<table>
<thead>
<tr>
<th>Title</th>
<th>Piecewise-polynomial associated transform macromodeling algorithm for fast nonlinear circuit simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Zhang, Y; Fong, N; Wong, N</td>
</tr>
<tr>
<td>Citation</td>
<td>The 18th Asia and South Pacific Design Automation Conference (ASP-DAC 2013), Yokohama, Japan, 22-25 January 2013. In Conference Proceedings, 2013, p. 515-520</td>
</tr>
<tr>
<td>Issued Date</td>
<td>2013</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10722/189852">http://hdl.handle.net/10722/189852</a></td>
</tr>
<tr>
<td>Rights</td>
<td>Asia and South Pacific Design Automation Conference Proceedings. Copyright © IEEE.</td>
</tr>
</tbody>
</table>
Abstract — We present a piecewise-polynomial based associated transform algorithm (PWPAT) for macromodeling nonlinear circuits in system-level circuit design. The generated reduced model can provide both global and local accuracies with the most compact dimension. Numerical examples compare it with existing algorithms and verify its superior accuracy in higher order harmonics simulation over traditional Trajectory Piecewise-Linear (TPWL) approach.

I. INTRODUCTION

The latest electronic systems are integrating more and more complex functions into a small package. To fulfill the strict design requirements, comprehensive system-level circuit simulation is necessary during the design phase. This is usually performed in a hierarchical style: each functional block is firstly abstracted as a macromodel, then all the macromodels are interconnected to carry out the full-system simulation. Whatever circuit simulator is employed, a circuit macromodel can generally be expressed as a nonlinear ordinary difference equation (ODE),

\[ E \frac{dx}{dt} = f(x(t)) + Bu(t) \]
\[ y(t) = Cx(t) \]  

where \( E \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \), \( C \in \mathbb{R}^{p \times n} \), \( x(t) \in \mathbb{R}^{n} \) is the state variable, \( u(t) \in \mathbb{R}^{m} \) is the input and \( y(t) \in \mathbb{R}^{p} \) is the output. Since the order \( n \) is determined by the number of nodes in the circuit and the simulation runtime is generally proportional to \( n^3 \), if a circuit consists of a large number of passive and active devices (e.g., a post-layout simulation with back-annotated parasitics), it will consume vast computation resources and take a long time to simulate. Under these circumstances, a model order reduction (MOR) procedure is preferred to downsize the state dimension of the circuit model, i.e., to generate another ODE,

\[ \hat{E} \frac{dz}{dt} = \hat{f}(z(t)) + \hat{B}u(t) \]
\[ y(t) = \hat{C}z(t) \]  

where the reduced matrices \( \hat{E} \in \mathbb{R}^{q \times q} \), \( \hat{B} \in \mathbb{R}^{q \times m} \), \( \hat{C} \in \mathbb{R}^{p \times q} \) and \( z(t) \in \mathbb{R}^{q} \), with \( q \ll n \). Moreover, the electrical relationships between input nodes and output nodes should be preserved as much as possible.

Linear and weakly nonlinear MOR techniques have been studied extensively in recent years [1–6]. However, most practical circuits show strongly nonlinear relations. To accurately capture these systems, trajectory piecewise methods are proposed to divide the whole solution state space into multiple segments. In each of these segments, the system is approximated by a linear or weakly nonlinear ODE. And then MOR is performed in each of these ODEs to generate their corresponding reduced ODEs. Finally, the combination of these segments in a proper way constitutes the solution of the original problem. This is the intuitive understanding of the Trajectory Piecewise-Linear (TPWL) [7] or Piecewise-Polynomial (PWP) [8] approaches to achieve strongly nonlinear macromodeling.

In these two piecewise methods, TPWL employs linear MOR techniques to model each segment, and PWP needs more complex weakly nonlinear MOR (NMOR) techniques to generate the macromodels of each segment. Most popular NMOR techniques employ moment matching MOR, i.e., finding a projection matrix which projects the original bigger coefficient matrices \( E, f(\cdot), B \) and \( C \) into smaller coefficient matrices \( \hat{E}, \hat{f}(\cdot), \hat{B} \) and \( \hat{C} \). Representative algorithms are NORM [6] and the algorithms in [9] and [10]. In [11], a novel NMOR algorithm was proposed based on associated transform of multivariate Laplace-domain high-order Volterra transfer functions, resulting in generally much more compact reduced models. In this paper, we successfully adapt associated transform algorithm into PWP method, named piecewise-polynomial associated transform (PWPAT), to find more efficient reduced macromodels. Compared to TPWL, the model reduced by the proposed algorithm demonstrates much higher accuracy in higher order harmonics, which is critical for analog or radio-frequency (RF) circuit design. In the mean time, dimension of the reduced model is the most compact via PWPAT.

