A convergence analysis of iterative macromodeling methods using Whitfield's estimator*

D. Deschrijver, T. Dhaene
Computer Modeling and Simulation
Dept. of Mathematics and Computer Science
University of Antwerp, Middelheimlaan 1
2020 Antwerp, Belgium

G. Antonini
UAq EMC Laboratory
Dept. of Electrical Engineering
University of L'Aquila
67100 Poggio di Roio, Italy

Abstract

It is well-known that the Vector Fitting algorithm suffers from convergence problems when fitting noisy frequency responses. These difficulties are caused by the use of a Sanathanan-Koerner iteration. This paper presents an alternative iteration framework, applied to the Orthonormal Vector Fitting technique. The convergence properties of this method are analysed and compared to the classical Vector Fitting iteration scheme.

1 Introduction

Vector Fitting [1] is a reliable macromodeling technique, which synthesises the transfer function of passive electrical and electronical structures. The method is essentially a reformulation of the Sanathanan-Koerner iteration [2], provided that the numerator and denominator are expanded in a basis of partial fractions [3]. This approach was shown to be robust, and it has been widely applied to many engineering applications.

One limitation of the algorithm, is that the accuracy of the method is highly dependent on the initial choice of starting poles. This problem has been tackled, by using orthonormal rational functions, which has lead to the Orthonormal Vector Fitting (OVF) algorithm [4][5]. The OVF technique significantly improves the numerical conditioning of the system equations (and limits the number of SK-iterations) if the real part of the starting poles is non-negligible.

Another limitation of the VF algorithm is that the convergence of the pole-relocation sometimes fails if the approximation data is contaminated with noise. This difficulty is caused by the use of a Sanathanan-Koerner iteration, which doesn't guarantee convergence to the true fundamental least squares solution. In this paper, this problem will be analyzed and a new technique will be proposed which attempts to address some of these issues. Its convergence properties are analysed and compared to the classical Vector Fitting scheme by some examples. The results in this paper are based on Whitfield's achievements for polynomial bases [6], and are extended for orthonormal rational functions.

2 Analysis of the estimator

The goal of the identification problem, is to approximate the frequency domain data $(s_k, H(s_k))$ by a rational function $R(s_k)$ in a least-squares sense.

$$R(s_k) = \frac{N(s_k)}{D(s_k)} = \frac{\sum_{p=1}^{P} c_p \phi_p^n(s_k)}{\tilde{c}_0 + \sum_{p=1}^{P} \tilde{c}_p \phi_p^d(s_k)} \quad (1)$$

*This work was supported by the Fund for Scientific Research Flanders (FWO Vlaanderen)

The basis functions ϕ are chosen to be the following orthonormal rational functions, based on a prescribed set of poles $-a_1, ..., -a_P$. They are defined as

$$\phi_p(s_k) = \frac{\kappa_p \sqrt{2\Re e(a_p)}}{s_k + a_p} \left(\prod_{j=1}^{p-1} \frac{s_k - a_j^*}{s_k + a_j} \right) \quad (2)$$

if the pole $-a_p$ is real, and

$$\phi_p(s_k) = \frac{\sqrt{2\Re e(a_p)}(s_k - |a_p|)}{(s_k + a_p)(s_k + a_{p+1})} \left(\prod_{j=1}^{p-1} \frac{s_k - a_j^*}{s_k + a_j} \right) \quad (3)$$

$$\phi_{p+1}(s_k) = \frac{\sqrt{2\Re e(a_p)}(s_k + |a_p|)}{(s_k + a_p)(s_k + a_{p+1})} \left(\prod_{j=1}^{p-1} \frac{s_k - a_j^*}{s_k + a_j} \right) \quad (4)$$

if $-a_p = -a_{p+1}^*$. For more details about these basis functions, and their application in the Vector Fitting scheme, the reader is referred to the OVF paper [4]. If the following vectors are defined as

$$C = [c_1 \ldots c_p]^T, \ \tilde{C} = [\tilde{c}_0 \ldots \tilde{c}_p]^T \quad (5)$$

$$\phi^n(s_k) = [\phi_1^n(s_k) \ldots \phi_p^n(s_k)]^T \quad (6)$$

$$\phi^d(s_k) = [1 \ \phi_1^d(s_k) \ldots \phi_p^d(s_k)]^T \quad (7)$$

$$N(s_k) = C^T \phi^n(s_k), \ D(s_k) = \tilde{C}^T \phi^d(s_k) \quad (8)$$

then the identification problem reduces to minimizing the non-linear cost function E in terms of the indeterminates C and $\tilde{C}$.

