Lehrstuhl für Hochfrequenztechnik
der Technischen Universität München

Application of System Identification (SI) to Full–Wave Time Domain Characterization of Microwave and Millimeter Wave Passive Structures

Fabio Coccetti

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor–Ingenieurs

genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Dr.-Ing. Wolfgang Utschick
Prüfer der Dissertation:	1. Univ.-Prof. Dr. techn. Peter Russer 2. Prof. Christos Christopoulos, University of Nottingham, UK

Die Dissertation wurde am 12.05.2004 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 15.07.2004 angenommen.

Abstract

Numerical time–domain methods for electromagnetic field simulations typically provide very broad band frequency–domain characterizations as well as transient response with a single simulation and without in general requiring any pre–processing. However long simulation times and large memory requirements arise for the case of electromagnetic structures characterized by low loss (high quality factor) and high aspect ratios (complex three–dimensional structures), since the first yields long transient responses and the second small time discretization intervals. Passive network impulse responses can be characterized by the singularity expansion method theory, implying that they can be efficiently described by means of exponentially damped oscillating components corresponding to the network natural frequencies. In principle, the entire time behavior of an electromagnetic structure can therefore be predicted from a few time samples by applying high resolution parametric model estimation techniques, based on system identification (SI) methods. These methods allow the determination of the network equivalent model directly from the simulated results. The number of a model's parameters, also called model order, and the parameters themselves, typically represented by complex natural frequencies or poles, significantly effect this methodology since they are indicators of the complexity and the accuracy of the model respectively. Once correctly identified these parametric analytical descriptions can replace more cumbersome and demanding full–wave numerical models, in network level (SPICE like) simulators, enabling a much faster analysis. Although SI techniques are a quite well known topic in electromagnetic numerical applications, a systematic and efficient approach is still missing. The aim of the present work is to develop an improved approach first, by re-examining the theoretical background of the network oriented modelling (NOM) in order to justify the use of a poles series model (Prony model) as the more obvious choice for describing passive electromagnetic structures, and second by reviewing some of the most common and efficient SI techniques for the model order selection and model parameters estimation. The intention is to formulate an algorithm that allows

for entire network modelling to be carried out in a completely autonomous and automatic fashion. The methodology is to estimate the model's parameters from the time–domain responses generated by means of a full–wave analysis, be it the Transmission Line Matrix (TLM) method or the finite difference time–domain (FDTD) method, and by adaptively refining them, fit the model recovered responses, to the numerically simulated ones. This algorithm runs in parallel with a full-wave analysis which is discontinued as soon as the model accuracy becomes satisfactory. In this way a time demanding numerical simulation may be reduced by one order of dimension. Since the model taken in consideration is Prony's and the parameter estimation procedures are Prony based, the algorithm is called Prony Model based System Identification (PMSI). Once the network responses are available they may be used for identifying the network natural frequencies of the impedance (admittance) Foster representation, enabling the direct implementation of the corresponding lumped element equivalent circuit. Since the Foster representation for the impedance (admittance) is practically a Prony model this operation may be carried out again by means of the PMSI algorithm.

Acknowledgments

First of all I would like to thank Prof. Peter Russer for the opportunity he offered me to undertake this experience and for his guidance. My deep gratitude goes moreover to a number of persons who have shared with me these past years in the good as well as in the not so good times enriching my professional and private life with their opinion and view of the world. Among them my partner Hariet Mieskes, for her patience and moral support, my many friends and colleagues among whom Vitali Hertzuvsky, and Mark Casciato, deserve special thanks for the numerous stimulating discussions, and last but, off course not least, my parents without whom all this would not exist at all.

Munich April 1*st* 2004

Contents

Chapter 1

Introduction

The analysis of electromagnetic (EM) structures involves the solution of Maxwell's equations and the identification of appropriately chosen physical parameters, in order to define a global and efficient description. This may be established by the impedance **Z**, admittance **Y**, scattering **S** matrices, or any other suitable representation, which will be eventually used in a network level solver such as SPICE, for high level system analysis. Among the several type of electromagnetic field solvers, time–domain numerical methods such as the Transmission Line Matrix (TLM) method and the finite difference time–domain (FDTD) method constitute powerful tools able to handle structures of arbitrary three–dimensional geometries and composed of arbitrary materials, yielding with a single simulation, a broad band characterization and the complete system time–domain behavior (i.e. network transient and driving responses). These characteristics make time–domain techniques the proper tool for the analysis of an entire class of novel devices, based on three–dimensional designs (as the Micro–electro–mechanical Systems (MEMS), and Low Temperature Co–fired Ceramic (LTCC) technologies). These devices are very promising for applications in future handheld communication equipment, working at very wide frequency bands (e.g. $> 20\%$ as for the Ultra Wide Band (UWB) systems such as the wireless personal area network WPAN). Although long computational times and large memory requirements have so far limited the use of time–domain techniques making frequency–domain methods the preferred choice, in the last couple of years, progress in more efficient time–domain methods seems to be reversing this trend. Beside improved processor speed and memory availability, much effort is currently being devoted to improving time–domain methods, with the goal of making of them competitive tools for analysis and optimization.

Table 1.1: Example of model complexity in the case of high aspect ratio and low losses devices (* Performances for a PA8000 Risk CPU 360MHz).

Structure	Δl $[m]$	dt $[sec]$	Aspect Ratio	Memory $[Mb]$	Run Time* $[hour]$
MEMS Switch [14]	$2.5\ 10^{-6}$	$2.06\ 10^{-15}$	240	153	~ 18
LTCC antenna [42]	$1.5\ 10^{-5}$	$11.6\ 10^{-15}$	135	114	~ 12.5

1.1 Definition of the problem

The present work focuses on the TLM algorithm limits which arise in the modelling of low loss and high aspect ratio electromagnetic structures. These characteristics are quite common in planar microwave and millimeter wave applications, but are also becoming extremely relevant with regard to an entire class of innovative devices, as for instance those based on MEMS and LTCC technologies. If on one hand planar structures can be easily analyzed in the frequency–domain via two dimensional based tools (in general by Method of Moments), for arbitrary three dimensional structures full-wave field solvers are recommended. In these latter methods accurate analysis by means of discrete space and time techniques (be it TLM or FDTD), demand space and time discretization intervals able to resolve the dimension of the smallest feature to be taken into account. Now if high aspect ratio objects and low losses materials (i.e. high quality factor structures) have to be modelled, fine space resolution yields a large number of discretization elements and small time discretization intervals yield a very high number of simulation time steps (overall simulation duration) in order to observe the transient, which is typically very long. The immediate consequences are large memory requirements and long simulation run times. An example is given in the Tab. 1.1, where the TLM model complexity for a MEMS and an LTCC device are given. Three of the main strategies applied for reducing model complexity are: 1) at the spatial discretization level (subgridding) the aim is to optimize the field resolution to keep down the memory requirements [77] [27]; 2) at the algorithm kernel level, to reduce the correspondent state–space equation complexity [54] [63] [9]; and 3) at the post processing level where by means of digital signal processing analysis tools (SI methods) to predict and model the structure behavior [12]. This latter will be the focus of the present work.

1.2 State of the art

An electromagnetic structure at microwave and millimeter wave frequencies behaves as a passive distributed circuit which may be modelled as an abstract multiport network [72]. The characterization of a passive network, excited be means of impulse driving function, can be efficiently carried out by an exponential or poles series model using the singularity expansion method (SEM – C.E. Baum 1991) [8]. The entire network response characterization therefore, becomes a problem of pole identification and the fitting of an exponential series model. These properties will be exploited for two different purposes:

- **prediction**: modelling of the TLM responses for simulation time reduction
- **equivalent circuit extraction**: Z– (Y –) matrix model identification in equivalent Foster representation (EFR).

The prediction of a system response by analytical modelling is a necessary procedure for the drastic reduction of the observation time. By focusing on the very early results of a time–domain analysis, which is typically very time consuming, coupled with system identification (SI) high resolution signal processing tools, a parametric description of physical phenomena or systems can be provided. This description is adaptively refined and as soon as becomes accurate enough it provides a synthetic analytic model of the response itself.
Once the response is available the complete high frequency network characterization is usually given by means of scattering parameters. However, since the sought goal in most applications is to produce an electrical equivalent lumped element representation, to be directly applied in a network level SPICE like simulator, approaches which exploit the scattering parameters to derive such elements have been developed. In the present work the **Z**– (**Y**–) matrix descriptions are computed from the scattering parameters and analyzed in the time–domain by exploiting the above mentioned SI techniques. The result of this analysis is a mathematical description of such matrices, closely related to network synthesis representations such as the Foster or Cauer canonic representations. These latter have the advantage of providing a very general network description in terms of passive lumped elements and ideal transformers. The entire synthesis procedure can be automatized and once the frequency validity range has been given, from the time–domain responses, it yields the final equivalent network.
Network–oriented modelling is revisited in the chapter 2 in order to provide a motivational background for the present work. In electromagnetic field analysis, network oriented methods contribute significantly to the problem formulation and solution methodology [20]

[21] [73]. In network theory, systematic approaches for circuit analysis are based on the separation of the circuit into the circuit elements and the connection circuit [13]. The connection circuit represents the topological structure of the circuit and contains only the connections, including ideal transformers. In the connection circuit neither energy storage nor energy dissipation occurs. The connection circuit, governed by the Tellegen's theorem [80] [81], and Kirchoff's law [13], connects the circuit elements that may consist of be one–port or multiports. Electromagnetic field theory and network theory are related via method of moments [25]. In the method of moments the electromagnetic field functions are represented by an expansion in series of basis functions. The linear system of equations relating the expansion coefficients can be seen as a set of linear circuit equations, and if a rational expansion of the circuit equations exists then a lumped element equivalent circuit representation can be given. In analogy with network theory individual subdomains are characterized via subdomain relations obtained either analytically or numerically and described in an unified format by using the generalized network formulation [73]. For each subdomain the impedance (admittance) matrix allows for a canonical representation, for example the Foster or Cauer [24] [6] representation. For any linear, reciprocal lossless distributed circuit equivalent Foster or Cauer descriptions exist [18] [46]. The canonic form of each subregion is embedded into a network representation of the entire domain by a connection circuit representing the subregion boundary surfaces. The Foster or Cauer representations can be obtained via an analytical solution of the field problem or by applying pole extraction procedures to the numerical results of the field problem. Among the available numerical techniques, the finite–difference time–domain (FD) method [86] [79] and the transmission line matrix (TLM) method [10] [71] [28], are very well suited for the impulse response modelling of general three dimensional structures characterized by arbitrary geometry and materials.

Electromagnetic passive structures can be characterized by a mathematical description fundamentally based upon parametric expressions or models. Be it coefficients of ordinary differential equations, as is typically done in systems theory, or the pole residues of a poles series model, the use of parameters for describing a physical system such as a passive network has numerous advantages. The first is that a parametric description is far more efficient than a non parametric one. An example of this is the speech digital processing, where only a few parameters are necessary for speech synthesis of an audio signal to be transmitted over a communication channel (shared resource), instead of the more cumbersome technique of a sampled and quantized description of the same signal [55]. Another advantage in dealing with parametric descriptions of signal or systems is the possibility of

model simplification, enabled by model order reduction techniques, which can within an acceptable tolerance or error reduce the number of parameters [49]. A further advantage is the possibility of establishing bijection between the original parametric model and a second physical one, which has the characteristic of being immediately realizable and therefore of practical interest. The model based parameter estimation approach is the subject of the third chapter. The sought parameters are the structure natural frequencies (poles) or eigen-frequencies. After providing a brief introductory background in SI, the focus will be on the different model based parameter estimation techniques. Since, as shown in chapter 2, the most natural model to be considered in writing the impedance (admittance) representation of the Green's function is the series poles expansion, also known as Prony model, the entire chapter will be devoted to techniques for efficiently seeking the model's, poles and the corresponding residues. Prony model related pole extrapolation and model order selection techniques are introduced and evaluated.

Chapter 4 is dedicated to the introduction of the Prony model base System Identification (PMSI) algorithm and its use first for network response prediction and second for the identification of the impedance (admittance) Foster representation. Results are shown for some planar as well as three-dimensional structures in terms of the improvements in accuracy, due to application of the PMSI algorithm, for both the network response models and the network impedance (admittance) models.

Chapter 2

Network-Oriented Modelling (NOM)

For analysis purposes a complex electromagnetic structure may be subdivided into spatial subdomains. Comparing a distributed circuit, represented by a geometric structure, with a lumped element circuit, represented by a network, the spatial subdomain may be considered as a subcircuit and the complete set of boundary surfaces separating the subcircuits correspond to the connection circuit [73]. The generalized network formulation has been already successfully applied to electromagnetic structures such as waveguide N–furcation in combination with partial wave synthesis [51].

Fig.2.1 shows the segmentation of a generic electromagnetic structure into different regions R_l, $l = 1, 2, \ldots, N$, separated by boundary surface B_{lk} from the generic neighboring region R_k with $l \neq k$. Each of the subregions may either contain an arbitrary substructure or simply be filled with any material including perfect electric (PEC) or perfect magnetic (PMC) conductors. In the network analogy the set of all boundary surfaces represent the connection circuit whereas the subdomain R_l represents the circuit elements. On the boundary surface the tangential components are related to each other via the Green's function. The Green's function can be expressed by means of an expansion upon a basis of eigenfunctions. This representation is similar to the Foster's representation, and allows a full–wave characterization of the region by providing a closed form between the boundary tangential field components. Expanding the tangential electric and magnetic fields into basis functions provides an equivalent circuit representation for the boundary surfaces. This equivalent representation is based on the Tellegen's theorem which describes a fundamental relationship between voltage and current in a network [80], [81], and can be generalized to field problems [59]. The eigenfunction expansion of the dyadic Green's function allows one to highlight the central role of the parametric modelling in synthetically describing

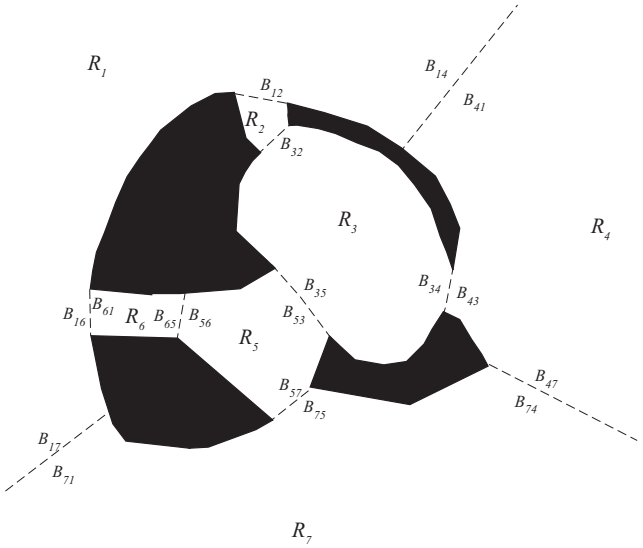


Figure 2.1: Abstract electromagnetic structure partitioned into N regions R_l ($l = 1, 2, \ldots, N$, $N = 8$ for this example) and for which connection boundary surfaces are designated B_{lk} (dark regions may be either PEC or PMC).

electromagnetic structures.

2.1 Characterization of the connection circuit

The connection network is established via the relation of the tangential field components on both sides of the boundary surfaces. Since the boundary or connection region, exhibits zero volume, no energy is stored in it and no power loss occurs therein.

2.1.1 The field theoretic formulation of Tellegen's theorem

Let V be an arbitrarily small volume enclosing the entire boundary surface and any of the PEC or PMC regions in the partitioned structure, i.e. $V \supset \bigcup_{l=1}^{N} \partial R_l$ with N number of subregions. Referring to Fig.2.1 the volume V is depicted in Fig.2.2. Starting directly from Maxwell's equations and considering the electric and magnetic field components in

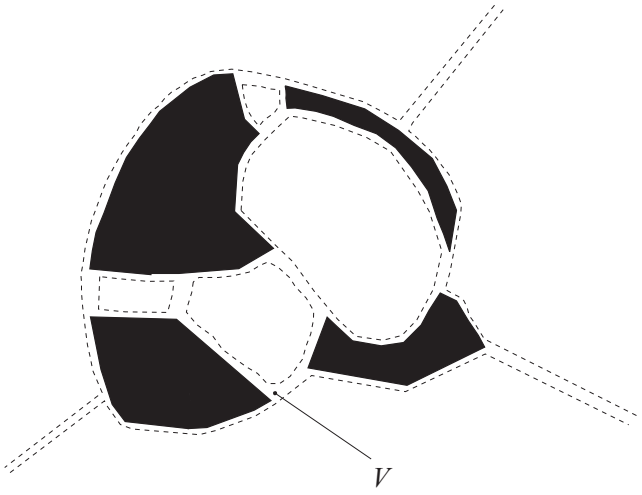


Figure 2.2: Volume V enclosing the segmentation boundaries including the PEC or/and PMC objects.

the Laplace domain $\underline{\mathcal{E}}(\rho, s)$ and $\underline{\mathcal{H}}(\boldsymbol{\rho}, s)$, the following relationship applies [72]

$$\begin{aligned}\int_{\partial V} \underline{\mathcal{E}}(\boldsymbol{\rho}, s) \wedge \underline{\mathcal{H}}(\boldsymbol{\rho}, s) &= -\int_V \underline{\mathcal{E}}(\boldsymbol{\rho}, s) \wedge \underline{\mathcal{J}}(\boldsymbol{\rho}, s) \\ &- \int_V \underline{\mathcal{E}}(\boldsymbol{\rho}, s) \wedge (s\underline{\mathcal{D}}(\boldsymbol{\rho}, s) - \int_V (s\underline{\mathcal{B}}(\boldsymbol{\rho}, s)) \wedge \underline{\mathcal{H}}(\boldsymbol{\rho}, s)\end{aligned} \tag{2.1}$$

where each term to the right of the equal sign constitutes an integral over a volume which in the limit be reduced to zero. This allows Tellegen's theorem for the field problem to be written as

$$\int_{\partial V} \underline{\mathcal{E}}(\boldsymbol{\rho}, s) \wedge \underline{\mathcal{H}}(\boldsymbol{\rho}, s) = 0. \tag{2.2}$$

Consider a single portion of the entire boundary surface, $B_{li} = R_l \bigcap R_i$, the Tellegen's theorem for this interface becomes

$$\int_{B_{li}} \underline{\mathcal{E}}(\boldsymbol{\rho}, s) \wedge \underline{\mathcal{H}}(\boldsymbol{\rho}, s) + \int_{B_{il}} \underline{\mathcal{E}}(\boldsymbol{\rho}, s) \wedge \underline{\mathcal{H}}(\boldsymbol{\rho}, s) = 0 \tag{2.3}$$

which implies that all power incident on the boundary B_{il} from one side flows entirely to the other side. If either of the two regions l or i is PEC or PMC both integral terms of (2.3) identically vanish. Tellegen's theorem also holds for the case of general electric and magnetic fields in an arbitrary material body (different source, filling material and

geometries but same structure partitioning) [72].

2.1.2 Discretized connection network

The electromagnetic field can be expanded in a series of basis functions. The assumption of orthogonality may be introduced in order to simplify the computation. Consider a set of expansion functions of dimension N_{li} on the boundary B_{li} and a set of dimension N_{il} on the boundary B_{il}, the corresponding transverse field components can be written as follows

$$\begin{cases} \underline{\tilde{\mathcal{E}}}_t^{li}(\rho, s) = \sum_{n=1}^{N_{li}} V_n^{li}(s) e_n^{li}(u, v) \\ \underline{\tilde{\mathcal{H}}}_t^{li}(\rho, s) = \sum_{m=1}^{N_{li}} I_m^{li}(s) h_m^{li}(u, v) \end{cases} \qquad \begin{cases} \underline{\tilde{\mathcal{E}}}_t^{il}(\rho, s) = \sum_{n=1}^{N_{il}} V_n^{il}(s) e_n^{il}(u, v) \\ \underline{\tilde{\mathcal{H}}}_t^{il}(\rho, s) = \sum_{m=1}^{N_{il}} I_m^{il}(s) h_m^{il}(u, v) \end{cases} \tag{2.4}$$

where the tilde on $\underline{\tilde{\mathcal{E}}}_t^{\xi}(\rho, s)$ and $\underline{\tilde{\mathcal{H}}}_t^{\xi}(\rho, s)$, $\xi = li, il$, denotes that they are a finite series expansion approximation, and the variable s denotes the Laplace domain. The set of curvilinear coordinate (u, v, w) are defined so that an arbitrary point on the boundary, ρ, is given by $\rho \equiv (u, v, w = const.)$. The quantities $V_n^{\xi}(s)$ and $I_n^{\xi}(s)$, $\xi = li, il$, denote the amplitudes of the electric and magnetic fields in the Laplace domain, respectively. The frequency independent terms $e_n^{\xi}(u, v)$ and $h_n^{\xi}(u, v)$, $\xi = il, il$, are the so called basis structure functions in their differential forms, for the electric and magnetic fields respectively [72]. Rewriting (2.3) as a finite series expansion approximation gives,

$$\int_{B_{li}} \underline{\tilde{\mathcal{E}}}_t^{li}(\boldsymbol{\rho}, s) \wedge \underline{\tilde{\mathcal{H}}}_t^{li}(\boldsymbol{\rho}, s) + \int_{B_{il}} \underline{\tilde{\mathcal{E}}}_t^{il}(\boldsymbol{\rho}, s) \wedge \underline{\tilde{\mathcal{H}}}_t^{il}(\boldsymbol{\rho}, s) = 0. \tag{2.5}$$

Inserting the decomposition (2.4) into the (2.5), yields:

$$\begin{aligned} &\sum_{n=1}^{N_{li}} \sum_{m=1}^{N_{li}} V_n^{li}(s) I_m^{li}(s) \int_{B_{li}} e_n^{li}(u, v) \wedge h_m^{li}(u, v) \\ &- \sum_{n=1}^{N_{il}} \sum_{m=1}^{N_{il}} V_n^{il}(s) I_m^{il}(s) \int_{B_{li}} e_n^{il}(u, v) \wedge h_m^{il}(u, v) = 0 \end{aligned} \tag{2.6}$$

where the minus sign in the second term is due to the inversion of the boundary integral surface. By introducing the matrix $\Lambda = \left[\Lambda^{li}\ \Lambda^{il}\right]^T$ whose elements are given as

$$\Lambda^{\xi}_{nm} = \int_{B_{li}} e^{\xi}_n(u,v) \wedge h^{\xi}_m(u,v), \quad \text{with} \quad \xi = li, il, \tag{2.7}$$

and by defining the following vectors

$$\begin{aligned} \mathbf{V}^{li} &= \left[V^{li}_1, V^{li}_2, \ldots, V^{li}_{N_{li}}\right]^T \qquad \mathbf{V}^{il} = \left[V^{il}_1, V^{il}_2, \ldots, V^{il}_{N_{il}}\right]^T \\ \mathbf{I}^{li} &= \left[I^{li}_1, I^{li}_2, \ldots, I^{li}_{N_{li}}\right]^T \qquad \mathbf{I}^{il} = \left[I^{il}_1, I^{il}_2, \ldots, I^{il}_{N_{il}}\right]^T \end{aligned} \tag{2.8}$$

the general matrix notation for Tellegen's theorem for the boundary B_{li} or B_{il} becomes

$$\begin{bmatrix} \mathbf{V}^{li} \\ \mathbf{V}^{il} \end{bmatrix}^T \Lambda \begin{bmatrix} \mathbf{I}^{li} \\ \mathbf{I}^{il} \end{bmatrix} = 0 \tag{2.9}$$

In case of orthonormal electric and magnetic basis structure functions the matrices Λ^{ξ}, $\xi = li, il$, become the identity matrix and (2.9) reduces to

$$\begin{bmatrix} \mathbf{V}^{li} \\ \mathbf{V}^{il} \end{bmatrix}^T \begin{bmatrix} \mathbf{I}^{li} \\ \mathbf{I}^{il} \end{bmatrix} = 0 \tag{2.10}$$

The relation between the field formulation of Tellegen's theorem (2.3) and its network formulation expressed by (2.10), can be established in terms of MOM solution by considering the expansion coefficients in (2.8) as generalized voltages and currents. (2.10) implies that the two vectors are linearly independent on a space of dimension $(N_{li} + N_{il})$. This means that by arbitrarily choosing $\bar{N} = N_{li} + N_{il}$ independent variables, among the voltage and current elements of $\mathbf{V}^{li}$,$\mathbf{V}^{il}$ and $\mathbf{I}^{li}$,$\mathbf{I}^{il}$, the rest of the $\bar{N}$ elements may be considered as the dependent variables, obtained from the first by means of a linear transformation. If the independent variables are chosen by taking $\mathbf{V}^{il}$ and $\mathbf{I}^{li}$, the dependent ones are given by the elements of $\mathbf{V}^{li}$ and $\mathbf{I}^{il}$ and the relation between the two sets can be given by introducing the matrix of turns ratios $\mathbf{N}^i_l$ (the indexes highlight that this transformation takes place

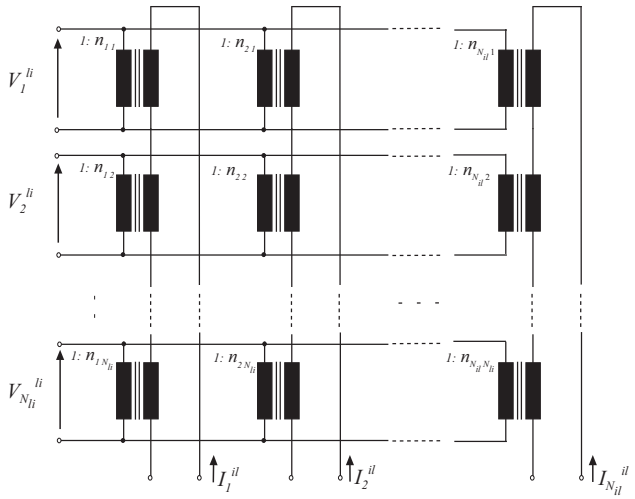


Figure 2.3: Canonical form of the connection network between regions R_i and R_l through B_{il} and B_{li}.

between the boundary portion B_l and B_i) as follows

$$\begin{bmatrix} \mathbf{V}^{li} \\ \mathbf{I}^{il} \end{bmatrix} = \begin{bmatrix} \mathbf{N}_l^i & \mathbf{0} \\ \mathbf{0} & \mathbf{N}_l^{i\,T} \end{bmatrix} \begin{bmatrix} \mathbf{V}^{il} \\ \mathbf{I}^{li} \end{bmatrix}, \tag{2.11}$$

where $\mathbf{N}_l^i$ is rectangular $N_{li} \times N_{il}$ given by

$$\mathbf{N}_l^i = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1N_{il}} \\ n_{21} & n_{22} & \cdots & n_{2N_{il}} \\ \vdots & \vdots & \ddots & \vdots \\ n_{N_{li}1} & n_{N_{li}2} & \cdots & n_{N_{li}N_{il}} \end{bmatrix}. \tag{2.12}$$

The lumped element equivalent circuit interpretation of (2.11) can be given as a network of ideal transformer connected as shown in Fig. 2.3, whereas the real coefficients n_{mk}, $m = 1, 2, \ldots, N_{li}$, $k = 1, 2, \ldots, N_{il}$, denotes the turns ratio.

2.2 The characterization of circuit and subcircuit

2.2.1 Green‘s function representation by series expansion of eigenfunctions

In an electromagnetic problem the separation of the Helmholtz equations yields ordinary differential equation of the Sturm–Lioville form, [15] [32]. In the one dimensional case this assumes the following expression in the given range of validity $x \in [a, b]$

$$\frac{d}{dx}\left[p(x)\frac{d\varphi(x)}{dx}\right] + [q(x) + \lambda\sigma(x)]\,\varphi(x) = g(x). \tag{2.13}$$

The functions $p(x)$, $q(x)$ and $\sigma(x)$ are positive and real functions of x, λ is a separation constant and $g(x)$ is the excitation distribution function. The Green‘s function $G(x, x')$ is the solution for the Sturm–Lioville equation for the field at same arbitrary location x caused by a point source located at x'. So (2.13) can be written as

$$\frac{d}{dx}\left[p(x)\frac{dG(x, x')}{dx}\right] + [q(x) + \lambda\sigma(x)]\,G(x, x') = \delta(x - x') \tag{2.14}$$

which satisfies the boundary conditions, given as follows

$$G(x, x') + k_a\frac{dG(x, x')}{dx}\bigg|_{x=a} = G(x, x') + k_b\frac{dG(x, x')}{dx}\bigg|_{x=b} = 0 \tag{2.15}$$

where k_a, k_b are real constants. Letting $p(x)$, $q(x)$ and $\sigma(x)$ identically equal 1 for $x \in [a, b]$ (2.14) becomes the scalar Helmoltz wave equation. The Green’s function $G(x, x')$ is therefore the impulse response of the given electromagnetic structure (linear system interpretation) described by (2.14) and (2.15). In the case of a generic excitation $g(x)$ the solution $\varphi(x)$ may be obtained by the spatial convolution of the Green’s function and the excitation function itself or

$$\varphi(x) = \int_a^b G(x, x')g(x')dx' \tag{2.16}$$

Call $\varphi_n(x)$ the eigenfunctions of the associated homogeneous equation of (2.13), with λ_n the corresponding eigenvalues then, a possible representation of the Green’s function may

be given by expanding it in terms of the eigenfunctions $\varphi_n(x)$ as follows

$$G(x, x') = \sum_n A_n(x')\varphi_n(x) \tag{2.17}$$

with the expansion coefficients dependent upon the source location x'. Now by substituting the (2.17) into (2.14), this latter can be rewritten in terms of eigenfunction expansion as

$$\sum_n A_n(x')(\lambda - \lambda_n)\sigma(x)\varphi_n(x) = \delta(x - x'). \tag{2.18}$$

Recognizing that the eigenfunctions $\varphi_n(x)$ form a complete set of orthogonal basis functions, both sides of (2.18) can be multiplied by $\varphi_m^*(x)$ and integrating in $[a, b]$ gives

$$\sum_n A_n(x')(\lambda - \lambda_n)\int_a^b \sigma(x)\varphi_m^*(x)\varphi_n(x)dx = \int_a^b \varphi_m^*(x)\delta(x - x')dx. \tag{2.19}$$

This by the orthogonality property

$$\int_a^b \sigma(x)\varphi_m^*(x)\varphi_n(x) = \begin{cases} 0 \text{ for } m \neq n \\ \eta_n \text{ for } m = n \end{cases}$$

(2.18) reduces to

$$A_n(x')(\lambda - \lambda_m)\eta_n = \varphi_m^*(x') \tag{2.20}$$

where η_n is a normalization constant which become unitary for the case of orthonormal eigenfunctions. From (2.20) the coefficients $A_n(x')$ can be derived and substituted into (2.17) in order to obtain

$$G(x, x', \lambda) = \sum_n \frac{\varphi_n^*(x')\varphi_n(x)}{\eta_n(\lambda - \lambda_n)} \tag{2.21}$$

Recognizing that (2.18) is a form of wave equation, the eigenvalues correspond to the propagation coefficients, and are in general complex values, i.e. $\lambda_n = \sigma_n + j\omega_n$. For (2.21) these values represent the system poles for $G(x, x')$ and the expression may be referred to as the poles series expansion of the Green's function.

2.2.2 Impedance and admittance representation of the Green's function

Let us consider a subregions R_l of the original structure (see Fig. 2.1). Assuming that the field solution is described by ordinary differential equations of the Sturm–Lioville type (2.13), it may be expressed by the integral relationship (2.16) which in a three-dimensional curvilinear coordinate system and in Laplace domain, becomes

$$\underline{\mathcal{E}}(\boldsymbol{\rho}, s) = \int_{R_l}^{'} \underline{\mathcal{G}}_e^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s) \wedge \underline{\mathcal{J}}(\boldsymbol{\rho}', s), \tag{2.22}$$

where $\underline{\mathcal{J}}(\boldsymbol{\rho}', s)$ is the excitation electric density current distribution within the region R_l and $\underline{\mathcal{G}}_e^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s)$ is the electric double form dyadic Green's function (defined in A.3 see appendix A) associated with the considered region R_l, and the prime in the integral denote that this operation is carried out with respect to the source point $\boldsymbol{\rho}'$[72]. The current density can be express by means of a surface density current $\underline{\mathcal{J}}_{eA}(\boldsymbol{\rho}', s)$ flowing on the surface $\boldsymbol{\rho}' = (u', v', w' = w_0)$ which can be related to ∂R_l, as follows

$$\underline{\mathcal{J}}(\boldsymbol{\rho}', s) = \delta(w' - w_0)\, n' \wedge \underline{\mathcal{J}}_{eA}(\boldsymbol{\rho}', s) \quad \boldsymbol{\rho}' \in \partial R_l. \tag{2.23}$$

where the n is the unit differential form corresponding to the vertical coordinate w whose orientation is normal outward with respect to ∂R_l. Inserting (2.23) in (2.22) yields

$$\underline{\mathcal{E}}(\boldsymbol{\rho}, s) = \int_{\partial R_l}^{'} \underline{\mathcal{G}}_e^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s) \wedge \underline{\mathcal{J}}_{eA}(\boldsymbol{\rho}', s). \tag{2.24}$$

Now by imposing the continuity condition of the tangential components, and applying the equivalence principle the surface ∂R_l is replaced by a perfect magnetic conductor and the equivalent electric surface current defined as,

$$\underline{\mathcal{J}}_{eA}(\boldsymbol{\rho}', s) = \underline{\mathcal{H}}_t^l(\boldsymbol{\rho}', s). \tag{2.25}$$

Also the tangential component of the electric field can be obtained by recognizing that $\underline{\mathcal{E}}_t^l = n \lrcorner n \wedge \underline{\mathcal{E}}$. Applying this relationship together with (2.25), in (2.24) results in

$$\underline{\mathcal{E}}_t^l(\boldsymbol{\rho}, s) = \int_{\partial R_l}^{'} n \lrcorner \left(n \wedge \underline{\mathcal{G}}_e^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s)\right) \wedge \underline{\mathcal{H}}_t^l(\boldsymbol{\rho}', s). \tag{2.26}$$

The superscript l in (2.26) implies that the corresponding quantity belongs to the region R_l, so that $\underline{\mathcal{E}}_t^l$ and $\underline{\mathcal{H}}_t^l$, for instance, represent the electric and magnetic field components tangential to ∂R_l, transferred into the region R_l. The operation $n \lrcorner n \wedge$ applies only to the observation point $\boldsymbol{\rho}$ while the integral is over $\boldsymbol{\rho}'$. This allows to define

$$\underline{\mathcal{Z}}^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s) = n \lrcorner \left(n \wedge \underline{\mathcal{G}}_e^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s)\right) \tag{2.27}$$

as the double differential form for the impedance representation of the dyadic Green's function. The substitution of (2.27) into (2.26) yields

$$\underline{\mathcal{E}}_t^l(\boldsymbol{\rho}, s) = \int_{\partial R_l}^{'} \underline{\mathcal{Z}}^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s) \wedge \underline{\mathcal{H}}_t^l(\boldsymbol{\rho}', s). \tag{2.28}$$

which provides an integral relationship between the tangential electric and magnetic components on the considered subdomain surface ∂R_l.
Similarly introducing the magnetic double form dyadic Green's function $\underline{\mathcal{G}}_M^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s)$ associated with the considered region R_l, and applying duality principle the (2.24) can be immediately rewritten as

$$\underline{\mathcal{H}}(\boldsymbol{\rho}, s) = \int_{\partial R_l}^{'} \underline{\mathcal{G}}_m^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s) \wedge \underline{\mathcal{J}}_{mA}(\boldsymbol{\rho}', s). \tag{2.29}$$

The continuity condition of the tangential components applied to the surface ∂R_l replaced by a perfect electric conductor yields an equivalent magnetic surface current defined as,

$$\underline{\mathcal{J}}_{mA}(\boldsymbol{\rho}', s) = -\underline{\mathcal{E}}_t^l(\boldsymbol{\rho}', s), \tag{2.30}$$

while the tangential component of the magnetic field are given by $\underline{\mathcal{H}}_t^l = n \lrcorner n \wedge \underline{\mathcal{H}}$. Applying this relationship together with (2.30), in (2.29) results in

$$\underline{\mathcal{H}}_t^l(\boldsymbol{\rho}, s) = \int_{\partial R_l}^{'} \underline{\mathcal{E}}_t^l(\boldsymbol{\rho}', s) \wedge n \lrcorner \left(n \wedge \underline{\mathcal{G}}_m^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s)\right). \tag{2.31}$$

