

A Bayesian Approach to Adaptive Frequency Sampling

Simon De Ridder, Dirk Deschrijver, Domenico Spina, Tom Dhaene and Dries Vande Ginste

Abstract—This paper introduces an adaptive frequency sampling scheme, based on a Bayesian approach to the well-known vector fitting algorithm. This Bayesian treatment results in a data-driven measure of intrinsic model uncertainty. This uncertainty measure can in turn be leveraged to sample sequentially in an efficient and robust way. A realistic example is used to visualize the proposed scheme, and to confirm its proficiency.

I. INTRODUCTION

Over the course of the past decades, computer aided design simulations have become essential to designers of high-speed circuits, as an efficient alternative to prototyping, to deal with increasing complexity and bandwidth. Nevertheless, this increase in complexity and bandwidth drives such simulations to ever more intricate and computationally expensive calculations.

For many passive and linear electromagnetic systems (filters, connectors, ...), analysis in the frequency domain is the norm, and in these cases an adaptive sampling scheme is of considerable importance. Adaptive Frequency Sampling (AFS), namely, allows the characterization of systems over the desired frequency range using a minimal amount of such expensive simulations [1]–[4]. At the same time, the number of frequencies at which simulations are carried out, must be sufficient to accurately model the dynamic behavior of the system. Customarily, this is achieved by means of a rational macromodel, most often Vector Fitting (VF) [5]–[8], combined with a heuristic measure of model uncertainty by comparing models of different orders.

In this paper, a Bayesian approach to VF is introduced, termed Linear Bayesian VF (LBVF). This framework produces an intrinsic measure of model uncertainty, most suitable for AFS schemes. The application of this framework to a representative example will confirm its aptness and efficiency. A prototypical version of this framework was presented in [9], on a one-port example.

II. LINEAR BAYESIAN VECTOR FITTING

A. Vector fitting

The VF algorithm is a well established method for constructing rational macromodels of transfer functions. At its core, it uses so-called Sanathanan-Koerner (SK) iterations [10] to relocate the poles of the rational model until convergence. In the following, the VF algorithm is briefly detailed.

S. De Ridder, D. Deschrijver, D. Spina, T. Dhaene and D. Vande Ginste (IDLab, Department of Information Technology, Ghent University-imec, 9000 Gent, Belgium)

E-mail: simon.deridder@ugent.be

The VF algorithm fits a vector-valued transfer function $\overline{F}(s)$ of length N_F to a number of partial fractions

$$\overline{F}(s) \approx \sum_{k=1}^K \frac{\overline{R}_k}{s - a_k} + \overline{D} + s\overline{E}, \quad (1)$$

where s is the Laplace variable, the \overline{R}_k are the residues, the a_k are the poles, and \overline{D} and \overline{E} are optional terms describing the asymptotic behavior of the transfer function. The residues and poles are either real-valued, or come in complex conjugate pairs. Stability is guaranteed if $\Re(a_k) < 0$ for all k [5].

As the identification of the poles in the least squares sense is a non-linear problem, (1) is first rewritten as:

$$\overline{F}(s) = \frac{\overline{p}(s)}{\sigma(s)} = \frac{\sum_{k=1}^K \frac{\overline{r}_k}{s - q_k} + \overline{d} + s\overline{e}}{\sum_{k=1}^K \frac{\widehat{r}_k}{s - q_k} + \widehat{d}}, \quad (2)$$

The q_k defined here are called starting poles, as they are the starting point of the SK iteration scheme. In the relaxed version of VF [6], \widehat{d} is a free variable, otherwise it is set to 1.

The SK iteration is now introduced in the following manner. The linearized system $\sigma(s)\overline{F}(s) = \overline{p}(s)$ is solved in a least squares sense for \widehat{r}_k , \widehat{d} and (superfluously) \overline{r}_k , \overline{d} and \overline{e} . As $\sigma(s)$ and $\overline{p}(s)$ share the same poles q_k , the only remaining poles of $\overline{F}(s)$ are the zeros of $\sigma(s)$. These can be computed by solving a suitable eigenvalue problem [6]. The new poles are then used as the starting poles for the next iteration. This process is repeated until the final poles a_k are obtained.