$$E_{NL} = \sum_{k=0}^{K} \left| \frac{N(s_k)}{D(s_k)} - H(s_k) \right|^2 \quad (9)$$

If the minimizing solution, say C_{est} and $\tilde{C}_{est}$, is truly a minimum (stationary point) of E_{NL}, then the gradient $\nabla E = [(\delta E_{NL}/\delta C)^T \ (\delta E_{NL}/\delta \tilde{C})^T]$ will be the 0-vector for this configuration.

A possible option to solve this non-linear problem, is the use of a Sanathanan-Koerner iteration. Based on a set of prescribed poles, the following cost function is minimized

$$E_{SK} = \sum_{k=0}^{K} \frac{\left| N^{(t)}(s_k) - D^{(t)}(s_k) H(s_k) \right|^2}{|D^{(t-1)}(s_k)|^2} \quad (10)$$

Clearly, the solution of (10) reduces to (9) asymptotically, if $D^{(t-1)}$ approaches D. Note however, that this iteration scheme doesn't necessarily guarantee convergence to the true fundamental least-squares solution, even though the cost function itself tends asymptotically to the true fundamental least-squares

1-4244-0454-1/06/$20.00 ©2006 IEEE

$$\frac{\delta E_{NL}}{\delta C} = \sum_{k=0}^{K} \frac{1}{|D(s_k)|^2} \left(\phi^n(s_k)(N(s_k) - D(s_k)H(s_k))^* + (\phi^n(s_k))^* (N(s_k) - D(s_k)H(s_k))\right) \tag{11}$$

$$\frac{\delta E_{NL}}{\delta \tilde{C}} = -\sum_{k=0}^{K} \frac{1}{|D(s_k)|^2} \left(\phi^d(s_k)\frac{N(s_k)}{D(s_k)}(N(s_k) - D(s_k)H(s_k))^* + (\phi^d(s_k))^* \frac{N(s_k)^*}{D(s_k)^*}(N(s_k) - D(s_k)H(s_k))\right) \tag{12}$$

$$\frac{\delta E_{SK}}{\delta C} = \sum_{k=0}^{K} \frac{1}{|D^{(t-1)}(s_k)|^2} \left(\phi^n(s_k)(N^{(t)}(s_k) - D^{(t)}(s_k)H(s_k))^* + (\phi^n(s_k))^* (N^{(t)}(s_k) - D^{(t)}(s_k)H(s_k))\right) \tag{13}$$

$$\frac{\delta E_{SK}}{\delta \tilde{C}} = -\sum_{k=0}^{K} \frac{1}{|D^{(t-1)}(s_k)|^2} \left(\phi^d(s_k)H(s_k)(N^{(t)}(s_k) - D^{(t)}(s_k)H(s_k))^* + (\phi^d(s_k))^* H(s_k)^*(N^{(t)}(s_k) - D^{(t)}(s_k)H(s_k))\right) \tag{14}$$

$$\begin{aligned}\frac{\delta E_{WF}}{\delta C} = \sum_{k=0}^{K} \frac{1}{|D^{(t-1)}(s_k)|^2} \Bigg[&\phi^n(s_k)\left(N^{(t)}(s_k) - \frac{N^{(t-1)}(s_k)}{D^{(t-1)}(s_k)}D^{(t)}(s_k) + N^{(t-1)}(s_k) - D^{(t-1)}(s_k)H(s_k)\right)^* \\ &+ (\phi^n(s_k))^* \left(N^{(t)}(s_k) - \frac{N^{(t-1)}(s_k)}{D^{(t-1)}(s_k)}D^{(t)}(s_k) + N^{(t-1)}(s_k) - D^{(t-1)}(s_k)H(s_k)\right)\Bigg]\end{aligned} \tag{15}$$