$\underline{\mathcal{H}}_t^l$ and $\underline{\mathcal{E}}_t^l$, represent the electric and magnetic field components tangential to ∂R_l, transferred into the region R_l. Is possible now to define

$$\underline{\mathcal{Y}}^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s) = n \lrcorner \left(n \wedge \underline{\mathcal{G}}_m^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s)\right) \tag{2.32}$$

as the double differential form for the admittance representation of the dyadic Green's function. From substituting (2.32) into (2.31) follows

$$\underline{\mathcal{H}}_t^l(\boldsymbol{\rho}, s) = \int_{\partial R_l}^{'} \underline{\mathcal{E}}_t^l(\boldsymbol{\rho}', s) \wedge \underline{\mathcal{Y}}^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s). \tag{2.33}$$

which provides an integral relationship between the tangential magnetic and electric components on the considered subdomain surface ∂R_l.
Now applying an expansion similar to that applied to (2.2) which resulted in the form of (2.21) the impedance and the admittance representation can be decomposed into a set of orthonormal basis functions such as

$$\underline{\mathcal{Z}}^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s) = \frac{1}{s}\mathrm{z}_0^l(\boldsymbol{\rho}, \boldsymbol{\rho}') + \sum_p \frac{\mathrm{z}_p^l(\boldsymbol{\rho}, \boldsymbol{\rho}')}{s - s_p^l} \tag{2.34}$$

and

$$\underline{\mathcal{Y}}^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s) = \frac{1}{s}\mathrm{y}_0^l(\boldsymbol{\rho}, \boldsymbol{\rho}') + \sum_q \frac{\mathrm{y}_q^l(\boldsymbol{\rho}, \boldsymbol{\rho}')}{s - s_q^l}, \tag{2.35}$$

where z_p^l and y_q^l are complex double forms and s_p^l and s_q^l are the complex poles of the impedance and admittance expansions for the given subregion, respectively. The total number and the position of these poles, also called subregion natural frequencies, depends essentially on the subregion internal physical characteristics (geometry and materials). For the case of passive structures, the poles must satisfy the stability criterion, which implies that $\Re(s_p^l) \leq 0$ and $\Re(s_q^l) \leq 0$ for any p and q.
In the case of a loss–free domain, (2.34) and (2.35) become purely imaginary, and their poles lie on the imaginary axes, so that these equations become:

$$\underline{\mathcal{Z}}^l(\boldsymbol{\rho}, \boldsymbol{\rho}', \omega) = \frac{1}{j\omega}\mathrm{z}_0^l(\boldsymbol{\rho}, \boldsymbol{\rho}') + j\omega \sum_p \frac{\mathrm{z}_p^l(\boldsymbol{\rho}, \boldsymbol{\rho}')}{\omega^2 - (\omega_p^l)^2} \tag{2.36}$$

and

$$\underline{\mathcal{Y}}^l(\boldsymbol{\rho}, \boldsymbol{\rho}', \omega) = \frac{1}{j\omega}\mathrm{y}_0^l(\boldsymbol{\rho}, \boldsymbol{\rho}') + j\omega \sum_q \frac{\mathrm{y}_q^l(\boldsymbol{\rho}, \boldsymbol{\rho}')}{\omega^2 - (\omega_q^l)^2}. \tag{2.37}$$

The quantities $\mathrm{z}_0^l(\boldsymbol{\rho}, \boldsymbol{\rho}')$ and $\mathrm{y}_0^l(\boldsymbol{\rho}, \boldsymbol{\rho}')$ represent the static parts of the dyadic Green's function, while each term $\mathrm{z}_p^l(\boldsymbol{\rho}, \boldsymbol{\rho}')$ and $\mathrm{y}_q^l(\boldsymbol{\rho}, \boldsymbol{\rho}')$ represents the residue corresponding to the poles ω_p^l and ω_q^l. All these quantities are frequency (Laplace) independent, and are only a function of the spatial coordinate $\boldsymbol{\rho}$ and the source position $\boldsymbol{\rho}'$.
A matrix representation of (2.34) and (2.35) can be derived by applying a MOM expansion to the tangential fields on the boundary surface ∂R_l, by means of a complete set of orthonormal basis functions. Let us consider the expansion of the tangential electric and magnetic field components as given by

$$\begin{aligned} \underline{\tilde{\mathcal{E}}}_t^l(\rho, s) &= \textstyle\sum_{n=1}^{N_l} V_n^l(s) \underline{e}_n^l(u, v) \\ \underline{\tilde{\mathcal{H}}}_t^l(\rho, s) &= \textstyle\sum_{m=1}^{N_l} I_m^l(s) \underline{h}_m^l(u, v) \end{aligned} \tag{2.38}$$

where the tilde implies the approximation due to the expansion variable, s denotes the Laplace domain, (u, v, w) are the curvilinear coordinates defined so that the generic point on the boundary, be it ρ, is given by $\rho \equiv (u, v, w = const.)$. The quantities $V_n^l(s)$ and $I_m^l(s)$, denote the amplitudes of the electric and magnetic fields as function of the normal coordinate and the complex frequency s. The frequency independent terms $\underline{e}_n^l(u, v)$ and $\underline{h}_m^l(u, v)$, are the basis structure functions, for the electric and magnetic fields, respectively, and are equivalent to $\underline{e}_n^{li}(u, v)$ and $\underline{h}_n^{li}(u, v)$ in the expansion used in (2.4) on the portion B_{li} of ∂R_l. These structure functions depend only on the transverse coordinates (u, v) and fulfill the following orthogonal relationship

$$\int_{\partial R_l} \underline{e}_n^{l\,*}(u, v) \wedge \underline{e}_m^l(u, v) = \delta_{mn} \tag{2.39}$$

$$\int_{\partial R_l} \underline{h}_n^{l\,*}(u, v) \wedge \underline{h}_m^l(u, v) = \delta_{mn} \tag{2.40}$$

where δ_{mn} denotes the Kronecker's delta (function which assumes the value 1 only and only if $m = n$ and is zero otherwise). If now the expansions of (2.38) are inserted into (2.28), it results in

$$\sum_{n=1}^{N_l} V_n^l(s) \underline{e}_n^l(u, v) = \sum_{n=1}^{N_l} I_n^l(s) \int_{\partial R_l}' \underline{\mathcal{Z}}^l(\rho, \rho', s) \wedge \underline{h}_n^l(u', v'). \tag{2.41}$$

Now multiplying both sides of (2.41) by ${e_m^l}^*(u, v)$ and integrating over ρ and the domain ∂R_l, and applying the orthogonality property of the basis functions yields

$$V_m^l(s) = \sum_{n=1}^{N_l} I_n^l(s) \int_{\partial R_l} \int_{\partial R_l}^{'} {e_m^l}^*(u, v) \wedge \underline{\underline{\mathcal{Z}}}^l(\rho, \rho', s) \wedge h_n^l(u', v'), \tag{2.42}$$

where the rest of the right hand side of (2.42) is defined as

$$Z_{mn}^l(s) = \int_{\partial R_l} \int_{\partial R_l}^{'} {e_m^l}^*(u, v) \wedge \underline{\underline{\mathcal{Z}}}^l(\rho, \rho', s) \wedge h_n^l(u', v') \tag{2.43}$$

the mn-th element of the $N_l \times N_l$ square impedance matrix for the given region R_l. Equation (2.42) can be written in matrix notation as

$$\mathbf{V}^l(s) = \mathbf{Z}^l(s)\mathbf{I}^l(s) \tag{2.44}$$

with $\mathbf{V}^l(s)$ are the $\mathbf{I}^l(s)$ the $N_l \times 1$ vectors of the electric and magnetic field expansion coefficients. By inserting (2.34) into (2.43) the pole expansion of the impedance matrix elements become

$$Z_{mn}^l(s) = \frac{1}{s} z_{0mn}^l + \sum_p \frac{z_{p_{mn}}^l}{s - s_p^l} \tag{2.45}$$

whereas z_{0mn}^l and $z_{p_{mn}}^l$ are complex coefficients corresponding to the residues of the poles in zero and in s_p^l, respectively, given by

$$z_{0mn}^l = \int_{\partial R_l} \int_{\partial R_l}^{'} {e_m^l}^*(u, v) \wedge \mathrm{z}_0^l(\rho, \rho') \wedge h_n^l(u', v') \tag{2.46}$$

and

$$z_{p_{mn}}^l = \int_{\partial R_l} \int_{\partial R_l}^{'} {e_m^l}^*(u, v) \wedge \mathrm{z}_p^l(\rho, \rho') \wedge h_n^l(u', v'). \tag{2.47}$$

The expression in (2.45) is also known as the Foster representation of the impedance parameter. Applying duality the mn-th element of the $N_l \times N_l$ admittance matrix can be

immediately written as:

$$Y^l_{mn}(s) = \int_{\partial R_l} \int_{\partial R_l}{}' \; {h^l_m}^*(u, v) \wedge \underline{\mathcal{Y}}(\rho, \rho', s) \wedge e^l_n(u', v'). \tag{2.48}$$

This allows the magnetic field coefficient vector $\mathbf{I}^l(s)$ to be expressed in terms of the electric field coefficient vector $\mathbf{V}^l(s)$ as

$$\mathbf{I}^l(s) = \mathbf{Y}^l(s)\mathbf{V}^l(s) \tag{2.49}$$

where the pole expansion for the admittance matrix elements are written as

$$Y^l_{mn}(s) = \frac{1}{s} y^l_{0mn} + \sum_q \frac{y^l_{q\,mn}}{s - s^l_q} \tag{2.50}$$

where y^l_{0mn} and $y^l_{q\,mn}$ are residues corresponding to the poles in zero and in s^l_q, respectively, given by

$$y^l_{0mn} = \int_{\partial R_l} \int_{\partial R_l}{}' \; {h^l_m}^*(u, v) \wedge \mathcal{Y}^l_0(\rho, \rho') \wedge e^l_n(u', v') \tag{2.51}$$

and

$$y^l_{q\,mn} = \int_{\partial R_l} \int_{\partial R_l}{}' \; {h^l_m}^*(u, v) \wedge \mathcal{Y}^l_q(\rho, \rho') \wedge e^l_n(u', v') \tag{2.52}$$

The expression in (2.50) is also known as the Foster representation of the admittance parameter. The expressions in (2.44) and (2.49) represent the generalized voltage and current amplitude relationship given by means of the discretized impedance and admittance matrices.

2.3 Equivalent lumped element description of the impedance and admittance representation.

In order to describe the considered subregion by means of a network of lumped elements with the same port description as (2.45) and (2.50) a very general approach is used to

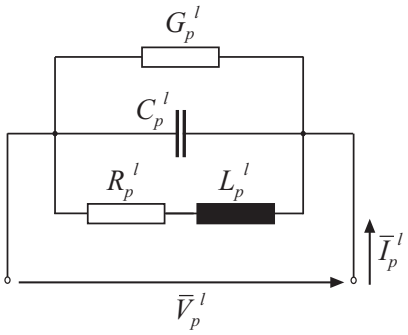


Figure 2.4: Lossy parallel resonant circuit corresponding to the pole s_p^l

exploit the Foster canonical representations [22] [73]. This method makes use of a network of ideal transformers such as the one already shown in Fig. 2.3, in combination with passive lumped elements connected so to build resonant groups which yield the natural frequencies of the considered region.
Let P_l be the number of natural frequencies of the region R_l and form a lossy parallel group of lumped elements as shown in Fig. 2.4, which can be characterized by the voltage current relationship

$$\bar{V}_p^l(s) = \bar{Z}_p^l(s)\ \bar{I}_p^l(s) \quad p = 1, 2, \ldots, P_l \tag{2.53}$$

where $\bar{Z}_p^l(s)$ is the group equivalent impedance given as

$$\bar{Z}_p^l(s) = \frac{R_p^l + sL_p^{\ l}}{s^2 L_p^{\ p} C_l^p + s(G_p^l L_p^l + R_p^l C_p^l) + 1 + R_p^l G_p^l}. \tag{2.54}$$

The resonant group formed in this way represents an electrical lumped element interpretation of an electromagnetic resonance taking place within the subregion R_l and which corresponds to the natural frequency s_p^l introduced in (2.45). In order to account for the influence of these resonances at the N_l access ports, a connection network able to transform the $P_l + 1 \times 1$ vectors of voltages and currents $\bar{\mathbf{V}}^l = [\bar{\mathbf{V}}_0^l, \ldots, \bar{\mathbf{V}}_{P_l}^l]^T$ and $\bar{\mathbf{I}}^l = [\bar{\mathbf{I}}_0^l, \ldots, \bar{\mathbf{I}}_{P_l}^l]^T$, is necessary. For this purpose a network of ideal transformers such as those of Fig. 2.3 may be considered. In this way the P_l resonate group impedances are transformed to the N_l

ports by

$$\mathbf{Z}^l(s) = \sum_{p=0}^{P_l} \bar{Z}^l_p(s)\mathbf{T}^p \tag{2.55}$$

where $\mathbf{T}^p$, $p = 0, 1, \ldots, P_l$ is a rank one matrix given by:

$$\mathbf{T}_p = \begin{bmatrix} {n_{1p}}^2 & n_{1p}n_{2p} & \cdots & n_{1p}n_{N_l p} \\ n_{2p}n_{1p} & {n_{2p}}^2 & \cdots & n_{2p}n_{N_l p} \\ \vdots & \vdots & \vdots & \vdots \\ n_{N_l p}n_{1p} & n_{N_l p}n_{2p} & \cdots & {n_{N_l p}}^2 \end{bmatrix}. \tag{2.56}$$

The matrix of (2.55) establishes the relationship between the voltage and current coefficients of the expansion of (2.38) as

$$\mathbf{V}^l(s) = \mathbf{Z}^l(s)\mathbf{I}^l(s) \tag{2.57}$$

Noting that the expression of (2.57) is the same as (2.44), (2.45) can be written in terms of the corresponding element of (2.55) yielding the following

$$z^l_{p_{mn}} = n_{mp}\bar{Z}^l_p(s)n_{np} \quad \text{with} \quad p = 0, 1, \ldots, P_l, \tag{2.58}$$

also so the $z^l_p(s)$ can be rewritten as

$$z^l_{mn}(s) = \sum_{p=1}^{P_l}\left[\frac{z^l_{p_{mn}}}{s - s^l_p} + \frac{{z^l_{p_{mn}}}^*}{s - {s^l_p}^*}\right] = \sum_{p=1}^{P_l}\left[\frac{K^l_p}{s - s^l_p} + \frac{{K^l_p}^*}{s - {s^l_p}^*}\right] k^l_{p_{mn}} \tag{2.59}$$

where:

$$k^l_{p_{mn}} = |z^l_{p_{mn}}|, \quad \text{and} \quad K^l_p = e^{j\measuredangle z^l_{p_{mn}}} \quad \forall\, m, n = 1, 2, \ldots, N_l \tag{2.60}$$

Now (2.58) together with (2.54) and (2.60) yields the following relationships between the lumped element values of the equivalent circuit and the parameters (poles and residues) of

the discrete impedance representation for the dyadic Green's function of the region R_l:

$$C_0^l = \frac{1}{K_0^l} \tag{2.61}$$

$$R_p^l = \frac{-2\Re\{s_p^l K_p^l\}\Re\{s_p^l\}^2}{\Delta_z} \tag{2.62}$$

$$L_p^l = \frac{2\Re\{K_p^l\}^3}{\Delta_z} \tag{2.63}$$

$$G_p^l = \frac{\Re\{s_p^l K_p^l\}}{\Re\{K_p^l\}^2} - \frac{\Re\{s_p^l\}}{\Re\{K_p^l\}^2} \tag{2.64}$$

$$C_p^l = \frac{1}{2\Re\{K_p^l\}} \tag{2.65}$$

where

$$\Delta_z = |s_p^l|^2\Re\{K_p^l\}^2 + \Re\{s_p^l K_p^l\}^2 + 2\Re\{s_p^l\}\Re\{K_p^l\}\Re\{s_p^l K_p^l\} \tag{2.66}$$

and the turn ratio n_{pn}^l is given by:

$$n_{pn}^l = \sqrt{k_{p_{nn}}^{\ l}}, \tag{2.67}$$

with subscripts $n = 1, 2, \ldots, N_l$ and $p = 0, 1, \ldots, P_l$. An example of this modelling is shown in Fig. 2.5. For the dual case, the $Q_l + 1 \times 1$ current and voltage vectors $\bar{\mathbf{I}}^l = [\bar{\mathbf{I}}_0^l, \ldots, \bar{\mathbf{I}}_{Q_l}^l]^T$ and $\bar{\mathbf{V}}^l = [\bar{\mathbf{V}}_0^l, \ldots, \bar{\mathbf{V}}_{Q_l}^l]^T$ are related to each other as

$$\bar{I}_q^l(s) = \bar{Y}_q^l(s)\ \bar{V}_q^l(s), \quad q = 0, 1, \ldots, Q_l \tag{2.68}$$

where Q_l is the number of natural frequencies for the admittance representation of the dyadic Green's function and $\bar{Y}_q^l(s)$ represents the impedance of a lossy lumped elements series resonant circuit, as shown in Fig. 2.6, which is given by

$$\bar{Y}_q^l(s) = \frac{G_q^l + sC_q^l}{s^2 C_q^l L_q^l + s(R_q^l C_q^l + G_q^l L_q^l) + 1 + G_q^l R_q^l}. \tag{2.69}$$

The Q_l admittances of (2.69) are transformed to the N_l ports by means of a network of

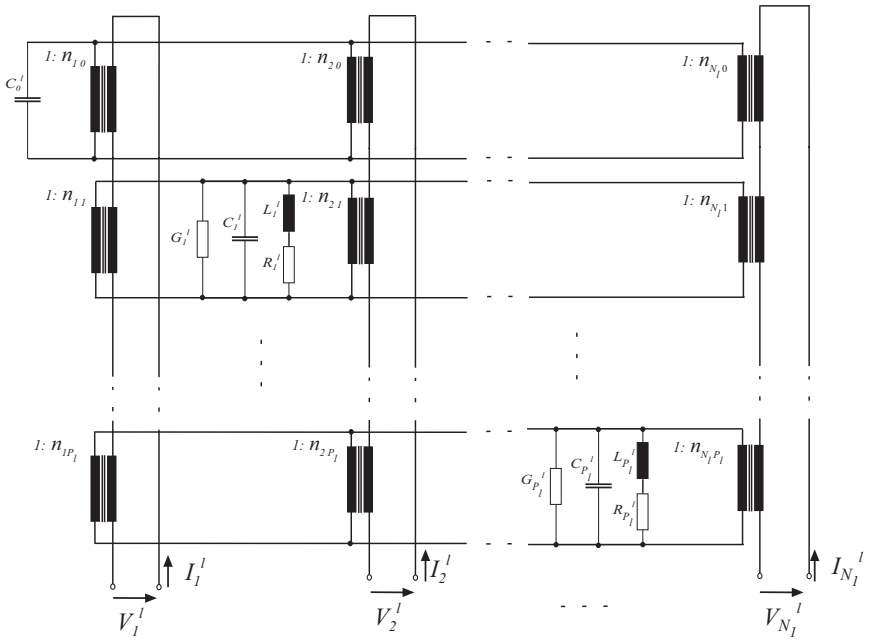


Figure 2.5: Foster realization of the $\mathbf{Z}^l$-matrix by lumped elements.

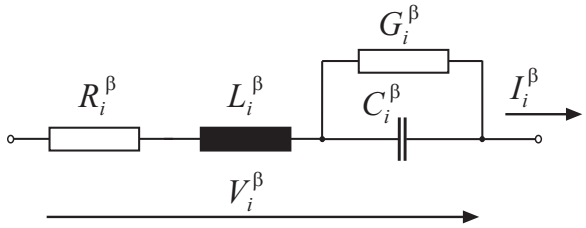


Figure 2.6: Lossy series resonant circuit corresponding to the pole s_q^l

ideal transformers forming the following matrix equation

$$\mathbf{Y}^l(s) = \sum_{q=1}^{Q_l} \bar{Y}_q^l(s)\mathbf{T}^q, \tag{2.70}$$

with $\mathbf{T}^q$, $q = 0, 1, \ldots, Q_l$, a rank one matrix given by:

$$\mathbf{T}_q = \begin{bmatrix} {n_{q1}}^2 & n_{q1}n_{q2} & \cdots & n_{q1}n_{qN_l} \\ n_{q2}n_{q1} & {n_{q2}}^2 & \cdots & n_{q2}n_{qN_l} \\ \vdots & \vdots & \vdots & \vdots \\ n_{qN_l}n_{q1} & n_{qN_l}n_{q2} & \cdots & {n_{qN_l}}^2 \end{bmatrix}. \tag{2.71}$$

Equation (2.70) establishes the relationship between the current and voltage coefficients of the expansion of (2.38) as

$$\mathbf{I}^l(s) = \mathbf{Y}^l(s)\mathbf{V}^l(s) \tag{2.72}$$

Again noting that the expression in (2.72) is equal to that of (2.49), (2.50) can be written in terms of (2.70), yielding the following equation

$$y^l_{q_{mn}} = n_{qm}\bar{Y}^l_q(s)n_{qn} \quad \text{with} \quad q = 0, 1, \ldots, Q_l \tag{2.73}$$

where in a similar fashion to (2.59) y^l_{mn} can be written as

$$y^l_{mn}(s) = \sum_{q=1}^{Q_l}\left[\frac{y^l_{q_{mn}}}{s - s^l_q} + \frac{{y^l_{q_{mn}}}^*}{s - {s^l_q}^*}\right] = \sum_{q=1}^{Q_l}\left[\frac{K^l_q}{s - s^l_q} + \frac{{K^l_q}^*}{s - {s^l_q}^*}\right]k^l_{q_{mn}} \tag{2.74}$$

where:

$$k^l_{q_{mn}} = |y^l_{q_{mn}}|, \quad \text{and} \quad K^l_q = e^{j\angle(y^l_{q_{mn}})} \quad \forall\, m, n = 1, 2, \ldots, N_l, \tag{2.75}$$

applying (2.73) together with (2.69) and (2.75) yields the following relationships between the lumped element values of the equivalent circuit and the parameters (poles and residues)

of the discrete admittance representation for the dyadic Green's function of the region R_l:

$$L_0^l = \frac{1}{K_0^l} \tag{2.76}$$

$$R_q^l = \frac{\Re\{s_q^l K_q^l\}}{\Re\{K_q^l\}^2} - \frac{\Re\{s_q^l\}}{\Re\{K_q^l\}^2} \tag{2.77}$$

$$L_q^l = \frac{1}{2\Re\{K_q^l\}} \tag{2.78}$$

$$G_q^l = \frac{-2\Re\{s_q^l K_q^l\}\Re\{s_q^l\}^2}{\Delta_y} \tag{2.79}$$

$$C_q^l = \frac{2\Re\{K_q^l\}^3}{\Delta_y} \tag{2.80}$$

where

$$\Delta_y = |s_q^l|^2\Re\{K_q^l\}^2 + \Re\{s_q^l K_q^l\}^2 + 2\Re\{s_q^l\}\Re\{K_q^l\}\Re\{s_q^l K_q^l\} \tag{2.81}$$

and for the turn ratio n_{qn}^l:

$$n_{qn}^l = \sqrt{k_q nn} \tag{2.82}$$

with $n = 1, 2, \ldots, N_l$ and $q = 0, 1, \ldots, Q_l$. The result of this modelling is shown in Fig. 2.7. The realizability of an equivalent circuit using lumped elements as parallel resonant circuits (2.54) and series resonant circuits (2.69) must be verified [24]. In order to do so the partial fraction decompositions of (2.59) and (2.74) must be real and positive (yielding positive values for the resistances (2.62) (2.77) and the conductances (2.64) (2.79)) which holds only if the following conditions are true [24]:

$$\frac{\Re\{z_{p_{mn}}^l\}}{|\Im\{z_{p_{mn}}^l\}|} \geq \frac{\Im\{s_p^l\}}{\Re\{s_p^l\}} \tag{2.83}$$

and

$$\frac{\Re\{y_{p_{mn}}^l\}}{|\Im\{y_{p_{mn}}^l\}|} \geq \frac{\Im\{s_q^l\}}{\Re\{s_q^l\}}, \tag{2.84}$$

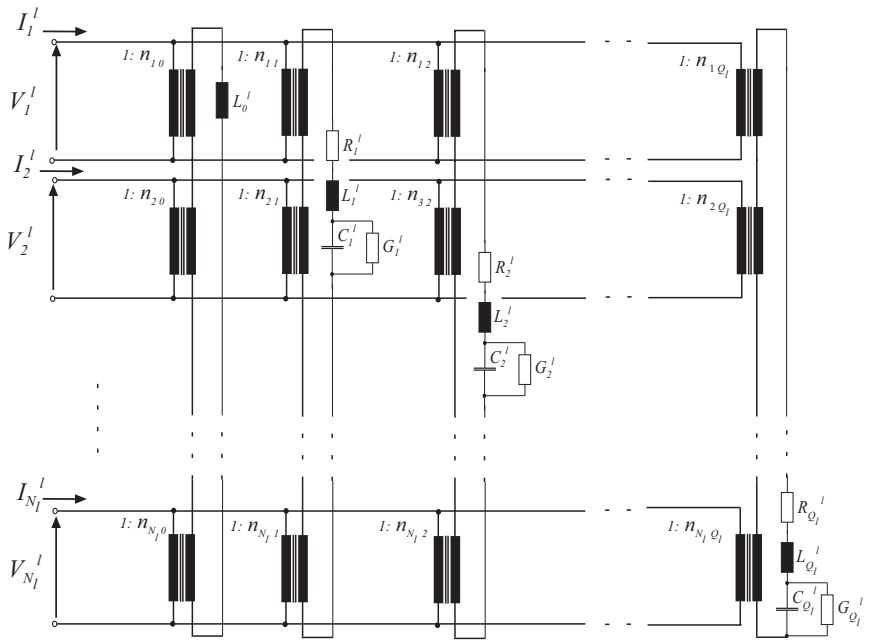


Figure 2.7: Foster realization of the $\mathbf{Y}^l$-matrix by lumped elements.

for any $p = 0, 1, \ldots, P_l$, $q = 0, 1, \ldots, Q_l$, $m, n = 1, \ldots, N_l$. The number of lumped elements required to model the $P_l + 1$ ($Q_l + 1$) natural frequencies and the connection matrix turns ratios may be minimized by applying a linear matrix transformation to reduce the number of elements different from zero in T_p and T_q [46].

2.4 Numerical Implementation of NOM

In order to accomplish a NOM of structures of arbitrary complexity and over a large frequency range, time–domain full–wave numerical analysis tools become an obvious choice. Two of the most commonly used methods are the Transmission Line Matrix (TLM) method [36, 29, 10, 71] and the Finite–Difference Time–Domain (FDTD) method [86, 79, 40]. Both of them are space and time discretizing techniques for the solution of electromagnetic field problems including lossy and non linear materials.

The TLM and FDTD methods allow to relate the wave pulses incident on the boundary of the a given spatial domain and the wave pulses scattered from the boundary providing, a very straight forward computation of the discrete Green's function [39]. Each subregion R_l of a partitioned electromagnetic structure may be characterized by the pole expansion of

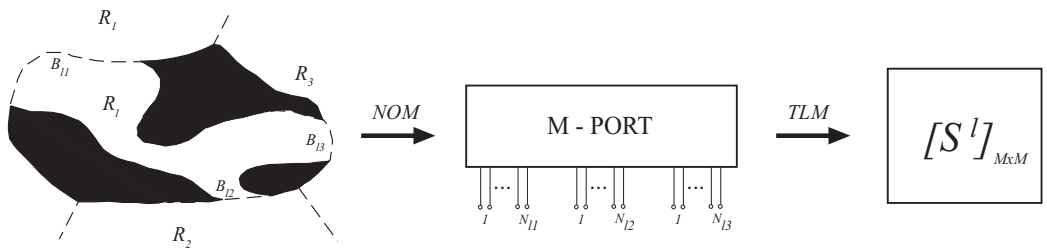


Figure 2.8: Modelling principle: from the general subregion to the M–port network by NOM, (with $M = N_{l1} + N_{l2} + N_{l3}$), and from the network to the matrix representation $\mathbf{S}^l$ (equivalently $\mathbf{Z}^l$ or $\mathbf{Y}^l$) by the TLM method.

the scattering matrix similarly to that for the impedance in (2.45) and admittance in (2.50). For the mn-th element of the $N_l \times N_l$ scattering matrix, the following expression can be written as

$$S^l_{mn}(s) = \frac{1}{s} c^l_{0mn} + \sum_r \frac{c^l_{r\,mn}}{s - s^l_r} \tag{2.85}$$

whereas c^l_{0mn} and $c^l_{r\,mn}$ are the residues corresponding to the poles at zero and at s^l_r, respectively. Hence the given electromagnetic structure can be described by a parametric model which in the complex frequency and in the time–domain is expressed by the following

$$\mathbf{S}^l(s) = \frac{1}{s}\mathbf{c}^l_0 + \sum_{r=1}^{R} \frac{\mathbf{c}^l_r}{s - s^l_r} \quad \overset{\mathscr{L}}{\longleftrightarrow} \quad \mathbf{c}^l_0 + \sum_{r=1}^{R} \mathbf{c}^l_r e^{s^l_r t} = \mathbf{S}^l(t) \tag{2.86}$$

where the $\mathbf{S}^l(s)$, $\mathbf{S}^l(t)$ are the discrete scattering matrices of size $N_l \times N_l$, in the Laplace and time–domain, respectively, and $\mathbf{c}^l_r$ and s^l_r, with $r = 1, 2, \ldots, R$ are the $R + 1$ complex residue matrices and natural frequencies respectively. Typically R depends from the number of natural frequencies within the considered frequency band. The identification of these parameters, starting from time–domain simulated data, by means of SI techniques, as described in the following chapter, enables a systematic network oriented model of the original EM structure. In Fig. 2.8 an example is shown where the considered subregion R_l is modelled as an M–port network, with M given by the sum of the number expansion functions N_{lk} used at each of the three access boundary surfaces B_{lk}, with $k = 1, 2, 3$. In general the discrete Green's function may assume the expression of an impedance, admittance or scattering representation depending an whether the boundary surfaces are replaced

by a PEC, PMC or absorbing boundary condition (ABC), respectively. The numerical implementation of magnetic and electric boundary surfaces is very simple and efficient, but becomes of poor practical application since first, for the case of low loss structures, bounding box resonances with very long transients covers up the structure's intrinsic resonances, and second, it is not possible to arbitrarily and selectively impose a PEC and/or a PMC at the same boundary where more ports may be defined. Attempts to do it have been successfully carried out only by defining a single expansion function (port) at each access boundary surface [65] [11].

The number of ports necessary to describe a structure is strictly related to the number of expansion functions necessary to describe the tangential components of the field upon the boundary surfaces. For the very general case, and by using the TLM method, with decomposition of the structure space in a rectangular grid, a possible choice of expansion functions is represented by the TLM node field components lying on the boundary surface for both orthogonal polarizations. In figure 2.9 the TLM model for the structure is depicted together with the rectangular boundary surface B_{l3}. The expansion functions introduced in (2.38), correspond here to the TLM node field components (electric and magnetic) which yield in this case $N_{l3} = 2mn$, where m and n are the number of nodes by means of which the area B_{l3} has been discretized. Similar representation can be applied for the other boundary surfaces B_{l1} and B_{l2}. Since as many time–domain simulations are required as the number of ports defined, by this choice of the expansion functions, the number of modes, e.g. of ports, can become prohibitive. The knowledge of the field distribution (carried out via a previously conducted modal analysis), may help in reducing this number by selecting suitable expansion functions. The typical case is represented from the structure geometry for which the existence of a well known fundamental mode may be considered. Examples of this are the boundary surface taken orthogonally with respect to a microstrip (MS) or a coplanar (CPW) transmission line as depicted in figure 2.10. In this case the description of the transverse field components for the fundamental mode, can be carried out by using a single ($N_{lk} = 1$) expansion function e for the electric component and a single expansion function h for the magnetic component.

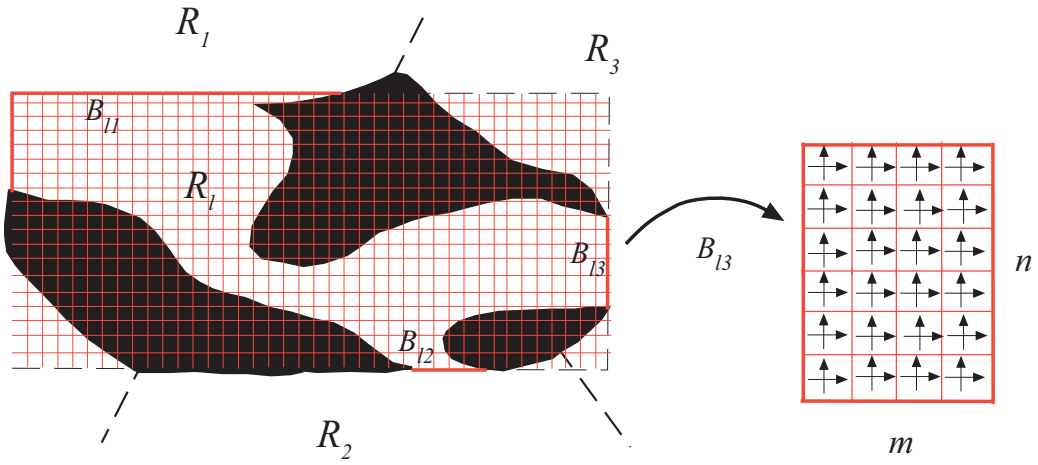


Figure 2.9: TLM modelling with an expanded view of the boundary surface B_{l3} discretized into m by n cells upon which the expansion functions are taken as the TLM fields components.

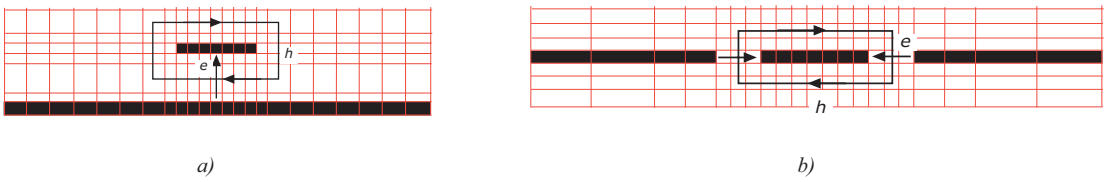


Figure 2.10: TLM modelling and port definition for the case of a boundary surface orthogonal to a MS (a) and a CPW (b) transmission line.

Chapter 3

Prony Model Based System Identification

An understanding of the physical phenomena to observe is an USEFUL insight in choosing the analytical description by means of which they may be more efficiently modelled. In the case of time domain responses of electromagnetic structures a very ubiquitous model, as shown in the first chapter (see expressions (2.45), and (2.50)), is the exponential or poles series model also known as Prony model. This model becomes therefore the candidate model upon which SI procedures will focus in the course of the present chapter. The Prony model is an analytical representation which describes sampled data as linear combination of exponential components. Gaspard Riche, Baron de Prony (1795) was led to believe that laws governing the expansion of various gases could be represented by a summation of damped exponentials [61]. Prony proposed a method for interpolating experimental data by fitting an exponential model to few equally spaced measured data points and then computing the additional values by evaluating the model at intermediate points. In his original paper Prony presents a method for exactly fitting as many purely damped exponentials as needed to fit the N available data samples. Prony's method is closely related to the least square (LS) linear prediction algorithms used for autoregressive (AR) and autoregressive moving average (ARMA) parameter estimation. However, whereas AR and ARMA methods try to fit a "random" model to the second–order data statistic (spectrum), Prony's method seeks to fit a "deterministic" exponential model to the available data. Since the data to be considered is the result of a numerical simulation, and consists of discrete time sequences, the poles (or natural frequencies) will be treated and analyzed in the Z-domain.

3.1 Prony model

Given a set of data to be modelled $x[0]$, $x[1], \ldots, x[N-1]$ the corresponding Prony' model can be expressed as follows

$$\hat{x}[n] = \sum_{k=1}^{p} A_k e^{[\alpha_k + j2\pi f_k]nT + j\theta_k} \quad 0 \leq n \leq N-1 \tag{3.1}$$

where T is the sample interval in seconds, A_k is the initial (at time zero) amplitude of the exponential and has the same physical dimension as the original data $x[n]$, α_k is the damping factor in s^{-1}, f_k is the oscillating frequency in Hz, and θ_k is the initial phase in radians. For the case of real data samples, the complex exponential in (3.1) must occur in complex conjugate pairs of equal amplitude thus reducing the exponential representation to:

$$\hat{x}[n] = \sum_{k=1}^{p} 2A_k e^{\alpha_k nT} \cos(2\pi f_k nT + \theta_k) \quad 0 \leq n \leq N-1. \tag{3.2}$$

Equation (3.1) can also be rewritten as:

$$\hat{x}[n] = \sum_{k=1}^{p} c_k z_k^n \quad 0 \leq n \leq N-1 \tag{3.3}$$

where the complex constants c_k and z_k $k = 1, 2, \ldots, p$ are defined as:

$$c_k = A_k e^{i\theta_k} \tag{3.4}$$

$$z_k = e^{(\alpha_k + j2\pi f_k)T}. \tag{3.5}$$

In order to obtain a good prediction the goal is to minimize the square error:

$$\rho = \sum_{n=N_1}^{N_2} |e[n]|^2 \tag{3.6}$$

where

$$e[n] = x[n] - \hat{x}[n] = x[n] - \sum_{k=1}^{p} c_k z_k^n \tag{3.7}$$

and $[N_1, N_2]$ define an interval arbitrarily chosen, and upon which operate the parameter estimation (prediction interval). Equation (3.6) should be minimized with respect to the parameter vectors $\mathbf{c} = [c_1, c_2, \ldots, c_p]$ and $\mathbf{z} = [z_1, z_2, \ldots, z_p]$ and their number p. This results in a difficult non linear problem with no analytic solutions [48] [19]. The original Prony's approach to this problem was to embed the non linearity aspects of the exponential model into a polynomial factorization for which relatively fast algorithms are available.