Based on these poles, the residues \overline{R}_k can then be identified by solving (1).

B. Bayesian approach

In the LBVF framework, both linear systems above are solved using Bayesian linear regression [11], and uncertainty is propagated from \widehat{r}_k and \widehat{d} to the poles by means of sampling.

Suppose we have a number of data points (s_i, \overline{y}_i) for $i \in \{1, \dots, N\}$, distributed as a Gaussian around $\overline{F}(s)$:

$$P(\overline{y}_i | \overline{F}(s_i)) \sim \mathcal{N}(\overline{y}_i | \overline{F}(s_i), \sigma^2 \overline{I}). \quad (3)$$

If we write the linear system used to arrive at \widehat{r}_k and \widehat{d} as $\overline{A}\overline{x} = \overline{b}$, where \overline{x} contains the residues \widehat{r}_k and \widehat{d} to be found, and \overline{b} contains a linear combination of the \overline{y}_i in vectorized form (see [7]), then \overline{b} is also distributed as in (3), with $\overline{A}\overline{x}$ as mean. Since σ^2 is a priori unknown, it is treated as a (nuisance) random variable.

A conjugate prior can then be formulated as the product of a Gaussian and an Inverse-Gamma distribution:

$$P(\bar{x}, \sigma^2) \sim \mathcal{N}\left(\bar{x} \mid \bar{x}_0, \sigma^2 \bar{\Lambda}_0^{-1}\right) \text{IG}(\sigma^2 \mid \alpha_0, \beta_0). \quad (4)$$

Here, \bar{x}_0 , $\bar{\Lambda}_0$, α_0 and β_0 are prior parameters, either determined by prior information, or set to zero to apply an uninformative (but improper) Jeffrey's prior. The posterior distribution of \bar{x} , after marginalizing σ^2 , becomes a Student T -distribution:

$$P(\bar{x} \mid \bar{b}) = t_{2\alpha_f} \left(\bar{x} \mid \bar{x}_f, \left(\frac{\alpha_f \bar{\Lambda}_f}{\beta_f} \right)^{-1} \right), \quad (5)$$

where the parameters are calculated as

$$\begin{aligned} \bar{\Lambda}_f &= \bar{\Lambda}_0 + \bar{A}^T \bar{A}, \\ \bar{x}_f &= \bar{\Lambda}_f^{-1} \left(\bar{\Lambda}_0 \bar{x}_0 + \bar{A}^T \bar{b} \right), \\ \alpha_f &= \alpha_0 + \frac{N N_F}{2}, \\ \beta_f &= \beta_0 + \frac{\bar{b}^T \bar{b} - \bar{x}_f^T \bar{\Lambda}_f \bar{x}_f + \bar{x}_0^T \bar{\Lambda}_0 \bar{x}_0}{2}. \end{aligned} \quad (6)$$

The sampling of the residues from a Bayesian solution to (1) with fixed poles is very analogous, though matrix-variate. The equivalent of (3) is now a matrix normal distribution.

$$P(\bar{B} \mid \bar{A} \bar{X}) \sim \mathcal{MN}_{N, N_F}(\bar{B} \mid \bar{A} \bar{X}, \bar{I}, \bar{\Sigma}), \quad (7)$$

where row i of \bar{B} contains \bar{y}_i and row k of \bar{X} contains \bar{R}_k . For more information on this linear system, see [7].

After a similar procedure as the one leading to (5), the posterior distribution of \bar{X} becomes a matrix- T -distribution:

$$\text{T}_{K, N_F} \left(\bar{X} \mid \nu_f - N_F + 1, \bar{X}_f, \bar{\Lambda}_f^{-1}, \bar{V}_f \right), \quad (8)$$

in which

$$\begin{aligned} \bar{\Lambda}_f &= \left(\bar{\Lambda}_0 + \bar{A}^T \bar{A} \right), \\ \bar{X}_f &= \bar{\Lambda}_f^{-1} \left(\bar{\Lambda}_0 \bar{X}_0 + \bar{A}^T \bar{B} \right), \\ \bar{V}_f &= \bar{V}_0 + \bar{B}^T \bar{B} - \bar{X}_f^T \bar{\Lambda}_f \bar{X}_f + \bar{X}_0^T \bar{\Lambda}_0 \bar{X}_0, \\ \nu_f &= \nu_0 + N. \end{aligned} \quad (9)$$

In this, \bar{X}_0 , $\bar{\Lambda}_0$, \bar{V}_0 and ν_0 are prior parameters, which again can be set according to prior information, or set to zero to apply an uninformative prior. In this paper, the uninformative priors are adopted.