$$\begin{aligned}\frac{\delta E_{WF}}{\delta \tilde{C}} = \sum_{k=0}^{K} \frac{1}{|D^{(t-1)}(s_k)|^2} \Bigg[&\phi^d(s_k)\frac{N^{(t-1)}(s_k)}{D^{(t-1)}(s_k)}\left(N^{(t)}(s_k) - \frac{N^{(t-1)}(s_k)}{D^{(t-1)}(s_k)}D^{(t)}(s_k) + N^{(t-1)}(s_k) - D^{(t-1)}(s_k)H(s_k)\right)^* \\ &+ (\phi^n(s_k))^* \frac{(N^{(t-1)}(s_k))^*}{(D^{(t-1)}(s_k))^*}\left(N^{(t)}(s_k) - \frac{N^{(t-1)}(s_k)}{D^{(t-1)}(s_k)}D^{(t)}(s_k) + N^{(t-1)}(s_k) - D^{(t-1)}(s_k)H(s_k)\right)\Bigg]\end{aligned} \tag{16}$$

$$E_{WF} = \sum_{k=0}^{K} \frac{1}{|D^{(t-1)}(s_k)|^2} \left| N^{(t)}(s_k) - \frac{N^{(t-1)}(s_k)D^{(t)}(s_k)}{D^{(t-1)}(s_k)} + N^{(t-1)}(s_k) - D^{(t-1)}(s_k)H(s_k)\right|^2 \tag{17}$$

$$= \sum_{k=0}^{K} \left| \frac{N^{(t)}(s_k)}{D^{(t-1)}(s_k)} - \frac{N^{(t-1)}(s_k)D^{(t)}(s_k)}{(D^{(t-1)}(s_k))^2} + \frac{N^{(t-1)}(s_k)}{D^{(t-1)}(s_k)} - H(s_k)\right|^2 \tag{18}$$

$$= \sum_{k=0}^{K} \left| \frac{\frac{\prod_{p=1}^{P-1}(s_k+z_{p,n}^{(t)})}{\prod_{p=1}^{P}(s_k+a_p)}}{\frac{\prod_{p=1}^{P}(s_k+z_{p,d}^{(t-1)})}{\prod_{p=1}^{P}(s_k+a_p)}} - \frac{\frac{\prod_{p=1}^{P-1}(s_k+z_{p,n}^{(t-1)})}{\prod_{p=1}^{P}(s_k+a_p)}}{\frac{\prod_{p=1}^{P}(s_k+z_{p,d}^{(t-1)})}{\prod_{p=1}^{P}(s_k+a_p)}} \frac{\frac{\prod_{p=1}^{P}(s_k+z_{p,d}^{(t)})}{\prod_{p=1}^{P}(s_k+a_p)}}{\frac{\prod_{p=1}^{P}(s_k+z_{p,d}^{(t-1)})}{\prod_{p=1}^{P}(s_k+a_p)}} + \frac{\frac{\prod_{p=1}^{P-1}(s_k+z_{p,n}^{(t-1)})}{\prod_{p=1}^{P}(s_k+a_p)}}{\frac{\prod_{p=1}^{P}(s_k+z_{p,d}^{(t-1)})}{\prod_{p=1}^{P}(s_k+a_p)}} - H(s_k)\right|^2 \tag{19}$$

$$= \sum_{k=0}^{K} \left| \frac{\prod_{p=1}^{P-1}(s_k + z_{p,n}^{(t)})}{\prod_{p=1}^{P}(s_k + z_{p,d}^{(t-1)})} - \frac{\prod_{p=1}^{P-1}(s_k + z_{p,n}^{(t-1)})}{\prod_{p=1}^{P}(s_k + z_{p,d}^{(t-1)})} \frac{\prod_{p=1}^{P}(s_k + z_{p,d}^{(t)})}{\prod_{p=1}^{P}(s_k + z_{p,d}^{(t-1)})} + \frac{\prod_{p=1}^{P-1}(s_k + z_{p,n}^{(t-1)})}{\prod_{p=1}^{P}(s_k + z_{p,d}^{(t-1)})} - H(s_k)\right|^2 \tag{20}$$

$$= \sum_{k=0}^{K} \left| \frac{\prod_{p=1}^{P-1}(s_k + z_{p,n}^{(t)})}{\prod_{p=1}^{P}(s_k + z_{p,d}^{(t-1)})} - H^{(t-1)}(s_k)\frac{\prod_{p=1}^{P}(s_k + z_{p,d}^{(t)})}{\prod_{p=1}^{P}(s_k + z_{p,d}^{(t-1)})} + H^{(t-1)}(s_k) - H(s_k)\right|^2 \tag{21}$$