3.2 The Original Prony's Approach

When the number of data samples is equivalent to the number of exponential terms of the Prony's model, then an exact fitting of the data can be carried out. Consider the exponential time series:

$$x[n] = \sum_{k=1}^{p} c_k z_k^n \quad 0 \leq n \leq N-1 \tag{3.8}$$

where $x[n]$ represents the exact value of the data points to be modelled. Equation (3.8) can also be written in matrix notation as follows:

$$\mathbf{x} = \hat{\mathbf{Z}}\,\mathbf{c} \tag{3.9}$$

where $\hat{\mathbf{Z}}$ is a $p \times p$ matrix with a Vandermonde structure given as:

$$\hat{\mathbf{Z}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_p \\ \vdots & \vdots & & \vdots \\ z_1^{N-1} & z_2^{N-1} & \cdots & z_p^{N-1} \end{bmatrix} = \begin{bmatrix} \mathbf{z}^0 \\ \mathbf{z}^1 \\ \vdots \\ \mathbf{z}^{N-1} \end{bmatrix} \tag{3.10}$$

with the vectors **z**, **c** and **x** given by:

$$\mathbf{z} = \left[z_1, z_2, \ldots, z_p\right]^T \tag{3.11}$$

$$\mathbf{c} = \left[c_1, c_2, \ldots, c_p\right]^T \tag{3.12}$$

$$\mathbf{x} = [x[0], x[1], \ldots, x[N-1]]^T . \tag{3.13}$$

Since the unknown parameters in **c** and $\hat{\mathbf{Z}}$ are in total $2p$ in number (p in **z** and p in **c**), the solution by means of a system of linear equations requires the knowledge of at least $2p$ values, given by the original data samples $x[0], \ldots, x[p], x[p+1], \ldots, x[2p-1]$. If the parameters in **z** can be found a–priori, then (3.9) represents a linear system of p equations in the p unknown elements of **c**, which can be solved by means of LS techniques (see Appendix B) as follows

$$\mathbf{c} = \left(\hat{\mathbf{Z}}^H \hat{\mathbf{Z}}\right)^{-1} \hat{\mathbf{Z}}^H \mathbf{x}. \tag{3.14}$$

Prony's contribution has been to recognize that (3.8) can be seen as the solution of a homogeneous linear differential equation with constant coefficients. In order to find the form of this differential equation, first the polynomial that has z_k, with $k = 1, 2, \ldots, p$, as roots is defined as

$$\phi(z) = \prod_{k=1}^{p} (z - z_k) \tag{3.15}$$

which in polynomial form is given by

$$\phi(z) = \sum_{m=0}^{p} a[m] z^{p-m} \tag{3.16}$$

where $a[m]$ represents a set of complex coefficients with $a[0] = 1$. Shifting the index of (3.8) from n to $n - m$ and multiplying it by the parameters $a[m]$ yields

$$a[m]x[n-m] = a[m] \sum_{k=1}^{p} c_k z_k^{n-m}. \tag{3.17}$$

Now summing both side of (3.17) over $k = 0, 1, \ldots, p$, gives

$$\sum_{m=0}^{p} a[m]x[n-m] = \sum_{k=1}^{p} c_k \sum_{m=0}^{p} a[m]z_k^{n-m} \tag{3.18}$$

which is valid for $p \leq n \leq 2p-1$. Now making the substitution $z_k^{n-m} = z_k^{n-p} z_k^{p-m}$ (3.18) gives the following

$$\sum_{m=0}^{p} a[m]x[n-m] = \sum_{k=1}^{p} c_k z_k^{n-p} \sum_{m=0}^{p} a[m]z_k^{p-m} = 0 \tag{3.19}$$

where the right hand side of the summation in (3.19) goes to zero as it corresponds to the root of the polynomial (3.16) for z_k with $k = 1, 2, \ldots, p$. Equation (3.19) is the linear differential equation of degree p whose homogeneous solution is given by (3.8) and whose associated characteristic equation is given by the polynomial (3.16). Equation (3.19) can be reorganized into a system of p linear equations which in matrix form is given by

$$\mathbf{R}_x \, \mathbf{a} = -\mathbf{x}_p \tag{3.20}$$

where $\mathbf{R}_x$ is a $p \times p$ Toeplitz matrix

$$\mathbf{R}_x = \begin{bmatrix} x[p-1] & x[p-2] & \cdots & x[0] \\ x[p] & x[p-1] & \cdots & x[1 \\ \vdots & \vdots & & \vdots \\ x[2p-2] & x[2p-3] & \cdots & x[p-1] \end{bmatrix} \tag{3.21}$$

$$\mathbf{a} = [a[1], a[2], \ldots, a[p]]^T \tag{3.22}$$

and

$$\mathbf{x}_p = [x[p], x[p+1] \ldots, x[2p-1]]^T. \tag{3.23}$$

Solved for **a** the parameters **z** are directly obtained as the roots of the polynomial (3.16). Expression (3.20) demonstrates that with $2p$ complex data samples it is possible to decouple the **h** and **z** parameters search in two separate problems.

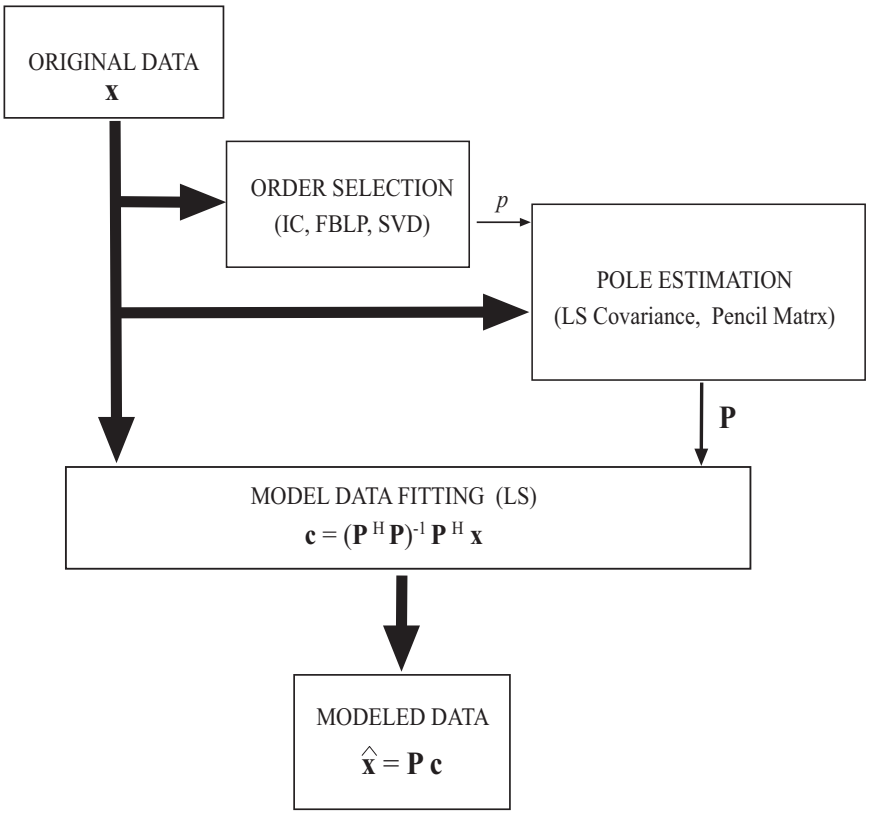


Figure 3.1: Prony modelling principle

In practice however the number of data points N usually exceeds the minimum number needed to fit the model of p exponential (or the model order), i.e. $N > 2p$. In this case the given sequence $x[n]$ can only be approximated by the Prony's exponential sequence $\hat{x}[n]$ and equation (3.9) becomes an overdeterminated system which deserve particular care when it is solved (by normal equation or SVD). The basic procedure for systematically carrying out the Prony model is shown in Fig. 3.1. The given data is used for the model order selection by means of different techniques such as those based on information criteria (IC), or on forward backward linear prediction (FBLP) or on singular value decomposition (SVD). The data is also used for pole $\mathbf{z}$ estimation, via LS linear prediction (covariance method) or the pencil matrix based approach. Eventually the model residues $\mathbf{c} = [c_1, c_2, \ldots, c_p]$ are computed by fitting, in the LS sense, the model to the original data as in (3.14).

3.3 Pole Estimation Methods

In order to extract the system poles from the time–domain transient response two main approaches will be considered, the polynomial method [47] and the pencil matrix method [30]. The first exploits the fact that (3.20) represents a typical linear prediction problem, where the coefficient vector **a** can be found via the LS linear prediction approach either solving the associated normal equation (NE) or via singular value decomposition (SVD). For linear prediction the forward and backward approaches can be used either separately or combined. The second technique, the pencil matrix method, is based on matrix shift techniques and on the properties of the underlying signals.

3.3.1 Linear Prediction Least Square Method

For practical applications, when the number N of available data points $x[0], x[1], \ldots, x[N-1]$ to fit a model of order p, exceeds the minimum number of required points i.e. $N > 2p$, the system is said to be overdetermined and the data can only be interpolated by applying an exponential sequences such as:

$$\hat{x}[n] = \sum_{k=1}^{p} c_k z_k^n \quad 0 \leq n \leq N-1 \tag{3.24}$$

where the approximation error is given by:

$$\tilde{e}[n] = \hat{x}[n] - x[n] \quad 0 \leq n \leq N-1 \tag{3.25}$$

and the parameters z_k and c_k, $k = 1, 2, \ldots, p$ must be found to minimize the total square error ρ given as:

$$\rho = \sum_{n=0}^{N-1} |\tilde{e}[n]|^2. \tag{3.26}$$

In this case the linear difference equation of(3.19) must be modified to

$$\sum_{m=0}^{p} a[m]x[n-m] = \sum_{i=1}^{p} c_i z_i^{n-p} \sum_{m=0}^{p} a[m] z_i^{p-m} = e[n] \quad 0 \leq n \leq N-1 \tag{3.27}$$

and since $a[0] = 1$ this could also be written as

$$\hat{x}[n] = -\sum_{m=1}^{p} a[m]x[n-m] + e[n] \quad 0 \leq n \leq N-1. \tag{3.28}$$

Forward LS Method

Equation (3.28) has a polynomial form and more exactly represents a forward linear prediction equation. This consists in a linear predictor of the forward sample $\hat{x}_p^f[n]$ obtained as the weighted sum of p prior samples by means of the coefficients $a_p^f[m]$, $m = 1, 2, \ldots, p$. The forward linear prediction error $e_p^p[n]$ given from:

$$e_p^f[n] = x[n] - \hat{x}_p^f[n] = x[n] + \sum_{m=1}^{p} a_p^f[m]x[n-m] \quad 0 \leq n \leq N-1 \tag{3.29}$$

and in matrix notation

$$\mathbf{e}_p^f = \mathbf{x}_p^f + \mathbf{X}_p^f \mathbf{a}_p^f. \tag{3.30}$$

Now since in general the available data can be the entire transient signal or simply a part of it, no a-priori knowledge can be assumed before the beginning of the data set at $n = 0$ and after $n = N-1$, and an $(N-p) \times (p+1)$ data matrix $\mathbf{T}_p$ can be introduced as follows

$$\mathbf{T}_p = \begin{bmatrix} x[p] & \cdots & x[0] \\ \vdots & \ddots & \vdots \\ x[N-p-1] & \cdots & x[p] \\ \vdots & \ddots & \vdots \\ x[N-1] & \cdots & x[N-p-1] \end{bmatrix}, \tag{3.31}$$

$$\mathbf{T}_p = \left[\mathbf{x}_p^f \ \mathbf{X}_p^f\right]. \tag{3.32}$$

Equation (3.30) can therefore be rewritten for forward linear prediction as

$$\mathbf{e}_p^f = \left[\mathbf{x}_p^f \ \mathbf{X}_p^f\right] \begin{bmatrix} 1 \\ \mathbf{a}_p^f \end{bmatrix} \tag{3.33}$$

where

$$\mathbf{a}_p^f = \left[a_p^f[1], \ldots, a_p^f[p]\right]^T,$$

and

$$\mathbf{e}_p^f = \left[e_p^f[p], \ldots, e_p^f[N-1]\right]^T.$$

Minimizing the square error in the LS sense means to minimize:

$$\rho_p^f = \sum_{n=p+1}^{N} |e_p^f[n]|^2 = (\mathbf{e}_p^f)^H \mathbf{e}_p^f \tag{3.34}$$

and the normal equation approach can be adopted (see Appendix B.1) thus obtaining:

$$\begin{bmatrix} {\mathbf{x}_p^f}^H \mathbf{x}_p^f & {\mathbf{x}_p^f}^H \mathbf{X}_p^f \\ {\mathbf{X}_p^f}^H \mathbf{x}_p^f & {\mathbf{X}_p^f}^H \mathbf{X}_p^f \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{a}_p^f \end{bmatrix} = \begin{bmatrix} \rho_p^f \\ \mathbf{0}_p \end{bmatrix} \tag{3.35}$$

where $\mathbf{0}_p$ is a $p \times 1$ all–zero vector. The matrix at first member of (3.35) is a $(p+1)\times(p+1)$ Hermitian matrix where: ${\mathbf{x}_p^f}^H \mathbf{x}_p^f$ is a scalar; ${\mathbf{X}_p^f}^H \mathbf{x}_p^f$ is a $p \times 1$ vector; and ${\mathbf{X}_p^f}^H \mathbf{X}_p^f$ is a $p \times p$ matrix. This matrix corresponds to the correlation matrix for the given sequence. The coefficients $\mathbf{a}_p^f$ are directly derived from (3.35) and are given by:

$$\mathbf{a}_p^f = \left(-{\mathbf{X}_p^f}^H \mathbf{X}_p^f\right)^{-1} {\mathbf{X}_p^f}^H \mathbf{x}_p^f. \tag{3.36}$$

The LS problem of (3.33) can be alternatively solved by a SVD approach which however yields the same results when the matrix $\mathbf{T}_p$ has full rank (see Appendix B.1). The desired poles are eventually obtained as the roots z_k of the associated polynomial (3.16) as:

$$\mathbf{z}_{LS}^f = \left\{ z_k \quad s.t. \quad 1 + \sum_{m=1}^{p} \mathbf{a}_p^f[m] z_k^{(p-m)=0}, \quad k = 1, 2, \ldots, p \right\}. \tag{3.37}$$

Once the coefficients $\mathbf{a}_p^f$ are determined from (3.36), the associated LS error can be found from (3.35):

$$\rho_p^f = {\mathbf{x}_p^f}^H \mathbf{x}_p^f + {\mathbf{x}_p^f}^H \mathbf{X}_p^f \mathbf{a}_p^f. \tag{3.38}$$

Backward LS Method

The same approach in (3.16) and (3.33) can also be applied for the backward linear prediction method. In this case the current data sample is given from a weighted sum of p future samples as follows:

$$\hat{x}_p^b[n] = -\sum_{m=1}^{p} \mathbf{a}_p^b[m] z[n+m] \tag{3.39}$$

where the linear prediction error is:

$$e_p^b[n] = x[n-p] + \sum_{m=1}^{p} \mathbf{a}_p^b[m] x[n-p+m] \quad p \le n \le N-1 \tag{3.40}$$

which can be rewritten in vector notation as:

$$\mathbf{e}_p^b = \left[\mathbf{X}_p^b \; \mathbf{x}_p^b\right] \begin{bmatrix} \mathbf{a}_p^b \\ 1 \end{bmatrix}, \tag{3.41}$$

and

$$\mathbf{T}_p = \left[\mathbf{X}_p^b \; \mathbf{x}_p^b\right], \tag{3.42}$$

with the backward linear prediction coefficients, and error vector determined from:

$$\mathbf{a}_p^b = \left[a_p^b[p], \ldots, a_p^b[1]\right]^T \tag{3.43}$$

$$\mathbf{e}_p^b = \left[e_p^b[0], \ldots, e_p^b[N-p-1]\right]^T \tag{3.44}$$

where the $(N-p) \times 1$ data vector $\mathbf{x}_p^b$ and the $(N-p) \times p$ data matrix $\mathbf{X}_p^b$ determine such that $\mathbf{T}_p = \left[\mathbf{X}_p^b \; \mathbf{x}_p^b\right]$. Applying the same considerations used for the forward linear

prediction method, the coefficient vector $\mathbf{a}_p^b$ can be found from the normal equation

$$\begin{bmatrix} \mathbf{X}_p^{b\,H}\mathbf{X}_p^b & \mathbf{X}_p^{b\,H}\mathbf{x}_p^b \\ \mathbf{x}_p^{b\,H}\mathbf{X}_p^f & \mathbf{x}_p^{f\,H}\mathbf{x}_p^f \end{bmatrix} \begin{bmatrix} \mathbf{a}_p^b \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_p \\ \rho_p^b \end{bmatrix} \tag{3.45}$$

where $\mathbf{0}_p$ is a $p \times 1$ all–zero vector. The matrix at first member of (3.45) is exactly the same as in (3.35) but differently decomposed. In particular: $\mathbf{x}_p^{b\,H}\mathbf{x}_p^b$ defines a scalar; $\mathbf{X}_p^{f\,H}\mathbf{x}_p^f$ is a $p \times 1$ vector; and $\mathbf{X}_p^{f\,H}\mathbf{X}_p^f$ is a $p \times p$ matrix. From (3.45),

$$\mathbf{a}_p^b = -\left(\mathbf{X}_p^{b\,H}\mathbf{X}_p^b\right)^{-1}\mathbf{X}_p^{b\,H}\mathbf{x}_p^b. \tag{3.46}$$

In the backward prediction scheme the desired poles are obtained as the inverse roots z_k of the associated polynomial (3.16):

$$\mathbf{z}_{LS}^b = \left\{ \tilde{z}_k^{-1} \quad s.t. \quad 1 + \sum_{m=1}^{p} \mathbf{a}_p^b[m]\tilde{z}_k^m = 0, \quad k = 1, 2, \ldots, p \right\} \tag{3.47}$$

and the associated LS error can be determined from (3.45) as:

$$\rho_p^b = \mathbf{X}_p^{b\,H}\mathbf{x}_p^b\mathbf{a}_p^b + \mathbf{x}_p^{b\,H}\mathbf{x}_p^b. \tag{3.48}$$

A necessary but not sufficient condition for the non singularity of the data matrix in (3.35) and (3.45) is that $p \leq \frac{N}{2}$ This represents an upper limit for the model order selection. The solution of the normal equation in both cases (forward) (3.35) and (backward) (3.45) requires general matrix solutions with a computational complexity proportional to p^3 and memory requirements proportional to p^2. In the literature, (3.33) is also called the LS linear prediction method, from its historical use in speech processing [45].

Combined Forward and Backward LS Method

For linear prediction methods, be it the forward or the backward method, the statistical information contained in the observed data is the same in both directions. It is therefore reasonable to combine the two techniques in order to generate more errors points, and as a consequence a better estimate. The $N - p$ forward and $N - p$ backward linear prediction errors given from (3.33) and (3.41) may be concisely written by using the matrix vector

Table 3.1: LS linear prediction pole estimation algorithm forward (f) and backward (b)

	Given a response $x[n]$ with $n = 0, 1, \ldots, N-1$ and a model order p:
1.	The data matrix $\mathbf{T}_p$ is formed as in (3.31) and decomposed as in (3.32) (f) or as in (3.42) (b)
2.	The polynomial coefficients are computed as in (3.36) (f) or as (3.46) (b)
3.	The poles are obtained as the roots of the polynomial (3.37) (f), or (3.47) (b)

product:

$$\begin{bmatrix} \mathbf{T}_p \\ \mathbf{T}_p^* \mathbf{J} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{a}_p^{fb} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_p^f \\ \mathbf{e}_p^{b*} \end{bmatrix} = \mathbf{e}_p. \tag{3.49}$$

The error $\mathbf{e}_p$ is a $2(N-p) \times 1$ vector made of the following forward and backward linear prediction errors

$$\mathbf{e}_p^f = \left[e_p^f[p], \ldots, e_p^f[N-1] \right]^T$$

$$\mathbf{e}_p^b = \left[e_p^b[p], \ldots, e_p^b[N-1] \right]^T$$

and the linear prediction coefficient vector $\mathbf{a}_p^{fb}$ is given by

$$\mathbf{a}_p^{fb} = \left[a_p[1], \ldots, a_p[p] \right]^T.$$

$\mathbf{T}_p$ is the $(N-p) \times (p+1)$ matrix already introduced in (3.31) and $\mathbf{J}$ is the $(p+1) \times (p+1)$ reflection matrix with 1s only on the cross diagonal, so that:

$$\mathbf{T_p}^* \mathbf{J} = \begin{bmatrix} x^*[0] & \cdots & x^*[p] \\ \vdots & \ddots & \vdots \\ x^*[p] & \cdots & x^*[N-p-1] \\ \vdots & \ddots & \vdots \\ x^*[N-p-1] & \cdots & x^*[N-1] \end{bmatrix}. \tag{3.50}$$

This latter is an $(N-p) \times (p+1)$ Hankel matrix (elements on the cross diagonals all identical). Minimizing the average of the forward and backward square error over the

available data yields:

$$\rho_p^{fb} = \frac{1}{2}\left[\sum_{n=p}^{N-1} |e_p^f[n]|^2 + \sum_{n=p}^{N-1} |e_p^b[n]|^2\right] = \frac{1}{2}\left[\mathbf{e}_p^{f\,H}\mathbf{e}_p^f + \mathbf{e}_p^{b\,H}\mathbf{e}_p^b\right] = \frac{1}{2}\mathbf{e}_p{}^h\mathbf{e}_P, \qquad (3.51)$$

which leads to the following normal equation

$$\mathbf{R}_p \begin{bmatrix} 1 \\ \mathbf{a}_p^{fb} \end{bmatrix} = \begin{bmatrix} 2\rho_p^{fb} \\ \mathbf{0}_p \end{bmatrix}, \qquad (3.52)$$

where $\mathbf{0}_p$ is a $p \times 1$ all–zeros vector and

$$\mathbf{R}_p = \begin{bmatrix} r & \mathbf{r}^H \\ \mathbf{r} & \mathbf{R} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_p \\ \mathbf{T}_p{}^*\mathbf{J} \end{bmatrix}^H \begin{bmatrix} \mathbf{T}_p \\ \mathbf{T}_p{}^*\mathbf{J} \end{bmatrix} = \mathbf{T}_p{}^*\mathbf{T}_p + \mathbf{J}\mathbf{T}_p{}^T\mathbf{T}_p{}^*\mathbf{J}, \qquad (3.53)$$

where r is a scalar, $\mathbf{r}$ is a $p \times 1$ data vector and $\mathbf{R}$ is a $p \times p$ data matrix. This method is known in the literature as the modified covariance method [7] [83] [53]. From (3.52) and (3.53), the combined forward and backward linear prediction coefficient vector $\mathbf{a}_p^{fb}$ is given from:

$$\mathbf{a}_p^{fb} = -\mathbf{R}^{-1}\mathbf{r} \qquad (3.54)$$

and the corresponding minimum square error

$$\rho_p^{fb} = r + \mathbf{r}^H\mathbf{a}_p^{fb}. \qquad (3.55)$$

Finally the poles are obtained in a similar fashion to the forward case as:

$$\mathbf{z}_{LS}^{fb} = \left\{ z_k \quad s.t. \quad 1 + \sum_{m=1}^{p} \mathbf{a}_p^{fb}[m] z_k^{p-m} = 0, \quad k = 1, 2, \ldots, p \right\}. \qquad (3.56)$$

A necessary but not sufficient condition for $\mathbf{R}_p$ to be non singular is that $p \leq \frac{2}{3}N$. This provides an upper limit for the combined forward and backward linear prediction technique. The LS method in all cases so far observed (forward backward and the combination of the two) does not guarantees stability, i.e. the poles $\mathbf{z}_{LS}^f$, $\mathbf{z}_{LS}^b$ and $\mathbf{z}_{LS}^{fb}$ could lie also outside the unitary circle, giving rise to an unstable Prony model. The LS linear prediction pole

extrapolation algorithm is oulined in Tab. 3.1.

3.3.2 Pencil Matrix Method

An alternative approach to the described polynomial or LS linear prediction method is the pencil matrix (PM) method [30] [56]. The PM method is representative of a technique which exploits the structure of a matrix of pencil of the (noiseless) underlying sought signal instead of solving a prediction equation as in the LS linear prediction approach (3.3.1). The method presented here was originally introduced in the late 80s [31] and is based on the Pencil of Function (PoF) approach [34] [35] [84], and previous work on estimation of signal parameters via a rotational invariant technique (ESPRIT) [68] [64] and the SVD Prony's method [82] [41]. In contrast with these the PM exploits in greater extend the free–moving window length, referred to the pencil parameter.
Given a sequence of time discrete signal samples $x[n]$ corrupted by additive noise $e[n]$ and expressed as:

$$\hat{x}[n] = x[n] + e[n] \quad 0 \leq n \leq N-1 \tag{3.57}$$

the resulting approximated signal can be modelled by the exponential sequence of (3.24). In order to find the damping coefficients and the frequency of the exponential components which characterize it (Prony'model poles) the following properties of the exponential signal are recalled.
Define an $(N-L) \times 1$ data vector:

$$\mathbf{x}_t = [x[t], x[t+1], \ldots, x[N-L+t-1]]^T \quad 1 \leq t \leq L \tag{3.58}$$

and two $(N-L) \times L$ data matrices

$$\mathbf{X}_0 = \left[\mathbf{x}_{L-1}, \mathbf{x}_{L-2}, \ldots, \mathbf{x}_0\right], \tag{3.59}$$

$$\mathbf{X}_1 = \left[\mathbf{x}_L, \mathbf{x}_{L-1}, \ldots, \mathbf{x}_1\right], \tag{3.60}$$

where $L \in I\!N$ is the pencil parameter. The quantities belonging to the set $\mathbf{z}_{PM} = [z_1, z_2, \ldots, z_p]$ are the sought poles and are found as the rank reducing values of the so called pencil matrix defined as

$$\mathbf{X}_1 - z\,\mathbf{X}_0 \quad \text{with} \quad z \in \mathbb{C} \tag{3.61}$$

where $p \leq L \leq N - p$. None of the values of $\mathbf{z}_{PM}$ are rank reducing number of (3.61) if $L \leq p$ or $L \geq N - p$. The very interesting physical meaning of this is the following. If the rank of matrix (3.61) reduces this means that at least one of its columns (portion of the signal) can be expressed by means of another column (another different portion of the signal) within a difference of magnitude and phase given by the rank reducing complex value. The proof of this is given by using the following decompositions for $\mathbf{X}_0$ and $\mathbf{X}_1$

$$\mathbf{X}_0 = \mathbf{Z}_L \, \mathbf{H} \, \mathbf{Z}_R \tag{3.62}$$

$$\mathbf{X}_1 = \mathbf{Z}_L \, \mathbf{H} \, \mathbf{Z} \, \mathbf{Z}_R \tag{3.63}$$

where

$$\mathbf{Z}_L = \begin{bmatrix} 1 & \cdots & 1 \\ z_1 & \cdots & z_p \\ \vdots & & \vdots \\ z_1^{N-L-1} & \cdots & z_p^{N-L-1} \end{bmatrix} \quad \text{and} \quad \mathbf{Z}_R = \begin{bmatrix} z_1^{L-1} & z_1^{L-2} & \cdots & 1 \\ \vdots & & & \vdots \\ z_p^{L-1} & z_p^{L-2} & \cdots & 1 \end{bmatrix}, \tag{3.64}$$

are two $(N - L) \times p$ and $p \times L$ Vandermonde matrices and where

$$\mathbf{H} = diag\left\{h_1, h_2, \ldots, h_p\right\} \tag{3.65}$$

and

$$\mathbf{Z} = diag\left\{z_1, z_2, \ldots, z_p\right\} \tag{3.66}$$

are $p \times p$ full–rank parameter matrices. Inserting (3.62) and (3.63) into (3.61) yields

$$\mathbf{X}_1 - z\,\mathbf{X}_0 = \mathbf{Z}_L \, \mathbf{H} \, \left(\mathbf{Z} - z\,\mathbf{I}_p\right) \, \mathbf{Z}_R \quad \text{with} \quad z \in \mathbb{C} \tag{3.67}$$

where $\mathbf{I}_p$ is the $p \times p$ identity matrix. Since for $p \leq L \leq N - p$ the matrices $\mathbf{Z}_L$ and $\mathbf{Z}_R$ have rank p, it follows that:

$$rank\left(\mathbf{X}_1 - z\mathbf{X}_0\right) = \begin{cases} p & \text{for} \quad z \notin \mathbf{z}_{PM} \\ p - 1 & \text{for} \quad z \in \mathbf{z}_{PM} \end{cases} \tag{3.68}$$

For the case of $L < p$ then $rank\,(\mathbf{Z}_R) = L$ and from this it follows that

$$rank\left[\left(\mathbf{Z} - z\mathbf{I}_p\right)\mathbf{Z}_R\right] = L \quad \forall\ z\ \in\ \mathbb{C} \tag{3.69}$$

and therefore $z \in \mathbf{z}_{PM}$. Similarly for $L > N - p$, the $rank\,(\mathbf{Z}_L) = N - L$ and hence

$$rank\left[\mathbf{Z}_L\mathbf{H}\left(\mathbf{Z} - z\mathbf{I}_p\right)\right] = N - L \quad \forall\ z\ \in\ \mathbb{C}\ . \tag{3.70}$$

This means that the value z_k with $k\ =\ 1, 2, \ldots, p$ can be found as the rank reducing quantities of the pencil matrix (3.61) if and only if $p \leq L \leq N - p$.
If for such values of L the rank of (3.61) decreases by one unit, for $z_k\ \in\ \mathbf{z}_{PM}$, then the solution of the associated eigenvalues problem is given by:

$$\begin{array}{rl} (\mathbf{X}_1\ - z\,\mathbf{X}_0)\ \mathbf{q} & = 0 \\ \mathbf{p}^H\ (\mathbf{X}_1\ - z\,\mathbf{X}_0) & = 0 \end{array} \tag{3.71}$$

which unique together with the corresponding eigenvectors $\mathbf{q}$ and $\mathbf{p}^H$ (which belong to the column space of $\mathbf{X}_0^H$ and to the row space of $\mathbf{X}_0$ respectively), within a scalar constant. These solutions are given as:

$$\mathbf{q}_k \quad \text{s.t} \quad k^{th}\text{column of} \quad \mathbf{Z}_R{}^{\#} = \mathbf{Z}_R^H\left(\mathbf{Z}_R\mathbf{Z}_R^H\right)^{-1} \tag{3.72}$$

$$\mathbf{p}_k^H \quad \text{s.t} \quad k^{th}\text{row of} \quad \mathbf{Z}_L{}^{\#} = \left(\mathbf{Z}_L^H\mathbf{Z}_L\right)^{-1}\mathbf{Z}_L^H \tag{3.73}$$

for $k = 1, 2, \ldots, p$ and where $\mathbf{Z}_R{}^{\#}$, $\mathbf{Z}_L{}^{\#}$ and $\mathbf{Z}_R^H$, $\mathbf{Z}_L^H$ are the corresponding pseudoinverse and Hermitian matrices of $\mathbf{Z}_R$ and $\mathbf{Z}_L$ respectively. It is of note that the equations (3.71) are the generalized singular eigenvalue problems, since for $p \leq L \leq N - p$ the matrices $\mathbf{X}_0$ and $\mathbf{X}_1$ are not full–rank. This can be easily shown by considering the decomposition (3.62) and (3.63) and by noticing that matrices $\mathbf{Z}_R$ and $\mathbf{Z}_L$ are of rank p. Therefore $rank\,(\mathbf{X}_0) = rank\,(\mathbf{X}_1) = p$ while its dimension is $N - L \times L$.
Once it is shown that the poles of $\mathbf{z}_{PM}$ are unique, they are calculated from the data matrices $\mathbf{X}_0$ and $\mathbf{X}_1$. Multiplying the first of (3.71) by $\mathbf{X}_0{}^{\#}$ gives:

$$\begin{array}{rcl} \mathbf{X}_0{}^{\#}\,(\mathbf{X}_1\ - z_k\,\mathbf{X}_0)\ \mathbf{q}_k & = & 0 \\ \mathbf{X}_0{}^{\#}\mathbf{X}_1\mathbf{q}_k = z_k\mathbf{X}_0{}^{\#}\mathbf{X}_0\,\mathbf{q}_k & = & z_k\,\mathbf{q}_k \end{array} \quad \text{with} \quad z_k \in \mathbf{z}_{PM} \tag{3.74}$$

which implies that the elements of $\mathbf{z}_{PM}$ are p non zero eigenvalues of the $L \times L$ matrix $\mathbf{X}_0^{\#}\mathbf{X}_1$. Moreover since $rank\left(\mathbf{X}_0^{\#}\mathbf{X}_1\right) = p \leq L$, (3.74) also has $L - p$ zero eigenvalues.
If now a noisy data sample series $\hat{x}[n]$ is considered, in place of the noiseless sequence $x[n]$, the same procedure in building the data matrices $\hat{\mathbf{x}}_t$, $\hat{\mathbf{X}}_0$ and $\hat{\mathbf{X}}_1$ applies as in (3.58) (3.59) and (3.60), the only difference being that in place of the exact pseudoinverse ${\mathbf{X}_0}^{\#}$ and ${\mathbf{X}_1}^{\#}$, the rank p truncated pseudoinverse matrices $\hat{\mathbf{X}}_0^{\#}$ and $\hat{\mathbf{X}}_1^{\#}$, of size $N - L \times L$ must be used. For $\hat{\mathbf{X}}_0^{\#}$ this yields:

$$\hat{\mathbf{X}}_0^{\#} = \sum_{n=1}^{p} {\sigma_0}_n^{-1} \mathbf{v}_{0n} {\mathbf{u}_{0n}}^H = \mathbf{V}_0 \mathbf{A}^{-1} \mathbf{U}_0, \tag{3.75}$$

where $\mathbf{A}$ is a diagonal matrix formed from the p largest singular values σ_{0n}, $n = 1, 2, \ldots, p$ of $\hat{\mathbf{X}}_0$ and $\mathbf{V}_0$ and $\mathbf{U}_0$ are $N - L \times p$ and $N \times p$ unitary matrices of eigenvectors of $\hat{\mathbf{X}}_0^H \hat{\mathbf{X}}_0$ and $\hat{\mathbf{X}}_0 \hat{\mathbf{X}}_0^H$ respectively (see Appendix B). An equation similar to (3.75) is derived for $\hat{\mathbf{X}}_1^{\#}$. The matrix $\hat{\mathbf{X}}_0^{\#}$ represents the p rank truncated noiseless pseudoinverse of the noisy data matrix $\hat{\mathbf{X}}_0$ in the sense that it is equal to ${\mathbf{X}_0}^{\#}$ only for the case of zero noise i.e. $e[n] = 0$, for $n = 1, 2, \ldots, N - 1$. Now since the $L \times L$ matrix $\hat{\mathbf{X}}_0^{\#}\hat{\mathbf{X}}_1$ has $L - p$ null eigenvalues its size can be reduced before the p non-zero eigenvalues are extracted. By using (3.74) and inserting $\hat{\mathbf{X}}_0^{\#}$ in place of $\mathbf{X}_0^{\#}$ in (3.75), the latter becomes:

$$\mathbf{V}_0 \mathbf{A}^{-1} \mathbf{U}_0^H \hat{\mathbf{X}}_1 \mathbf{q}_k = z_k \mathbf{q}_k \quad k = 1, 2, \ldots, p \tag{3.76}$$

and given that $\mathbf{V}_0^H \mathbf{V}_0 = \mathbf{I}_p$ and $\mathbf{q}_k = \mathbf{V}_0^H \mathbf{V}_0 \mathbf{q}_k$, and left multiplying (3.76) by $\mathbf{V}_0^H$ yields:

$$\mathbf{A}^{-1} \mathbf{U}_0^H \hat{\mathbf{X}}_1 \mathbf{V}_0 \mathbf{V}_0^H \mathbf{q}_k = z_k \mathbf{V}_0^H \mathbf{q}_k \quad k = 1, 2, \ldots, p \tag{3.77}$$

and from this the poles z_k can be computed as the eigenvalues of the square $p \times p$ matrix

$$\mathbf{Z}_E = \mathbf{A}^{-1} \mathbf{U}_0^H \hat{\mathbf{X}}_1 \mathbf{V}_0 \tag{3.78}$$

obtaining:

$$\mathbf{z}_{PM} = \left\{ z_k \quad s.t. \quad \det\left(\mathbf{Z}_E - z_k \mathbf{I}_p\right) = 0, \quad k = 1, 2, \ldots, p \right\}. \tag{3.79}$$

In the literature a perturbation analysis is used to study the noise sensitivity of the pencil matrix method. Investigation results have shown that in a first approximation a value for

Table 3.2: PM pole estimation algorithm.