Drawing samples from the LBVF model outlined above now occurs as follows. Unfortunately, it is not analytically possible to propagate (5) to the poles, as their calculation involves an eigenvalue problem. As a workaround, samples of \hat{r}_k and \hat{d} are drawn from (5). From each such sample, a set of poles is then calculated, which constitutes an empirical approximation of the poles' posterior distribution. For each of these pole sets, samples from (8) are then drawn to obtain several sets of residues. Each pair of pole set/residue set drawn in this way

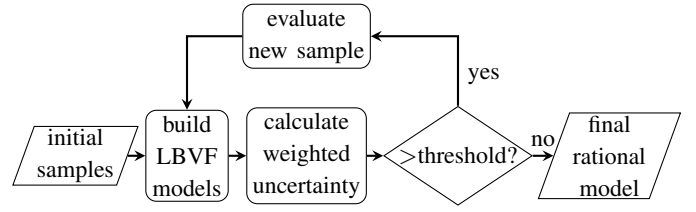


Figure 1. Flowchart of the proposed AFS strategy.

forms a rational model according to (1). Such models, denoted $\bar{F}_{\text{LBVF}}^j(s)$, can be interpreted as samples from the posterior distribution of fits to the data.

Another outcome of the LBVF framework is the analytic form for the marginal likelihood of the data under the pole relocation system:

$$p(\bar{b}) = (2\pi)^{-\frac{N}{2}} \sqrt{\frac{|\bar{\Lambda}_0|}{|\bar{\Lambda}_f|}} \frac{\beta_0^{\alpha_0} \Gamma(\alpha_f)}{\beta_f^{\alpha_f} \Gamma(\alpha_0)}, \quad (10)$$

using the parameters in (6). It can be interpreted as a measure of how likely it is that (2) could have produced the data. As such, it can be used as a measure of the quality of the starting poles and their number. This then offers an intrinsic way to evaluate how many poles are needed for a good fit to the data, where there were no such means in classical VF.

III. PROPOSED AFS STRATEGY

The sampled models $\bar{F}_{\text{LBVF}}^j(s)$ introduced above can now be used to construct a measure of uncertainty as a function of frequency, which would naturally form the basis of an adaptive frequency sampling scheme. One way to do this is to calculate the standard deviation of the $\bar{F}_{\text{LBVF}}^j(s)$ at any frequency. It is beneficial, however, to consider LBVF models of different orders, as the model is conditioned on the starting poles, and would otherwise suffer from overconfidence. As models of different orders do not necessarily possess the same quality, they are combined using their marginal likelihoods (10) as weights, to calculate a weighted standard deviation. To encourage exploration and to avoid sampling where frequency points are already evaluated, a small Gaussian penalty is added to this weighted standard deviation to finally obtain a fully fledged uncertainty measure.

Fig. 1 lays out the proposed adaptive sampling scheme. Firstly, a small number (here: four) of initial data points are evaluated, spread equidistantly over the desired frequency range. Secondly, LBVF models of up to four different orders are built, and a large number of samples (here: 500) are drawn from each of them. Thirdly, these samples are used to construct the uncertainty measure detailed above. Then, if this uncertainty is below a chosen threshold, the model with the highest marginal likelihood is accepted as a suitable rational model for the transfer function. If the threshold is exceeded, the frequency point where the uncertainty is highest is selected as the next to evaluate. After this evaluation, new LBVF models are built, the uncertainty is recalculated, and so on.

