quad (7) \]

Solving the minimization problem of (6) is equivalent to solving the eigenvalue problem \([9], [10]\)

\[ \frac{1}{N} X X^T \Phi_n = \Phi_n \Lambda_n. \quad (8) \]

The term \( \frac{1}{N} X X^T \) equals the state covariance matrix and the POD basis equals the eigenvectors of the covariance matrix. The most important POD basis function is the eigenvector corresponding to the first eigenvalue. The truncation degree is determined from the eigenvalue distribution in \( \Lambda_n = \text{diag}(\lambda_1, \ldots, \lambda_n) \). Based on the commonly adopted ad-hoc criterion, the truncation degree \( n \) should at least capture 99\% of the total energy. The fraction of the total energy is defined as

\[ P_n = \frac{\sum_{i=1}^{n} \lambda_i}{\sum_{i=1}^{N} \lambda_i} \]

For circuit models the snapshots can be collected from e.g. a transient simulation with fixed parameters and sources. The reduced model also can be used to approximate the model for different parameters or sources as long as the solution still approximately lies in the projected space. For circuit models with a lot of redundancy the reduced model can have a much smaller dimension. Unfortunately, direct application of POD to circuit models does not work well in practice. Firstly, for Differential Algebraic Equations the Galerkin projection scheme may yield an unsolvable reduced order model. Secondly, the computational effort required to solve the reduced order model and the original model is relatively the same in nonlinear cases. This is due to the fact that the evaluation costs of the reduced model (5) are not reduced at all because \( \Phi_n \) will be a dense matrix in general. The next two sections will present techniques which solve these problems.

### IV. Model Order Reduction for DAE Model in the Least Squares Sense

For linear time-invariant DAEs of the type

\[ Cx + Gx = s \quad (9) \]

the solvability depends on the matrix pencil \( \lambda C + G \). If the set of generalized eigenvalues \( \{ \lambda \in C : \det(\lambda C + G) = 0 \} \) is a finite set, it follows that (9) has an unique solution. Because the solvability of a nonlinear DAE is hard to analyze, we restrict ourselves to the solvability of the discretized model. We assume that the numerical scheme solves the following nonlinear equation at every time step \( t_i \):

\[ \lambda_i q(t_i, x_i) + j(t_i, x_i) = r_i, \quad (10) \]
where $x_i$ is the numerical approximation of $x(t_i)$ and $r_i$ is a known vector which may include the values of the state $x$ at previous time step(s). The coefficient $\lambda_i$ depends on the used discretization method and stepsize control. Notice that all Linear Multistep Methods have this property. The solvability condition of the nonlinear numerical model (10) depends on the implemented numerical method. In this paper, we employ the Newton method [15]. At every time step, the nonlinear equation (11) is solved iteratively for every time step.

$$f(x_i) = \lambda_i q(t_i, x_i) + j(t_i, x_i) - r_i = 0.$$  (11)

Let $x^l$ denote the value of $x_i$ at $l$-th iteration step and $J(x^l) = \lambda_i C(t_i, x_i^l) + G(t_i, x_i^l)$ be the Jacobian of the nonlinear function $f$. In each iteration (12) is solved until convergence criteria are met.

$$J(x^l)(x^{l+1} - x^l) = -f(x^l).$$  (12)

Clearly, if the Jacobian $J(x^l)$ is invertible for all iteration steps, the nonlinear equation (10) is solvable for the time step $t_i$. For the reduced order model derived by the Galerkin projection method the reduced Jacobian is $J_r(a^l) = \Phi^T_n J(x^l_n)\Phi_n$ where $x_n = \Phi_n a_n$. Since the reduced basis is generally a rectangular matrix, the reduced Jacobian is not equivalent to a similarity transformation of $J(x^l_n)$. Consequently, $\det(J(x^l_n)) \neq \det(J_r(a^l))$. Thus, $J_r(a^l)$ may not be invertible even though $J(x^l_n)$ is invertible. Notice that this phenomenon is typical for DAEs. Because $C = I$ for ODEs, it follows that $J_r(a^l) = \Phi^T_n (I + G(t_i, x_i^l))\Phi_n = I + \Phi^T_n G(t_i, x_i^l)\Phi_n$ is always invertible.