	Given a response $x[n]$ with $n = 0, 1, \ldots, N-1$ and a model order p:
1.	The pencil length is defined as $L = \lfloor \frac{N}{2} \rfloor$
2.	The data matrices $\hat{\mathbf{X}}_0$ and $\hat{\mathbf{X}}_1$ are built as in (3.62) and (3.63)
3.	The p –rank truncated pseudoinverse of $\hat{\mathbf{X}}_0$ is computed as in (3.75)
4.	The $p \times p$ matrix $\mathbf{Z}_E$ is computed as in (3.78)
5.	The p poles are found as eigenvalues of $\mathbf{Z}_E$

the pencil parameter between $\frac{N}{3} \leq L \leq \frac{2}{3}N$ gives the best results [30].
Compared with the linear prediction LS method of section (3.3.1), where a $2(N-p) \times p$ data matrix was decomposed via SVD, and the p roots of a p degree polynomial were found, the PM method requires the SVD of a $(N-L) \times L$ matrix and moreover to solve a $p \times p$ eigenvalue problem (3.79). For $L \leq p \leq N-L$ the linear prediction LS method, always has a complexity always higher than the PM method. The PM pole extrapolation algorithm is described in Tab. 3.2.

3.3.3 Pole Estimation Method Comparison

In physical applications the performances of pole extrapolation algorithms are significantly effected by strongly attenuated and/or closely spaced poles in noisy environments, (for example the case of selective filters built on lossy materials and operating in an environment of high thermic noise). A parametric analysis of the effects produced from signal damping, frequency separation (resolution) and noise power will be presented here in order to evaluate and compare the performances of the previously introduced pole estimation algorithms. In order to carry out the analysis a single pole signal, synthesized as in (3.1), with $p = 1$, and corrupted by additive white noise with a given SNR, will be evaluated. The signal is described by

$$s_0[n] = x[n] + e[n] = \sum_{k=1}^{p} A_k e^{[-\alpha_k + j(2\pi f_k + \theta_k)]nT} + e[n]. \tag{3.80}$$

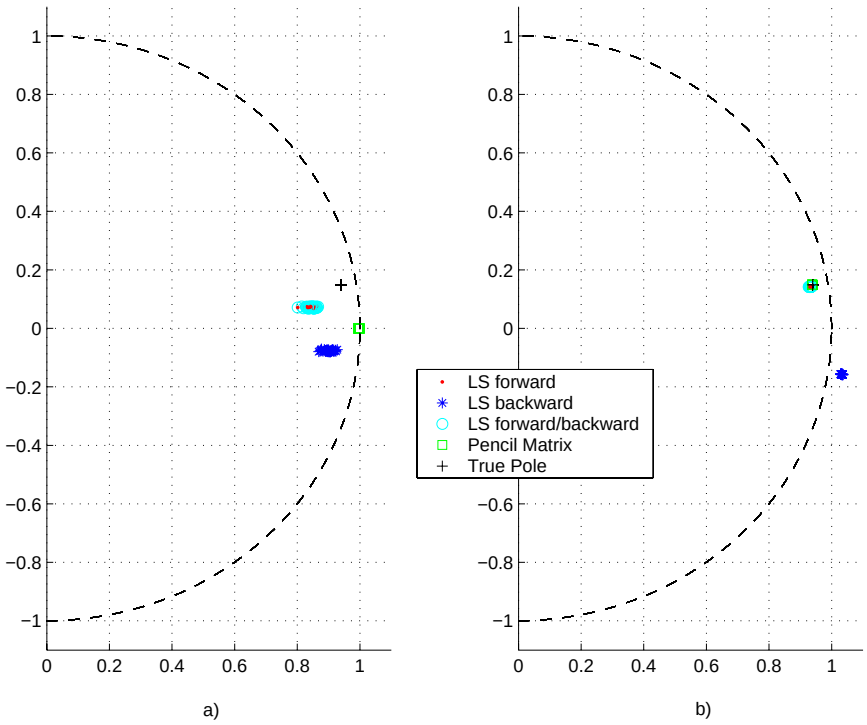


Figure 3.2: Effect of an additive white noise on a single poles signal: a) for $SNR = 0\ dB$, and b) for $SNR = 15\ dB$

The first sequence taken in consideration is a a single pole signal $s_0[n]$ with $p = 1$ and an initial amplitude of $A_1 = 1$ is considered. The initial phase is $\theta_1 = 0$, and the discretization interval is assumed to be unitary ($T = 1$). The pole parameters α_1, f_1 will be varied together with the SNR. In order to quantify the accuracy of the estimated pole a variance parameter called mean error distance (MED) is introduced as

$$MED_p = \frac{1}{R}\sum_{k=1}^{R}\left[\frac{|z_p - \hat{z}_p^{(k)}|}{|z_p|}\right]^2 \tag{3.81}$$

where z_p is the exact p^{th} pole and $\hat{z}_p^{(k)}$ the corresponding estimate at the k^{th} estimation trial from an overall number of trial R. The first comparison is given in Fig.(3.2). It shows the effects on the position of the estimated poles, for $R = 50$ of the sequence in (3.80), for two SNR levels of 0 and $15dB$, respectively. Since the signals are discrete in time, the representation of the poles is given on the z–plane. For a SNR of $0dB$ all considered methods produce similarly inaccurate results. For a SNR of $15dB$ however all methods

Table 3.3: Comparison of methods in terms of MED and pole position. True pole: $z_1 = 0.9395 + j\ 0.1488i$

Method	MED @ 0 dB	Poles @ 0 dB	MED @ 15 dB	Poles @ 15 dB
LS forw	0.1312	0.8417 + j 0.0717	0.0112	0.9319 + j 0.1421
LS back	0.1807	0.8995 - j 0.0767i	0.0093	1.0321 - j 0.1574
Pencil Matrix	0.1679	0.9968 - j 0.0003	0.0018	0.9379 + j 0.1485

except the LS backward yield good results. The numerical results are listed in Tab.(3.3), where the extracted pole accuracy is estimated by computing the MED. The noise level effects the performance of the pole identification especially in case of strong damping. An example of this is given in Fig.3.3, where two values of α, equal to $0.05\,s^{-1}$ and $0.5\,s^{-1}$ are taken in consideration. The estimation improves with the SNR as expected, and for the lower value of damping the PM method gives very good results with a $MED < 10^{-4}$ for a SNR of $14dB$, while the LS forward and backward methods obtain the same accuracy only when the SNR is increased by 10 and $15dB$ respectively. For the higher damping value, the three methods give similar accuracy, within few dBs, and get below a MED of 10^{-4} only for a $SNR > 38dB$. Fig.3.4 shows the behavior of the MED as a function of the damping factor α. Depending on the noise level the PM becomes worse than the LS based method, after α has reached a certain value (see Fig.3.4). This is due to the fact that the algorithm is not able to correctly recover the pole, and yields in this case a constant value of $z = 1$ (see Table 3.3). Another important characteristic to be considered is the ability to resolve poles which are close to each other. In order to show this a two pole signal will be considered as given in (3.80) with $p = 2$. In the first test the two poles are kept at a constant frequency spacing of 1% (i.e. $f_1 = 0.125Hz$ and $f_2 = 1.01\ f_1$), and the damping α is swept in the range $\alpha_1 = \alpha_2 \in [0\ \ 1]$. The results are depicted in Fig.3.5, where the PM method shows an improved ability to accurately resolve both poles over the LS forward and LS backward methods, for values of $\alpha < 0.35\ s^{-1}$. A further test can be carried out by keeping constant the damping factor $\alpha_1 = \alpha_2 = 0.25\ s^{-1}$, and introducing a frequency deviation Δf. The results with $f_1 = 0.125Hz$, and $f_2 = f_1(1 + \Delta f)$ where Δf is swept from 0 to 6%, are given in Fig.3.6 for an $SNR = 100dB$. Comparing the methods the PM shows a strong improvement over the other two methods above a frequency deviation of 0.5% and through the rest of the test range.

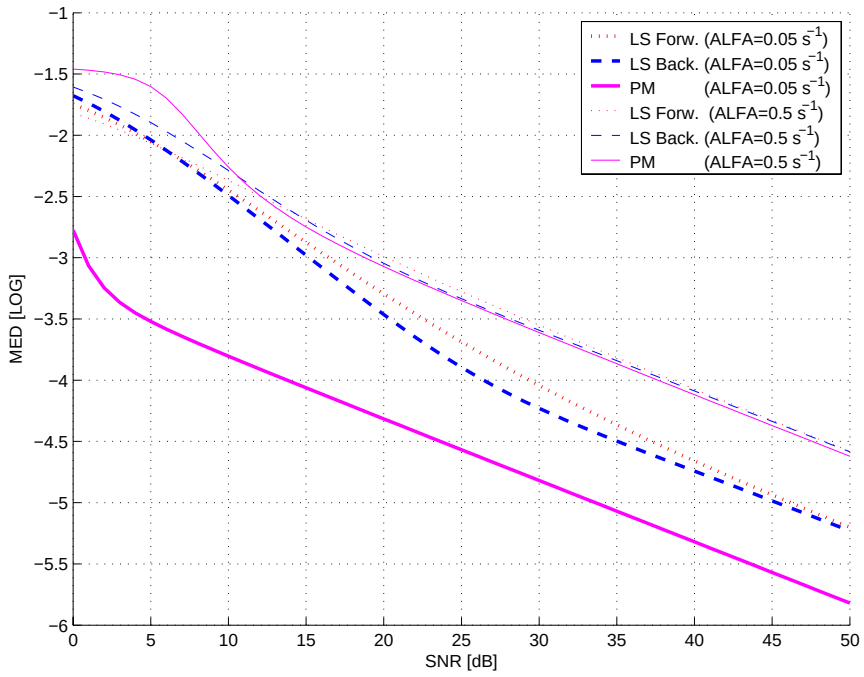


Figure 3.3: MED for the three pole estimation method (LS forward LS backward and PM) as a function of the SNR for the case of two damping constants: $\alpha = 0.05s^{-1}$ and $\alpha = 0.5s^{-1}$

Fig.3.7 shows a set of two dimensional plots of the MED, for all three methods, as function of both α and the percent frequency deviation.

3.4 Model Order Selection

The model order p is in general not known a–priori. In practice several model orders are defined and by means of some error criterion the optimum (minimum model error) is chosen. A low model order yields a simple system description, but also a smoothed spectral estimate and lack of accuracy. Too high an order increases the resolution, but may introduce spurious components (likely to be unstable) into the spectrum. Thus model order selection represents a trade off between minimum complexity and maximum accuracy. An intuitive approach would be to increase the model order progressively until a minimum threshold value for the error variance is reached. However, since most of the estimation procedures yield a prediction error variance which decreases monotonically with the order p, such an approach becomes inefficient unless a sudden reduction of in the error occurs. Normally

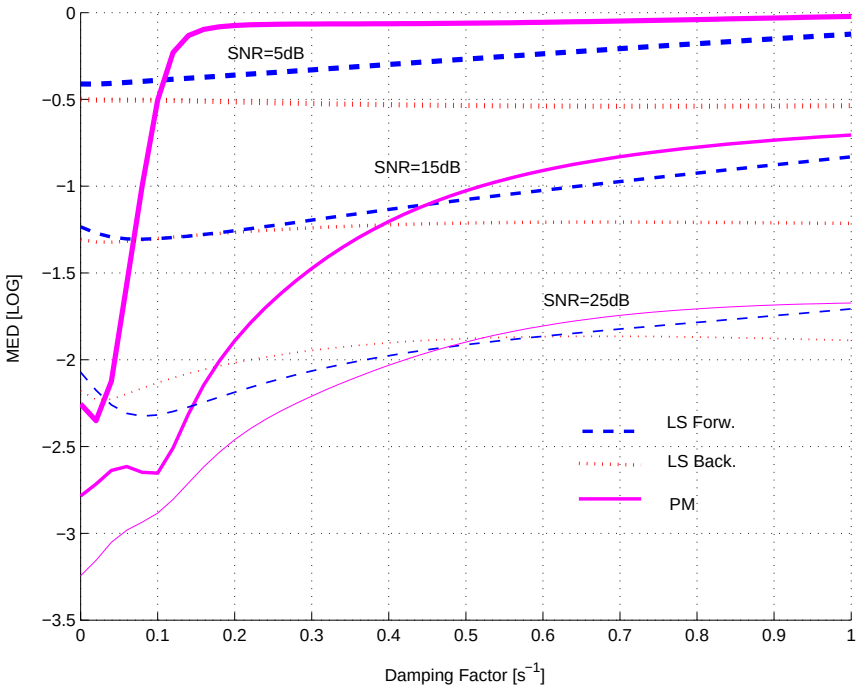


Figure 3.4: MED as a function of the damping factor α, for three different noise levels: $SNR = 5dB$, $SNR = 15dB$ and $SNR = 25dB$

the model order selection is based on the preliminary data analysis and falls into the following categories. The first approach is carried out by examining the spectral analysis estimate of the signal or transfer function and observing the resonance peaks, the high frequency roll–off, and the phase shift. A second class of techniques for model order estimation use an information criterion based on autoregressive linear prediction. Techniques based on the linear prediction of polynomial zeros are also possible. Last, a techniques based on the separation of the signal space from the noise space and the analysis of the covariance matrix rank can also be applied. Since the first mentioned technique is based on visual inspection of the spectrum it is very imprecise, especially in the presence of strongly damped resonances. Techniques based on information criterion are traditionally very common in SI but have very limited practical applications and are only briefly mentioned. Only the last two methods are of practical use and are hereby extensively discussed and applied here.

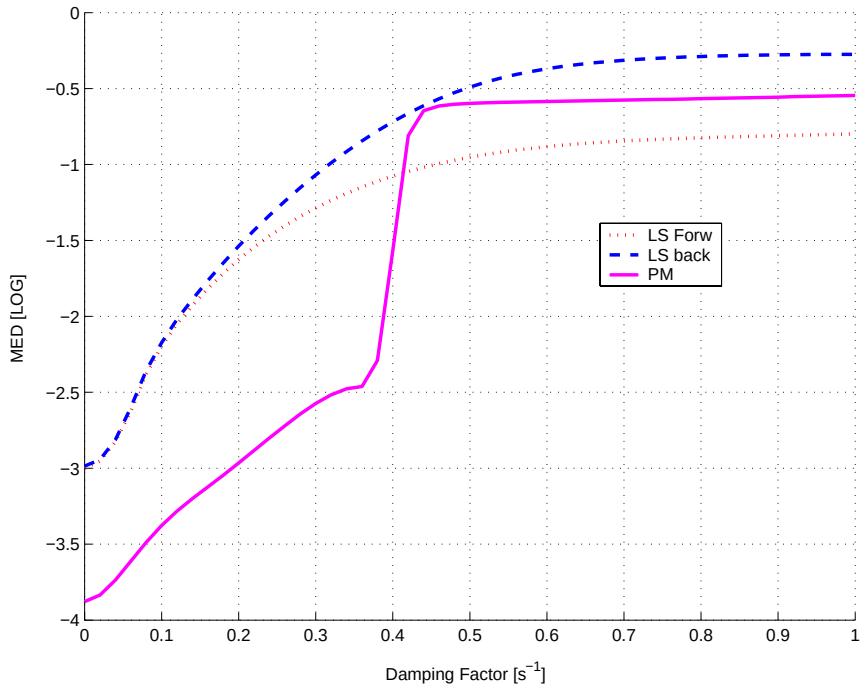


Figure 3.5: MED of a signal with two closely spaced poles ($f_2 = 1.01\ f_1$) as function of the damping ($SNR = 100dB$)

3.4.1 Method based on AR information criteria

The LS Prony's method (see Section 3.2) requires the solution of a LS linear prediction normal equation which incidentally corresponds with the same equation to be solved by autoregressive (AR) modelling ([47], Chap.8). This commonality suggests that an entire set of order selection criteria, originally developed for AR models, can also be applied to the Prony's covariance method. Akaike ([2]) has provided two criteria. The first, called the final prediction error (FPE), selects the order so that the average error variance for a one step prediction (prediction on the next single sample) is minimized. The FPE is defined as:

$$FPE(p) = \rho_p \left(\frac{N + p + 1}{N - p + 1} \right) \tag{3.82}$$

where N is the number of data samples, p is the model order, and ρ_p is the linear prediction noise variance already encountered in (3.38), (3.48) and (3.55) for the forward, backward and the combination of the forward and backward, respectively. Equation (3.82) assumes that the sample mean has been subtracted from the data. The selected order p is the value

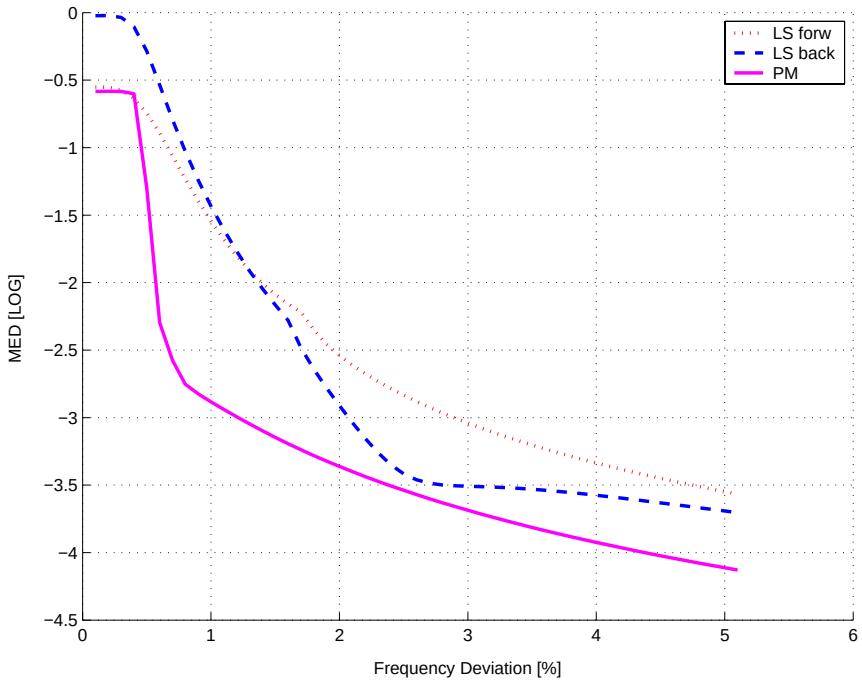


Figure 3.6: MED as function of the frequency deviation ($SNR = 100dB$)

that minimizes the $FPE(p)$. Akaike ([3]) also suggested another selection criterion known as the Akaike information criterion (AIC), based on a maximum likelihood approach. The AIC determines the model order p by minimizing an information theoretic function, which for the case of additive gaussian white noise is given as follows:

$$AIC(p) = N \ln\left(\rho_p\right) + 2p. \tag{3.83}$$

The term $2p$ represents the penalty for the use of extra parameters that do not contribute to the reduction of the prediction error variance. The selected order p is again given by the value which minimizes the $AIC(p)$. As $N \to \infty$ AIC and FPE are asymptotically equivalent. The AIC criterion however, has been proven to be statistically inconsistent since the error statistic does not go zero as N goes to infinity, yielding an overestimated order for data record increasing length [37]. In order to correct this estimate a variant of the AIC was introduced as minimum description length (MDL) [66] and expressed as follow:

$$MDL(p) = N \ln\left(\rho_p\right) + p \ln(N). \tag{3.84}$$

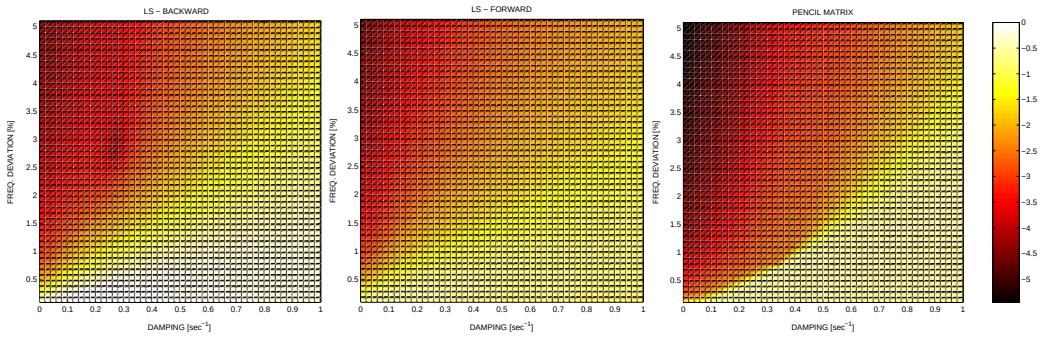


Figure 3.7: MED as function of the frequency deviation and damping for two poles signal in white noise ($SNR = 100dB$), for the LS Backward LS Forward, and Pencil Matrix method.

The model order selection criterion presented so far may be used as guidelines for initial order selection, but in general do not produces satisfactory results for experimental data that in general cannot be modelled as an AR process. One more disadvantage of these criteria is the limited application range to low noise data.

3.4.2 Forward and backward polynomial LP based method

The methods for the identification of an exponential signals from a noisy transient signal, as seen in sections (3.3.1) exploits the following principle. For a deterministic process represented as an exponentially damped signal, and mixed with stationary additive white noise, its statistical properties do not change for the case of time inversion. This principle can therefore be exploited for linear prediction once the forward and the backward LS linear prediction algorithms are applied. Recalling the LS linear prediction method, the forward and backward solutions (3.37) and (3.47) for the case of model order r can be rewritten respectively as:

$$\mathbf{z}^f = \left\{ z_k \quad s.t. \quad 1 + \sum_{m=1}^{r} \mathbf{a}_r^f[m] z_k^{(r-m)} = 0, \quad k = 1, 2, \ldots, r \right\} \tag{3.85}$$

and

$$\mathbf{z}^b = \left\{ \tilde{z}_k^{-1} \quad s.t. \quad 1 + \sum_{m=1}^{r} \mathbf{a}_r^b[m] \tilde{z}_k^m = 0, \quad k = 1, 2, \ldots, r \right\}. \tag{3.86}$$

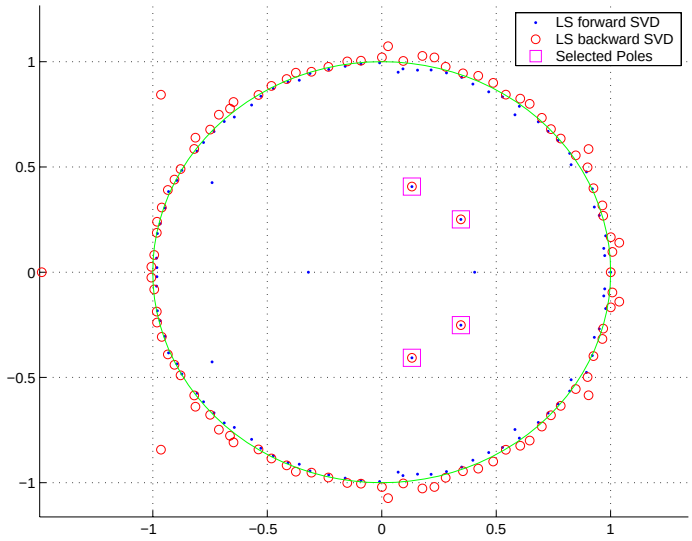


Figure 3.8: Order selection by the forward and backward polynomial LP based method for a two–tone signal (4 complex conjugated poles) corrupted by additive white noise ($SNR = 150dB$).

Therefore the true signal order p with $p \leq r$ can be found as the number of poles in $\mathbf{z}^f$ and $\mathbf{z}^b$ such as satisfy the following relation:

$$\left\{ |z_j| < 1 \quad \text{and} \quad |\tilde{z}_j| > 1 \quad \text{s.t.} \quad |z_j - \tilde{z}_j^{-1}| \leq e_z \quad \text{with} \quad j = 1, 2, \ldots, p \right\}. \tag{3.87}$$

An example of this is shown in Fig. 3.8 where a signal composed of two sinusoids (of normalized frequency $f_1 = 0.1$ and $f_2 = 0.2$, and damping factor $\alpha_1 = \alpha_2 = 0.85\,s^{-1}$) is summed with white noise for a SNR $150dB$. The total number of extrapolated poles is $r = 100$ and the true order $p = 4$ is detected by selecting the pair of poles satisfying (3.87), where $e_z = 10^{-4}$. These poles are depicted as square marks in Fig. 3.8. The rest of the poles correspond to noise poles and lies mainly on the unit circle. Low values for the damping factor would bring the true poles close to the noise ones, and too high a number for r of extrapolated poles may introduce interference between the two sets, with wrong order selected. In real applications, where this scenarios may occur quite often, this technique is not very reliable as is shown in Fig. 3.9, where systems poles and noise poles may lay also very close to each other. The forward and backward LS based order estimation algorithm is summarized in Tab. 3.4

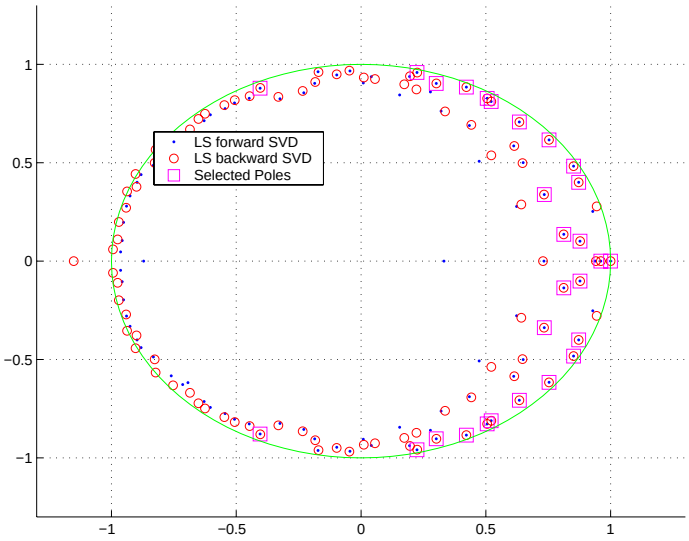


Figure 3.9: Poles selected by the forward and backward polynomial LP based method: p=28 among an initial r=100, for a real structure transient (patch antenna). The $e_z = 10^{-4}$ and with SNR above $100dB$.

Table 3.4: FB LS based model order selection algorithm

	Given a time discrete signal $x[n]$ with $n = 0, 1, \ldots, N-1$:
1.	A starting value for r is selected ($r \leq \lfloor \frac{N}{2} \rfloor$)
2.	A value for the threshold e_z is defined (10^{-4} is a reasonable value)
3.	The number of poles r poles is estimated via LS forward ($\mathbf{z}^f$) and backward ($\mathbf{z}^b$)
4.	The model order is selected as that value p satisfying (3.87)

3.4.3 SVD based Method

The SVD technique is a powerful tool in digital signal processing and SI because it allows to extract the main information about a system or a signal, from a corresponding system or signal data matrix (typically corrupted by noise) [52]. It enables the estimation of the associated model complexity, i.e. model order, by selecting the data matrix rank which specifies the dimension of the underling signal space. In this way the SVD method allows the separation of the signal from the coexisting noise. In many applications a data matrices $\hat{\mathbf{X}}$ of measured or simulated data samples is defined by $N \times p$ real quantities where $N \gg p$

(see (3.8)) and can be decomposed in an exact data matrix $\mathbf{X}$ and an additive perturbation noise matrix $\mathbf{E}$ so that:

$$\hat{\mathbf{X}} = \mathbf{X} + \mathbf{E}. \tag{3.88}$$

Typically the $rank(E) = r$ and therefore the $rank(\hat{\mathbf{X}}) = r$. For applications that are based upon linearity the exact data matrix is rank deficient, i.e.

$$rank(\mathbf{X}) = p < r. \tag{3.89}$$

Let us suppose that the SVD of $\mathbf{X}$ is known a–priori and is given by:

$$\mathbf{X} = [\mathbf{U}_{x1}\mathbf{U}_{x2}] \begin{bmatrix} \mathbf{S}_{x1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{x1}^T \\ \mathbf{V}_{x2}^T \end{bmatrix} \tag{3.90}$$

where $\mathbf{U}_{x1}$ and $\mathbf{U}_{x2}$ are matrices of $N \times p$ and $N \times (q-p)$ dimension and $\mathbf{V}_{x1}$ and $\mathbf{V}_{x2}$ are matrices of dimension $r \times p$ and $r \times (r-p)$, composed from the unitary eigenvectors of the matrices $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$, respectively. The matrix $\mathbf{S}_{x1}$ is an $p \times p$ diagonal matrix of singular values for $\mathbf{X}$ (see Appendix B). From this it is evident that the row space of $\mathbf{X}$ coincides with the column space of $\mathbf{V}_{x1}$ while the columns of $\mathbf{V}_{x2}$ generate an orthogonal basis for the null space of $\mathbf{X}$. Using $\mathbf{V}_{x1}$ and $\mathbf{V}_{x2}$ the matrix $\hat{\mathbf{X}}$ can be rewritten as

$$\begin{aligned} \hat{\mathbf{X}} = &\quad \mathbf{X} + \mathbf{E} = \mathbf{X} + \mathbf{E}\mathbf{V}_{x1}\mathbf{V}_{x1}{}^T + \mathbf{E}\mathbf{V}_{x2}\mathbf{V}_{x2}{}^T = \\ = &\quad (\mathbf{X}\mathbf{V}_{x1} + \mathbf{E}\mathbf{V}_{x1})\,\mathbf{V}_{x1}{}^T + (\mathbf{E}\mathbf{V}_{x2})\,\mathbf{V}_{x2}{}^T. \end{aligned} \tag{3.91}$$

Applying the following decompositions

$$\mathbf{X}\mathbf{V}_{x1} + \mathbf{E}\mathbf{V}_{x1} = \mathbf{P}_1\mathbf{S}_1\mathbf{Q}_1{}^T \tag{3.92}$$

$$\mathbf{E}\mathbf{V}_{x2} = \mathbf{P}_2\mathbf{S}_2\mathbf{Q}_2{}^T \tag{3.93}$$

(3.91) may be rewritten as

$$\begin{aligned} \hat{\mathbf{X}} = &\quad \mathbf{P}_1\mathbf{S}_1\mathbf{Q}_1{}^T\mathbf{V}_{x1}{}^T + \mathbf{P}_2\mathbf{S}_2\mathbf{Q}_2{}^T\mathbf{V}_{x2}{}^T = \\ = &\quad [\mathbf{P}_1\mathbf{P}_2] \begin{bmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_2 \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1{}^T\mathbf{V}_{x1}{}^T \\ \mathbf{Q}_2{}^T\mathbf{V}_{x2}{}^T \end{bmatrix}, \end{aligned} \tag{3.94}$$

which is an SVD of $\hat{\mathbf{X}}$ if and only if ${\mathbf{P}_1}^T\mathbf{P}_2 = 0$ or equivalently if and only if

$$ {\mathbf{V}_{x1}}^T \left(\mathbf{X}^T + \mathbf{E}^T\right)\mathbf{E}\mathbf{V}_{x2} = \mathbf{0}. \tag{3.95} $$

This expression can be split into the following two:

$$ \mathbf{X}^T\,\mathbf{E} = \mathbf{0} \tag{3.96} $$

$$ {\mathbf{V}_{x1}}^T\mathbf{E}^T\mathbf{E}\mathbf{V}_{x2} = \mathbf{0}. \tag{3.97} $$

The physical meaning of (3.96) is that the signal space and the noise space are orthogonal while for (3.97) the noise is white gaussian (i.e. $\mathbf{E}^T\mathbf{E} = \sigma^2\mathbf{I}_p$ where σ is the noise variance). This follows by recalling that $\mathbf{V}_{x1}$ and $\mathbf{V}_{x2}$, as mentioned in (3.90), are orthogonal, which is found to be consistent with most engineering applications. Applying these assumptions, (3.91) and the SVD of the matrix $\hat{\mathbf{X}}$ can be rewritten as follows:

$$ \hat{\mathbf{X}} = \mathbf{X} + \mathbf{E} = [\mathbf{U}_{\hat{x}1}\mathbf{U}_{\hat{x}2}]\begin{bmatrix} \mathbf{S}_{\hat{x}1} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{\hat{x}2} \end{bmatrix}\begin{bmatrix} {\mathbf{V}_{\hat{x}1}}^T \\ {\mathbf{V}_{\hat{x}2}}^T \end{bmatrix} \tag{3.98} $$

where

$$ \mathbf{U}_{\hat{x}1} = (\mathbf{U}_{x1}\mathbf{S}_{x1} + \mathbf{E}\mathbf{V}_{x1})\sqrt{\left(\mathbf{S}_{x1}^2 + \sigma^2\mathbf{I}_r\right)}, \tag{3.99} $$

$$ \mathbf{U}_{\hat{x}1} = \mathbf{E}\mathbf{V}_{x2}\sigma^{-1} \tag{3.100} $$

and (3.99) and (3.100) are of dimension $N \times p$ and $N \times (r-p)$, respectively and $\mathbf{V}_{\hat{x}1} = \mathbf{V}_{x1}$, $\mathbf{V}_{\hat{x}2} = \mathbf{V}_{x2}$. The square matrices $\mathbf{S}_{\hat{x}1}$ and $\mathbf{S}_{\hat{x}2}$, of size $p \times p$ and $(r-p) \times (r-p)$, respectively are defined as

$$ \mathbf{S}_{\hat{x}1} = \sqrt{\left(\mathbf{S}_{x1}^2 + \sigma^2\mathbf{I}_r\right)} \tag{3.101} $$

$$ \mathbf{S}_{\hat{x}1} = \sigma\mathbf{I}_{r-p}. \tag{3.102} $$

The same decomposition between signal and noise subspaces can be observed by SVD of the correlation matrix $\mathbf{R}$ related to the data matrix defined as

$$ \mathbf{R} = \hat{\mathbf{X}}\,\hat{\mathbf{X}}^T \tag{3.103} $$

where by substituting (3.94) and using the unitary properties of the matrices $\mathbf{V}_{\hat{x}_1}$ and $\mathbf{V}_{\hat{x}_2}$ it is possible to decompose $\mathbf{R}$ as:

$$\mathbf{R} = [\mathbf{U}_{\hat{x}1}\mathbf{U}_{\hat{x}2}] \begin{bmatrix} \mathbf{S}^2_{\hat{x}1} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^2_{\hat{x}2} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\hat{x}1}{}^T \\ \mathbf{U}_{\hat{x}2}{}^T \end{bmatrix}. \tag{3.104}$$

Now concerning a noise corrupted signal, the assumption of the existence of a gap in the singular spectrum applies. This means that the smallest singular value in $\sqrt{\left(\mathbf{S}^2_{x1} + \sigma^2\mathbf{I}_r\right)}$, be it σ^x_p, is bigger than the largest singular value of $\sigma\mathbf{I}_{r-p}$. In reality this gap is very small and in general the signal singular values are almost contiguous to the noise ones. The separation in such cases can be carried out only by setting a threshold value defined as follows

$$D_T = 20\log_{10}\frac{\sigma^x_p}{\sigma}. \tag{3.105}$$

For many practical cases where the matrix $\mathbf{S}_{\hat{x}2}$ has singular values which are not constant (valid only for white noise), but decreasing as $\sigma^e_1 \geq \sigma^e_2 \geq \ldots \geq \sigma^e_{r-p} \geq 0$, and a more reasonable value for the threshold (3.105) may be chosen as follows:

$$D_T = 10\log_{10}\frac{\sum_{1=1}^{p}\sigma_i^{x\,2}}{\sum_{j=1}^{r-p}\sigma_j^{e\,2}}. \tag{3.106}$$

D_T can be interpreted as a SNR of the signal values space. Worth of note is that taking the square of the singular values in (3.106) is equivalent to using the singular values for the correlation matrix in the same formula. This reasoning yields a model order selection criterion which can be formulated as follows. The model order of the system or signal of associated data matrix is $\hat{\mathbf{X}}$ is given from the dimension of the underling signal space, or equivalently from the approximated rank of $\hat{\mathbf{X}}$. This order is determined from the value p such that:

$$SNR(p) = 10\log_{10}\frac{\sum_{i=1}^{p}\sigma_i^{\hat{x}\,2}}{\sum_{i=p+1}^{r}\sigma_i^{\hat{x}\,2}} \;\geq\; D_T \tag{3.107}$$

where $\sigma_i^{\hat{\mathbf{x}}}$ are the singular values of the data matrix $\hat{\mathbf{X}}$. Since D_T in (3.107) is given in terms of the unknown singular values of the exact data matrix $\mathbf{X}$ and the error matrix $\mathbf{E}$, it is in

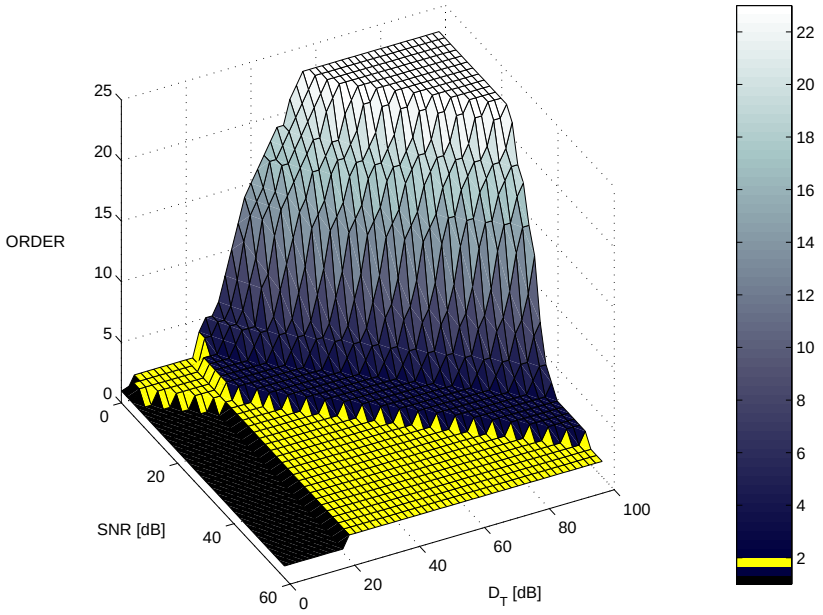


Figure 3.10: Order selection by SVD as function of the threshold level D_T and the SNR.

general unknown and starting value must be chosen. In Fig 3.10 the estimated order for a two–pole signal corrupted by additive white noise (as in (3.80)) is taken into consideration by showing the dependency from the D_T and the noise level (SNR). The higher the SNR is the wider can be the threshold level range assumed by D_T in order to get the exact order ($p = 2$). For this case of high values for the threshold D_T and low SNR the criteria does not converge and the estimated order goes to the upper limit of order of $N/2$, where N is the length of the available signal (here $N = 50$). In Fig 3.11 three different slices at three correspondent different SNR values are given. For the case of real system transients, as in Fig.3.12, the order is proportional to D_T and varies linearly for high SNR ($SNR > 50dB$ – dashed line line) until it saturates for $D_T > 100bB$ settling down to an asymptotic value (likely candidate for the model order here $p = 32$). At the other extreme for a very noisy signal ($SNR < 10dB$) the order diverges (continuous line) since the algorithm cannot distinguish the signal from the added noise. For this example a value of $D_T = 110bB$ seems to be the best compromise in defining the correct model order. If on one hand the implementation of this algorithm is very time consuming, since it is based on SVD (its complexity is proportional to p^3), on the other hand it provides a very stable technique to separate the signal from the noise space. The SVD based model order estimation algorithm

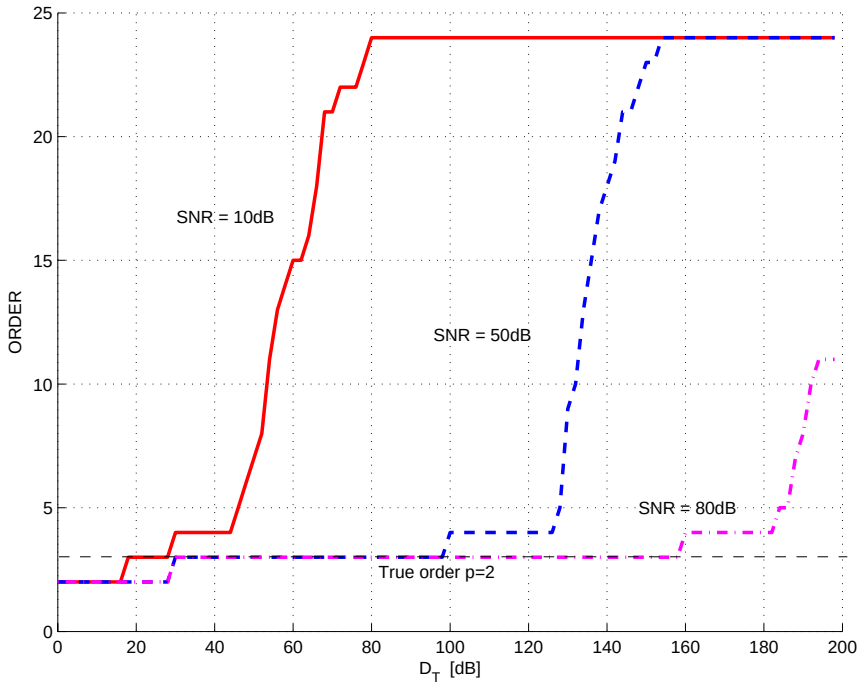


Figure 3.11: Order selection by SVD as function of the threshold level D_T for three level of SNR.

is summarized in Tab. 3.5.