The paper is organized as following. In Section II, backgrounds of TPWL, PWP, associated transform etc., are reviewed. And then the detailed PWPAT algorithm is introduced in Section III. Section IV presents numerical examples to contrast with existing algorithms. Finally, Section V draws the conclusion.

II.BACKGROUND

A. TPWL and PWP algorithms

TPWL algorithm, such as [7, 8, 12, 13], is widely used in strongly NMOR in which linear MOR is performed in each piecewise segment. The algorithm is outlined as follows:
1. Given a nonlinear system of (1), linearize it at various expansion states points \(x_i, i = 1, 2, \cdots, k\),
\[
E \dot{x} = f(x_i) + A_i (x - x_i) + Bu
\]
\[
y = Cx.
\]
(3)

2. Generate a projection basis \(V\) through Arnoldi process for each segment and calculate a common orthonormal basis \(V\) via \([V_1 V_2 \cdots V_k]\), where \(V \in \mathbb{R}^{n \times q}\).

3. Perform linear MOR on each segment using \(V\), which reduces (3) into
\[
\dot{\hat{z}} = \hat{f}(x_i) + \hat{A}_i (z - z_i) + \hat{B} u
\]
\[
y = \hat{C} z
\]
where the reduced matrices \(\hat{E} \in \mathbb{R}^{q \times q}, \hat{A}_i \in \mathbb{R}^{q \times q}, \hat{B} \in \mathbb{R}^{q \times m}\) and \(\hat{C} \in \mathbb{R}^{p \times q}\).

4. The final reduced model is the weighted combination of all the reduced models
\[
\hat{z} = \sum_{i=1}^{s} \omega_i (z) \left( \hat{f}(x_i) + \hat{A}_i (z - z_i) + \hat{B} u \right)
\]
\[
y = \hat{C} z
\]
where \(\omega_i (z)\) is the weight function, which can be obtained with the approach in [7].

However, since TPWL algorithm divides the whole state space into small segments and uses first-order approximation in each segment, this algorithm can only provide global accuracy. In some practical simulation with interests in higher order harmonics, the macromodels generated by TPWL algorithm are insufficient and produce large errors. Therefore a method which provides both global and local accuracies is desired. In [8], a general form of TPWL, the PWP, is proposed. As shown in Fig.1, its improvement compared to traditional TPWL is that in each segment of the state space, higher order approximation is employed to replace the first-order approximation. Hence, both global and local accuracies are preserved, and the higher order harmonics can be accurately captured in the simulations by the generated macromodels.

**B. NMOR by associated transform**

In each segment of the state space, the input-output relationships can be described by an ODE. For example, consider one segment with a quadratic relation
\[
E \dot{x} = G_1 x + G_2 x \otimes x + Bu
\]
\[
y = Cx,
\]
(6)
where \(G_i = \frac{\partial^2 f}{\partial x_i^2} \bigg|_{x=x_i} \in \mathbb{R}^{n \times n}\) and \(\otimes\) denotes the Kronecker product. By Volterra theory, the 1st- to 3rd-order transfer functions of (6) can be obtained,
\[
H_1(s) = (sE - G_1)^{-1} b
\]
(7a)
\[
H_2(s_1, s_2) = \frac{1}{2} ((s_1 + s_2) E - G_1)^{-1} \{G_2 [H_1(s_1) \otimes H_1(s_2) + H_1(s_2) \otimes H_1(s_1)]\}
\]
(7b)
\[
H_3(s_1, s_2, s_3) = \frac{1}{3} ((s_1 + s_2 + s_3) E - G_1)^{-1} \{G_2 [H_1(s_1) \otimes H_2(s_2, s_3) + H_2(s_2, s_3) \otimes H_1(s_1) + H_1(s_2) \otimes H_2(s_1, s_3) + H_2(s_1, s_3) \otimes H_1(s_2) + H_1(s_3) \otimes H_2(s_1, s_2) + H_2(s_1, s_2) \otimes H_1(s_3)]\}
\]
(7c)

Traditionally, with the existing weakly NMOR algorithms such as [6,9,10], a projection matrix \(V \in \mathbb{R}^{n \times q}\) in which \(q < n\) can be generated by moments matching of all the Laplace coefficients \(1, s_1, s_2, s_3, s_1^2, s_1 s_2, s_1 s_3, s_2^2, s_2 s_3, s_3^2, \cdots\). And the reduced system will be constructed via this projection matrix. However, these Laplace coefficients \(s_1, s_2, s_3, \cdots\) are corresponding to multi-time coefficients \(t_1, t_2, t_3, \cdots\) by multidimensional inverse Laplace transform. The time-domain results are obtained from the unification of all these multi-time coefficients, i.e., \(h_n(t) = h_n(t_1, \cdots, t_n)\bigg|_{t_1=t_2=\cdots=t_n=t}\). Therefore, there are some redundancies if moments matching are performed with all the Laplace coefficients taken into account.

