Robust Macromodeling of Frequency Responses with Outliers

Dirk Deschrijver, Luc Knockaert, Tom Dhaene
Dept. of Information Technology, Ghent University - IBBT,
Sint Pietersnieuwstraat 41, 9000 Ghent, Belgium
email: {dirk.deschrijver, luc.knockaert, tomm.dhaene}@intec.ugent.be

Abstract
This paper introduces a robust Vector Fitting algorithm for macromodeling of measured or simulated frequency responses with outliers. The use of a new frequency-dependent weighting scheme ensures that the complex fitting error is minimized in the $L_1$ sense, rather than the $L_2$ sense. Numerical results indicate that this approach leads to more accurate results.

1 Introduction
Robust macromodeling techniques are of paramount importance for efficient time domain and frequency domain simulations of passive linear microwave systems and devices. Vector Fitting (VF) has proved to be a fast and reliable method that calculates high-order transfer function from measured or simulated frequency responses. It has been adopted in many societies of applied engineering, including power systems and microwave systems. A recent survey of the methodology is found in [1].

The standard Vector Fitting algorithm computes a macro-model by minimizing a weighted iterative cost function in the $L_2$ sense [2][3]. However, in real-life situations, it is possible that outliers in the data strongly degrade the quality of the $L_2$ fitting model. Outliers are values in the frequency response that deviate strongly from the other values, and they can be caused by various measurement or instrumentation errors [4].

This paper discusses a modified Vector Fitting algorithm that minimizes the $L_1$ norm of the complex fitting error instead of the $L_2$ norm [5, 6]. It gives a more elaborate explanation of the derivations in [7] and presents a new example which demonstrates that the approach is more robust towards outliers [8].

2 $L_1$-norm Macromodeling Algorithm
Based on a discrete set of S-parameter data samples \( \{\mathbf{s}_k, H(s_k)\}_{k=0}^K \), VF computes a rational macro-model with numerator $N(s)$ and denominator $D(s)$ in an iterative way \( (t = 1, \ldots, T) \) by successively solving least squares problems

$$
\arg\min_{\{s_k\}} \sum_{k=0}^{K} |W^t(s_k)|^2 \left| (\sigma H)^t(s_k) - \sigma^t(s_k) H(s_k) \right|^2
$$

(1)

As shown in [1] and [9], both $N(s)$ and $D(s)$ belong to a linear span of $p = 1, \ldots, P$ rational basis functions $\Psi_p(s, a^{t-1})$ that are based on the previously identified set of poles $a^{t-1}$. In the first iteration step ($t = 1$) these initial poles $a^{0}$ are chosen by the standard heuristic scheme to ensure a good numerical conditioning [2]. The basis functions $\Psi$ are typically chosen as partial fractions or orthonormal rational functions [10]

$$
(\sigma H)^t(s) = \frac{N^t(s)}{D^{t-1}(s)} = c_0^t + \sum_{p=1}^{P} c_p^t \Psi_p(s, a^{t-1})
$$

(2)

$$
\sigma^t(s) = \frac{D^t(s)}{D^{t-1}(s)} = 1 + \sum_{p=1}^{P} c_p^t \Psi_p(s, a^{t-1})
$$

(3)

Note that the cost function (1) also contains a user-defined weighting function $W^t(s)$. This weighting function should not be confused with the Sanathanan-Koerner weighting, in this paper denoted as $w^t(s)$, which occurs implicitly by pole relocation. In a standard application of Vector Fitting, the value of $W^t(s)$ is set to 1 by default. However, it is well known that an additional frequency-dependent weighting factor can provide a powerful way to control the accuracy of the macromodel.

In the following sections, it will be shown that the $L_1$ norm of the complex fitting error can be minimized by selecting the user-defined weighting factor $W^t(s)$ in (1) as follows

$$
W^t(s) = \frac{\sqrt{H(s) - H^{t-1}(s)}}{H(s) - H^{t-1}(s)}
$$

(4)

where $H^{t-1}(s) = N^{t-1}(s)/D^{t-1}(s)$ denotes the frequency response of the macromodel at previous iteration step $t-1$.