$$= \sum_{k=0}^{K} \left| \sum_{p=1}^{P} \alpha_p^t \phi_p^{n,t}(s_k) - H^{(t-1)}(s_k)\left(\sum_{p=1}^{P} \tilde{\alpha}_p^t \phi_p^{d,t}(s_k) + \tilde{\alpha}_0\right) + H^{(t-1)}(s_k) - H(s_k)\right|^2 \tag{22}$$

$$= \sum_{k=0}^{K} \left| \sum_{p=1}^{P} \alpha_p^t \phi_p^{n,t}(s_k) - H^{(t-1)}(s_k)\left(\sum_{p=1}^{P} \tilde{\alpha}_p^t \phi_p^{d,t}(s_k)\right) - H(s_k)\right|^2 \tag{23}$$

criterion. This can easily be seen by calculating the gradient of the cost function. After applying some basic vector differentiation rules, the following gradients (11)(12) and (13)(14) are obtained for the non-linear estimator and for the Sanathanan-Koerner cost function respectively.

Note that the gradients (11)(12) and (13)(14) are asymptotically equivalent ($D^{(t-1)} \to D$), only if $N(s_k)/D(s_k) = H(s_k)$, which is the situation when interpolating data. This equivalence does not hold when solving an overdetermined problem, or when the data is contaminated with noise ($N(s_k)/D(s_k) \approx H(s_k)$). As the true minimum of E_{NL} is a solution of $\delta E_{NL}/\delta C = 0$ and $\delta E_{NL}/\delta \tilde{C} = 0$, then it is clear to see that the minimizing solution of (10) will be different. Note that this corresponds exactly to the kind of situations when convergence problems with Vector Fitting occur.

Whitfield proposed minimizing the cost function (17) which is obtained by a Taylor series expansion of the non-linear estimator E_{NL} with respect to the system parameters.

If $N^{(t-1)} \to N$ and $D^{(t-1)} \to D$, then this cost function has several advantages :

1) The cost function E_{WF} (17) tends asymptotically to the non-linear estimator E_{NL} (9)

2) The gradient of the cost function E_{WF} (15)(16) tends asymptotically to the gradient of the non-linear estimator E_{NL} (11)(12).

3) The cost function E_{WF} (17) provides the same minimizing solution as the non-linear estimator E_{NL} (9).

As the second and third criteria are not necessarily satisfied by SK-iteration, it is to be expected that the Whitfield iteration outperforms the SK-iteration if the data is contaminated with noise, or when the modelling errors are relatively large.

In the next section, the proposed cost function is reduced to a linear set of equations which can be solved efficiently.

3 Whitfield's Iteration using ORFs

The numerator and denominator of (1) can be written in factorized form as follows

$$N^{(t)}(s_k) = \sum_{p=1}^{P} c_p^t \phi_p^n(s_k) = \frac{\prod_{p=1}^{P-1}(s + z_{p,n}^{(t)})}{\prod_{p=1}^{P}(s + a_p)} \quad (24)$$

$$D^{(t)}(s_k) = \tilde{c}_0 + \sum_{p=1}^{P} \tilde{c}_p^t \phi_p^d(s_k) = \frac{\prod_{p=1}^{P}(s + z_{p,d}^{(t)})}{\prod_{p=1}^{P}(s + a_p)} \quad (25)$$

The constant term in the denominator, $\tilde{c}_0$, is fixed to 1, since numerator and denominator can be divided by the same constant value without loss of generality. Based on this expansion of numerator and denominator, Whitfield's cost function can be simplified in a very elegant way as shown in equations (17-23).

Basically, this estimator reduces to a simple form, which corresponds to solving the following set of linear least-squares equations

$$\sum_{p=1}^{P} \alpha_p^t \phi_p^{n,t}(s_k) - H^{(t-1)}(s_k) \left(\sum_{p=1}^{P} \tilde{\alpha}_p^t \phi_p^{d,t}(s_k) \right) = H(s_k) \quad (26)$$

Note that the basisfunctions $\phi_p^{n,t}$ and $\phi_p^{d,t}$ of iteration t are based on the previously identified poles $-z_{p,d}^{t-1}$. If these equations are compared to the pole-identification equations of the SK-scheme,

$$\sum_{p=1}^{P} \alpha_p^t \phi_p^{n,t}(s_k) - H(s_k) \left(\sum_{p=1}^{P} \tilde{\alpha}_p^t \phi_p^{d,t}(s_k) \right) = H(s_k) \quad (27)$$

it becomes clear that the right Cauchy block of the system matrices is multiplied with the function values from the previous iteration, rather than the actual function values.