3.4.4 Model order selection methods comparison

The estimation of the model order by means of the forward and backward linear prediction (FBLP) polynomial technique (3.87) and the SVD based technique (3.107) is tested in the following on a signal known a–priori. The four conjugated pairs of poles forming the signal, as in (3.80), are characterized by the following parameters: damping factor $\alpha_1 = \alpha_2 = 0.015\,s^{-1}$, $\alpha_3 = \alpha_4 = 0.051\,s^{-1}$, and normalized frequency $f_1 = \pm 0.2$, $f_2 = \pm 0.1$, $f_3 = \pm 0.5$, $f_4 = \pm 0.25$.The complex pole amplitudes are assumed to be $A_i = 1$ for $i = 1, 2, \ldots, 4$. As shown in Fig. 3.13, the two techniques calculate the same correct result $p = 8$, but while the SVD based technique is already converged for $SNR = 50\,dB$, the forward backward LP based methods need a minimum SNR of $70dB$. The better performance of the SVD based technique however requires more in terms of computational effort since its complexity is $(N - p)^3$ compared to $(N \times p)$ of the forward/backward LP technique.

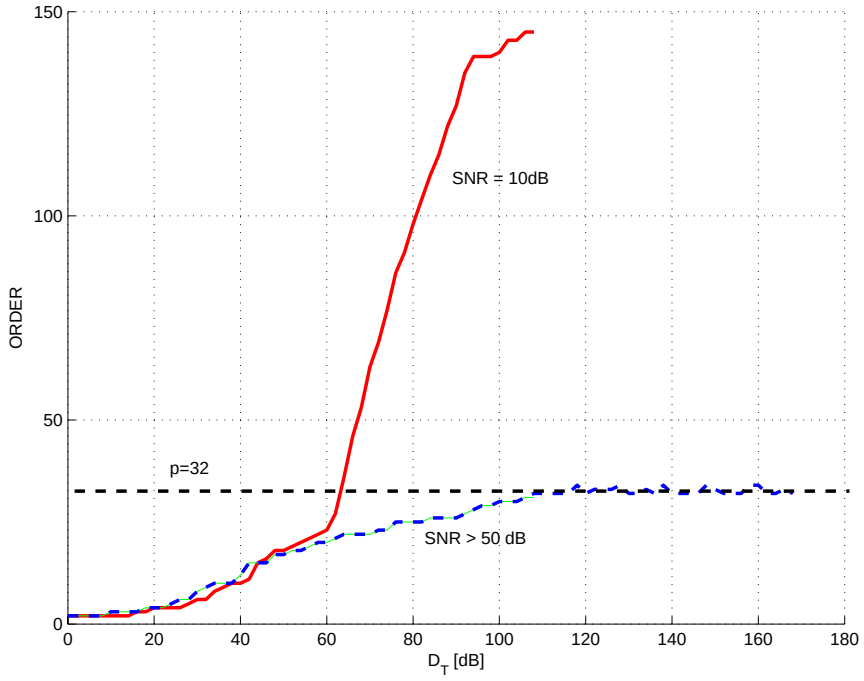


Figure 3.12: Effect of the noise level on the model order selection using the SVD algorithm for the transient response of a typical electromagnetic structure (patch antenna) corrupted with additive white noise.

Table 3.5: SVD based model order selection algorithm

	Given a time discrete signal $x[n]$ with $n = 0, 1, \ldots, N-1$:
1.	A starting value for r is initialized ($r \leq \lfloor \frac{N}{2} \rfloor$)
2.	The $(N-r) \times (r+1)$ data matrix T_r is built as in (3.31)
2.	The threshold value for the selection D_T is initialized (100 default value)
3.	The singular values for the data matrix T_r are computed
4.	The model order is selected as that value p satisfying the (3.107)

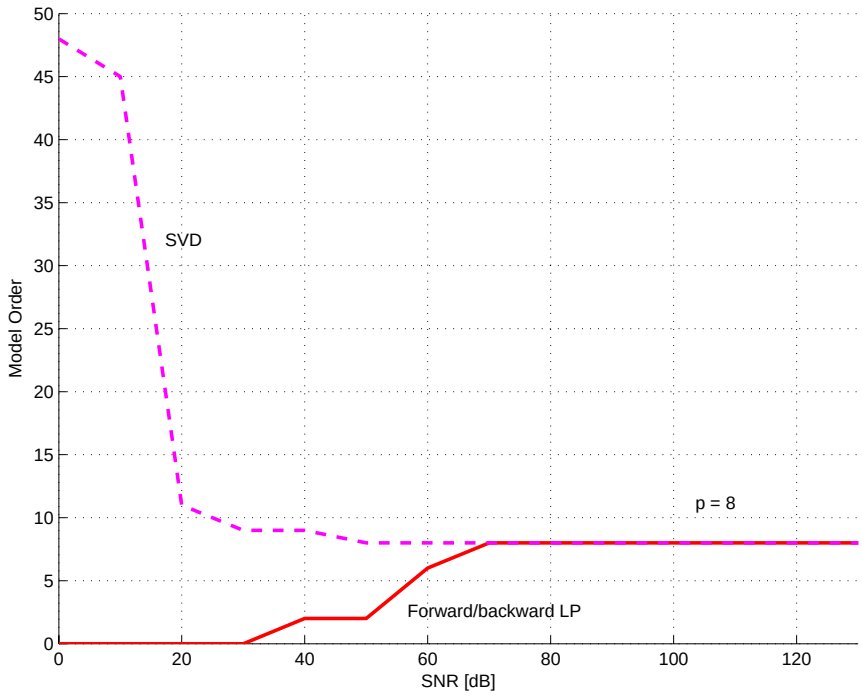


Figure 3.13: Comparison of the order selection techniques for a 4 tone signal (p=8), with forward backward LS based method (—) and SVD based method (- - -). The selection limits for the two methods are $e_z = 10^{-5}$ and $D_T = 100dB$.

Chapter 4

Modelling of Passive Electromagnetic Networks

Typically, experimental analysis of physical phenomena is based upon time domain update of observable data. The ability to predict the behavior of an observed quantity from an already available ensemble of data samples can play a key role in shortening down the overall analysis. If such an operation is accurately and successfully carried out the advantages can be significant: First the period of observation for the data, normally time and resource expensive (be it a measurement or a simulation), can be terminated as soon as the predicted data set differs from the experimental one by a selected error criteria, and second, since the prediction is normally based upon an adaptive parametric modelling, the observed data becomes immediately available in analytic form by means of few (compared with the data length) parameters and the corresponding model structure. Therefore, in general, signal prediction yields shorter experimental observations as well as compact representations of the observed data. There are several techniques applied to the modelling of data.
In the present chapter, the techniques of SI introduced in the previous chapter, will be applied for the prediction of the impulse responses obtained by means of a full–wave numerical analysis of passive electromagnetic network. A fully automated algorithm, able to adaptively select the model order and carry out the parameter identification will be presented and tested on some practical examples. Since the basic model structure is the Prony method, the algorithm is hereby referred at as Prony model based System Identification (PMSI). This algorithm will be used to recover the impulse response (prediction problem) and the impedance and admittance Foster representations (identification) of some electromagnetic structures.

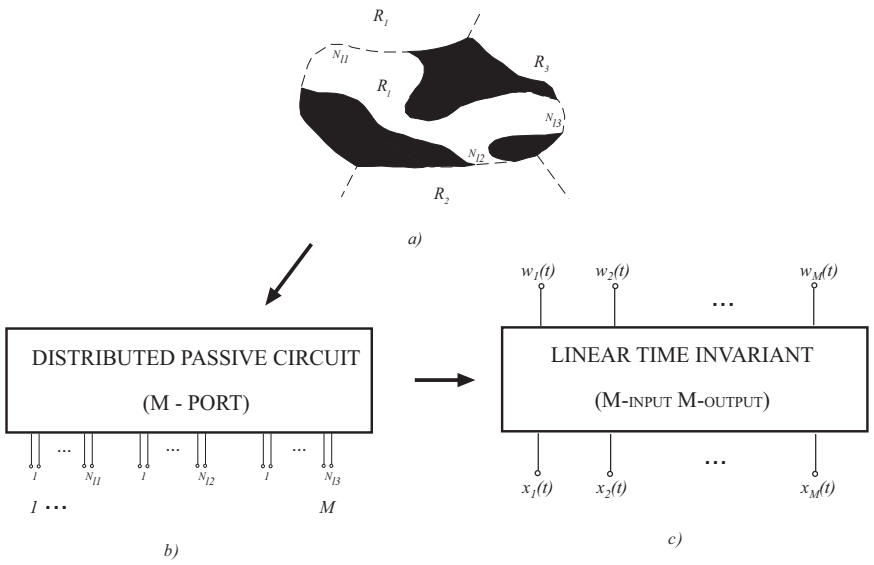


Figure 4.1: Passive structure modelling: a)Definition of the access ports (modes) on the access (boundary) surfaces; b) interpretation of the original structure as distribute M-port network; c) interpretation as an LTI system with M inputs and M outputs

4.1 Passive Network characterization

A passive distributed microwave and millimeter wave circuit may be considered as a linear time invariant (LTI) system, whose global description may be given either in time, with the impulse response, or in frequency, with the transfer function.
The case of a structure modelled as an M ports network is shown schematically in Fig. 4.1 (a) and (b). Let the $M \times 1$ vectors $\mathbf{w}(t)$ and $\mathbf{x}(t)$ be the input and output vectors of the corresponding LTI system given in Fig. 4.1 (c). The complete characterization of the network at system level is given by the well know convolution relationship:

$$\mathbf{x}(t) = \int_{-\infty}^{+\infty} \mathbf{h}(t-\tau)\mathbf{w}(\tau)\, d\tau \tag{4.1}$$

where

$$\mathbf{x}(t) = [x_1(t)\, x_2(t) \ldots x_M(t)]^T,$$
$$\mathbf{w}(t) = [w_1(t)\, w_2(t) \ldots w_M(t)]^T,$$

and

$$\mathbf{h}(t) = \begin{bmatrix} h_{11}(t) & h_{12}(t) & \cdots & h_{1M}(t) \\ h_{21}(t) & h_{22}(t) & \cdots & h_{2M}(t) \\ \vdots & \vdots & & \vdots \\ h_{M1}(t) & h_{M2}(t) & \cdots & h_{MM}(t) \end{bmatrix}$$

is the $M \times M$ impulse response matrix of the system. By transforming (4.1) into the Laplace domain the matrix convolution simplifies into a simple matrix product, given as:

$$\mathbf{X}(s) = \mathbf{H}(s)\mathbf{W}(s), \tag{4.2}$$

where: $\mathbf{W}(s) = \mathscr{L}\{\mathbf{w}(t)\}$, $\mathbf{X}(s) = \mathscr{L}\{\mathbf{x}(t)\}$ and $\mathbf{H}(s) = \mathscr{L}\{\mathbf{h}(t)\}$ with:

$$\mathbf{X}(s) = [X_1(s)\, X_2(s), \ldots, X_M(s)]^T,$$

$$\mathbf{W}(s) = [W_1(s)\, W_1(s), \ldots, W_M(s)]^T,$$

and where

$$\mathbf{H}(s) = \begin{bmatrix} H_{11}(s) & H_{12}(s) & \cdots & H_{1M}(s) \\ H_{21}(s) & H_{22}(s) & \cdots & H_{2M}(s) \\ \vdots & \vdots & & \vdots \\ H_{M1}(s) & H_{M2}(s) & \cdots & H_{MM}(s) \end{bmatrix}.$$

$\mathbf{H}(s)$ is the $M \times M$ network transfer function matrix which relates input $\mathbf{W}(s)$ and output $\mathbf{X}(s)$. Depending on the choice of these latter the transfer function can assume the physical meaning of the impedance (admittance) matrix $\mathbf{Z}(s)$ ($\mathbf{Y}(s)$) for the case of the input $\mathbf{W}(s)$ corresponding to the open circuit port current $\mathbf{I}(s)$ (short circuit port voltage $\mathbf{V}(s)$) and the output $\mathbf{X}(s)$ the corresponding port voltages $\mathbf{V}(s)$ (port current $\mathbf{I}(s)$). $\mathbf{H}(s)$ become also the scattering matrix $\mathbf{S}(s)$ when the input and output are chosen as the incident and the reflected wave amplitude vectors $\mathbf{A}(s)$ and $\mathbf{B}(s)$ [72].

In order to completely identify the $M \times M$ matrix $\mathbf{S}(s)$ the network must be analyzed in M separated simulations each exploring the M port responses once the excitation separately applied to one of them. Recalling what has been mentioned in section 2.4, this operation may be carried out with TD simulation (TLM or FDTD). The scattering matrix, is found by observing the M responses once each of the M ports are separately excited. If the excitation is a Dirac Delta impulse function (flat constant spectrum of amplitude 1), applied at the m-th port, the m-th row of the matrix results is directly given from the M responses in

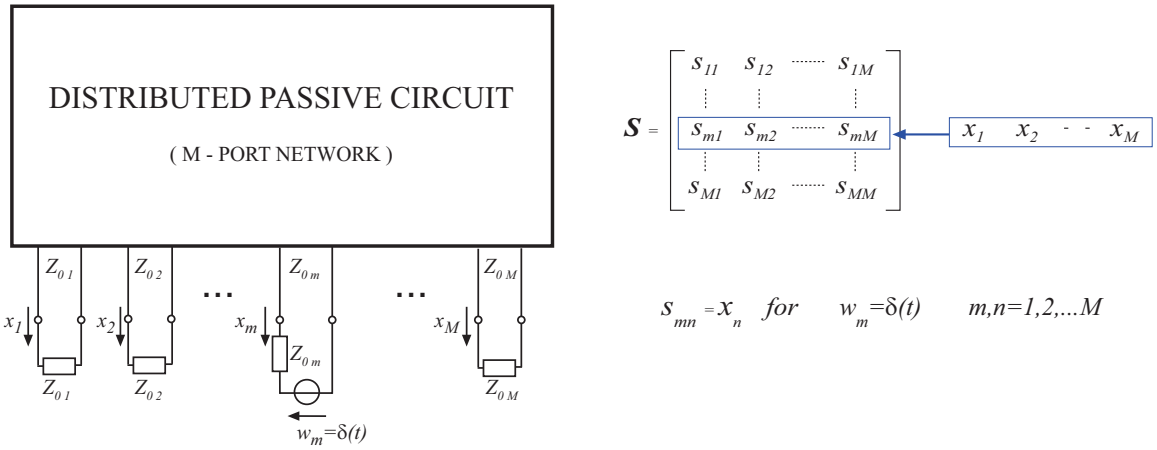


Figure 4.2: Scattering parameters analysis for a M port network by means of impulse excitation.

the vector $\mathbf{x}(t)$ (see Fig. 4.2). Hence, recalling the model for the scattering matrix element $\mathbf{S}_{mn}(s)$ given in equation (2.86), this means that each response $x_n(t)$ may be decomposed exactly in the same manner, by means of a poles series (Prony's) model as:

$$x_n(t) = S_{mn}(t) = \sum_{r=1}^{R} c_{r\,mn} e^{-s_r t} \quad \xrightarrow{\mathscr{L}} \quad \sum_{r=1}^{R} \frac{c_{r\,mn}}{s - s_r} = S_{mn}(s) = X_n(s). \tag{4.3}$$

For a passive network, the natural frequencies must be real or complex conjugate pairs, allowing (4.3) to be rewritten as

$$\mathbf{x}_n(t) = \sum_{r=1}^{R} c_{r\,mn} e^{-\alpha_r t} \cos(\omega_r t + \phi_r) \tag{4.4}$$

$$\mathbf{X}_n(s) = \frac{1}{2} \sum_{r=1}^{R} \left[\frac{c_{r\,mn}}{s - s_r} + \frac{c_{r\,mn}{}^*}{s - s_r^*} \right] \tag{4.5}$$

where R is the number of structure's natural frequencies; $s_r = \alpha_r + \jmath\omega_r$, and ϕ_r are the r-th structure's natural frequency and initial phase respectively; and $c_{r\,mn}$ is the corresponding residue. For a passive network in order to ensure the stability $\alpha_r \leq 0$, for any $r = 1, 2, \ldots, M$. In order to accomplish the entire network characterization M impulse response analysis are needed, unless some of them can be omitted because of symmetry properties. In radar theory, expression (4.3) plays a very relevant role with regard to target

identification, a topic which is directly related to the present work, and is also extensively treated by the singularity expansion method theory.

4.2 Singularity Expansion Method

The theory of the so called singularity expansion method (SEM) is based on the analytical properties of the electromagnetic system response as a function of the Laplace transform variable or complex frequency s. Development of the SEM was stimulated by the observation of the generic characteristics of typical transient responses in experiments on various complex scatterers (e.g. aircraft) [5]. The transient response waveforms of objects are dominated by few damped sinusoids. These are not the only type of functions appearing in such transient data, but they are quite prevalent, especially for long observation times (steady state systems). Since the Laplace transform of a damped sinusoid corresponds to one pole or a pair of conjugate poles in the complex frequencies plane, the scattering object can be expected to have a long time response at frequencies in the vicinity of such poles. A broadband pulse excites such poles or natural frequencies. These natural frequencies correspond to natural system modes which are a field, charge or current solution of the free oscillation problem. The modal distributions are not a functions of the incident field, but rather depend on the object structure itself (geometry and materials). The complex amplitude coefficients of the natural modes are dependent on the source function. They correspond to the strength of the resonances and are also called coupling coefficients. Natural frequencies have been discussed and investigated in electromagnetic theory since 1897, with Pocklington, who gave an analytical expression for the natural frequencies of linear and circular scatterers [60]. Page and Adams investigated the case of a prolate spheroid and thin wires [57] [58], and Stratton treated a spherical scatterer [78]. The theory of natural frequencies applies also to cavities or more in general to network scattering problems, where the natural modes form a complete set and have convenient orthogonality properties [76]. In the present work, the SEM theory will be applied to the impulse response of passive networks obtained from the full–wave simulation of electromagnetic structures. When SEM is applied to scattering from objects, the response can be decomposed into two separate responses, which are typically distributed over contiguous, (when not overlapping) time regions. The first one, which is directly related to the input excitation function, is called the driven response; the second has the characteristics of a transient response, and immediately follows the input.

An expression for describing a such behavior is given by:

$$x(t) = x_D(t) + x_T(t) \tag{4.6}$$

where $x_D(t)$, is the the driven response, and $x_T(t)$ is the transient response. This latter may be written in the form of (4.3) as,

$$x_T(t) = \sum_{p=1}^{P} c_p e^{-\alpha_p t} \cos\left(\omega_p t + \phi_p\right). \tag{4.7}$$

P is in general the number of all possible resonances taking place in the structure (scattering object modes), and the quantities ω_p and α_p are parameters denoting the angular frequency and the damping of the oscillations both of which are directly related to the physical properties (dimensions and material) of the scattering object they refer to. The quantity c_p and ϕ_p depend on the spatial relationship between scattering objects and observation points (ports), and represent a sort of weighting coefficient for the amplitude of a resonance at a specific output.

The driven response may be considered as the input excitation itself after propagating through the structure from the excitation port to the considered output port (see Fig. 4.3(a)), and is typically located in the very first portion (early time) of the impulse response. In the first approximation it may be describe as an attenuated and delayed function (considering multipath reflection and dispersion) of the original excitation. The transient response on the other hand is given as the superposition (see 4.3) of damped oscillations, whose delay depends on the distance between the input and output ports. Typically the transient response is located at the end (late time) of the impulse response. This principle is qualitatively depicted in Fig. 4.3(b). In general when dispersion and multipath are taken into account the identification of the driven response becomes a quite complicated task, as the two responses may overlap and cover-up each other. Each object, or part of the entire structure (including the enclosure) can give rise to a discrete, infinite number of modes each characterized by a natural frequency, whose value may be however, out of the range of interest. These modes represent field, charge or current solutions of the electromagnetic problem. The structure natural frequencies are evidently the most representative parameters for an analytical description (Prony's or poles series model type) of the network responses being closely related to the physical characteristics of the given structure. These complex frequencies or poles can only be extracted from the transient part $x_T(t)$ and become the

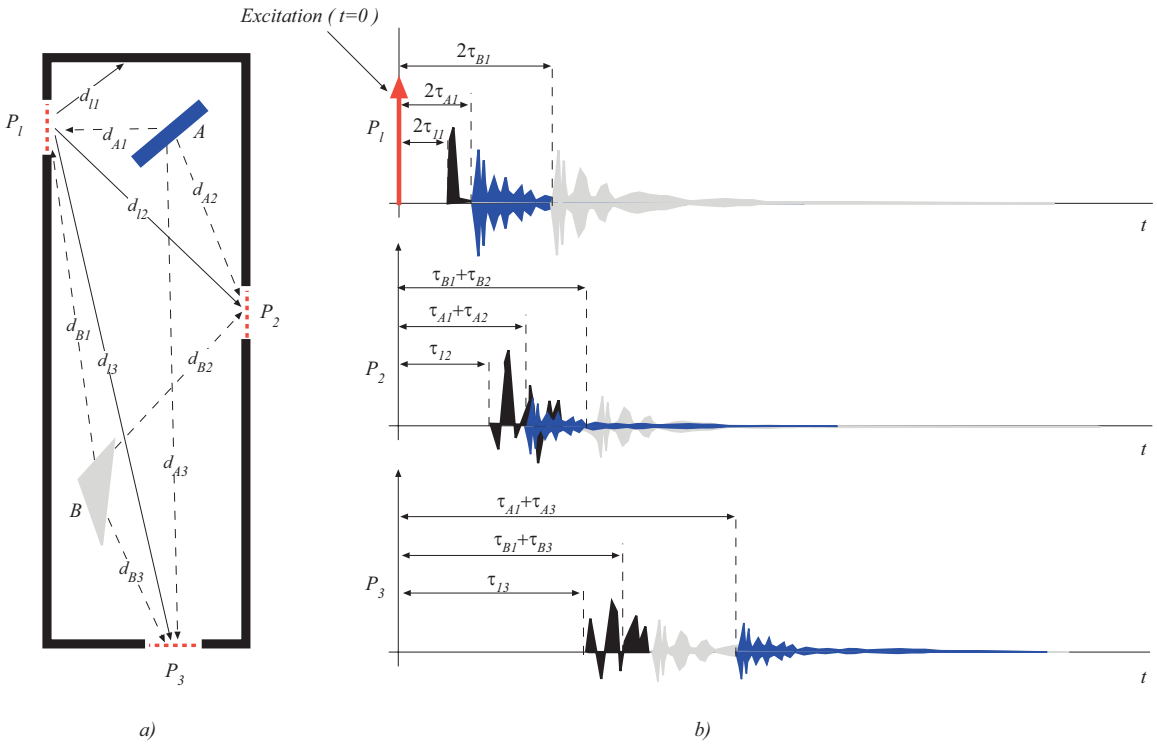


Figure 4.3: Description of the SEM applied to internal scattering: a) a closed electromagnetic structure with three access ports and two scattering objects (A and B); b)impulse response at the three ports as decomposed into driven (black) and transient responses (blue for A and light gray for B) for a Dirac impulse excitation at port P_1 (delay $\tau_{1i} = d_{1i}/c$, with $i = 1, 2, 3, A, B$ and c the phase velocity of the media).

structure identification parameters. The same principle is exploited in radar applications where the aim is to recover the target signature by identifying the object's unique set of poles [1].

The analysis of electromagnetic structures is often carried out by means of a band limited impulse excitation to avoid the generation of spurious modes in numerical computation. In Fig. (4.4) the impulse (gaussian pulse) response of an electromagnetic structure (microstrip stop band filter [16]) is depicted. The two regions corresponding to the driven and transient part are highlighted. Although as mentioned above, the separation of the two can be theoretically carried out by estimating propagation delay times, a more reliable method consists of the identification of the natural frequencies. Indeed, while the transient response may be characterized with a few frequency components, once a frequency limit has been imposed, the driven part, typically as broad band as the excitation, would need many of them. This is evident in Fig. (4.5), where the poles are extracted, from the impulse response starting from

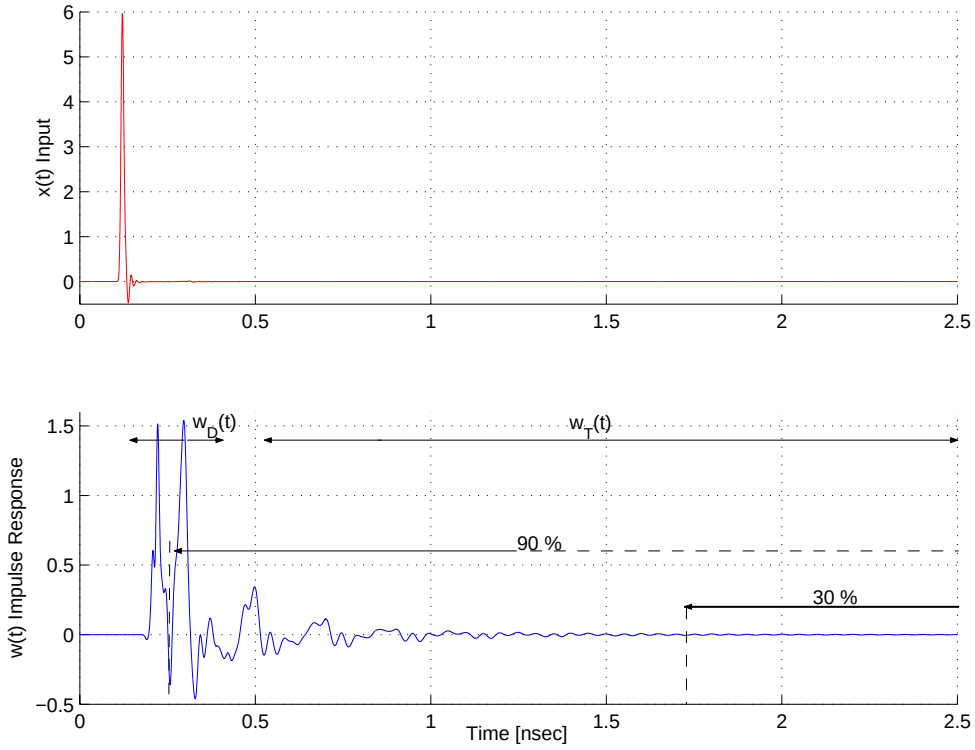


Figure 4.4: Impulse excitation (above) and impulse response of a passive microwave network (below). The qualitative separation between driven and transient components is shown.

the end of the 2.5 *ns* impulse response of Fig. (4.4), and by taking an increasing portion of it. When up to 30% of the late response is used to extract the system poles, the number of natural frequencies is stable at 4, i.e. it can be decomposed in as few as only 2 oscillating components. By taking more time samples in the earlier region of the time response (above 30% of the response queue) the pole content increases until it reaches the region in the time response where the driving response is likely to be. This is observed in Fig. (4.4) as the number of poles (needed to rebuild the signal) quickly increases to 29.

4.3 Systematic Network Response Modelling and Prediction

The problem of response prediction is essentially a problem of system identification (SI) [43] [26] and more properly speaking it consists of data modelling. The data to be modelled

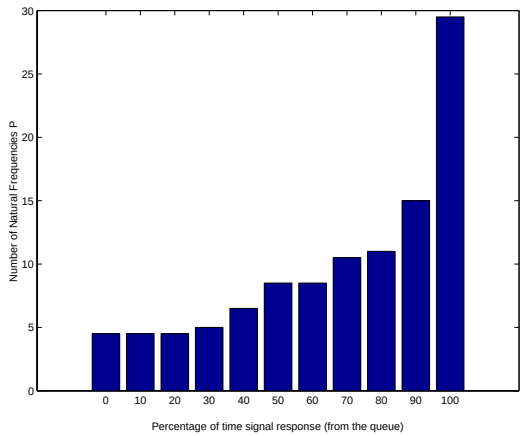


Figure 4.5: Number of natural frequencies extracted as function of the percentage of the considered late time domain response given in Fig. 4.4.

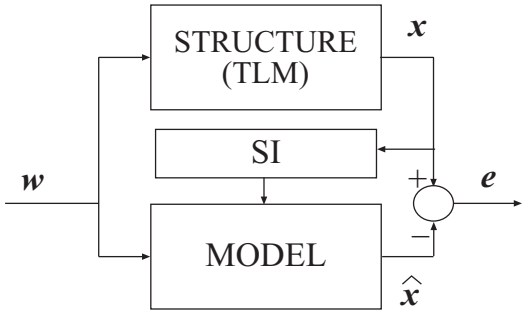


Figure 4.6: Block diagram for the SI modelling approach of Electromagnetic structures.

in the present work are ensembles (**x**) of time discrete signals generated from full-wave simulations. Since these latter may be even computationally very demanding, it makes sense to process them while they are running, for extracting a model (analytical description) able to recover such responses ($\hat{\mathbf{x}}$). This operation allows the system response prediction, based on the recovered system response model. The principle is shown in the block diagram of Fig. 4.6 where the model, is build and adaptively refined by exploiting the new response updates provided from the TD simulation of the structure under test, running in parallel. As soon as the comparison between SI model reaches a satisfactory accuracy (error **e** below a given threshold) the full–wave simulation may be disrupted. At this point the model represents a synthetic accurate analytical description of the original network.

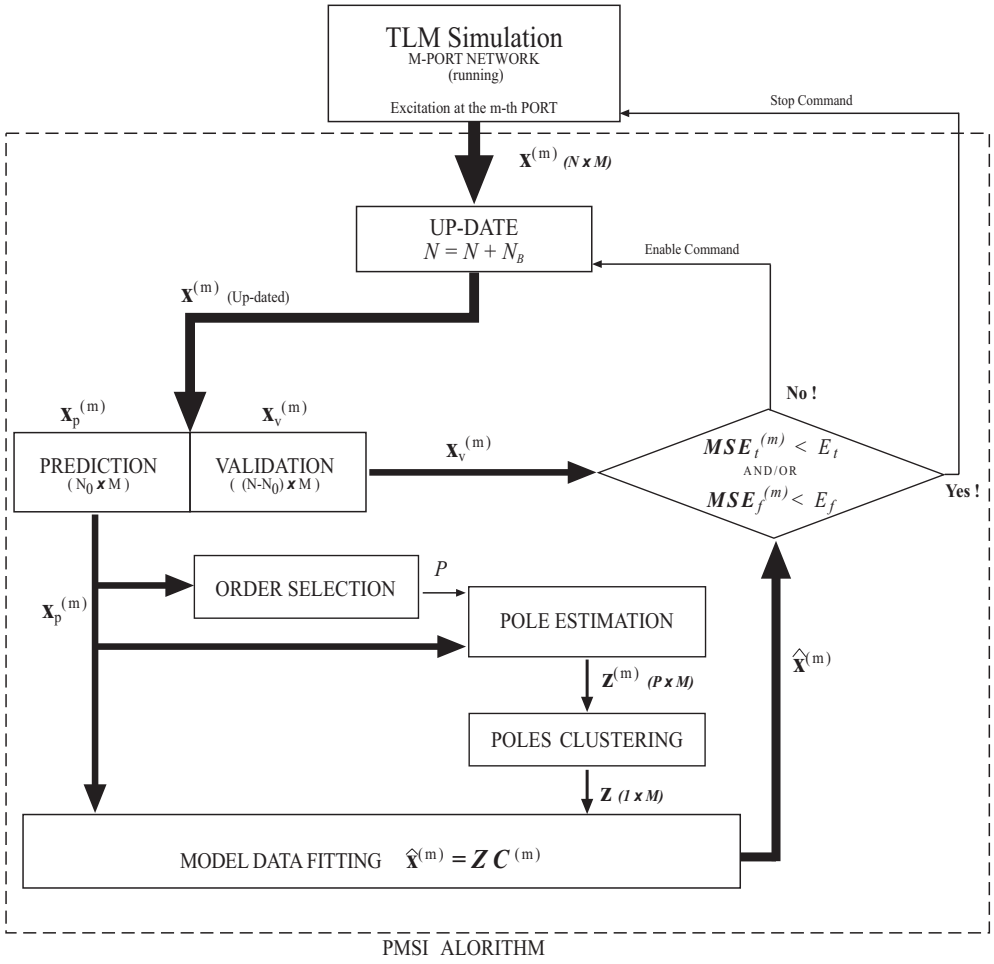


Figure 4.7: Block diagram for the prediction of TLM simulated impulse responses, by means of the Prony model based SI (PMSI) algorithm.

4.3.1 Prony model based System Identification

Based on the result of the investigation on SI techniques, carried out in the previous chapter, an algorithm called Prony model based system identification (PMSI), has been developed. The PMSI algorithm will be presented here by applying it directly to the problem of modelling and prediction of TLM simulated responses, as shown in Fig. 4.7. The algorithm, is conceived as a completely autonomous module which requires as input only the impulse responses provided by a TD full–wave solver. A detailed explanation of the method is described in the following.