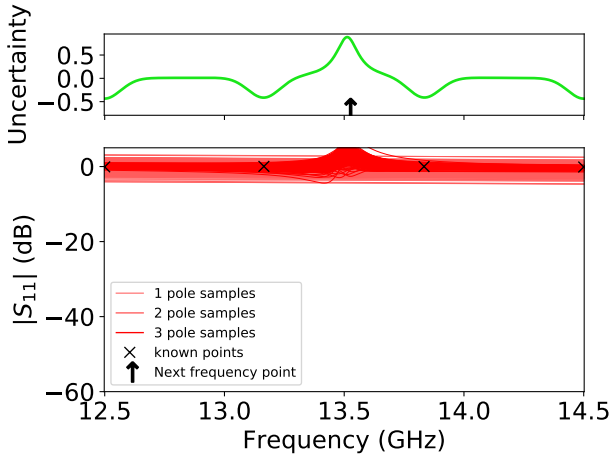


Figure 2. The first step in the adaptive sampling scheme, after four initial simulations, for the S_{11} parameter.

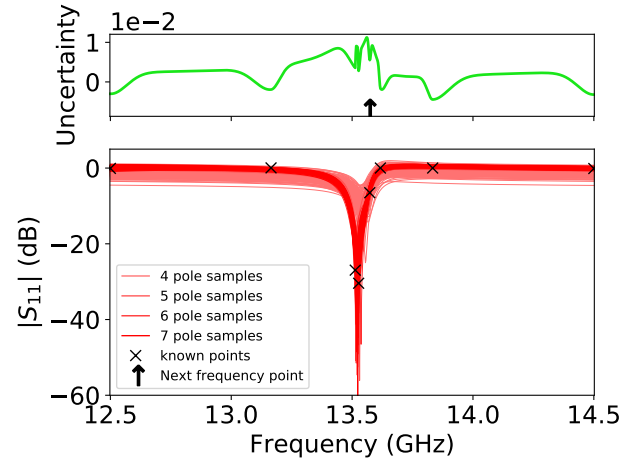


Figure 4. The fifth step in the adaptive sampling scheme, for the S_{11} parameter.

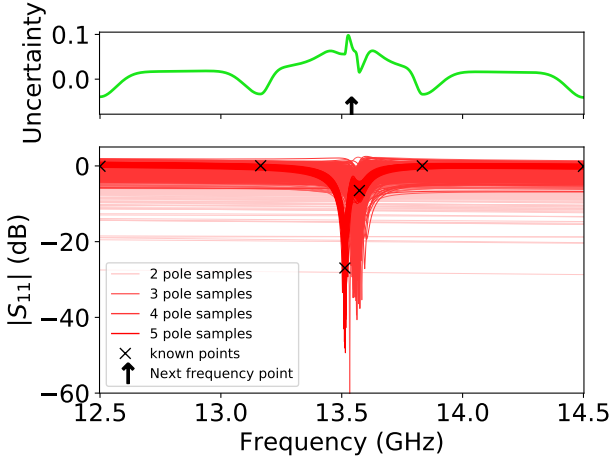


Figure 3. The third step in the adaptive sampling scheme, for the S_{11} parameter.

In contrast to classical AFS schemes [1]–[4], the proposed AFS scheme does not merely rely on disagreement between models, but also and most importantly, where individual models are uncertain in their interpolation. This avoids premature halting of the algorithm when several models are uncertain but agree in their mean prediction. This added stability usually outweighs the added computational cost, especially so when the underlying electromagnetic simulation is expensive.

IV. EXAMPLE

To illustrate the process of the proposed AFS scheme, it is now applied to the example of a two-port hairpin filter [12]. Some of the steps in the AFS process are shown in Figs. 2–4. Note that this happens simultaneously for all elements of the transfer function matrix, with priority given to the diagonal elements' uncertainty. Only the element with the highest uncertainty determines the next evaluation point. The previously simulated points are portrayed as black crosses, while the LBVF samples $\bar{F}_{\text{LBVF}}^j(s)$ are shown in shades of red, proportional to their respective models' marginal likelihood.