The theory of association of variables \(\mathcal{E}_n\) was firstly developed to unify these Laplace coefficients in the frequency domain [14], i.e., by converting the Laplace coefficients \(s_1, s_2, s_3, \cdots\) into a single \(s\) ahead of the moment matching, the generated reduced system has the most compact dimension. This is especially desirable in piecewise based methods since the assembled system consists of multiple subsystems.

**III. PWPAT**

As mentioned, PWP algorithm approximates each state space segment via a higher order ODE. The choice of segments is carried out in a “training” procedure: from \(t = 0\), the original system is reduced at current states point and simulated along the time axis. If the output deviates by a certain distance from the existing trajectory, a new expansion state point is picked up and a set of new reduction parameters are updated. Suppose finally \(k\) expansion points \(\{x_1, x_2, \cdots, x_k\}\) are chosen from the state space of the original system, each of which has an expression
\[
E \dot{x} = f(x_i) + G_{1i} (x - x_i) + G_{2i} (x - x_i) \otimes (x - x_i)
\]
\[
+ G_{3i} (x - x_i) \otimes (x - x_i) \otimes (x - x_i) + \cdots + Bu
\]
\[
y = C x
\]
(8)
\( \mathcal{A}_2(H_2(s_1, s_2)) = (sE - G_1)^{-1} \left( G_2(sE \otimes E + G_1 \otimes E \otimes G_1)^{-1} \right) (E^{-1}b) \otimes (E^{-1}b) \)

\[
= \left\{ E, \begin{bmatrix} G_1 & I_n \\ I_n & 0 \end{bmatrix} \right\} \bigg\{ E \otimes E, \begin{bmatrix} G_1 \otimes E + E \otimes G_1 & (E^{-1}b) \otimes (E^{-1}b) \\ G_2 & 0 \end{bmatrix} \bigg\} \\
= \begin{bmatrix} E & 0 \\ 0 & E \otimes E \end{bmatrix}, \begin{bmatrix} G_1 \otimes E + E \otimes G_1 & (E^{-1}b) \otimes (E^{-1}b) \\ 0 & 0 \end{bmatrix} \bigg\} = \left\{ P_1, \begin{bmatrix} M_1 & F_1 \\ N_1 & 0 \end{bmatrix} \right\}
\]

(13)

where \( G_{1_i}, G_{2_i}, \ldots \) are the first-, second- and higher order derivatives of \( f(x) \) at \( x_i \). For ease of illustration, we consider up to the second order. Therefore at each of the expansion point \( x_i \), we have

\[ E \dot{x} = f(x_i) + G_{1i}(x - x_i) + G_{2i}(x - x_i) \otimes (x - x_i) + Bu, \quad (9) \]

By defining \( \dot{x} = x - x_i \), the transfer functions of (9) are the same as (7) except the DC point described by the constant \( E \). Therefore at each of the expansion point \( x_i \), we have

\[ E \dot{x} = f(x_i) + G_{1i}(x - x_i) + G_{2i}(x - x_i) \otimes (x - x_i) + Bu, \quad (9) \]

Applying similar mechanism, the two-variable associated of the univariate transfer function \((s_1E - A)^{-1}b\) is simply \((E^{-1}b, E^{-1}b)\).

\[ \mathcal{A}_2((s_1E - A)^{-1}b) = E^{-1}b. \]

(11)

Subsequently, using the often used transfer function notation

\[ \begin{bmatrix} E, \begin{bmatrix} A & B \\ C & D \end{bmatrix} \end{bmatrix} = C(sE - A)^{-1}B + D \]

and (11), we get (13) on the top of this page. In (13), obviously \( \mathcal{A}_2(H_2) \) is recast into a higher order \((n^2 + n)\) linear state space. Using similar mechanism, \( \mathcal{A}_3(H_3) \) can be carefully derived to be

\[ \mathcal{A}_3(H_3) = (sE - G_1)^{-1}G_2 \hat{H}_3(s), \]

(14)

which can again be put into a linear state space as in (13). The resulting dimension of \( H_3(s) \) is \((2n^3 + 2n^2 + n)\).