### A. Example

Consider problem (9), where

$$C = G = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$  

If we take the basis

$$\Phi = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$  

the application of Galerkin projection as in (5) leads to the following unsolvable reduced DAE:

$$0\dot{a} + 0a = s_1.$$  

To preserve the solvability of the numerical model, the reduced Jacobian $J_r(a^l)$ must be defined differently. Instead of the Galerkin projection (5) we consider the DAE (4) itself. Discretizing this continuous problem leads for each Newton iteration to the linear system:

$$J(x^l)(\Phi_n a^{l+1} - \Phi_n a^l) = -f(\Phi_n a^l),$$  (13)

which can be abbreviated by

$$M\Phi_n a^{l+1} = b.$$  (14)

Of course this linear algebraic system is not solvable if $\Phi_n$ is not invertible. Because $\Phi_n$ is a matrix with full column rank $r = n$, it is still possible to solve the least square problem posed in (15):

$$\min \|M\Phi_n a^{l+1} - b\|_2$$  (15)

Solving the least square problem (15) is equivalent to solving (16)

$$\Phi^T_n M^T M\Phi_n a^{l+1} = \Phi^T_n M^T b.$$  (16)

This linear system can be efficiently solved by a QR factorization of $M\Phi_n$. The model (16) is a low dimensional model, the dimension of mass matrix $M_n$ is equal to the dimension of the reduced state $a$. If $M$ is invertible, the invertibility is automatically preserved in the least-square reduced order model (16) since $M_n$ is a symmetric positive definite matrix and therefore invertible.

### V. THE MISSING POINT ESTIMATION

Despite the resulting low dimensional model, the computational effort required to solve the reduced order model and the original model is relatively the same in nonlinear cases. It may even occur that the original model is cheaper to evaluate than the reduced order model. The low dimensionality is obtained by means of projection, either by the Galerkin projection method or the least square method. In the projection schemes, the original numerical model must be projected onto the projection space. It implies that the original model must be re-evaluated when the original numerical model is time-varying, which is the general case for nonlinear systems. A consequence is that the evaluation costs for the reduced model are not reduced at all.

For both presented projection approaches each Newton iteration the Jacobian $J = J(x^l)$ and the right-hand-side $f = f(x^l)$ of the large-dimensional original DAE (1) have to be evaluated. The Missing Point Estimation (MPE) was proposed in [1] as a method to reduce the computational cost of reduced order, nonlinear, time-varying model. The method is inspired by the Gappy-POD approach which was introduced by Everson and Sirovich in [5].

Provided that the original, high-dimensional state $x \in \mathbb{R}^K$ can be approximated closely by $n$ POD basis functions, where $n \ll K$, it follows that the POD coefficients $a = (a_1, \ldots, a_n)$ can also be estimated from the knowledge of $n \ll K$ state variables or data points only [10]. Suppose that the POD coefficients are estimated from $G < K$ state variables.

Let $P \in \{0, 1\}^{G \times K}$ be a permutation matrix where $G \ll K$ and define the restricted basis $\hat{\Phi}_n$ as

$$\hat{\Phi}_n = P\Phi_n.$$  (17)

Note that it always holds that $P P^T = I_G$.

Corresponding to the restricted basis $\hat{\Phi}_n$, introduce the restricted state $\tilde{x} \in \mathbb{R}^G$, which is defined as

$$\tilde{x} = Px.$$  (18)
Accordingly, the restricted state $\hat{x}_n$ can be approximated by the expansion of the restricted basis $\tilde{\Phi}_n$:

$$\hat{x} \approx \tilde{x}_n = \tilde{\Phi}_n \tilde{a}_n$$

The paper [5] posed the following problem: Given the restricted state $\hat{x}$ and the restricted basis $\tilde{\Phi}$, is it possible to estimate the coefficients $a_n$? This problem is solved in [5] by minimizing (19).

$$\| \hat{x} - \tilde{\Phi}_n \tilde{a}_n \|_2$$

The coefficients $\tilde{a}_n$ are the POD coefficients estimated from the knowledge of the restricted state $\hat{x}$. This idea is then extended for dynamical systems in [1]. The extended approach is referred to as the Missing Point Estimation (MPE) method.

In MPE, the model of the original state $x$ is given. The dynamics of the restricted state $\hat{x}$ is also given. Note that the Gappy-POD approach, the restricted state $\hat{x}$ is exact, while here we only consider the dynamical model of the restricted state.

Recall that in order to solve the original model, we solve the following equation at every iteration step $l$:

$$J(x') (x^{l+1} - x') = -f(x')$$

In MPE, we only solve a part of (20). Similar to the formulation of the restricted basis and the restricted state, we multiply both sides of (20) by the permutation matrix $P \in \mathbb{R}^{G \times K}$.

$$PJ(x') (x^{l+1} - x') = -Pf(x')$$

Note that in the MPE case, we only need the evaluations of the restricted Jacobian $PJ$ and forcing term $Pf$.