Model parameters estimation

Let $\mathbf{x}^{(m)}$ be an $N \times M$ data matrix, corresponding to the m-th simulation (carried out by exciting the m-th port) and consisting of the first N samples of the M responses. As later on, in order to check the model accuracy independently from the sample used to build it, the so called model validation should be carried on a different set of samples, the matrix $\mathbf{x}^{(m)}$ is split in two parts. The first $\mathbf{x}_p^{(m)}$, of dimensions $N_0 \times M$, is used for the prediction and contains the early N_0 samples, and the second $\mathbf{x}_v^{(m)}$, of dimensions $N - N_0 \times M$, is used later on for validation. $\mathbf{x}^{(m)}$ is given by

$$\mathbf{x}^{(m)} = \begin{bmatrix} \mathbf{x}_p^{(m)} \\ \mathbf{x}_v^{(m)} \end{bmatrix}, \tag{4.8}$$

wheras

$$\mathbf{x}_p^{(m)} = \begin{bmatrix} x_1^{(m)}[0] & \dots & x_M^{(m)}[0] \\ \vdots & \dots & \vdots \\ x_1^{(m)}[N_0-1] & \dots & x_M^{(m)}[N_0-1] \end{bmatrix}, \quad \mathbf{x}_v^{(m)} = \begin{bmatrix} x_1^{(m)}[N_0] & \dots & x_M^{(m)}[N_0] \\ \vdots & \dots & \vdots \\ x_1^{(m)}[N] & \dots & x_M^{(m)}[N] \end{bmatrix}. \tag{4.9}$$

The quantity N_0 is normally chosen as a fixed percentage of the ensemble length N. Typically in SI theory values above 66% are considered appropriate. As mentioned the portion $\mathbf{x}_p^{(m)}$ is devoted to the prediction, and more specifically is exploited for three different purposes, the selection of the model order P, the estimation of the poles, and eventually the computation of the residues for the data fitting.

Model Order Selection

The model order selection is carried out by applying the SVD based method (see section 3.4.3) to each column of $\mathbf{x}_p^{(m)}$, yielding a vector $\mathbf{P}^{(m)} = \left[P_1^{(m)}, P_2^{(m)}, \dots, P_M^{(m)}\right]$. The model order is a single integer quantity which denotes the total number of the structure's natural frequencies. An approximated criteria to select it can be to choose the maximum between all the model orders estimated during the current simulation and all the previously determined ones as:

$$P = P^{(m)} = \max_{i=1,\dots,M} \left(P_i^{(m)}, P^{(m-1)}\right) \tag{4.10}$$

whereas $P^{(m-1)}$ denotes the maximum of the previously performed $m-1$ analysis ($P^{(0)}$ must be initialized to 1, since it corresponds to the first simulation in the row to take place). An overestimated model order does not represent a problem since, while it may yield higher complexity, it also ensures that the most significant natural frequencies have been taken into account. Poles which do not significantly affect the system response (very low value) may be easily modelled by setting to a zero value the corresponding residues in (3.8).

Model pole estimation

Once the order P is available the pole estimation may be carried out. Here the the linear prediction LS method (see section 3.3.1) or PM method (see section 3.3.2) may be applied. P poles are estimated for each of the M columns in $\mathbf{x}_p^{(m)}$ and stored in a $P \times M$ matrix $\mathbf{z}^{(m)}$, given as,

$$\mathbf{z}^{(m)} = \begin{bmatrix} z_1^{(m1)} & z_1^{(m2)} & \dots & z_1^{(mM)} \\ z_2^{(m1)} & z_2^{(m2)} & \dots & z_2^{(mM)} \\ \vdots & \vdots & \dots & \vdots \\ z_P^{(m1)} & z_P^{(m2)} & \dots & z_P^{(mM)} \end{bmatrix}, \tag{4.11}$$

where $z_i^{(mn)}$ denotes the i-th pole of the mn-th response, with $i = 1, 2, \dots, P$ and $m, n = 1, 2, \dots, M$. The final set of poles must be unique for all of the M^2 responses (M responses for each of the M simulations). In order to find this unique set, a technique called pole clustering is applied [44]. This operation consists of averaging poles which numerically differ from a given tolerance and that in theory should refer to the same natural frequency value. Since some of the responses may not contain one or more of the network P poles, which can happen that during the pole estimation, in their place, completely wrong values are computed. These values should therefore be disregarded by selecting only those M_i poles stored in each i-th row of $\mathbf{z}^{(m)}$, which lie within a given tolerance range (see Fig. 4.8). The poles obtained from the average between the selected quantities

$$z_i^{(m)} = \frac{1}{M_i} \sum_{n=1}^{M_i} z_i^{(mn)} \quad i = 1, 2, \dots, P \tag{4.12}$$

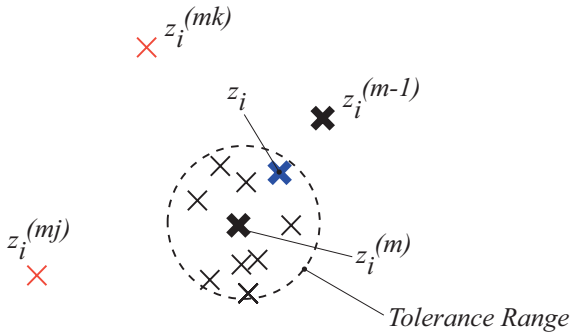


Figure 4.8: Example of pole clustering. Of the overall $P = 10$ poles two $z_i^{(mk)}$ and $z_i^{(mj)}$ are excluded ($M_i = 8$) since they like outside the defined tolerance range.

must be further averaged (see Fig. 4.8) with the previous $z_i^{(m-1)}$ (for $m = 1$, no further averaging is required) in order to find the updated value of the m-th simulation as

$$z_i = \frac{z_i^{(m)} + z_i^{(m-1)}}{2}. \tag{4.13}$$

The $z_i^{(m)}$ should be finally updated with the value $z_i^{(m)} = z_i$ for use in the next pole clustering operation. The final result, at the end of the M simulations, is a $P \times 1$ vector of clustered poles $\mathbf{z} = [z_1, z_2, \ldots, z_M]^T$. The pole clustering allows one to adaptively modify the network natural frequencies by seeking the most likely values among those found in the entire set of simulated responses. Pole clustering however, is a very critical operation which strongly effects the model accuracy.

Model Fitting Parameters (Residues) Computation

The next and final step in building the model is the fitting of the Prony model to the simulated data response $\mathbf{x}_p^{(m)}$ as given in (3.8). The more direct and efficient way is to use a least square (LS) approach (see Appendix B) and compute the residues from (3.14) obtaining:

$$\mathbf{C}^{(m)} = \left(\hat{\mathbf{Z}}^H \hat{\mathbf{Z}}\right)^{-1} \hat{\mathbf{Z}}^H \mathbf{x}_p^{(m)}. \tag{4.14}$$

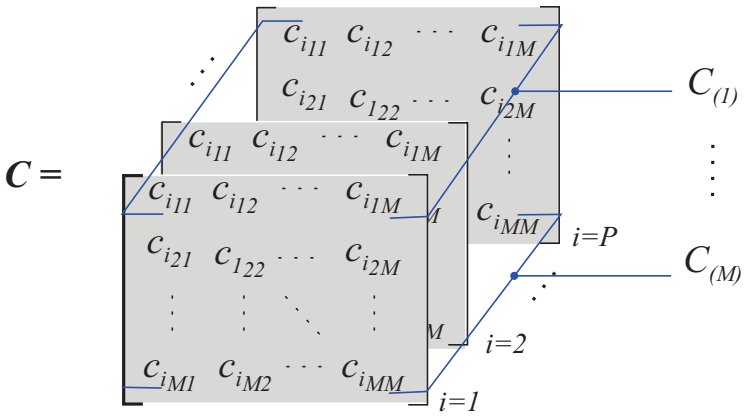


Figure 4.9: Residue matrix **C** for an M-port network.

with $\hat{\mathbf{Z}}$ the Vandermonde matrix made by the clustered poles, and $\mathbf{C}^{(m)}$ the $P \times M$ residues matrix, each given respectively by

$$\hat{\mathbf{Z}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_P \\ \vdots & \vdots & & \vdots \\ z_1^{N_0-1} & z_2^{N_0-1} & \cdots & z_P^{N_0-1} \end{bmatrix}, \quad \mathbf{C}^{(m)} = \begin{bmatrix} c_{1m1} & c_{1m2} & \cdots & c_{1mM} \\ c_{2m1} & c_{2m2} & \cdots & c_{2mM} \\ \vdots & \vdots & & \vdots \\ c_{Pm1} & c_{Pm2} & \cdots & c_{PmM} \end{bmatrix}. \tag{4.15}$$

where c_{imn} denote the residue corresponding to the i-th pole for the mn-th response. The recovered response denoted with $\hat{\mathbf{x}}^{(m)}$ can be eventually written as

$$\hat{\mathbf{x}}^{(m)} = \hat{\mathbf{Z}}\,\mathbf{C}^{(m)}. \tag{4.16}$$

At the end of the entire analysis, i.e. when $m = M$, the residues may be organized in a three dimensional data matrix **C** obtained by packing together P matrices, of size $M \times M$, as shown in Fig. 4.9.

Model Validation

In order to decide whether the available model is accurate enough to stop the current m-th simulation and start the next one the recovered responses $\hat{\mathbf{x}}$ must be compared with the simulated ones. This operation is called validation and is carried out by estimating the

mean square error (MSE). $MSEs$ are computed for each response and ordered in a vector $MSE_t^{(m)}$, and computed as follows:

$$MSE_t^{(m)} = \frac{1}{(N - N_0 - 1)} \sum_{k=0}^{N-1} \left| \mathbf{x}^{(m)}[k] - \hat{\mathbf{x}}^{(m)}[k] \right|^2, \tag{4.17}$$

where the t denotes time domain. Indeed, in most practical applications the final aim is to test the accuracy within a finite frequency band. It therefore makes sense to provide a further accuracy check in the frequency domain by introducing a second mean square error $MSE_f^{(m)}$. This is found by using the simulated $\mathbf{X}^{(m)}$ and modelled $\hat{\mathbf{X}}^{(m)}$ spectrums obtained from the discrete Fourier transform time domain data matrices $\mathbf{x}^{(m)}$ and $\hat{\mathbf{x}}^{(m)}$ as:

$$X^{(m)}[k] = \frac{1}{T} \sum_{n=0}^{N-1} x^{(m)}[n] e^{j2\pi \frac{k}{N} n}, \quad \hat{X}^{(m)}[k] = \frac{1}{T} \sum_{n=0}^{N-1} \hat{x}^{(m)}[n] e^{j2\pi \frac{k}{N} n} \tag{4.18}$$

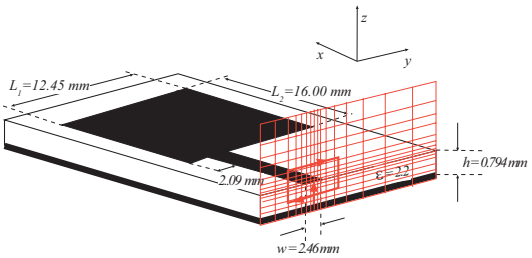
with $k = 0, 1, \ldots, N-1$ denoting the discrete frequency index. For this case this frequency range is bounded by the two values f_{min} and f_{max}, and the MSE computed on the corresponding discrete spectrum range is

$$MSE_f^{(m)} = \frac{1}{(k_{max} - k_{min})} \sum_{k=k_{min}}^{k_{max}} \left| X^{(m)}[k] - \hat{X}^{(m)}[k] \right|^2, \tag{4.19}$$

where $[k_{min},\ k_{max}] = [\, f_{min} NT,\ f_{max} NT, \,]$ is the frequency range, with T denoting the sampling interval.
The two vectors $MSE_t^{(m)}$ and $MSE_f^{(m)}$ are compared with the arbitrarily fixed values of accuracy, E_t and E_f. These quantities can vary separately and typically values below 10^{-4} are satisfactory for both cases. If at this point all responses satisfy the accuracy requirement in terms of MSE, in time and/or in frequency domain, then

$$\max\left(MSE_t^{(m)}\right) \leq E_t, \quad \text{and/or} \quad \max\left(MSE_f^{(m)}\right) \leq E_f. \tag{4.20}$$

When these requirements are met then the model is accurate enough and the current m-th simulation can be disrupted (stop command) and the next one started until $m = M$. If however (4.20) is not satisfied then a refining of the present model parameters (which means the order, the poles and the residues complex values) is necessary. In order to do so



TLM Parameters		Values
$N_x \times N_y \times N_z$	=	$100 \times 165 \times 53$
dl	=	0.2647 mm
dt	=	$2.18\ 10^{-13}\ sec$

Figure 4.10: MS Patch antenna and relative TLM model parameters

more information about the network response must be collected, extending the originally considered data matrix $\mathbf{x}^{(m)}$ (enabling command) and adding further N_B samples, and then the entire procedure must be repeated. N_B is normally a fixed integer quantity arbitrarily chosen and dependent on the speed of the TD solver in providing simulated data. The PMSI algorithm applied for modelling and prediction is summarized in Tab. 4.1.

4.3.2 Application of the method: Examples of prediction by PMSI Algorithm

The PMSI algorithm will now be applied and tested on some practical applications such as planar (miscrostrip) and three-dimensional structures using the same workstation (a PA8000 Risk based CPU 360MHz). The following values are used for the algorithm's parameters: the length of the data sample block is $N_B = 1024$, the SVD based model order selection a threshold is $D_T = 100dB$ (see section 3.4.3), for the N_0 a value of 85%, of the current simulated response length, is chosen, and the accuracy thresholds E_t and E_f have been definet at 10^{-4}. These values along with their description are listed in Tab. 4.2.

Rectangular microstrip patch antenna

The first structure is a rectangular microstrip patch antenna fabricated on alumina substrate and fed by 50 Ω microstrip transmission line [16]. In Fig. 4.10 the antenna design and specification are given. The structure has been modelled as a one port network ($M = 1$), and the frequency range of interest is from zero, (f_{min}) to $30\,GHz$ (f_{max}). The metallization thickness has been neglected and the the structure has been enclosed in a box of perfect matched layers [62], except at the bottom where the structure is bounded by a PEC, to realize the microstrip ground plane. The memory occupation needed for the chosen TLM

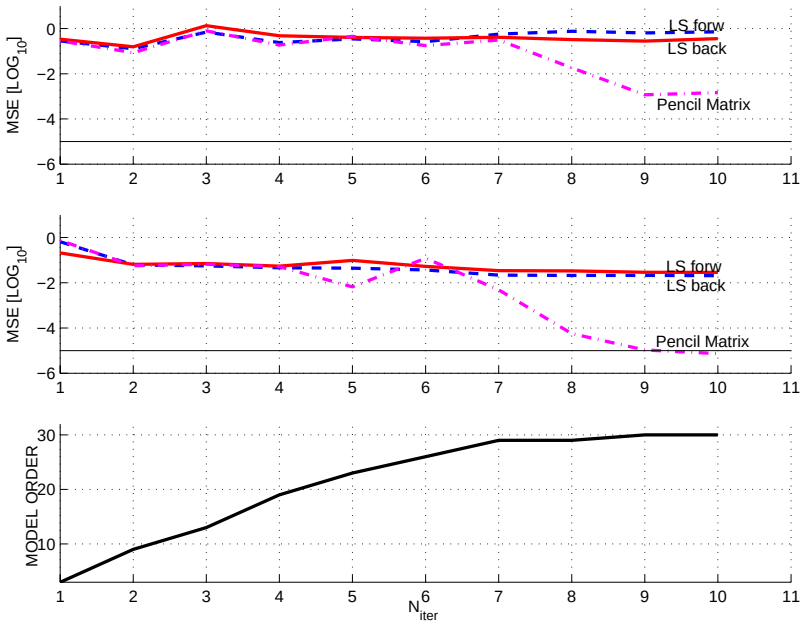


Figure 4.11: Performance of the PMSI algorithm as function of the TLM iteration for the patch antenna. The spectrum of the PM model response reaches acceptable accuracy when the model order converges to $P = 30$. The LS linear prediction models seems not to improve at all for the same amount of data samples (10 iterations).

mesh is approximately 63 Mb. The overall length of the time window originally analyzed is $N_{tot} = 30000$ timesteps, for an overall elapsed simulation time of about $4\,h\ 56'\ 14''$. This duration has been estimated so that the amplitude of the response was required to fall below the 0.001% of the maximum value at the last timestep. Applying the PMSI algorithm, the model order converges to $P = 30$ after only 9 simulation iterations of 1024 samples each. As shown in Fig. 4.11, the model obtained by LS linear prediction techniques, both back and forward, seems to improve very slowly with the progression of the simulation, while on the other hand, the model based on the PM method improves very quickly, and as soon as the model order is found, it reaches here the spectrum accuracy requirements (convergence). A smaller improvement of the same model is observed in the time domain since it is more likely to be effected by high frequency poles which are neglected in the SI analysis range. In order to quantify the efficiency of the algorithm the amount of samples from which the final model has been obtained is given by $N_{pre} = N_{iter} * N_B$ and must be compared with the length of the response when simulated without any parallel modelling

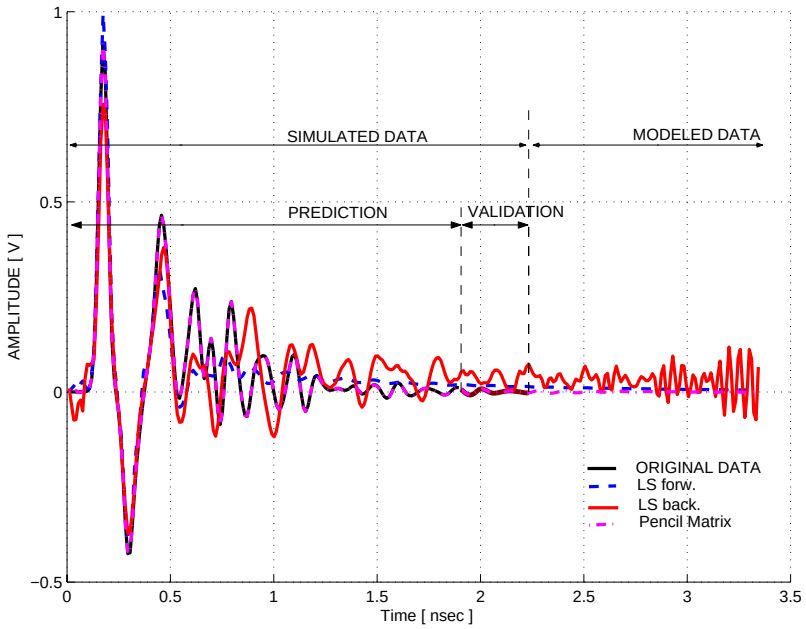


Figure 4.12: Comparison of the models with the original simulated response in the time domain for the the patch antenna

given by N_{tot}. This quantity may be expressed by the following relationship

$$\eta = 1 - \frac{N_{iter}\, N_B}{N_{tot}} \tag{4.21}$$

which estimates the percentage of saved time, while maintaining the same accuracy as achieved by the longer simulation. In the present example $\eta = 53.3\%$ has been calculated. This indicates that the simulation time has been reduced by more than half using the PMSI algorithm. The result of the modelling can be seen in the time domain in Fig. 4.12. The prediction and, the validation ranges are both highlighted. Although the PM based model has an MSE_t of around 10^{-3} (-3 in the log scale of Fig. 4.11), it agrees very well with the simulated response. Another important observation is that the LS linear prediction models, as clearly shown from the LS backward model results, may become unstable for an increasing number of poles, which explain the lack of improvement in the figure. The comparison of the spectrum for the different models and the original simulated ones are depicted in Fig. 4.13, confirms the very good performance of the PM model, which is almost indistinguishable from the simulated spectrum, in the entire analyzed frequency range from zero up to 20GHz. If the poles estimated with the different methods, are

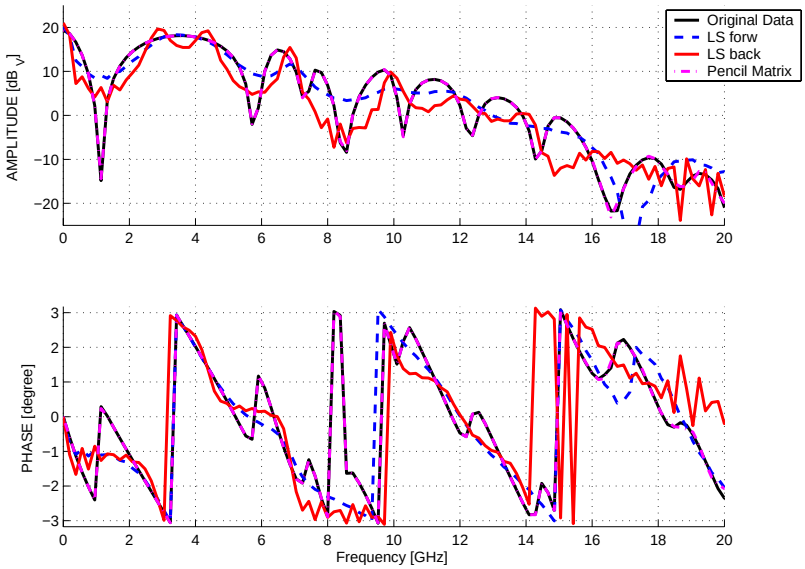


Figure 4.13: Comparison of the model with the spectrum of the original simulated data, for the patch antenna

depicted in the complex plane as shown in Fig. 4.14 the instability of the LS backward algorithm becomes clear, as six of the 30 poles lie outside the unit circle. Useful results of the modelling by means of PMSI are given in terms of scattering parameters in Fig. 4.15, where a comparison with reference measured data is depicted.

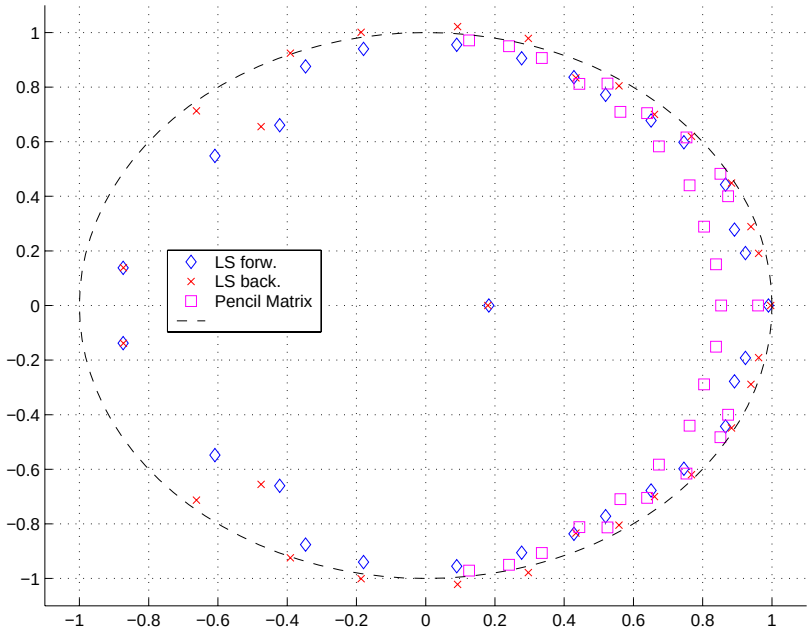


Figure 4.14: Pole positions in the Z–plane for the three models, for the patch antenna.

Table 4.1: PMSI Algorithm applied to modelling and prediction of TLM simulated responses. The final model parameters are the order P, the P residues matrices of dimension $(M \times M)$ in $\mathbf{C}$ and the $P \times 1$ poles (natural frequencies) vector $\mathbf{z}$.

Given the TLM model of the M-port network under test ($m = 1$, $P^{(0)} = 1$)
1. The m-th simulation (excitation at the m-th port) is started
2. The first N samples for the M network responses are generated and stored in $\mathbf{x}^{(m)}$ $(N \times M)$
2.1 $\mathbf{x}^{(m)}$ is separated into two matrices : $\mathbf{x}_p^{(m)}$ $(N_0 \times M)$ and $\mathbf{x}_v^{(m)}$ $(N - N_0 \times M)$
3. The initial network model order is computed:
3.1 The order for each response in $\mathbf{x}^{(m)}$ is computed (SVD) and stored in a vector $\mathbf{P}^{(m)}$ $(1 \times M)$;
3.2 The model order $P = P^{(m)}$ is computed by (4.10)
4. The P poles are estimated and clustered:
4.1 For each column (response) of $\mathbf{x}_p^{(m)}$ the P poles are estimated (PM or LS Covariance);
4.2 The poles are ordered and stored in $\mathbf{z}^{(m)}$ $(P \times M)$;
4.3 The M_i poles, on the i-th row of $\mathbf{z}^{(m)}$, closed to each other and within a given tolerance are selected;
4.4 The M_i poles are averaged (clustered) by (4.12), obtaining $z_i^{(m)}$;
4.5 The pole $z_i^{(m)}$ and $z_i^{(m-1)}$ are averaged by (4.13) obtaining z_i;
4.6 The value $z_i^{(m)}$ is updated to z_i;
5. The model is fit to the data via the LS methodand the residue matrix $\mathbf{C}^{(m)}$ is computed by (4.14);
6. The M recovered responses $\hat{\mathbf{x}}^{(m)}$ $(N \times M)$ are given by (4.16);
7. The model is validated based on the MSE between simulated and recovered data:
7.1 The $MSE_t^{(m)}$ and $MSE_f^{(m)}$ both $(1 \times M)$, are computed from (4.17) and (4.19);
7.2 If (4.20) is satisfied convergence is achieved and then the current (m-th) simulation may be stopped ;
7.3 If not, N_B samples are added to the M responses in $\mathbf{x}^{(m)}$ and the algorithm is repeated from 2.;
8. Check on whether all the network responses have been analyzed:
8.1 If $m < M$ the next simulation $m + 1$ is prepared and the algorithm repeated from 1. ;
8.2 If $m = M$ the analysis is completed.

Table 4.2: PMSI parameters (N is the number of samples currently being simulated).

Name	Typical Value	Description
N_B	1024	Number of samples (length) of each TLM iteration block
N_0	0.85 N	Number of simulated samples used for building the model
E_t	10^{-4}	Accuracy threshold for impulse response accuracy
E_f	10^{-4}	Accuracy threshold for spectrum accuracy
D_T	100 dB	Threshold for (SVD) model order selection (see section 3.4.3).

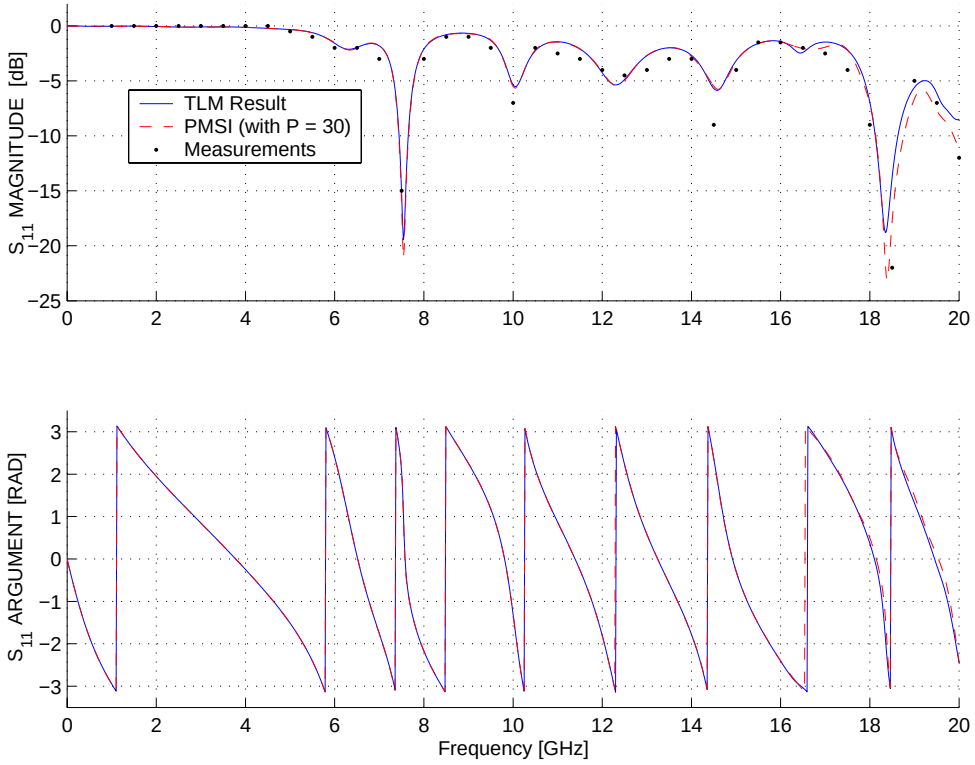


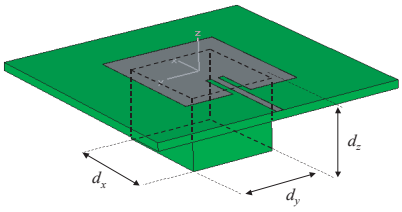
Figure 4.15: Comparison between scattering parameters obtained from the TLM simulation results without any modelling (—) and by applying the PMSI (- -), for the patch antenna. Measured results are depicted in the top diagram, for the magnitude of S_{11} by $(\cdots)$ [16].

Substrate Cavity SOI Resonator

The resonator considered hereby is obtained by forming a parallelepiped volume by means of deep–silicon etching of the 400 μm bulk side of a $400/0.3/47\ \mu m\ Si/Si0_2/Si$ wafer, followed by a metal deposition. On the other side of the wafer a microstrip transmission line of characteristic impedance 50 Ω is connected to a shielding metal patch (resonator lid) which is manufactured with by photolithography process. The result of this process is the device shown in Fig. 4.16 along with the TLM model specifications. Since the top metal patch is separated from the bottom cavity by just 47 μm, the structure behaves as a quasi–rectangular resonator filled with a silicon dielectric. By choosing the two cavity planar dimensions appropriately, the fundamental resonance mode can be arbitrarily established. For this case the TE_{011} mode is selected, and the corresponding resonance frequency may be determined by applying the semi–empirical formula

$$f_{res} = \frac{c_0}{2\sqrt{\varepsilon_{r\,Sil}}}\sqrt{\frac{1}{d_y^2} + \frac{1}{d_{z_{eff}}^2}} \tag{4.22}$$

where $\varepsilon_{r\,Sil}$ is the silicon dielectric constant whose value for this case is 11.9, d_y is one of the planar dimension for the cavity, and $d_{z_{eff}} = 550\ \mu m$ represents the effective vertical length obtained by a 20% extension (obtained empirically) of the physical substrate whose thickness is 450 μm. For the TE_{011} mode to be excited, the third dimension of the cavity must be smaller than the other two, i.e. $d_x < d_y$ and $d_x < d_{z_{eff}}$, where $d_x = 300\ \mu m$, for this structure. The magnetic coupling mechanism may be adjusted by varying the position of the connection between microstrip line with the metal patch. for this design, a choice



TLM Parameters		Values
$N_x \times N_y \times N_z$	=	$32 \times 66 \times 37$
dl	=	$10\ \mu m$
dt	=	$8.14\ 10^{-15}\ sec$

Figure 4.16: SOI cavity resonator and TLM model specifications.

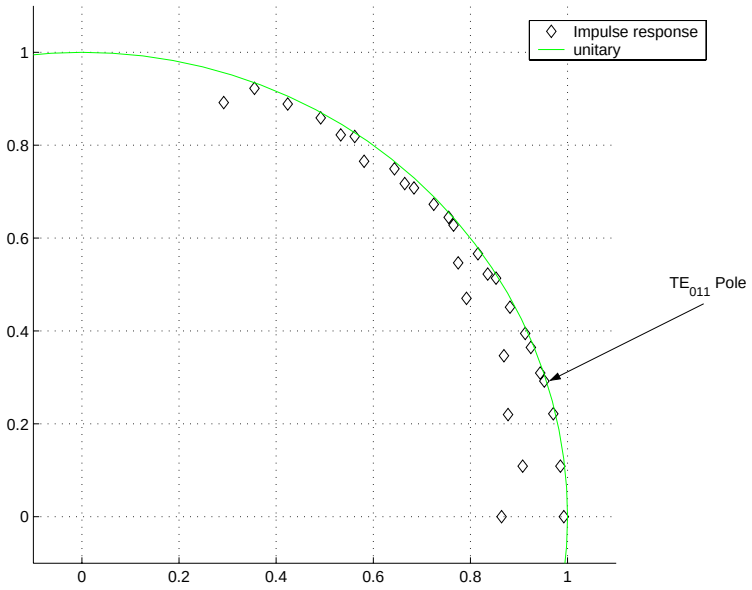


Figure 4.17: Poles distribution for the TLM impulse response of the SOI resonator. The pole corresponding to the fundamental mode $z_{(TE_{011})} = 0.9438 + j\ 0.3093$ is highlighted (frequency 126.5 GHz and damping factor $-0.31\ 10^{-11}$)

of $d_y = 450\ \mu m$, yields a resonator with fundamental resonance frequency of 124 GHz. The full–wave modelling of the fabricated one–port ($M = 1$) resonator has been done by using a TLM model whose parameters given in details in the table of Fig. 4.16. The simulation has been extended over a maximum number of 78934 time steps for an overall elapsed simulation time of $1h\ 0'\ 54''$ and a memory requirement of 5.62 Mb. The analyzed frequency range extends from zero up to 250 GHz (f_{max}). The TLM signal has been down sampled by a factor of 122 and processed by the PMSI algorithm, saving approximately 62% (by using (4.21)) of the overall simulation run time. The estimated pole distribution is shown in Fig. 4.17 where the pole corresponding to the fundamental resonant mode is highlighted. As shown in Fig. 4.18, the model order converges to a final value of $P = 58$, as the error MSE_F for the spectrum goes below the threshold value $E_f = 10^{-4}$. The top graph in Fig. 4.18 shows that MSE_T (error in the time domain) does not improve and remain above 10^{-3}. The results of the analysis are shown in Fig. 4.19 where the resonator input reflection S_{11} is given up to 400 GHz. The extracted model copies very well the full–wave simulation result in the entire frequency band of analysis. The measured results available in the neighborhood of the fundamental resonance mode between 117 and 130 GHz shows less than a 3 GHz off–set with respect to the designed central frequency,

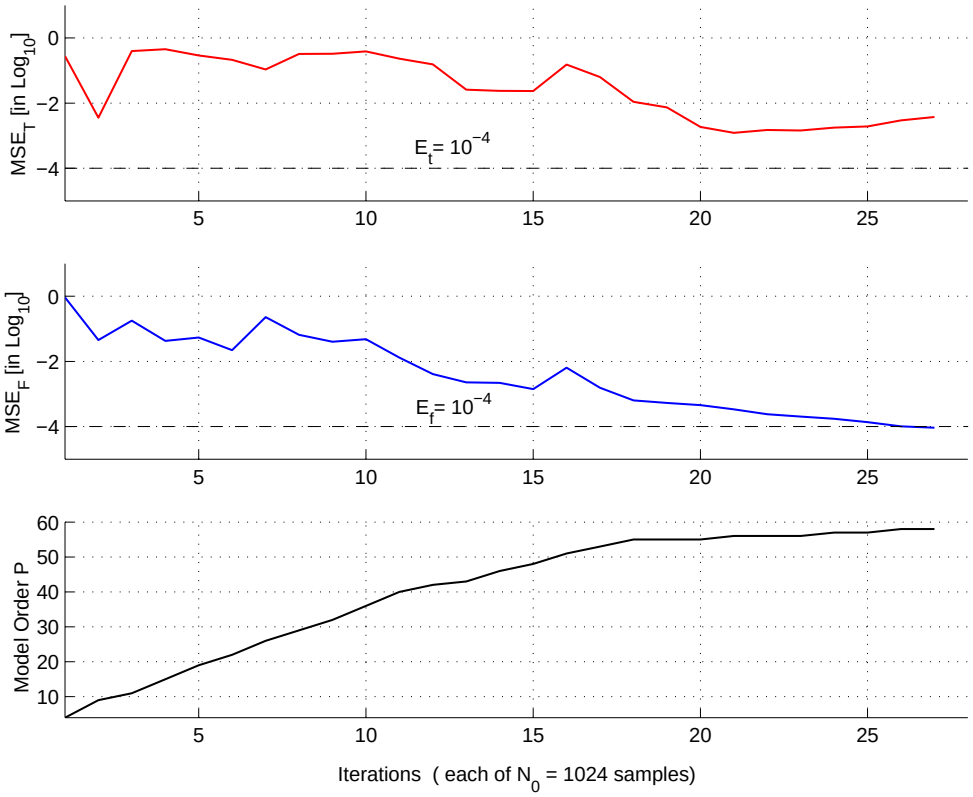


Figure 4.18: Results of the PMSI algorithm applied to the SOI resonator impulse response. From the top to the bottom graph the behavior of the MSE in time and frequency respectively and the estimated model order P, are depicted as a function of the number of iterations (number of simulated impulse response samples from the TLM simulation)

corresponding to an error of about 2.5% of shift. The other two resonances around 44 GHz and 90 GHz account for the metal patch resonant modes. A first guess of the quality factor, may be given by the theoretical formula which when applied to a metal parallelepiped cavity resonators, with metal losses only [72], yields a value of around 400. A more realistic value of 120 however has been obtained from the simulated results.