Accordingly, the quadratic ODE of (9) can be converted into multiple separated linear ODEs via associated transform. To reduce (9), Arnoldi process is employed on the Krylov subspace of these linear ODEs to generate several projection bases \( V_1 \in R^{n \times (k_1 + 1)}, V_2 \in R^{n \times (k_2 + 1)} \) and \( V_3 \in R^{n \times (k_3 + 1)} \), where \( k_1, k_2 \) and \( k_3 \) are the moments matching order numbers of each Krylov subspace.

\[
\text{Algorithm 1 PWPAT algorithm for nonlinear macromodeling.} \\
1. \text{Transient simulation of original system } E \dot{x} = f(x(t)) + Bu(t) \text{ for } x_1, x_2, \ldots x_{\text{end\_training}}. \\
2. \text{Initialize } x_0, k = 1, 0. A_0 \text{ and } W_0 \text{ are the Jacobian and Hessian of } f(x(t)) \text{ at } x = x_0, K_0 = f(x_0). \\
3. \text{Perform NMOR via associated transform on } E_0 \dot{\tilde{x}} = A_0 x + W_0 x \otimes x + K_0 + B_0 (u) \text{ for } V_0. \\
4. \text{Generate initial training model } E_0 \dot{\tilde{z}} = A_0 \tilde{z} + \tilde{W}_0 \tilde{z} \otimes \tilde{z} + K_0 + B_0 (u). \\
5. \text{while } i < \text{end\_training do} \\
6. \text{Transient simulation of training model in step (4) for } z_i. \\
7. \text{if } \|V_i \cdot z_i - z_i\| > \delta \text{ then} \\
8. \quad k = k + 1, \\
9. \quad \text{Choose } x_i \text{ as a new expansion point. Update } x_k = x_i, E_k, A_k, W_k, B_k \text{ and } K_k = f(x_k). \\
10. \quad \text{Perform nonlinear MOR via associated transform on } E_k \dot{\tilde{z}} = A_k \tilde{x} + W_k \tilde{x} \otimes x + K_k + B_k (u) \text{ for } V_k. \\
11. \quad \text{Generate training model } E_k \dot{\tilde{z}} = \tilde{A}_k \tilde{z} + \tilde{W}_k \tilde{z} \otimes \tilde{z} + \tilde{K}_k + \tilde{B}_k (u). \\
12. \quad \text{Store } z_k, E_k, A_k, W_k, \tilde{K}_k \text{ and } \tilde{B}_k. \\
13. \quad \text{end if} \\
14. \text{end while} \\
15. \quad V = \text{svd} ([V_0 V_1 \cdots]). \\
16. \text{Construct the reduced model of the system} \\
\[ \dot{\tilde{z}} = \sum_{i=1}^{k} (w_i(z)) (\tilde{A}_i(z - z_i) + \tilde{W}_i(z - z_i) \otimes (z - z_i) + \tilde{K}_i + \tilde{B}_u (u) \\
\]

in which \( w_i (z) = \exp \left( -\beta \|z - z_i\|_m \right) \) where \( \beta \) is some constant (typically we use 25), and \( m = \min_i \|z - z_i\| \).
Then, by aggregating the projection bases $[V_1 V_2 V_3]$, an orthonormal project basis $V_i \in \mathbb{R}^{n \times q}$ ($q \ll n$) can be formulated. The reduced order model (ROM) generated by associated transform is usually more compact than the ROM generated by conventional weakly NMOR algorithms such as [6, 9, 10]. By singular value decomposition (SVD) on all the segment bases $V_i$, a uniform projection matrix for the whole system $V$ is generated. Finally, the ROM is obtained by a weighted combination of all these segments such that

$$\hat{E} \dot{z} = \sum_{i=1}^{k} (\omega_i(z)) \left[ \hat{f}(x_i) + \hat{G}_{1i}(z - z_i) + \hat{G}_{2i}(z - z_i) \right] \otimes (z - z_i) + \hat{B}u$$

$$y = C \left[ \sum_{i=1}^{k} (\omega_i(z)) (x_i + V(z - z_i)) \right]$$

where $\hat{f}(x_i) = V^T f(x_i)$, $\hat{G}_{1i} = V^T G_{1i} V$, $\hat{G}_{2i} = V^T G_{2i} (V \otimes V)$, $\hat{B} = V^T B$ and $z_i = V^T x_i$.

Algorithm 1 in the last page summarizes the completed PW-PAT algorithm.