3 Proof Outline
First, let’s define the auxiliary function $f(s)$ as follows

$$
f(s) = \frac{|D^t(s)H(s) - N^t(s)|^2}{|D^{t-1}(s)H(s) - N^{t-1}(s)|^2}
$$

(5)

Applying the weighting factor $W^t(s)$ in (4) to the cost function (1) yields the following equivalent expression

$$
\arg\min_{\{s_k\}} \sum_{k=0}^{K} \left| H(s_k) - H^{t-1}(s_k) \right| f(s_k)
$$

(6)

Hence, upon convergence of the iterative scheme ($D^{t-1} \rightarrow D^t$ and $N^{t-1} \rightarrow N^t$), it follows that $f(s) \rightarrow 1$. Therefore, it is clear that cost function (6) effectively minimizes the $L_1$ norm of the complex fitting error $\|H(s) - H^{t-1}(s)\|_1$.

4 Analysis of the weighting function
In this section, it will be shown how the weighting function $W^t(s)$ in (4) was derived. The objective of the standard Vector Fitting procedure is to identify the model parameters in such a way that the $L_2$ norm of the complex error is minimized

$$
\arg\min_{\{s_k\}} \sum_{k=0}^{K} \left| H(s_k) - \frac{N(s_k)}{D(s_k)} \right|^2
$$

(7)
To avoid non-linear optimization techniques, a linear set of equations is obtained by minimizing Levi’s cost function [11]

$$\arg \min \sum_{k=0}^{K} |D(s_k)H(s_k) - N(s_k)|^2$$  \hspace{1cm} (8)

This cost function (8) can easily be generalized by introducing a frequency-dependent weighting function, denoted as $Q(s)$

$$\arg \min \sum_{k=0}^{K} |Q(s_k)|^2 |D(s_k)H(s_k) - N(s_k)|^2$$  \hspace{1cm} (9)

Note that (7) and (9) are equivalent if $Q(s)$ is set to $1/D(s)$. In order to minimize the $L_1$ norm of the complex fitting error, a different weighting function $Q(s)$ is chosen, such that

$$|Q(s)|^2 = \left| \frac{H(s) - N(s)/D(s)}{D(s)H(s) - N(s)} \right|$$  \hspace{1cm} (10)

Since $N(s)$ and $D(s)$ are not known in advance, they are replaced by the estimated numerator $N^{t-1}(s)$ and denominator $D^{t-1}(s)$ from the previous iteration step. Just like in the Sanathanan-Koerner iteration [12], inserting these values gives an estimate $Q^t(s)$ of the weight function $Q(s)$ that satisfies

$$|Q^t(s)|^2 = \frac{|H(s) - N^{t-1}(s)/D^{t-1}(s)|}{|D^{t-1}(s)H(s) - N^{t-1}(s)|^2}$$

$$= \frac{1}{|D^{t-1}(s)|^2} \left| \frac{H(s) - N^{t-1}(s)/D^{t-1}(s)}{|H(s) - N^{t-1}(s)/D^{t-1}(s)|^2} \right|$$

$$= \frac{1}{|D^{t-1}(s)|^2} \left| \frac{H(s) - H^{t-1}(s)}{|H(s) - H^{t-1}(s)|^2} \right|$$  \hspace{1cm} (11)

The use of weighting function $Q^t(s)$ leads to updated values of $N^t(s)$ and $D^t(s)$ in successive iteration steps ($t = 1, ..., T$)

$$\arg \min \sum_{k=0}^{K} |Q^t(s_k)|^2 |D^t(s_k)H(s_k) - N^t(s_k)|^2$$  \hspace{1cm} (12)