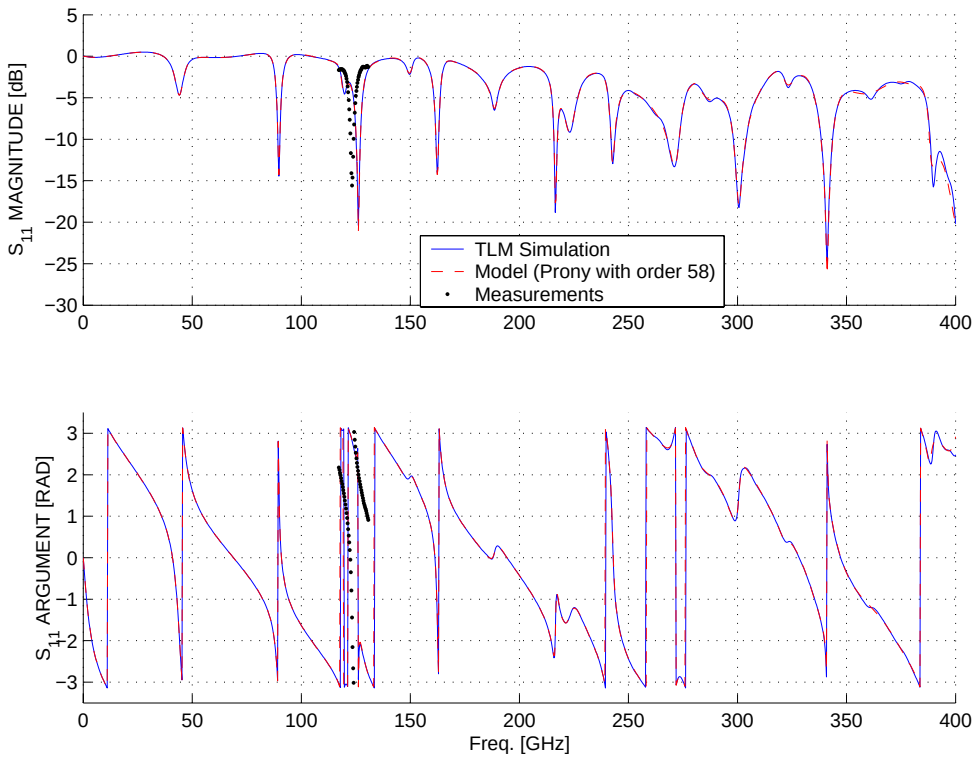
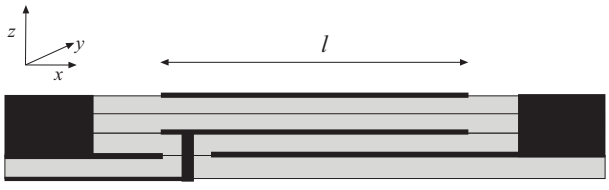


Figure 4.19: Input reflection coefficient S_{11} for the modelling of the SOI resonator.

Low Temperature Co–fired Ceramic (LTCC) coupled stacked patch antenna

The next structure to be analyzed is a broad band coupled patch antenna fabricated by means of LTCC process. Application of this antenna can include local multipoint distribution service (LMDS) systems. The antenna is realized by means of a four layered GL550 dielectric material with $\varepsilon_r = 5.6$ and a thickness of $h = 97\ \mu m$ and the square patches are placed on top of the last layer and between the second and third. This latter is connected by means of a via through a metal ground plane to the input microstrip line fabricated on



TLM Parameters		Values
$N_x \times N_y \times N_z$	=	$219 \times 125 \times 58$
dl	=	$15\ \mu m$
dt	=	$1.16^{-14}\ sec$

Figure 4.20: LTCC antenna cross section view and relevant TLM model parameters. Parameters are for a half structure (due to the symmetry)

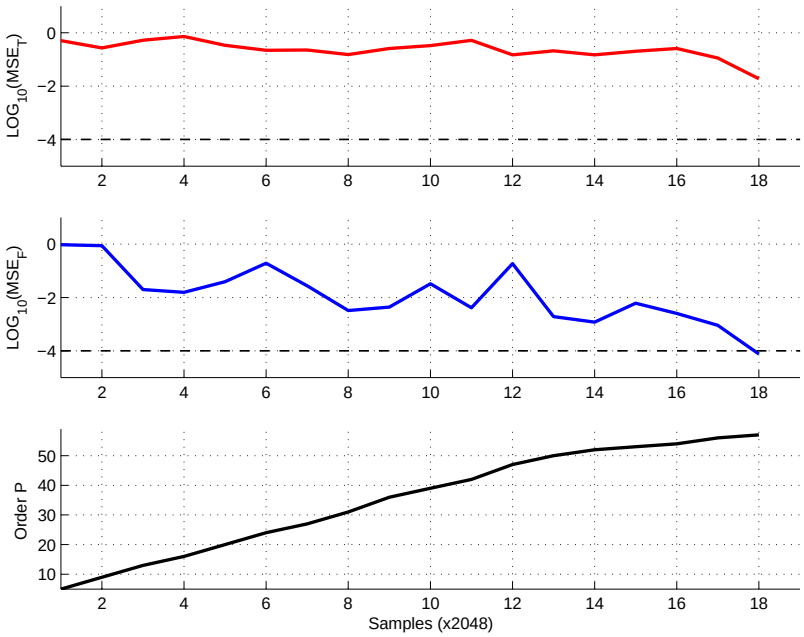


Figure 4.21: Performance of the PMSI algorithm as function of the TLM iteration for the LTCC antenna. The model need almost 85 % of the run time of the TLM simulated response and converges at $P = 58$ (the iterations here are chosen to be 2048 samples instead of 1024).

the other side of the multilayered structure. The metal patches are square with the length of each side $l = 1032\ \mu m$ and together with the rest of the metal connecting vias they are made of copper ($\sigma = 5.8\ 10^7\ S/m$). The antenna is depicted in Fig. 4.20 with the corresponding TLM model parameters. The structure has been modelled as one port network ($M = 1$), and the upper limit for the frequency range of interest is from zero up to $f_{max} = 50\,GHz$. Although only half of the structure has been modelled (since it is possible to define a symmetry PMC plane), the high aspect ratios of the structure (ratio between the largest planar dimension and the smallest layer thickness is around 135 for this structure) requires a TLM model with more than one million cells (see table in Fig. 4.20) with an overall memory requirement of around 115 Mb. The time window, originally analyzed using only the TLM method, without any SI prediction, spans over $N_{tot} = 50000$ timesteps, for an overall simulation run time of about 12 h 26′ 2″. The PMSI algorithm for this case yields a model order of $P = 58$ with a run time reduced to 9 h 26′ 59′. For this case the corresponding efficiency is of 24 %. This low value of reduction efficiency for the model prediction highlights that this structure has a very rich spectral contents. The spectrum for the port responses in terms of electric and magnetic fields are given in Fig. 4.22, where the

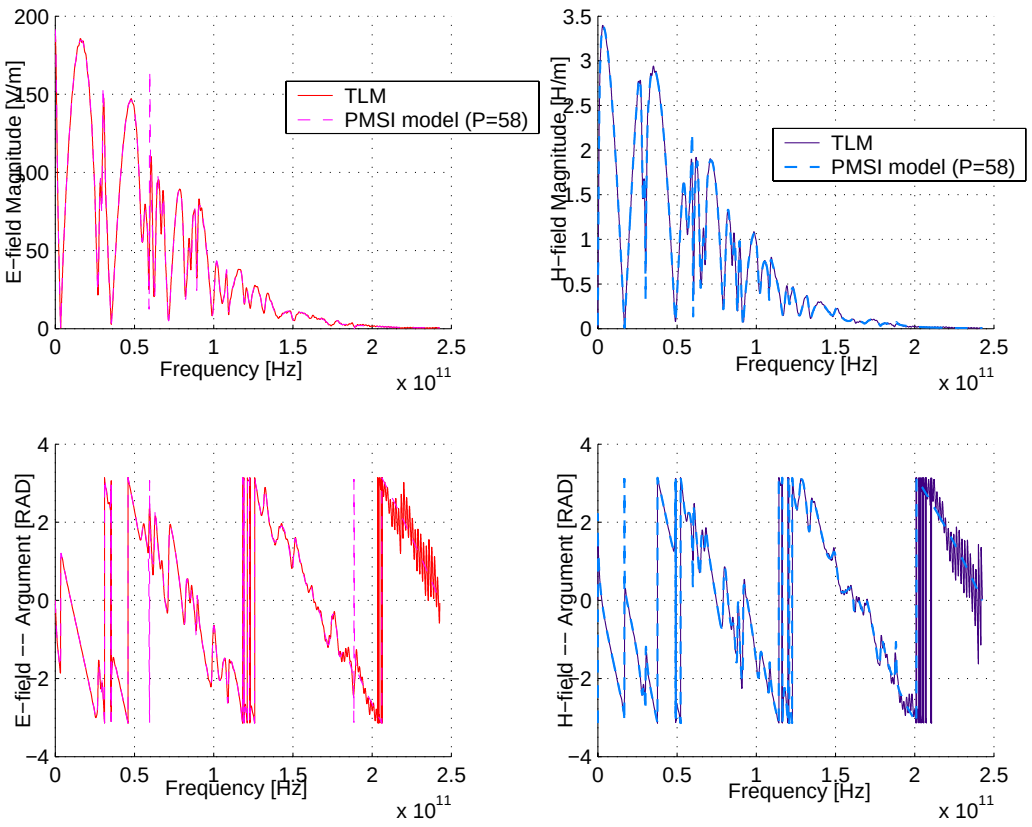


Figure 4.22: Comparison of the simulated and modelled spectrum for the E–field and H–field of the LTCC antenna

peaks of several resonances are observed over the 250 GHz of bandwidth taken in consideration. The results of the modelling can be furthermore observed in a comparison of the input reflection coefficient for the simulation and the model as shown in Fig. 4.23. As can be seen the results agree very well.

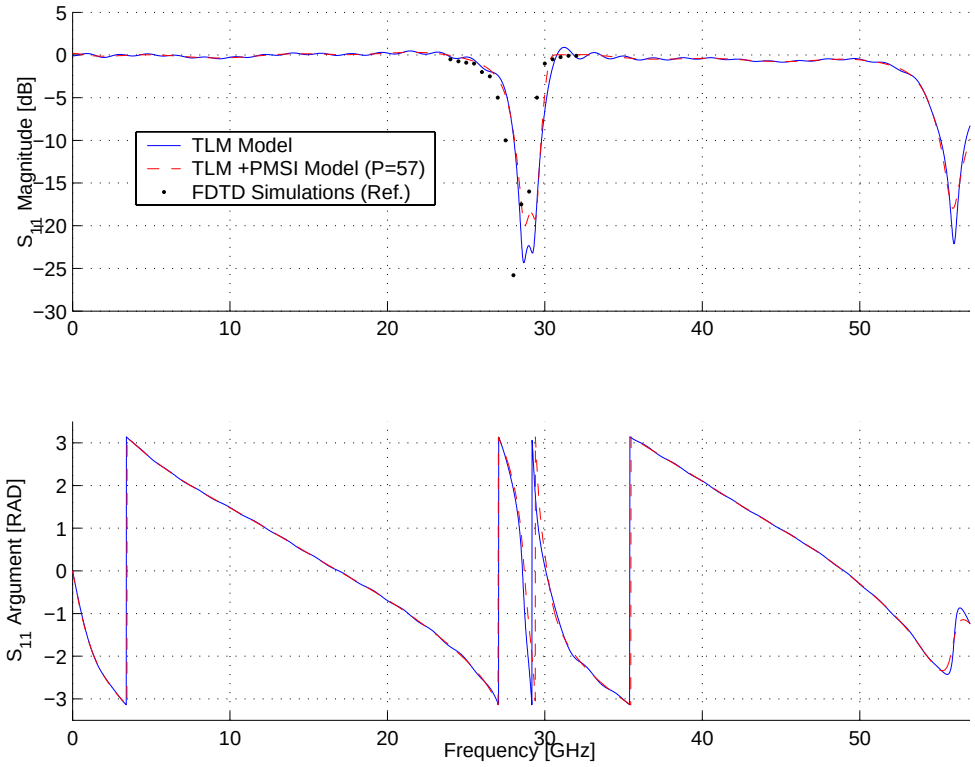


Figure 4.23: Comparison of the result for the input reflection of the LTCC antenna. TLM model with (- - -) and without PMSI (—) and available reference results from an FDTD simulation ($\cdots$) [67].

4.3.3 Comments on the PMSI algorithm applied to response prediction

The use of SI for estimation of model parameters has been shown to increase the performance of TD simulation tools by reducing the elapsed run time. This can be seen in Tab. 4.3, where elapsed run time and improvement efficiency for the application examples analyzed here are listed. The PMSI algorithm has also been shown to have a stable behavior and robustness. The model order is selected automatically and adaptively and it has been observed that it converges regularly once the accuracy in terms of the MSE in time and/or in frequency falls below the 10^{-4} threshold limit. Since, as commented in section 4.2, the driven response is located in the early part of the simulated data, the overall number of poles of the final model may be affected by the poles related to the input and this may result in an over dimensioned model.

Table 4.3: Summary of the PMSI Algorithm performance for the presented test structures.

Structure	TLM elapsed time	TLM and PMSI elapsed time	η
MS Antenna	4 h 56' 14"	$2\,h\ 37'\ 53''$	53.3%
SOI Resonator	$1h\ 0'\ 54''$	$0\,h\ 23'\ 8''$	62%
LTCC antenna	$12\,h\ 26'\ 2''$	$9\,h\ 26'\ 59'$	24 %

4.4 Systematic $Z-(Y-)$ matrix Foster Representation Modelling

The importance of deriving a synthetic description of a large and complex electromagnetic structure, for optimization and synthesis problems, is testified by the enormous amount of literature on this regard. Starting with the look–up table techniques, there are more sophisticated descriptions, by means of abstract parametric models as those based on space mapping (SM) [4] or the multidimensional adaptive parameter sampling (MAPS) method [33], and the artificial neural network (ANN) technique. All of these approaches are very abstract and lack of physical meaning in terms of electrical parameters (SPICE like model). The synthesis of equivalent circuits of passive networks from full-wave simulations however, has been widely investigated and provides SPICE compatible models. The most direct equivalent circuit synthesis, adopted in several CAD tools, exploits the interpolation of the experimental scattering parameters by means of transmission line and RLC lumped element networks of a given topology, typically of π or τ configuration. Although this technique yields very compact circuits, it is accurate within a limited frequency range, which may be extended at the cost of further lumped elements. Moreover it has the drawback of not being automatic as it does not guarantee the convergence for an arbitrary number of ports. A more general and well known technique is the partial elements equivalent circuit (PEEC), where an integral equation is used to establishe the electrical description of the physical geometry [69] [70] [23] [38] [75]. In the PEEC, the number of circuit nodes is related to geometrical complexity and, therefore, typically very large. This same problem is faced in another approach called the compact equivalent circuit (CEC) method, also based on the quasi–static mixed potential integral equation, but with the number of nodes depending on the electrical size rather than on the geometrical complexity [50]. The CEC however, lacks accuracy for electrically large problems and it is therefore restricted in frequency. A further class of techniques for the automatic synthesis of equivalent circuits are those based on the

Foster equivalent representation (EFR), using multipole expansion of the **Z** or **Y**–matrices [22]. These techniques are based on the estimation of the network **Z** or **Y**–matrix poles or network natural frequencies either directly in the frequency domain, or from the time domain simulated responses, as in [74] [46]. The method has drawbacks such as cumbersome representation, realizability conditions that are difficult to enforce, it relies on accurate pole estimation techniques, and lacks a direct correspondence between the extracted lumped elements and the physical structure. However, it has the enormous advantages of being a fully automatic procedure, which systematically applies to any arbitrary lossy multiport. Indeed if combined with a three dimensional TD solver such as TLM or FDTD, it may be extended to structures of arbitrary geometries over very broad frequency ranges. In the following the SI techniques introduced previously, will be applied to build the Foster representation for the impedance and admittance matrices, **Z** and **Y**, as given in (2.45) and (2.50).

4.4.1 The ZY–SI Algorithm

From knowledge of the impulse response, the structure under test may be entirely characterized by means of the scattering parameters (see section 4.1). These parameters are known in both the time and the frequency domain, over a typically very large frequency range. Moreover, the numerical model provides the access port characteristic impedances also. Hence the impedance and admittance parameters are easily computed by using the known transformations [72]. Once the impedance (admittance) parameters are known in the frequency domain, the identification of the Foster representation is equivalent to the identification of the corresponding impedance (admittance) Prony‘s model. This operation may be carried out directly on the spectrum by means of frequency domain SI methods [44], or by exploiting the PMSI algorithm (see section 4.3.1) as will be done here. In the latter case, since the PMSI works on time–domain signals, the impedance (admittance) matrix parameters must be obtained by inverse Fourier transform (here the Inverse Fast Fourier Transformation - IFFT is used). At this point, the PMSI algorithm may be applied by carrying out the identification of the impedance (admittance) natural frequencies and residues of (2.45) and (2.50). The knowledge of these parameters, i.e. the knowledge of the Prony model or equivalently, of the Foster representation in the complex Laplace domain, enables, as was shown in chapter 2, one to derive the lumped element equivalent circuit in the Foster canonical form. A flowchart diagram for this algorithm is shown in Fig. 4.24, and a detailed desciption is given in Tab. 4.4. The advantages of this approach are the systematic and completely automatic nature of the procedure, and the generation of

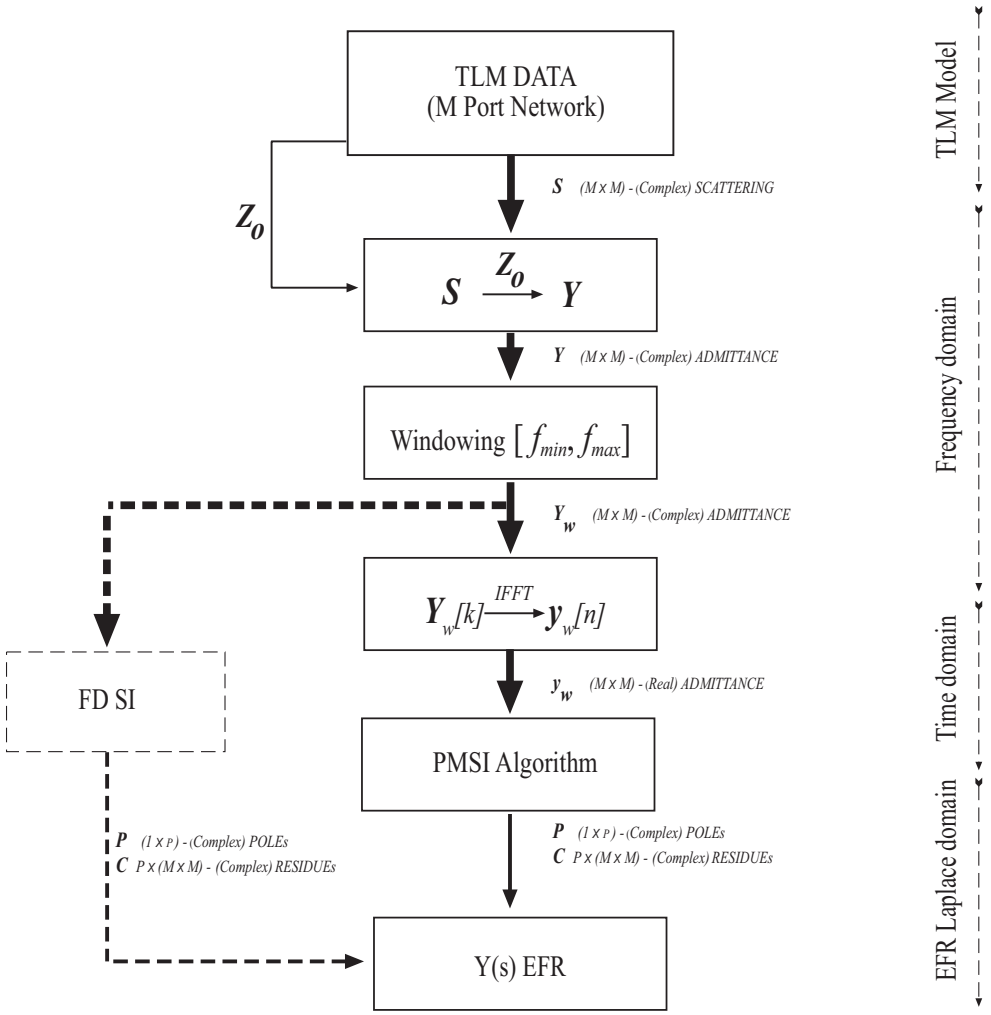


Figure 4.24: Flow chart for the modelling of the admittance impedance as EFR.

the Foster representations of the network impedance (admittance) matrix, directly from the TLM (or FDTD) simulated network responses. The windowing between f_{min} and f_{max} is necessary in order to filter out undesired portions of the spectrum where numerical noise and other unwanted spurious components may appear. The procedure described above is now applied to the rectangular microstrip patch antenna of Fig. 4.10. The results in terms of an equivalent Foster Representation of the admittance parameters are given in Fig. 4.25. The discrete frequency domain behavior for the admittance is directly obtained as the cross section at $|z| = 1$ of the most general expression $Y_{11}(z)$ depicted over the complex frequencies plane in Fig. 4.26. This latter consists exactly in the EFR for the input impedance. The

Table 4.4: Modelling procedure for the automatic extraction of the equivalent Foster representation (EFR) form full–wave numerical results. (* the poles are here considered in $P/2$ conjugate pairs).

	Given the complete set of M TLM impulse responses of the M-port network under test
1.	$\mathbf{S}$ $M \times M$ complex matrix and the M real port–impedances $\mathbf{Z}_0$ are computed;
2.	$\mathbf{Y}$ ($\mathbf{Z}$) $M \times M$ complex matrix is computed in the frequency–domain from $\mathbf{S}$ and $\mathbf{Z}_0$;
3.	The $\mathbf{Y}$ ($\mathbf{Z}$) matrix is windowed (filtered) between [f_{min}, f_{max}] obtaining the $M \times M$ complex matrix $\mathbf{Y}_w$;
4.	$\mathbf{Y}_w$ is inverse fourier transformed (IFFT), obtaining the time–domain $M \times M$ real matrix $\mathbf{y}_w$;
5.	The PMSI algorithm (see section 4.3.1) is applied to $\mathbf{y}_w$;
6.	The $P \times 1$ complex poles* $\mathbf{P}$ and the $P \times (M \times M)$ complex residues $\mathbf{C}$ for the EFR of $\mathbf{Y}$ ($\mathbf{Z}$) are now available;

model parameters obtained from the PMSI algorithm could be theoretically directly used for extracting the lumped elements of the corresponding Foster canonical equivalent circuit as shown in Fig. 4.27. The SPICE like model as given in Fig. 2.7 for the admittance and in Fig. 2.5 for the impedance can be generated by computing the 4 vectors $\mathbf{R}$, $\mathbf{L}$, $\mathbf{G}$ and $\mathbf{C}$ of $P/2 \times 1$ real lumped element values, directly from the complex poles $\mathbf{P}$ via (2.76) to (2.80) and (2.61) to (2.65) respectively. The matrix $\mathbf{n}$ of the P $(M \times M)$ matrices of real turns ratios are computed from $\mathbf{C}$ via the (2.82) and (2.66).
The use of the Foster lumped element equivalent circuit deserves at this point a comment. Indeed the corresponding representation is based on the realizability condition expressed in 2.83 and 2.84. The realizability condition state essentially that each partial fraction expansion component of the Foster representation must be real positive (i.e. positive with real part for an argument which has positive real part). This ensure that the lossy components (resistance and conductance) are positive too. Although for the entire impedance or admittance representation of passive structures this condition is normally satisfied, in general it is not satisfied for each partial fraction expansion component. Therefore may be concluded that the stability of the poles it is not a sufficient condition to enforce the realizability as Foster equivalent lumped element circuit [24]. Nevertheless other representation as the Cauer, may exist and be realizable (see Fig. 4.27) [24].

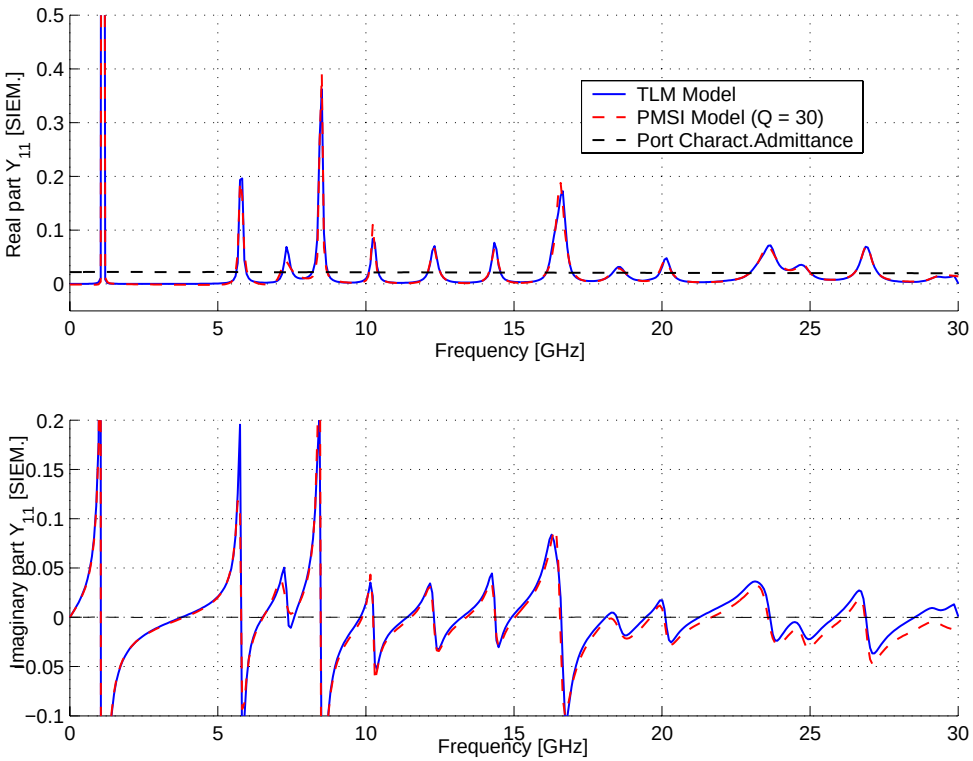


Figure 4.25: Comparison between the admittance parameters of the antenna ($M = 1$) for the Y_{11} obtained directly from the TLM simulation and from the PMSI model with 30 poles. The input port admittance is moreover depicted (- -).

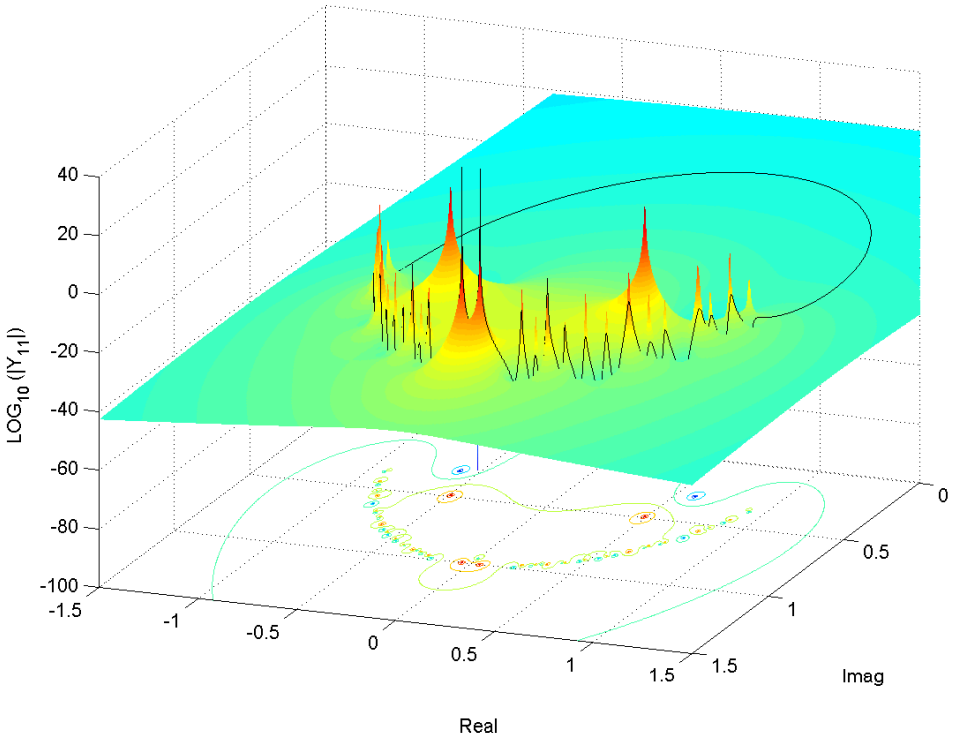


Figure 4.26: Behavior of then EFR for the admittance $|Y_{11}|$ in the z-domain. The cross section with the unit circle $z = e^{J\omega}$ corresponding to the spectrum is shown in black.

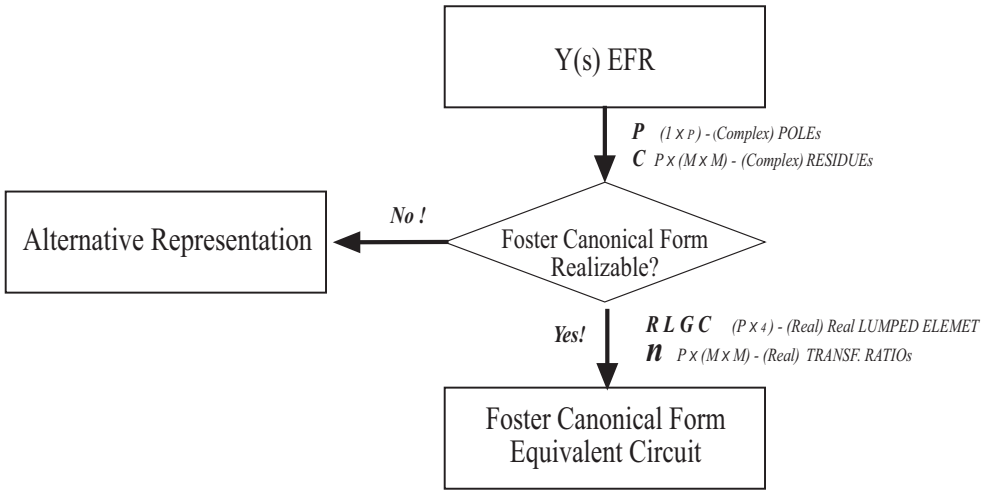


Figure 4.27: Flow chart for the modelling of the admittance impedance as EFR.

Chapter 5

Conclusions

Network oriented modelling (NOM) allows the analysis of complex electromagnetic structures by partitioning them into subregions, which can be described by parametric models. These parameters are strictly related to the structure‘s physical geometries and material properties. A parametric description of an electromagnetic structure has many advantages over more cumbersome models such as full-wave models. Although the latter are sometimes necessary for very accurate analysis of three-dimensional structures, the availability of parametric models enable compact descriptions, which can be further simplified by applying model reduction techniques. The model chosen in the present work is a pole and residue series or Prony model. In this work the Prony’s model has been preferred among others (i.e. AR or ARMA models) as it is perfectly equivalent to the eigenvalue expansion of the dyadic Green‘s function, derived by applying the Method of Moments (MoM). The key parameters of this model are complex quantities called the structure natural frequencies. Their time–domain behavior can be observed in the impulse response of a given passive electromagnetic structure as exponentially damped oscillating components. The number of these natural frequencies, defines model order, and this number may be systematically found together with the complex frequencies values by means of SI techniques. The combination of time-domain full-wave analysis with SI techniques has allowed the construction of an algorithm, here referred to as Prony‘s model based SI or PMSI. This is a signal processing module, able to perform on–the–fly analysis of TLM (or FDTD) simulated impulse responses. The algorithm runs in parallel with the time-domain simulation and autonomously refines the model, by adaptively improving the accuracy of the parameters and model order of the parameters. Numerical responses obtained from TLM

simulations are typically characterized by high SNR levels (above 100 dB), and their description represents a typical problem known in SI as output error (OE) modelling. In spite of this, even a very low level of noise (in these cases mainly due to finite numerical precision), may affect the reliability of the present methodology by strongly interfering with the parameter estimation methods. Pole estimation represents the most critical operation and is carried out by the pencil matrix (PM) method. This methos is preferred over other methods such as the least square (LS) linear prediction (LS-LP) technique, due to its efficiency and robustness with respect to the noise level. The model order, which determines the model complexity, is selected by exploiting the singular value decomposition (SVD) technique. SVD combined with a threshold decision criteria, allows one separation of the signal space from the underlying noise space. The dimension of the signal space corresponds to the sought model order. This approach has demonstrated good convergence characteristics within a given frequency bandwidth. Once the number and the value of the poles are known the response is recovered by applying the Prony‘s model, and computing the residues, by least square fitting of the simulated response. The accuracy of the resulting model is validated by calculating the mean square error (MSE) in both, time and frequency-domain. The first is computed over a validation time window (typically taken as the last 15% of the available response), and the second computed over the spectrum of interest. Although the modelling is performed in the time-domain, the desired accuracy may be more quickly reached in the spectral domain, especially in case of narrowband modelling (narrow band analysis). As soon as the model recovered response reaches the accuracy requirements, the algorithm may be stopped and the full-wave analysis disrupted.

When applied to the TLM simulated impulse responses of different electromagnetic structures such as antennas and resonators, the algorithm has allowed a significant reduction of the simulation run time. In particular, a 24% and a 53% simulation time saving have been observed for an LTCC and a microstrip antenna, respectively. Structures with higher quality factors, such as the SOI cavity resonator presented here, have shown a further time savings, achieving a reduction of 62% in simulation time.

Beside the network response prediction, the PMSI algorithm has been successfully applied to modelling of the network itself. By means of it, the corresponding equivalent Foster representation (EFR) can be generated from the time-domain network impedance and admittance parameters. Indeed since the EFR of a passive network is equivalent to a Prony‘s model description, the PMSI algorithm becomes a very straight forward tool for deriving it. An example of this is the modelling of the input admittance for the above mentioned microstrip antenna. The resulting model has been applied out up to a frequency of 30 GHz,

and the spectrum, when compared with the simulated results, shows very good agreement. The presented approach enables a very compact and general parametric description of arbitrary electromagnetic structures in the complex frequency plain. Once the range of accuracy and frequency bandwidth have been specified, it yields the model in a completely automatic and independent fashion.

Possible limits are represented by very lossy structures whose impulse responses are strongly attenuated and masked by noise. The identification of the corresponding natural frequencies may become in this case a very difficult task. This problem is more evident for multiport networks where a pole corresponding to a specific natural frequency can be erroneously estimated, at different ports, through different numerical values, yielding an incorrect final model order and thereby compromising the entire model. In order to avoid this possible identification flaw, more advanced pole clustering tools should be investigated. Another limit to the method presented may be in the lumped element equivalent circuit based on the Foster representation of the impedance (admittance) parameters. The procedure yields a relationship between the simulated data responses and the final lumped element values, but does not relate these latter to the properties (geometry and material) of the original electromagnetic structure. While this procedure works in a systematic fashion in one direction (from physical structure to model parameters and equivalent circuit elements), it does not allow reverse synthesis (from the set of parameters to the physical structure). This means that it is not helpful, where this reverse synthesis is required, as for instance for optimization purposes. For further development of this work, the relationship between the estimated model parameters and the originating electromagnetic phenomena (as resonating fields or currents), can be investigated. Time-domain methods, such as the TLM and FDTD, do indeed provide, local field solutions all over the modelled spatial domain, and in a given time window. This information, normally lost after performing the sampling of the responses at the access ports, may be exploited for determining a closed form relationship between the complex pole values and the physical characteristics of the structure component responsible for it. Something similar is already been done in target identification problems, where the natural frequency values are strictly related to the dimensions and the material properties of the scatterer, while the amplitude of the response depends on the observer position and the intermediate media.

Appendix A

Exterior Differential forms

Scalar and vector fields may be represented by exterior differential forms, which are essentially the expression under an integration symbol. In order to describe this differential notation, it is shown in Table A.1 along with the corresponding vector notation for some of the most relevant electromagnetic laws and relations.

A complete and detailed treatment of the differential forms can be found in [72].