IV. NUMERICAL EXPERIMENTS

A. Nonlinear transmission line

We first illustrate with a practical nonlinear transmission line circuit. The original circuit has 100 stages. All resistors and capacitors are set to 1 and the I-V characteristic of the diodes is $i_D = e^{40v_D} - 1$. Using modified nodal analysis (MNA), the circuit can be characterized by 100 differential equations with exponential relationships. The generated models can attain a 15X and 33X speedups for PW-PAT and TPWL, respectively, in time-domain simulation. From Fig. 3(b) and Fig. 3(d), although TPWL model shows a faster simulation speed, its accuracy with small signal input is not satisfactory. On the other hand, PW-PAT provides much better accuracies for both large and small signal inputs.

Local accuracy is important when performing simulations which reflects higher order characteristics, such as 2nd-order and 3rd-order intermodulation products and 1dB compression point. Fig. 4 sweeps both the fundamental and 2nd-order intermodulation products under different input magnitudes. The result verifies that PW-PAT model is necessary in performing circuit linearity simulations.

As a comparison, another PWP method via NORM algo-

![Fig. 2. Example of a nonlinear transmission line.](image-url)

![Fig. 3. Comparison between original and reduced models. (a) Large input in time domain. (b) Small input in time domain. (c) Large input in frequency domain. (d) Small input in frequency domain.](image-url)
TABLE I
RESULTS SUMMARY BETWEEN PWPA T, PWP VIA [6] AND TPWL

<table>
<thead>
<tr>
<th></th>
<th>Large input</th>
<th>Original</th>
<th>PWPAT</th>
<th>PWP via [6]</th>
<th>TPWL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>100</td>
<td>15x2</td>
<td>22x2</td>
<td>16x3</td>
<td></td>
</tr>
<tr>
<td>Time (Solving ODE)</td>
<td>315.57s</td>
<td>20.54s</td>
<td>25.29s</td>
<td>9.42s</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>—</td>
<td>0.0168</td>
<td>0.0175</td>
<td>0.0125</td>
<td></td>
</tr>
<tr>
<td>1st order harmonic</td>
<td>-23.0dBV</td>
<td>-23.1dBV</td>
<td>-23.1dBV</td>
<td>-23.0dBV</td>
<td></td>
</tr>
<tr>
<td>2nd order harmonic</td>
<td>-57.0dBV</td>
<td>-55.4dBV</td>
<td>-55.3dBV</td>
<td>-54.9dBV</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Small input</th>
<th>Original</th>
<th>PWPAT</th>
<th>PWP via [6]</th>
<th>TPWL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>100</td>
<td>15x2</td>
<td>22x2</td>
<td>16x3</td>
<td></td>
</tr>
<tr>
<td>Time (Solving ODE)</td>
<td>327.39s</td>
<td>20.79s</td>
<td>24.62s</td>
<td>9.57s</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>—</td>
<td>0.0128</td>
<td>0.0145</td>
<td>0.2197</td>
<td></td>
</tr>
<tr>
<td>1st order harmonic</td>
<td>-43.0dBV</td>
<td>-43.0dBV</td>
<td>-43.0dBV</td>
<td>-42.8dBV</td>
<td></td>
</tr>
<tr>
<td>2nd order harmonic</td>
<td>-94.4dBV</td>
<td>-89.5dBV</td>
<td>-89.3dBV</td>
<td>-67.8dBV</td>
<td></td>
</tr>
</tbody>
</table>

B. Analog buffer

In this example, a CMOS analog buffer circuit is simulated by the proposed algorithm. The small-signal circuit schematic is shown in Fig. 5 and it has totally 61 nodes. The transconductance of NMOS and PMOS, \( g_m \), is nonlinear with a polynomial expression

\[
i_{ds} = g_{m1}v_{gs} + g_{m2}v_{gs}^2 + g_{m3}v_{gs}^3 + \cdots.
\]

The circuit is reduced by PWPA T, PWP via [6] and TPWL modeling methods into ROM dimensions of 30, 60 and 33, respectively, and placed in the testbench for top-level simulations. Fig. 6(a) plots the spectrum of the final output node. From the zoomed 2nd- and 3rd-harmonic plots in Fig. 6(b) and Fig. 6(c), under similar ROM dimensions, the accuracy superiority of the PWPA T method over TPWL is obvious.

V. CONCLUSION

A novel trajectory piecewise based strongly nonlinear model order reduction algorithm, called PWPA T, has been proposed.
This algorithm combines the benefits of TPWL’s global accuracy and local higher order approximations to perform macro-modeling of nonlinear circuits. The generated macromodels are crucial when doing simulations calling for higher order characteristics.

ACKNOWLEDGMENTS

This work is supported in part by the Hong Kong Research Grant Council under GRF Project 718711E, and in part by the University Research Committee of The University of Hong Kong.

REFERENCES