4 Example

4.1 Underground cable system The proposed technique is used to model an underground cable system, where the phase angle has been compensated by an amount corresponding to the lossless time delay in the dielectric (coaxial mode). For more details, the reader is referred to section VI of [7].

The data samples and the initial poles are logarithmically spaced over the frequency range of interest [1 Hz - 100 MHz]. A 10-pole, strictly proper rational model was calculated using the classical OVF algorithm using 25 Sanathanan-Koerner and Whitfield iterations.

Figure 1 shows the RMS error of both approaches in terms of iteration count. Clearly in the noiseless situation, the results are quite similar and the improvement of the new approach is fairly negligible. If 1% of noise has been added to the data, the classical SK-iteration causes the fitting error to diverge, while the WF-cost function succeeds in minimizing the error. The final accuracy of the model is comparable to the noiseless case.

4.2 PEEC Example The method is applied to model S_{12} of a two-conductor transmission line with frequency dependent per unit length parameters [9]. The data samples and initial poles are linearly spaced over the frequency range of interest [0 Hz - 10 GHz].

A 60-pole proper transfer function was calculated, and the RMS error is considered in terms of iteration count for 3 particular situations :

1. 10 Whitfield iterations are calculated (VF/OVF). (In the first iteration, the previous function values are chosen to be the data values).
2. 10 SK-iterations are calculated (VF/OVF).
3. 5 SK-iterations are calculated (VF/OVF), succeeded by 5 Whitfield iterations (VF/OVF).

As can be seen in Figure 2 (dotted lines), minimization of the Whitfield cost function doesn't always guarantee better convergence results. It has been observed that a poor fitting model can be obtained, particularly if the initial set of starting poles is not well-chosen. This result shows that the approach doesn't fully solve all convergence problems, as a good estimate of the poles remains essential to ensure asymptotic convergence ($N^{(t-1)} \to N$ and $D^{(t-1)} \to D$). A practical solution is to apply the method after the convergence of the SK-iteration has stalled, and to terminate if the WF-iteration stalls or tends to diverge. Figure 2 shows that this combined approach (full lines) can give some improvement compared to the classical formulation of the VF or OVF methods (dashed lines).

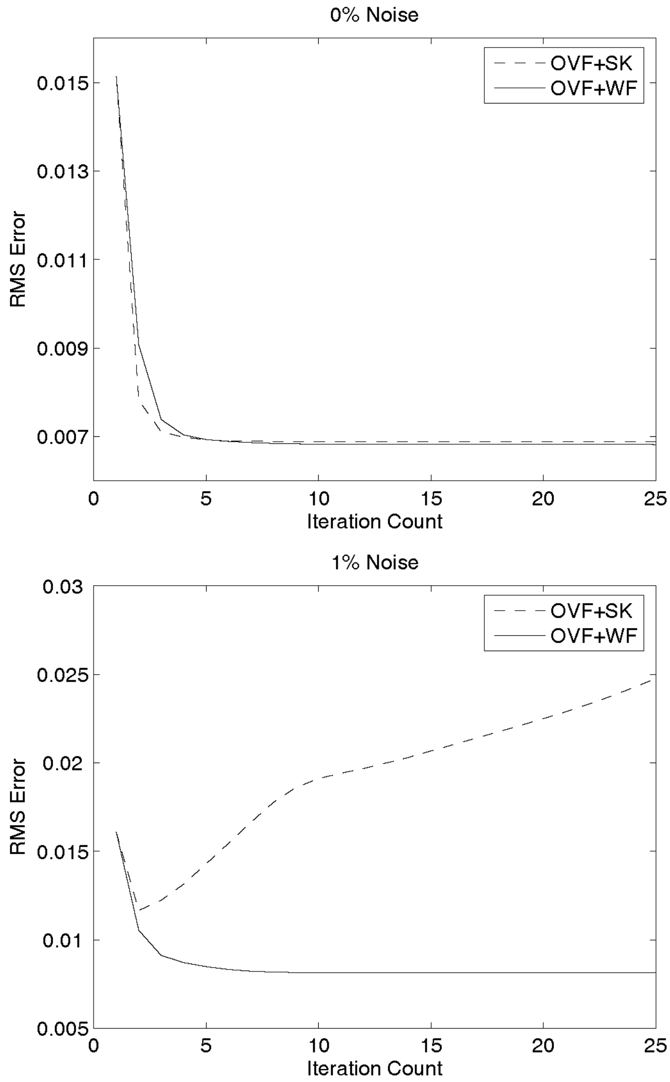

Figure 1: RMS Error as function of Iteration Count - 0% Noise (upper) 1% Noise (lower)