As before, the model (21) can be discretized into:

$$PM\tilde{\phi}_n^{l+1} = Pb$$

Introduce $P_{nb} = \{0, 1\}^{G_{nb} \times K}$ as another permutation matrix. The permutation matrix $P_{nb}$ is introduced in order to include the state variables which contribute to the dynamics of the restricted state $\hat{x}$ but do not belong to the restricted state $\hat{x}$.

Hence, let

$$\tilde{x}_{nb} = P_{nb} \tilde{x} = \{\tilde{x}, \tilde{x}_c\}, \quad \tilde{x}_c \notin \tilde{x}$$

where $\tilde{x}_c$ are the variables which contribute to the dynamics of $\tilde{x}$, $G \leq \dim \tilde{x}_{nb} \leq K$. Let

$$M = PM, \quad b = Pb, \quad \tilde{\Phi}_n^{nb} = P_{nb} \tilde{\Phi}_n$$

Solving (22) as a least square problem and rewriting the appropriate terms based on the redefinitions in (23) lead to the following MPE reduced order model:

$$\tilde{\phi}_n^{nb} M \tilde{\Phi}_n^{nb} \tilde{a}^{l+1} = \tilde{\phi}_n^{nb} b$$

Upon obtaining the estimated POD coefficients $\tilde{a}$ from (24), the complete state $x$ can be approximated using the complete POD basis $\tilde{\Phi}_n$ as $\hat{x}$ presented in (25).

$$x \approx \hat{x} = \tilde{\Phi}_n \tilde{a}$$

The model (24) is solvable if $M_{r}$ is invertible. The condition will imply that $M = PM$ should be full rank. It is difficult to verify this requirement as $M$ will change when the operating condition changes. Ideally, a new set of state variables should be chosen whenever the operating condition changes. This problem is simplified by assuming that the selected state variables $\hat{x}$ will still be representative for the operating region we are interested in. Instead of imposing the rank condition on $M$, we impose the condition of full row rank on $\tilde{\Phi}_n$. It means that the selected state $\hat{x}$ should comprise of state variables whose dynamics are independent of each other.

The problem of finding the suitable set of state variables is similar to the sensor placement problem. In this paper, we will use the criterion proposed in [4] and [2] on the minimization of the aliasing error resulted from using $\hat{x}$.

Based on [4] and [2], the minimization of the aliasing error is equivalent to the minimization of the following norm

$$|| (\tilde{\Phi}_n - \Phi_n)^{-1} - I_n ||$$

where $I_n$ is an identity matrix and

$$\tilde{\Phi} = (\tilde{\phi}_1 \ldots \tilde{\phi}_n) = P\Phi_n.$$  

Thus we get the following optimization problem:

$$\text{find} \quad P$$

such that

$$|| (\tilde{\Phi}_n^{nb} P\Phi_n)^{-1} - I_n || < \text{TOL}$$

subject to

$$P P^T = I$$

$$P \in \{0, 1\}$$

There exist various methods to solve this optimization problem (27). With regard to the reduction of overall computational complexities, preference will be generally given to non-combinatorial methods, although these methods tend to yield non optimal set of $\hat{x}$. We will use the iterative version of the greedy algorithm [16]. Note that the constraint ensures that the matrix $(\tilde{\Phi}_n^{nb} P\Phi_n)$ is well conditioned. Therefore, it is approximated by

$$\text{cond}(\tilde{\Phi}_n^{nb} P\Phi_n) < \text{TOL}.$$  

VI. APPLICATION

We will consider the reduced order modeling of an inverter chain model. An inverter chain is built from the combination of resistors and capacitors. The inverter chain is a concatenation of 100 inverters in this case. The function of an inverter is to invert and smoothen the incoming signal. The output of an inverter will be a delayed, inverted signal. The time scale of an inverter chain ranges between 30–200 nanoseconds due to the high frequency sampling.

The schematic diagram of an inverter chain is given in Figure 1.
The state of the inverter chain model comprises of 100 nodal voltages and some electrical currents. In total, there are 104 state variables. The excitation signal $U_{op}$ for the inverter chain is depicted in Figure 2. The dynamical response of the inverter chain is shown in Figure 3 for the first 20 ns, when there are about 30 nodes active. The total simulation time is 30 ns. The remaining simulation data will be used as validation data.