Inserting (11) in (12) yields the following cost function

$$\arg \min \sum_{k=0}^{K} \left| W^t(s_k) \right|^2 \left| \frac{N^t(s_k)}{D^{t-1}(s_k)} - \frac{D^t(s_k)}{D^{t-1}(s_k)}H(s_k) \right|^2$$  \hspace{1cm} (13)

Using the notations in (2) and (3), this is equivalent to

$$\arg \min \sum_{k=0}^{K} \left| W^t(s_k) \right|^2 \left| (\sigma H)^t(s_k) - \sigma H^{t-1}(s_k) \right|^2$$  \hspace{1cm} (14)

From (11), (12) and (13), $W^t(s)$ must satisfy the condition

$$|W^t(s)|^2 = \frac{|H(s) - H^{t-1}(s)|}{|H(s) - H^{t-1}(s)|^2}$$  \hspace{1cm} (15)

So, in order for (15) to hold, $W^t(s)$ is chosen as follows

$$W^t(s) = \sqrt{H(s) - H^{t-1}(s)}$$  \hspace{1cm} (16)

which corresponds to the user-defined weighting factor of Vector Fitting in (4). It is found that explicit weighting by $Q^t(s)$ in (12) is equivalent to explicit weighting by $W^t(s)$ in (14) + pole relocation (i.e. implicit weighting by $w^t(s) = 1/D^{t-1}(s)$) [1].

5 Example : Coplanar Waveguide

The reflection coefficient $S_{11}$ of a 2-port coplanar waveguide is considered over the frequency range [0.1 GHz - 50 GHz]. Suppose that, due to inaccuracies in the data acquisition process, the S-parameter response contains three outlying data samples which are marked by black arrows in Figs. 1 to 4.

All the data samples are modeled by a rational 6-pole proper transfer function using the proposed VFM methodology ($L_1$ norm) and the standard VFM algorithm ($L_2$ norm). The magnitude of the data and the magnitude response of the model is shown in Figs. 1 and 2 respectively. Also the phase of the data is compared to the phase response of the model, as shown in Figs. 3 and 4. It is clear that the $L_1$ norm approximation yields an overall accurate result, and is not much affected by the presence of the outliers. On the other hand, the outliers lead to an undesired degradation of the model quality for the $L_2$ norm.

Figure 1: Data samples of reflection coefficient $S_{11}$ (blue dots) versus magnitude response of the $L_2$ model (red line).

Figure 2: Data samples of reflection coefficient $S_{11}$ (blue dots) versus magnitude response of the $L_1$ model (red line).
approximation. This observation is also confirmed by Fig. 5 where the absolute fitting error of both models is shown.

\[ \text{Absolute fitting error} = |H(s) - H^{(1)}(s)| \]  \hspace{1cm} (17)

Fig. 6 visualizes the magnitude of the weighting function \(W^f(s)\), defined in (16), as a function of the frequency. It is seen that three sharp spikes occur at the exact frequencies where the outliers are located. This indicates that the algorithm automatically detects the occurrence of outliers, and gives them a very small weight when compared to the other data samples. In the general case, this leads to a better accuracy than the classical VF algorithm, where \(W^f(s)\) is constant for all frequencies.

As an additional test, the three outliers are manually removed from the data set and both modeling algorithms are applied to the remaining data samples. It is seen from Fig. 7 that the absolute error of both fitting models (\(L_1\) and \(L_2\)) is comparable.

This result confirms that \(L_1\) norm approximation is particularly useful in cases where the data is contaminated with outliers.

**Conclusions**

A modified Vector Fitting algorithm is proposed for \(L_1\) norm identification of broadband macromodels from S-parameter data. The effectiveness of the algorithm is illustrated by applying it to a coplanar waveguide example, and the results are compared to the standard VF approach. It is found that the method is more robust when the frequency response contains outliers.

**References**


Figure 7: Absolute error $L_1$ model (green) and $L_2$ model (blue). All the outliers are removed from the set of data samples.