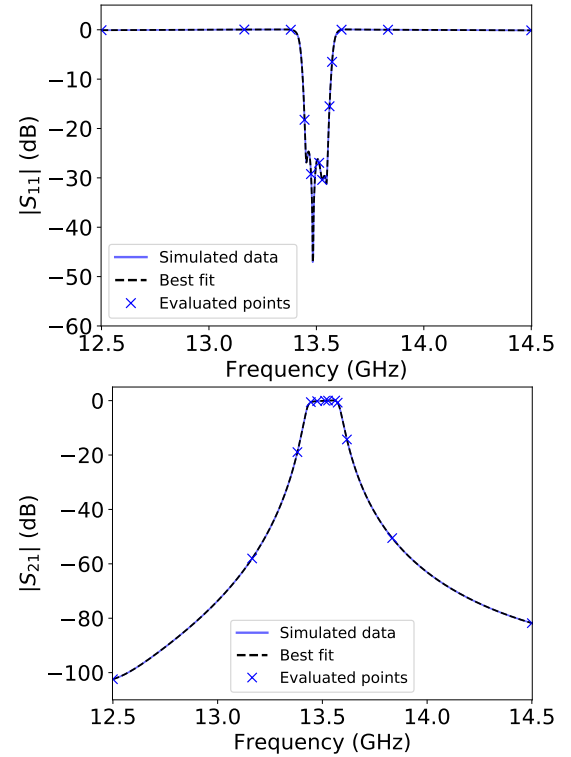


Figure 5. The final fit to the data, after 12 simulations in total. The actual simulated S parameters are shown in blue, while the fit is shown as a black dashed line. The evaluated points are rendered as blue crosses.

The upper plots show the uncertainty measure in green. An arrow shows the next point to be evaluated based on this uncertainty.

After 12 evaluations, the threshold (0.01) on the uncertainty is no longer surpassed, and the sampling scheme ends. The highest order model (with 11 poles) is chosen to represent the transfer function, see Figs. 5 and 6. Its mean can serve as a traditional macromodel. In Fig. 7, the evolution of the best mean model is shown.

With respect to the transfer function simulated in 1001

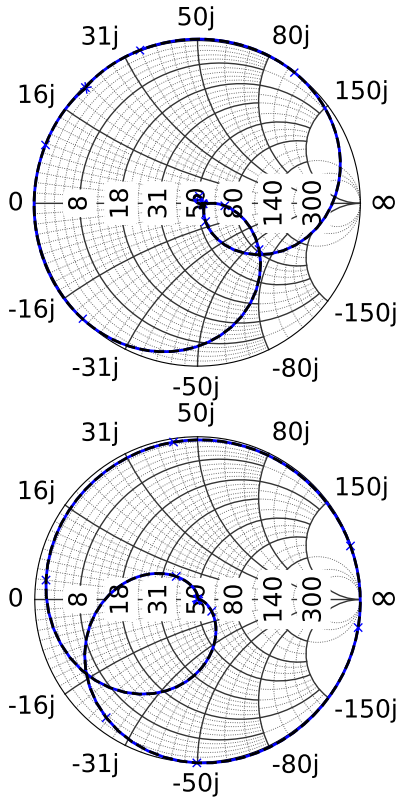


Figure 6. Smith charts of the final fit to the data, for the S_{11} parameter (top) and the S_{21} parameter (bottom). Colors are as in Fig. 5.

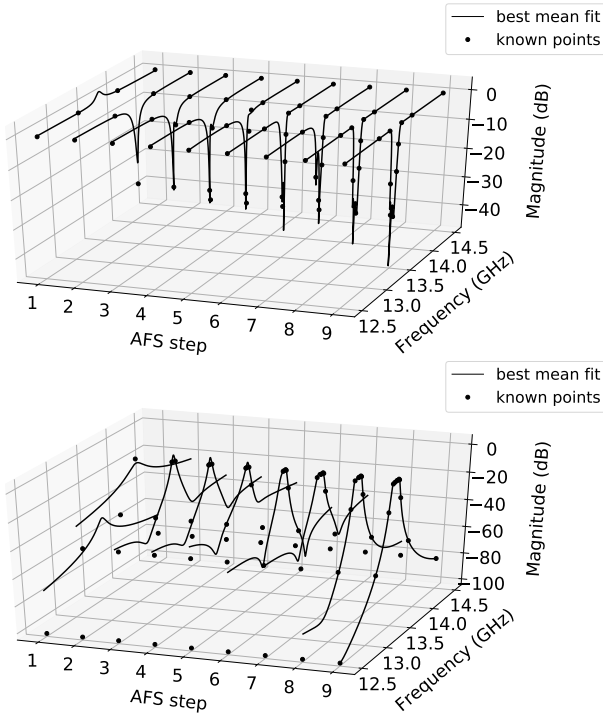


Figure 7. Progression of the mean of the model with the highest marginal likelihood at each step of the AFS scheme, for the S_{11} parameter (top) and the S_{21} parameter (bottom). Evaluated points are shown as black dots.