In a three dimensional curvilinear coordinate system with coordinates $\boldsymbol{\rho}$=(u, v, w), there are four differential forms given by

$$\begin{aligned}
\text{zero-form } (scalar): &\quad \varphi(\boldsymbol{\rho})\,, \\
\text{one-form } (vector): &\quad \underline{\mathcal{E}}(\boldsymbol{\rho}) = E_u\, s_1 + E_v\, s_2 + E_w\, s_3\,, \\
\text{two-form } (pseudovector): &\quad \underline{\mathcal{D}}(\boldsymbol{\rho}) = D_u\, s_2 \wedge s_3 + D_v\, s_3 \wedge s_1 + D_w\, s_1 \wedge s_2\,, \\
\text{three-form } (pseudoscalar): &\quad \underline{\mathcal{Q}}(\boldsymbol{\rho}) = Q\, s_1 \wedge s_2 \wedge s_3\,. \qquad \text{(A.1)}
\end{aligned}$$

whereas s_1, s_2 and s_3 are referred to as the unit forms. The double one form used for the dyadic Green's function is defined as follows:

$$\begin{aligned}
\underline{\mathcal{G}}(\boldsymbol{\rho}, \boldsymbol{\rho}') \;=\; & G_{uu}\, s_1\, s_1 + G_{vu}\, s_2\, s_1 + G_{wu}\, s_3\, s_1 + \\
& + G_{uv}\, s_1\, s_2 + G_{vv}\, s_2\, s_2 + G_{wu}\, s_3\, s_2 + \qquad \text{(A.2)} \\
& + G_{uw}\, s_1\, s_3 + G_{vw}\, s_2\, s_3 + G_{ww}\, s_3\, s_3, \\
& \qquad \text{(A.3)}
\end{aligned}$$

Table A.1: Correspondences between the main electromagnetic equations in conventional vector notation and exterior differential form notation in the Laplace domain

Equation	*Vector Notation*	*Differential Form Notation*
Ampère's law	$\text{curl}\,\underline{\boldsymbol{H}} = s\,\underline{\boldsymbol{D}} + \underline{\boldsymbol{J}}$	$d\underline{\mathcal{H}} = s\underline{\mathcal{D}} + \underline{\mathcal{J}}$
Faraday's law	$\text{curl}\,\underline{\boldsymbol{E}} = -s\,\underline{\boldsymbol{B}}$	$d\underline{\mathcal{E}} = -s\,\underline{\mathcal{B}}$
Gauss's law	$\text{div}\,\underline{\boldsymbol{D}} = \rho$	$d\underline{\mathcal{D}} = \underline{\mathcal{Q}}$
Magnetic flux continuity	$\text{div}\,\underline{\boldsymbol{B}} = 0$	$d\underline{\mathcal{B}} = 0$
Electric constitutive equation	$\underline{\boldsymbol{D}} = \underline{\underline{\varepsilon}}\,\underline{\boldsymbol{E}}$	$\underline{\mathcal{D}} = \star\underline{\underline{\varepsilon}}\,\underline{\mathcal{E}}$
Magnetic constitutive equation	$\underline{\boldsymbol{B}} = \underline{\underline{\mu}}\,\underline{\boldsymbol{H}}$	$\underline{\mathcal{B}} = \star\underline{\underline{\mu}}\,\underline{\mathcal{H}}$
Electric boundary condition	$\boldsymbol{n} \times (\underline{\boldsymbol{E}}_2 - \underline{\boldsymbol{E}}_1) = -\underline{\boldsymbol{J}}_{mA}$	$n \lrcorner n \wedge (\underline{\mathcal{E}}_2 - \underline{\mathcal{E}}_1) = -\underline{\mathcal{J}}_{mA}$
Magnetic boundary condition	$\boldsymbol{n} \times (\underline{\boldsymbol{H}}_2 - \underline{\boldsymbol{H}}_1) = \underline{\boldsymbol{J}}_{eA}$	$n \lrcorner n \wedge (\underline{\mathcal{H}}_2 - \underline{\mathcal{H}}_1) = \underline{\mathcal{J}}_{eA}$
Electric flux boundary condition	$\boldsymbol{n} \cdot (\underline{\boldsymbol{D}}_2 - \underline{\boldsymbol{D}}_1) = \rho_s$	$n \lrcorner n \wedge (\underline{\mathcal{D}}_2 - \underline{\mathcal{D}}_1) = \underline{\mathcal{Q}}_A$
Magnetic flux boundary condition	$\boldsymbol{n} \cdot (\underline{\boldsymbol{B}}_2 - \underline{\boldsymbol{B}}_1) = 0$	$n \lrcorner n \wedge (\underline{\mathcal{B}}_2 - \underline{\mathcal{B}}_1) = 0$
Source power density	$d\underline{P}_s = -\frac{1}{2}(\underline{\boldsymbol{E}} \cdot \underline{\boldsymbol{J}}^*_{eA} + \underline{\boldsymbol{E}} \cdot \underline{\boldsymbol{J}}^*_{eA})$	$\underline{\mathcal{P}}_s = -\frac{1}{2}(\underline{\mathcal{E}} \wedge \underline{\mathcal{J}}^*_{eA} + \underline{\mathcal{H}} \wedge \underline{\mathcal{J}}^*_{mA})$
Power loss density	$d\underline{P}_L = \frac{\omega\underline{\underline{\varepsilon}}''}{2}\underline{\boldsymbol{E}} \cdot \underline{\boldsymbol{E}}^* + \frac{\omega\underline{\underline{\mu}}''}{2}\underline{\boldsymbol{H}} \cdot \underline{\boldsymbol{H}}^*$	$\underline{\mathcal{P}}_L = \frac{\omega\underline{\underline{\varepsilon}}''}{2}\underline{\mathcal{E}} \wedge \star\underline{\mathcal{E}}^* + \frac{\omega\underline{\underline{\mu}}''}{2}\underline{\mathcal{H}} \wedge \underline{\mathcal{H}}^*$
Electric energy density	$d\underline{W}_e = \frac{\underline{\underline{\varepsilon}}}{4}\underline{\boldsymbol{E}} \cdot \underline{\boldsymbol{E}}^*$	$\underline{\mathcal{W}}_e = \frac{\underline{\underline{\varepsilon}}}{4}\underline{\mathcal{E}} \wedge \underline{\mathcal{E}}^*$
Magnetic energy density	$d\underline{W}_m = \frac{\underline{\underline{\mu}}}{4}\underline{\boldsymbol{H}} \cdot \underline{\boldsymbol{H}}^*$	$\underline{\mathcal{W}}_m = \frac{\underline{\underline{\mu}}}{4}\underline{\mathcal{H}} \wedge \underline{\mathcal{H}}^*$
Complex Poynting's vector	$\underline{\boldsymbol{T}} = \frac{1}{2}\underline{\boldsymbol{E}} \times \underline{\boldsymbol{H}}^*$	$\underline{\mathcal{T}} = \frac{1}{2}\underline{\mathcal{E}} \wedge \underline{\mathcal{H}}^*$
Complex Poynting's theorem	$\underline{\boldsymbol{T}} \cdot \boldsymbol{n} dS = -2s(d\underline{W}_m - d\underline{W}_e) - d\underline{P}_L + d\underline{P}_0$	$d\underline{\mathcal{T}} = -2s(\underline{\mathcal{W}}_m - \underline{\mathcal{W}}_e) - \underline{\mathcal{P}}_L + \underline{\mathcal{P}}_0$

where G_{ij} with $i, j = u, v, w$, is the component relating the ith component of the field vector to the jth component of the source vector [17].

Appendix B

Least Square Problem

Consider the approximation $\hat{\mathbf{y}}[n]$ of the complex discrete sequence, $\mathbf{y}[n]$ by a linear combination of p complex sequences $X_1[n], X_2[n], \ldots, X_p[n]$ (in general non linear functions of n), such that

$$\hat{\mathbf{y}}[n] = \sum_{m=1}^{p} a_m X_p[n] \quad 0 \leq n \leq N-1. \tag{B.1}$$

One method to uniquely assign value to the unknown complex parameters $\mathbf{a} = [a_1, a_1, \ldots, a_p]$, is to minimize the real sum of square error:

$$\rho = \sum_{n=0}^{N} |\mathbf{y}[n] - \hat{\mathbf{y}}[n]|^2 = \sum_{n=0}^{N} |\boldsymbol{e}[n]|^2 \tag{B.2}$$

This problem is known as the least square (LS) approach and in matrix notation can be expressed as follows,

$$\left\{\hat{\mathbf{a}} \in \mathbb{C}^{\,p} \; s.t. \; |\mathbf{y} - \mathbf{X}\hat{\mathbf{a}}| \leq |\mathbf{y} - \mathbf{X}\mathbf{a}| \; \forall \mathbf{a} \in \mathbb{C}^{\,p}\right\} \tag{B.3}$$

where $\mathbf{X}$ is an $N \times p$ data matrix, $\mathbf{a}$ is an $p \times 1$ coefficients vector, and $\mathbf{y}$ and $\mathbf{e}$ are $N \times 1$ data vectors given respectively by:

$$\mathbf{X} = \begin{bmatrix} X_1[1] & \cdots & X_p[1] \\ \vdots & \cdots & \vdots \\ X_1[1] & \cdots & X_p[N] \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_p \end{bmatrix}. \tag{B.4}$$

Equation (B.2) can also be rewritten as:

$$[\mathbf{y}\ \mathbf{X}] \begin{bmatrix} 1 \\ -\mathbf{a} \end{bmatrix} = \mathbf{e} \tag{B.5}$$

This equation represents a system of N linear equations, for p unknown coefficients of $\mathbf{a}$, whose solution may be carried out in the LS sense by means of two well known approaches, the normal equation based technique, or the SVD based technique.

B.1 LS solution by Normal Equation

The square error ρ in (B.2) may be expressed as,

$$\rho = \mathbf{e}^H \mathbf{e} = \mathbf{y}^H\mathbf{y} - \mathbf{y}^H\mathbf{X}\mathbf{a} - \mathbf{a}^H\mathbf{X}^H\mathbf{y} + \mathbf{a}^H\mathbf{X}^H\mathbf{X}\mathbf{a} \tag{B.6}$$

in which the vector inner product $\mathbf{y}^H\mathbf{y}$ forms a scalar, the matrix vector product $\mathbf{X}^H\mathbf{y}$ form a $p \times 1$ column vector and the matrix product $\mathbf{X}^H\mathbf{X}$ forms a $p \times p$ squar matrix. The minimization of ρ may be obtained by isolating in (B.6) a matrix squared form as follows:

$$\begin{aligned} \rho &= \mathbf{y}^H\mathbf{y} - \mathbf{y}^H\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{y} + \\ &+ \left(\mathbf{X}^H\mathbf{y} - \mathbf{X}^H\mathbf{X}\mathbf{a}\right)^H \left(\mathbf{X}^H\mathbf{X}\right)^{-1} \left(\mathbf{X}^H\mathbf{y} - \mathbf{X}^H\mathbf{X}\mathbf{a}\right) \end{aligned} \tag{B.7}$$

where the last term is a positive quadratic form and it is the only part of (B.7) to be dependent on $\mathbf{a}$. To minimize ρ means to chose an $\mathbf{a} = \hat{\mathbf{a}}$ such that:

$$\mathbf{X}^H\mathbf{y} - \mathbf{X}^H\mathbf{X}\hat{\mathbf{a}} = \mathbf{0}_p \tag{B.8}$$

with $\mathbf{0}_p$ the $p \times 1$ all–zero vector. Solving (B.8) for $\hat{\mathbf{a}}$ gives:

$$\hat{\mathbf{a}} = \left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\mathbf{y}. \tag{B.9}$$

The same solution can be obtained by solving a problem of minimum, i.e. by equating to zero the derivative of ρ with respect to $\mathbf{a}$

$$\frac{\partial\rho}{\partial\mathbf{a}} = -2\mathbf{y}^H\mathbf{X} + 2\mathbf{X}^H\mathbf{X}\mathbf{a}\,|_{\mathbf{a}=\hat{\mathbf{a}}} = 0 \tag{B.10}$$

from which follows (B.9). after Substituting (B.8) and (B.9) into(B.7) gives the minimum value for the square error:

$$\rho_{min} = \mathbf{y}^H\mathbf{y} - \mathbf{y}^H\mathbf{X}\hat{\mathbf{a}}. \tag{B.11}$$

Equations (B.9) and (B.11) can be rewritten into a single matrix expression

$$\begin{bmatrix} \mathbf{y}^H\mathbf{y} & \mathbf{y}^H\mathbf{X} \\ \mathbf{X}^H\mathbf{y} & \mathbf{X}^H\mathbf{X} \end{bmatrix} \begin{bmatrix} 1 \\ -\hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \rho_{min} \\ \mathbf{0}_m \end{bmatrix} \tag{B.12}$$

called the normal equation associated to the LS problem (B.3).

B.2 LS solution by SVD

The solution of (B.5) can also be found by using the singular value decomposition approach based upon the following theorem.
Given an $m \times n$ complex value matrix $\mathbf{A}$ of rank k, then, there exists positive real numbers $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_k > 0$, called singular values of $\mathbf{A}$. This matrix can be decomposed into an $m \times m$ unitary matrix $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2 \ldots \mathbf{u}_m]$, an $n \times n$ unitary matrix $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_n]$, and

an $m \times n$ matrix Σ, such as:

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^H = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^H \tag{B.13}$$

where the $m \times n$ matrix Σ has the structure:

$$\Sigma = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{B.14}$$

and $\mathbf{D} = diag(\sigma_1, \sigma_2, \ldots, \sigma_k)$ is a $k \times k$ diagonal matrix. Noting that $\mathbf{A}^H = \mathbf{V}^H\Sigma\mathbf{U}$, $\mathbf{U}\mathbf{U}^H = \mathbf{I}_m$ and $\mathbf{V}^H\mathbf{V} = \mathbf{I}_n$, then,

$$\mathbf{A}^H\mathbf{A} = \mathbf{V}(\Sigma^H\Sigma)\mathbf{V}^H, \quad \mathbf{A}\mathbf{A}^H = \mathbf{U}(\Sigma\Sigma^H)\mathbf{U}^H \tag{B.15}$$

and

$$\mathbf{A}^H\mathbf{A}\mathbf{v}_i = \sigma_i^2\mathbf{v}_i, \quad \mathbf{A}\mathbf{A}^H\mathbf{u}_i = \sigma_i^2\mathbf{u}_i \quad \text{with} \quad 1 \leq i \leq k. \tag{B.16}$$

The matrices $\Sigma^H\Sigma$ and $\Sigma\Sigma^H$ are diagonal matrices of size $n \times n$ and $m \times m$ respectively, whose diagonal elements are σ_i^2 for $1 \leq i \leq k$. The matrices $\mathbf{A}^H\mathbf{A}$ and $\mathbf{A}\mathbf{A}^H$ are Hermitian matrices of size $n \times n$ and $m \times m$ respectively. From (B.16) is evident that the columns of $\mathbf{V}$ are the n unitary eigenvectors of $\mathbf{A}^H\mathbf{A}$ and the columns of $\mathbf{U}$ are the m unitary eigenvectors of $\mathbf{A}\mathbf{A}^H$. Both share the same eigenvalues σ_i^2 for $1 \leq i \leq k$, which can be obtained from the positive square roots of the non zero eigenvalues of $\mathbf{A}^H\mathbf{A}$ or $\mathbf{A}\mathbf{A}^H$. Using the unitary properties of $\mathbf{U}$, $\mathbf{V}$ and (B.13), it can also be shown that:

$$\begin{gathered} \mathbf{A}\mathbf{V} = \mathbf{U}\Sigma\mathbf{V}^H\mathbf{V} = \mathbf{U}\Sigma \\ \text{and} \quad \mathbf{U}^H\mathbf{A} = \mathbf{U}^H\mathbf{U}\Sigma\mathbf{V}^H = \Sigma\mathbf{V}^H \end{gathered} \tag{B.17}$$

or

$$\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i, \quad \mathbf{A}^H\mathbf{u}_i = \sigma_i\mathbf{v}_i \quad \text{with} \quad 1 \leq i \leq k \tag{B.18}$$

which gives a relationship between the column components of $\mathbf{u}_i$ and $\mathbf{v}_i$ by the common singular values σ_i.

The SVD of (B.13) can be used for calculating the Moore–Penrose pseudoinverse of a matrix $\mathbf{A}$ given as:

$$\mathbf{A}^{\#} = \mathbf{V}\Sigma^{\#}\mathbf{U}^{H} = \sum_{i=1}^{k} \sigma_i^{-1}\mathbf{v}_i\mathbf{u}_i^{H} \tag{B.19}$$

where $\Sigma^{\#}$ is an $n \times m$ matrix given by

$$\Sigma = \begin{bmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{B.20}$$

$\mathbf{A}^{\#}$ has dimensions $n \times m$ and provides the minimum norm LS solution for the minimization of the square error equation. This can be constructed from (B.5), obtaining:

$$|\mathbf{y} - \mathbf{X}\mathbf{a}|^2 = |\mathbf{e}|^2 \tag{B.21}$$

and solved in order to minimize $\mathbf{e}$ by means of the pseudoinverse of $\mathbf{X}$ as follows:

$$\bar{\mathbf{a}} = \mathbf{X}^{\#}\ \mathbf{y} \tag{B.22}$$

where $\mathbf{X}^{\#}$ is the $p \times N$ pseudoinverse of $\mathbf{X}$. The vector $\bar{\mathbf{a}}$ is a $p \times 1$ column vector which minimizes the square error (B.21) and at the same time the solution vector length $|\mathbf{a}|$. Minimum length means that $\bar{\mathbf{a}}$ is the shortest norm vector in $\mathbb{C}^p$ among the other possible solutions for minimizing (B.21).
If $p = N$ and $rank(\mathbf{X}) = p$ (i.e. $\mathbf{X}$ is not singular) then its pseudoinverse reduces to the usual inverse matrix

$$\mathbf{X}^{\#} = \mathbf{X}^{-1} \quad \text{for} \quad p = N \tag{B.23}$$

If $N > p$ and $rank(\mathbf{X}) = p$ then (B.21) is an overdetermined system of equations and the pseudoinverse reduces to

$$\mathbf{X}^{\#} = (\mathbf{X}^{H}\mathbf{X})^{-1}\mathbf{X}^{H} \quad \text{for} \quad N > p, \tag{B.24}$$

and (B.22) becomes (B.9), i.e. $\hat{\hat{\mathbf{a}}} = \hat{\mathbf{a}}$.
If $N < p$ and $rank(\mathbf{X}) = N$ then

$$\mathbf{X}^{\#} = \mathbf{X}^H(\mathbf{X}\mathbf{X}^H)^{-1} \quad \text{for} \quad N < p \tag{B.25}$$

and $\tilde{\mathbf{a}} = \mathbf{X}^{\#}\mathbf{y}$ is a $p \times 1$ column vector often called the minimum norm solution for a set of underdetermined equations.
Computing the pseudoinverse directly from the SVD (B.19) has two advantages over direct computation by means of either (B.24) for $N > p$ or (B.25) for $N < p$. First the SVD assists in determining the rank of $\mathbf{X}$ by examining the number of non negligible singular values (giving value zero to those below a fixed threshold). Second the accuracy in calculating $\mathbf{X}^{\#}$ by SVD is approximately twice as much as the accuracy obtained by calculating it from $(\mathbf{X}^H\mathbf{X})^{-1}$ and $(\mathbf{X}\mathbf{X}^H)^{-1}$ [85].

List of Figures

List of Tables

List of acronyms and symbols

FDTD	Finite Difference Time–Domain
LP	Linear Prediction
LS	Least Square
LTI	Linear Time Invariant
LTCC	Low Temperature Co-fired Ceramic
MED	Mean Error Distance
MEMS	Micro-Electro-Mechanical Systems
MoM	Method of Moments
MSE	Mean Square Error
PMSI	Prony model based Signal Identification
SI	System Identification
SNR	Signal to Noise Ratio
SOI	Silicon-On-Insulator
SVD	Singular Value Decomposition
$s.t.$	such that
TLM	Transmission Line Matrix

$\measuredangle$	argument for a complex quantity
$\supset$	mathematical inclusion between sets (as $V \supset W$ means set V contained in set W)
$\bigcup$	mathematical union between two or more sets
$\bigcap$	mathematical intersection between two or more sets
$\wedge$	exterior differential product
$\lfloor \ \rfloor$	approximation to the lower integer value
$\star$	Hodge operator
A^*	conjugate of A
A^T	transpose of A
A^H	Hermitian of A
A^{-1}	inverse of A
$A^{\#}$	pseudoinverse of A
$\underline{\mathcal{B}}$	magnetic flux density differential form phasor
$\underline{\boldsymbol{B}}$	magnetic flux density phasor
$\mathbb{C}$	set of the complex numbers

$\underline{\mathcal{D}}$	electric flux density differential form phasor
$\underline{\boldsymbol{D}}$	electric flux density phasor
D_T	threshold in the SVD based model order estimation
$\underline{\mathcal{E}}$	electric field differential form phasor
$\underline{\boldsymbol{E}}$	electric field phasor
$\underline{\varepsilon}$	complex permettivity
$\underline{\mathcal{G}}_e^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s)$	electric double form dyadic Green's function
$\underline{\mathcal{G}}_m^l(\boldsymbol{\rho}, \boldsymbol{\rho}', s)$	magnetic double form dyadic Green's function
$\underline{\mathcal{H}}$	magnetic field differential form phasor
$\underline{\boldsymbol{H}}$	magnetic field phasor
$\underline{\mathcal{J}}$	electric density current differential form phasor
$\underline{\mathcal{J}}_{eA}$	superficial electric density current differential form phasor
$\underline{\mathcal{J}}_{mA}$	superficial magnetic density current differential form phasor
$\Im$	imaginary part of a complex quantity
$\mathscr{L}$	laplace operator
$\underline{\mu}$	complex permeability
n	unit differential form correspondent to w (surface vertical coordinate)
$\mathbb{N}$	set of the integer numbers
$\underline{\mathcal{Q}}_A$	Electric area charge density differential form phasor
$\Re$	real part of a complex quantity
$\boldsymbol{\rho} = (u, v, w)$	Point in a curvilinear coordinate system
ρ_s	surface charge density
s	complex frequency in Laplace domain
$\hat{\mathbf{x}}$	recovered (modelled) response sequence
$\underline{\mathcal{Y}}$	admittance differential double form representation of the Dyadic Green's function
Y_q^l	q-th complex double form for the $\underline{\mathcal{Y}}$ series expansion
$\underline{\mathcal{Z}}$	impedance differential double form representation of the Dyadic Green's function
z_p^l	p-th complex double forms for the $\underline{\mathcal{Z}}$ series expansion

Printing in boldface refer at vector and underlining denote phasor quantities.

Bibliography

[1] Y. Kuznetsov A. Baev and A. Aleksandrov, *Ultra wideband radar target discrimination using the signature algorithm*, EuMC2003 (2003), 987–991.

[2] H. Akaike, *Power spectrum estimation through autoregression model fitting*, Ann. Inst. Stat. Math. **21** (1969), 407–419.

[3] _____, *A new look at the statistical model identification*, IEEE Trans. Autom. Control (1974), 716–723.

[4] J.W. Bandler, *Space mapping technique for electromagnetic optimization*, IEEE Trans. Microwave Theory Techn. (1994), no. 12.

[5] C. E. Baum, *The singularity expansion method for the solution of electromagnetic interaction problems*.

[6] V. Belevitch, *Classical network theory*, Holden-Day, San Francisco, California, 1968.

[7] J.P. Burg, *Maximum entropy spectral analysis*, Proceedings of the 37^{th} Meeting of the Society of Exploration Geophysicists (1967).

[8] K. M. Chen C. E. Baum, R. J. Rothwell and D. P. Nyquist, *The singularity expansion method and its application in target identification*, IEEE Proceeding **79** (1991), no. 9, 1481–1492.

[9] A.C. Cangellaris, M. Celik, S. Pasha, and L. Zhao, *Electromagnetic model order reduction for system level modeling*, IEEE Trans. Microwave Theory Techn. **47** (1999), 840–850.

[10] C. Christopoulos, *The transmission-line modeling method TLM*, IEEE Press, New York, 1995.

[11] V. Chtchekatourov, F. Coccetti, and P. Russer, *Direct Y-parameters estimation of microwave structures using TLM simulation and prony's method*, Proc. 17th Annual Review of Progress in Applied Computational Electromagnetics, Monterey, May 2001, pp. 580–586.

[12] V. Chtchekatourov, W. Fisch, F. Coccetti, and P. Russer, *Full-wave analysis and model-based parameter estimation approaches for s- and y- matrix computation of microwave distributed circuits*, 2001 Int. Microwave Symposium Digest, Phoenix, 2001, pp. 1037–1040.

[13] L.O. Chua, Ch.A. Desoer, and E.S. Kuh, *Linear and nonlinear circuits*, Mc Graw Hill, New York, 1987.

[14] F. Coccetti, L. Vietzorreck, V. Chtchekatourov, and P. Russer, *A numerical study of MEMS capacitive switches using TLM*, Proc. 16th Annual Review of Progress in Applied Computational Electromagnetics, Monterey (Monterey, CA), March 2000, pp. 580–586.

[15] R.E. Collin, *Field theory of guided waves*, second ed., IEEE Press, Inc., New York, 1991.

[16] M.D. Abouzahra D.M. Sheen, S.M. Ali and J.A. Kong, *Application of three–dimesional finite–difference time–domain method to the analysis of planar microstrip circiuts*, IEEE Trans. Microwave Theory Techn. **38** (1990), no. 7, 849–857.

[17] Robert S. Elliott, *Electromagnetics - history, theory, and applications*, IEEE Press, New York, 1991.

[18] C. Eswarappa and W.J.R. Hoefer, *Autoregressive (AR) and autoregressive moving average (ARMA) spectral estimation techniques for faster TLM analysis of microwave structures*, IEEE Trans. Microwave Theory Techn. **42** (1994), no. 12, 2407–2411.

[19] A.G. Evans and R. Fischl, *Optimal least–squared time–domain synthesis of recursive digital filters*, IEEE Transaction Audio Electroacoust.

[20] L.B. Felsen, M. Mongiardo, and P. Russer, *Electromagnetic field representations and computations in complex structures I: Complexity architecture and generalized network formulation*, Int. J. Numer. Modeling **15** (2002), 93–107.

[21] _____, *Electromagnetic field representations and computations in complex structures II: Alternative Green's functions*, Int. J. Numer. Modeling **15** (2002), 109–125.

[22] R.M. Foster, *A reactance theorem*, Bell System Tech. J. **3** (1924), 259–267.

[23] D. De Zutter G. Coen and N. Fache, *Automatic derivation of equivalent circuits for general microstrip interconnection discontinuities*, IEEE Trans. Microwave Theory Techn. (1996), no. 7.

[24] Ernst. A. Guillemin, *Synthesis of passive networks*, John Wiley & Sons, London, 1957.

[25] R. F. Harrington, *Field computation by moment methods,*, IEEE Press, San Francisco, 1968.

[26] Simon Haykin, *Adaptive filter theory. part xv. matched exponents for the representation of signals*, Prentice Hall, Upper Saddle River, NJ, 1996.

[27] J.L. Herring and C. Christopoulos, *Solving electromagnetic field problems using a multiple grid transmission–line modelling method*, ap (1994), no. 12, 1654–1658.

[28] W.J.R. Hoefer, *The transmission line matrix method-theory and applications*, IEEE Trans. Microwave Theory Techn. **33** (1985), no. 10, 882–893.

[29] _____, *The transmission line matrix (TLM) method*, Numerical Techniques for Microwave and Millimeter Wave Passive Structures (T. Itoh, ed.), John Wiley & Sons, New York, 1989, pp. 496–591.

[30] Y. Hua and T. Sarkar, *Matrix pencil method for estimating parameters of exponentially damped / undamped sinusoids in noise*, IEEE Trans. Acoust., Speech, Signal Processing **38** (1990), no. 5, 814–824.

[31] Y. Hua and T.K. Sarkar, *Further analysis of three modern techniques for pole retrieval from data sequence*, in Proc. 30^{th} Midwest Symp. Circuits Syst., Syracuse, NY (1987).

[32] Akira Ishimaru, *Electromagnetic wave propagation, radiation, and scattering*, Prentice Hall, Englewood Cliffs, NJ, 1991.

[33] N. Faché J. De Geest, T. Dhaene and D. De Zutter, *Adaptive CAD–model building algorithm for general planar microwavestructures*, IEEE Trans. Microwave Theory Techn. (1999), no. 9.

[34] V.K. Jain, *Filter analysis by use of the pencil–of–functions: Part i*, IEEE Trans. Circuits Syst. (1974), 574–579.

[35] _____, *Filter analysis by use of the pencil–of–functions: Part ii*, IEEE Trans. Circuits Syst. (1974), 580–583.

[36] P.B. Johns and R.L. Beurle, *Numerical solution of 2-dimensional scattering problems using a transmission-line matrix*, Proc. IEE **118** (1971), no. 9, 1203–1208.

[37] R.L. Kashyap, *Inconsistency of the AIC rule forestimating the order of autoregressive models*, IEEE Trans. Autom. Control (1980), 996–998.

[38] A.E. Ruehli K.M. Coperich and A. Cangellaris, *Enhanced skin effect for partial–element equivalent–circuit (PEEC) models*, IEEE Trans. Microwave Theory Techn. (2000), no. 9.

[39] M. Krumpholz, B. Bader, and P. Russer, *On the theory of discrete TLM Green's functions in three-dimensional TLM*, IEEE Trans. Microwave Theory Techn. **43** (1995), no. 7, 1472–1483.

[40] M. Krumpholz and P. Russer, *A field theoretical comparison of FDTD and TLM*, IEEE Trans. Microwave Theory Techn. **43** (1995), no. 8, 1935–1950.

[41] R. Kumaresan and D.W. Tufts, *Estimating the parameters of exponentially damped sinusoids and pole–zeros modelling in noise*, ASSP **30** (1982), 833–840.

[42] T. Rozzi F. Coccetti P. Russer N. Bushyager M. Tentzeris L. Pierantoni, M. Farina, *Comparison of electromagnetic solvers for the analysis of LTCC components*, Proc. of the 2003 IEEE AP-S Symposium - Columbus, Ohio, USA (2003).

[43] Lennart Ljung, *System identification theory for the user*, Prentice Hall, Upper Saddle River, NJ, 1999.

[44] K.L. Virga M. Elzinga and J.L. Prince, *Improved global rational approximation macromodeling algorithm for networks characterized by frequency–sampled data*, IEEE Trans. Microwave Theory Techn. **48** (2000), no. 9, 1461–1468.

[45] J. Makhoul, *Linear prediction: A tutorial review*, Proceedings of the IEEE **63** (1975), 561–580.

[46] T. Mangold and P. Russer, *Full-wave modeling and automatic equivalent-circuit generation of millimeter-wave planar and multilayer structures*, IEEE Trans. Microwave Theory Techn. **47** (1999), 851–858.

[47] S.L. Marple, *Digital spectral analysis*, Prentice Hall, Englewood Cliffs, NJ, 1997.

[48] R.N. McDonough, *Representation and analysis of signals*, PhD Dissertation, Department of Electrical engineering, John Hopkings University, Baltimore,Md.,, 1963.

[49] E.K. Miller and T.K. Sarkar, *An introduction to the use of model-based parameter estimation in electromagnetics*, Review of Radio Science 1996–1999, ed. by R. Stone, Oxford University Press (1999), 139–174.

[50] R.F. Milsom, *Efficient SPICE–compatible electromagnetic model of arbitrarily shaped integrated passive structure*, IEEE Trans. Microwave Theory Techn. (1999), no. 7.

[51] M. Mongiardo, P. Russer, C. Tomassoni, and L.B. Felsen, *Analysis of N-furcation in elliptical waveguides via the generalized network formulation*, 1999 Int. Microwave Symposium Digest, Anaheim (1999), 27–30.

[52] Bart De Moore, *The singular value decomposition and long and short spaces of noisy matrices*, IEEE Transaction on Signal Proicessing **41** (1993), no. 9, 2826–2838.

[53] A.H. Nuttall, *Spectral analysis ...*, Naval Underwater System Center Technical Report TR–5303 **12** (1976).

[54] A. Odabasioglu, M. Celik, and L.T. Pileggi, *PRIMA: Passive reduced-order interconnect macromodeling algorithm*, IEEE Trans. Computer-Aided Design **17** (1998), 645–653.

[55] Alan V. Oppenheim and Ronald W. Schafer, *Discrete–time signal processing*, Prentice–Hall International, Eaglewood Cliff,New Jersey, 1989.

[56] H. Ouibrahim, *A generalized approach to direction finding*, Ph.D Dissertation, Syracuse University, Syracuse, NY, December 1986.

[57] L. Page and N. Adams, *The electrical oscillation of a prolate spheroid; Paper I*, Phys. Rev. **33** (1938), 819–831.

[58] ———, *The electrical oscillation of a prolate spheroid; Paper I*, Phys. Rev. **33** (1944), 111–117.

[59] P. Penfield, R. Spence, and S. Duinker, *Tellegen's theorem and electrical networks*, MIT Press, Cambridge, Massachusetts, 1970.

[60] H.C. Pocklington, *Electrical oscillations in wire*, Cambridge Philos. Soc. Proc. **9** (1897), 324–332.

[61] R. Prony, *Essai experimental et analytique sur les lois de la dilatabilitè de fluides elastiques et sur celles del la force expansive de la vapeur de l'alkool, à differentes temperatures*, J. l'Ecole Polytech. 1795 (Paris).

[62] J.N. Rebel, *On the foundations of the transmission line matrix method*, Ph.D Dissertation, TU Mnchen, 2000.

[63] R. F. Remis and P. M. van den Berg, *A modified Lanczos algorithm for the computation of transient electromagnetic wavefields*, IEEE Trans. Microwave Theory Techn. (1997), 2139–2149.

[64] A. Paulraj R.H. Roy and T. Kailath, *Esprit – a subspace rotation approach to estimation of parameters of cisoids in noise*, ASSP **34** (1986), 1340–1342.

[65] M. Righi, C. Eswarappa, W.J.R. Hoefer, and P. Russer, *An alternative way of computing s-parameters via impulsive TLM analysis without using absorbing boundary conditions*, 1995 Int. Microwave Symposium Digest, Orlando (1995), 1203–1206.

[66] J.A. Rissanen, *A universal prior for the integers and estimation by minimum description length*, Ann. Stat. **11** (1983), 417–431.

[67] M. Maeng E. Tsai G. DeJean M. Tentzeris R.L. Li, K. Lim and J. Laskar, *Design of compact stacked-patch antennas on LTCC technology for wireless communication applications*, IEEE Antennas Propagat. Symposium (2002).

[68] R.H. Roy, *Esprit – estimation of signal parameters via rotational invariance techniques*, Ph.D Dissertation, Stanford Univ.,, August 1986.

[69] A.E. Ruehli, *Equivalent circuit models for three–dimensional multiconductor systems*, IEEE Trans. Microwave Theory Techn. (1974), no. 3.

[70] A.E. Ruehli and H. Heeb, *Circuit models for three–dimesnional geometries including dielectric*, IEEE Trans. Microwave Theory Techn. (1992), no. 7.

[71] P. Russer, *The transmission line matrix method*, Applied Computational Electromagnetics, NATO ASI Series, Springer, Cambridge, Massachusetts, London, England, 2000, pp. 243–269.

[72] P. Russer, *Electromagnetics, microwave circuit and antenna design for communications engineering*, Artech House, Boston, 2003.

[73] P. Russer, M. Mongiardo, and L.B. Felsen, *Electromagnetic field representations and computations in complex structures III: Network representations of the connection and subdomain circuits*, Int. J. Numer. Modeling **15** (2002), 127–145.

[74] P. Russer, M. Righi, C. Eswarappa, and W.J.R. Hoefer, *Circuit parameter extraction of distributed microwave circuits via TLM simulation*, 1994 Int. Microwave Symposium Digest, San Diego (1994), 887–890.

[75] S.T. Chew S.A. Teo, B.L. Ooi and M.S. Leon, *A fast PEEC technique for full–wave parameters extraction of distributed elements*, Microwave and Wireless Components Letters (2001), no. 5.

[76] J. Schwinger, *The theory of obstacles in resonant cavities and waveguides*, Tech. report, Boston, May 1943.

[77] D.B. Shorthouse and C.J. Railton, *The incorporation of static field solutions into the finite diffrence time domain algorithm*, IEEE Trans. Microwave Theory Techn. **40** (1992), no. 5, 986–990.

[78] Julius A. Stratton, *Electromagnetic theory*, McGraw-Hill, New York, 1941.

[79] A. Taflove, *Computational electrodynamics - the finite-difference time-domain method*, Artech House, Boston, London, 1995.

[80] B.D.H. Tellegen, *A general network theorem with applications*, Philips Research Reports **7** (1952), 259–269.

[81] _____, *A general network theorem with applications*, Proc. Inst. Radio Engineers **14** (1953), 265–270.

[82] D.W. Tufts and R. Kumaresan, *Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood*, ProcIEEE **70** (1982), 975–989.

[83] T.J. Ulrych and R.W. Clayton, *Time series modelling and maximum entropy*, Phys. Earth Planet. Inter. **12** (1976), 188–200.

[84] T.K. Sarkar V.K. Jain and D.D. Weiner, *Rational modelling by the pencil–of–functions method*, ASSP **31** (1983), 564–573.

[85] w. T. Vetterling B. P. Flannery W. A. Press, S. A. Teukolsky, *Numerical recipes in c. the art of science computing. second edition*, Cambridge University Press, Cambridge, 1992.

[86] K.S. Yee, *Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media*, IEEE Trans. Antennas Propagat. **14** (1966), 302–307.