5 Conclusions

Some considerations are given about the iterative framework of the (Orthonormal) Vector Fitting method. The use of a different cost function is proposed, and its convergence properties have been analyzed. It is observed that the behaviour of the Whitfield iteration can be less reliable than the SK-iteration, as a good choice of initial poles remains essential. Nevertheless, some practical examples show that the Whitfield iteration can be a useful tool to lower the RMS error if the SK-iteration diverges or stalls.

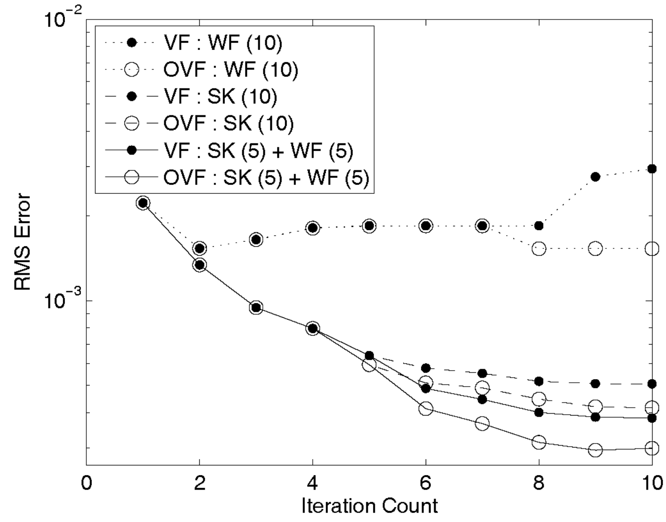

Figure 2: RMS Error as function of Iteration Count

References

[1] B. Gustavsen and A. Semlyen,"Rational Approximation of Frequency Domain Responses by Vector Fitting", IEEE Transactions on Power Delivery, vol. 14, no. 3, pp. 1052-1061, 1999.

[2] C.K. Sanathanan and J. Koerner, "Transfer Function Synthesis as a Ratio of Two Complex Polynomials", IEEE Transactions on Automatic Control, vol. AC-8, pp. 56-58, 1963.

[3] W. Hendrickx, D. Deschrijver and T. Dhaene, "Some Remarks on the Vector Fitting Iteration", Progress in Industrial Mathematics at ECMI 2004, Mathematics in Industry, vol. 8, Springer-Verlag, Post-conference proceedings of European Conference on Mathematics for Industry (ECMI 2004), pp. 134-138, 2006.

[4] D. Deschrijver, B. Haegeman and T. Dhaene, "Orthonormal Vector Fitting : A Robust Macromodeling Tool for Rational Approximation of Frequency Domain Responses", IEEE Transactions on Advanced Packaging, in print, May 2006.

[5] D. Deschrijver and T. Dhaene, "Broadband Macromodelling of Passive Components using Orthonormal Vector Fitting", IEE Electronics Letters, vol. 14, no. 21, pp. 1160-1161, 2005.

[6] A.H. Whitfield, "Asymptotic Behavior of Transfer Function Synthesis Methods", International Journal of Control, vol. 45, pp. 1083–1092, 1987.

[7] B. Gustavsen, "Improving the Pole Relocating Properties of Vector Fitting", 10^{th} IEEE Workshop on Signal Propagation on Interconnects, To be published, 2006.

[8] S. Grivet-Talocia, M. Bandinu, and F. Canavero, "An Automatic Algorithm for Equivalent Circuit Extraction from Noisy Frequency Reponses", IEEE International Symposium on EMC, pp. 163-168, 2005.

[9] A.E. Ruehli, "Equivalent Circuit Models for Three Dimensional Multiconductor Systems", IEEE Transactions on Microwave Theory and Techniques, vol. MTT-22, no. 3, pp. 216-221, 1974.