frequency points, the root mean square error (RMSE) and maximum absolute error of the best mean fit also drop below -60 dB, to -263 dB and -246 dB, respectively. On a server node with a Quad-Core AMD Opteron Processor (2 GHz) and 31.4 GiB DDR2 RAM, the total time spent building the LBVF uncertainty measure is 64.7 s

V. CONCLUSION

This paper introduces a Bayesian approach to the classical Vector Fitting method, termed LBVF. By sampling from the posterior distribution of these LBVF models, an intrinsic model uncertainty can be formulated. This uncertainty in turn forms the basis for an adaptive frequency sampling scheme, wherein simulations are carried out sequentially, at the frequencies where the uncertainty is maximal. As such, a reliable macromodel of the simulated transfer function is obtained with as few simulations as possible.

The AFS scheme is showcased by applying it to the illustrative example of a hairpin filter. Because the proposed scheme automatically focuses the sampling in the region of the resonance, a reliable macromodel is obtained. This establishes the efficiency of the proposed AFS scheme.

REFERENCES

- [1] G. Antonini, D. Deschrijver, and T. Dhaene, "Broadband rational macromodeling based on the adaptive frequency sampling algorithm and the partial element equivalent circuit method," *IEEE Transactions on Electromagnetic Compatibility*, vol. 50, no. 1, pp. 128–137, Feb 2008.
- [2] T. Dhaene, J. Ureel, N. Faché, and D. De Zutter, "Adaptive frequency sampling algorithm for fast and accurate s-parameter modeling of general planar structures," in *Proceedings of 1995 IEEE MTT-S International Microwave Symposium*, vol. 3. IEEE, May 1995, pp. 1427–1430.
- [3] D. Deschrijver and T. Dhaene, "Passivity-based sample selection and adaptive vector fitting algorithm for pole-residue modeling of sparse frequency-domain data," in *Proceedings of the 2004 IEEE International Behavioral Modeling and Simulation Conference, 2004. BMAS 2004*. IEEE, Oct 2004, pp. 68–73.
- [4] N. Mutonkole and D. I. L. de Villiers, "Adaptive frequency sampling for radiation patterns and s-parameters of antennas," in *2017 11th European Conference on Antennas and Propagation (EUCAP)*. IEEE, March 2017, pp. 3195–3199.
- [5] B. Gustavsen and A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Transactions on Power Delivery*, vol. 14, no. 3, pp. 1052–1061, July 1999.
- [6] B. Gustavsen, "Improving the pole relocating properties of vector fitting," *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1587–1592, July 2006.
- [7] D. Deschrijver, M. Mrozowski, T. Dhaene, and D. De Zutter, "Macro-modeling of multiport systems using a fast implementation of the vector fitting method," *IEEE Microwave and Wireless Components Letters*, vol. 18, no. 6, pp. 383–385, June 2008.
- [8] S. Grivet-Talocia and B. Gustavsen, *Passive macromodeling: Theory and applications*. John Wiley & Sons, 2015, vol. 239.
- [9] S. De Ridder, D. Deschrijver, D. Spina, T. Dhaene, and D. Vande Ginste, "Adaptive frequency sampling using linear Bayesian vector fitting," *Electronics Letters*, November 2018.
- [10] C. K. Sanathanan and J. Koerner, "Transfer function synthesis as a ratio of two complex polynomials," *IEEE Transactions on Automatic Control*, vol. 8, no. 1, pp. 56–58, January 1963.
- [11] C. M. Bishop, *Pattern recognition and machine learning*. New York, NY: Springer, 2006.
- [12] A. Lamecki and M. Mrozowski, "Equivalent SPICE circuits with guaranteed passivity from nonpassive models," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 3, pp. 526–532, March 2007.