Data collected during the first 15 ns is used to derive the POD basis $\Phi = (\varphi_1 \cdots \varphi_n)$. There are 251 samples, thus the snapshot matrix $X \in \mathbb{R}^{104 \times 53}$ reads:

$$X = \frac{1}{251} \begin{pmatrix} x(t_1) \cdots x(t_{251}) \end{pmatrix}$$

The POD basis $\Phi$ is found by solving the eigenvalue problem (8), the resulting eigenvalue plot is shown in Figure 4. The POD basis $\Phi$ comprises the eigenvectors corresponding to 20 largest eigenvalues.

The POD reduced order model is solved by implementing the least-square formulation in (16). To enhance the computational efficiency, the MPE method described in section V is implemented.

The state variables $\bar{x}$ are selected as follows. An iterative greedy algorithm is implemented. Several stages may need to be conducted.

**STAGE 1**

1) Set the obligatory set variables as the ones which directly relate to the input. In our case, it is the state variable $x_{104}$.
2) Set $\bar{x}^0 = x_{104}$
3) Suppose there are $K_c$ remaining state variables and re-index these state variables as $x^j$. For $j =
1, \ldots, K_c, \text{ set}
\tilde{x}_j = \begin{pmatrix} \tilde{x} \\ \tilde{x}_j^0 \end{pmatrix}

4) For \( j = 1, \ldots, K_c \), calculate the condition number of
\( c_j = \tilde{\Phi}^\top \tilde{\Phi} \)
where \( \tilde{\Phi} \) are the POD basis components correspond
to \( \tilde{x}_j \).

5) Choose the state variable \( x_m \) which yields the mini-
mum \( c_j \) from \( j = 1, \ldots, K_c \). Set
\( \tilde{x}_0 = \begin{pmatrix} \tilde{x}_0^0 \\ x_m \end{pmatrix} \)

6) Repeat step 3 until the condition number of \( \tilde{\Phi}^\top \tilde{\Phi} \) is
considered low enough.

STAGE 2

1) Suppose \text{dim}(\tilde{x}) \text{ obtained from executing the previous
steps is still too high, then reset the candidate points
as \( \tilde{x} \) from stage 1 and exclude the obligatory state
variable \( x_{104} \).}

2) For \( j = 1, \ldots, G - 1 \), run stage 1 again by setting
the initial obligatory mask for every step as
\( \tilde{x}_0 = \begin{pmatrix} \tilde{x}_{104}^0 \\ x_j \end{pmatrix} \)

Hence, we will get \( G - 1 \) options of the restriction
\( \tilde{x} \). Calculate the resulting dimension of \( \tilde{x} \) for every
choice of the initial mask.

3) Set the final mask \( \tilde{x} \) as the one with the minimum
dimension.

The application of this procedure to the inverter chain
model yields \text{dim}(\tilde{x}) = 29. The MPE reduced order model
will be built from less than 28\% of the original equations.

A. Reduced order model validation

Two comparisons are shown here. First, the reconstruc-
tion of the original dynamics by the POD and the MPE
models. The simulation conditions are the same as the
conditions when the POD basis is derived. Figure 5 shows
the plot of the original, POD, and MPE models. It is clear
that the simulation results are very close to each other.
The maximum of the average absolute error for the POD
model is 0.0019 while the maximum of the average absolute
error for the MPE model is 0.0020. It is obvious that the
MPE model can reconstruct the original dynamics very
well despite the fact that the model is built only from 29
equations.

The reduced order models are then validated by compar-
ing the dynamics simulated between 15 – 25 ns. Recall that
the POD basis is derived from the simulation data between
0 – 15 ns. The capability of the reduced models in capturing
the unknown dynamics will be investigated in this case.

Figure 6 shows the comparisons. The POD and MPE
models can adequately describe the dynamics of the nodes
which are already active during 0 – 15 ns. The models fail to
capture the dynamics of the nodes or state variables which
are inactive during the first 15 ns. This is understandable
as the inactive state variables will be viewed by the POD
basis as zero state variables only and it will be extremely
difficult, if not impossible for both reduced model to
simulate the conditions when these nodes become active.
During this time interval, both MPE and POD models have
the same maximum absolute error which occurs on the state
variable \( x_{23} \). The maximum absolute error is 3.2152 V. The
evaluation costs are decreasing for about 70\%.

VII. Conclusion

In conclusion, the MPE method yields a reduced order
model of the same quality as the POD model. From the
observations shown in this paper, there exists only a small performance deterioration when the MPE method is used while the reduced order model can be built from a significantly lower number of equations. The MPE method will be a potential reduction method for the future generation of nonlinear circuit design. The nonlinear characteristics can still be preserved and the nonlinear reduced order model can be built from significantly less number of original equations. The method is applicable to general projection-based reduction methods such as balanced truncation and Krylov subspace.

REFERENCES


