An iterative reweighting process for macromodel extraction of power distribution networks

An iterative reweighting process for macromodel extraction of power distribution networks

S. Grivet-Talocia, A. Ubolli
Politecnico di Torino
C. Duca degli Abruzzi 24, 10129 Torino, Italy
e-mail {grivet,andrea.ubolli}@polito.it

M. Bandinu, A. Chinea
IdemWorks s.r.l.
C. Trento 13, 10129 Torino, Italy
e-mail {m.bandinu,a.chinea}@idemworks.com

***Abstract*—This paper introduces a new algorithm for the generation of optimal time-domain macromodels of power distribution networks, starting from a set of tabulated scattering responses and given a nominal termination scheme for active blocks, decoupling capacitors, and voltage regulator module. The new concept being introduced is a modified metric to characterize and optimize the accuracy of the macromodel, which takes into account the operation conditions that will be applied to run transient simulations for power integrity assessment. This metric is applied through an iterative frequency-dependent reweighting scheme in a fully automated flow. Two examples illustrate the performance of the proposed algorithm.**

I. Introduction

The Power Distribution Network (PDN) is a fundamental block of any electronic system. The main purpose of the PDN is to connect a stablized voltage source, the Voltage Regulator Module (VRM), to all devices in the system, in order to supply power and allow their operation [1]. A stable voltage is guaranteed only at the VRM location, but any device will operate properly if its own supply voltage is within a prescribed small range. Therefore, the parasitics of the entire PDN at board, package and chip levels must be carefully controlled via proper design rules, in order to minimize their impact on the supply voltage variations at the device locations [2], [3].

The most common approach for designing PDN's relies on a frequency-domain characterization and optimization of a target impedance, which relates the supply voltage variations to the current loading from active devices [1], [2]. The smaller is this impedance, the less sensitive is the device voltage with respect to device supply currents. When this impedance is detected to be too large within some frequency range, sets of decoupling capacitors are connected in parallel to the PDN, in order to locally decrease the PDN impedance. The precise location and the actual values of the capacitors are often determined by complex optimization processes [4], [5].

Although design and the optimization of PDN's are performed in the frequency domain, the actual verification that the device supply voltages are within allowed ranges must be performed in the time domain, by running a transient simulation on a system-level circuit that includes suitable models for PDN, decoupling capacitors, VRM, and active device blocks. Macromodeling schemes are very useful for this task, since they allow to convert frequency-domain descriptions (typically in form of tabulated scattering responses computed by full-wave solvers) into a stable and passive time-domain macromodel in state-space form, ready for transient simulation [6]– [17].

Assuming a scattering form of the initial PDN frequency-domain responses, the state-space macromodel will represent with good accuracy the dynamics of the PDN in terms of the transient scattering outgoing waves at its ports, when excited by transient incident scattering waves. Equivalently, the "nominal" termination scheme for which the macromodel accuracy is optimized during the fitting is a set of resistances R_0 equal to the reference impedance used to define the scattering parameters. This termination scheme is very different from the actual terminations that will be applied to the macromodel during transient simulation, mainly consisting of decoupling capacitors, localized or distributed current sources, and at least one VRM model. It was verified for several application cases that even when the macromodel is very accurate in the scattering representation, the corresponding derived PDN impedance may differ significantly from the nomimal impedance computed from the raw scattering data. The main reason for this difference is due to the sensitivity of the transformation that converts the scattering PDN responses into the PDN impedance under realistic loading conditions. The main objective of this work is in fact to elminate this problem by optimizing the accuracy of the macromodel in terms of the PDN impedance, by explicitly including this sensitivity into account during model extraction. This is achieved by iteratively optimizing a set of frequency-dependent weights based on this sensitivity during the rational fitting stage.

This work is organized as follows. In Section II we provide evidence of PDN impedance sensitivity on a simple test case, and we formally state the problem under investigation. In Section III we describe our proposed iteratively reweighted rational fitting scheme. Finally, in Section IV we demonstrate the performance of the proposed algorithm on some application examples.

II. Problem statement

Our starting point is a set of frequency samples of the PDN scattering matrix

$$\hat{\mathbf{S}}_k = \hat{\mathbf{S}}(\mathrm{j}\omega_k), \quad k = 1, \ldots, K, \tag{1}$$

available from a field solver. Each of the samples $\hat{\mathbf{S}}_k$ is a complex-valued $P \times P$ matrix, where P denotes the number of ports of the PDN description. These ports are grouped into three different subsets $P = P_a + P_c + P_v$, where

- the first P_a ports are connected to suitable models of the active device blocks; the simplest model for these blocks are current sources drawing a fixed time-dependent signal representing the cumulative transient supply current of the active devices switching within the block;
- the second P_c ports are connected to suitable models of the decoupling capacitors, including their series resistance and inductance parasitics;
- the last P_v ports are connected to suitable VRM models, often simply represented by a constant voltage source with a small series impedance. We will replace this voltage source with a short circuit to characterize only the voltage fluctuations around the nominal value.

The above described network loading the PDN can be represented by its frequency-dependent non-homogeneous admittance representation

$$-\boldsymbol{I}(s) = \mathbf{Y}_L(s)\boldsymbol{V}(s) - \boldsymbol{J}(s), \tag{2}$$

where $\boldsymbol{I}(s)$ collects the currents entering each PDN port, $\boldsymbol{V}(s)$ collects PDN port voltages, and where $\mathbf{Y}_L(s)$ collects in its diagonal entries the admittances connected to each port. The source vector $\boldsymbol{J}(s)$ collects the active device currents $\boldsymbol{J}_a(s)$ in its first P_a entries and is vanishing otherwise. Combining (1) with (2) and solving for the port voltages at each frequency point ω_k leads to

$$\boldsymbol{V}_a(\mathrm{j}\omega_k) = \hat{\mathbf{Z}}_{a,k}\boldsymbol{J}_a(\mathrm{j}\omega_k), \tag{3}$$

where $\hat{\mathbf{Z}}_{a,k}$ denotes the upper-left $P_a \times P_a$ block of matrix $[\hat{\mathbf{Y}}_k + \mathbf{Y}_L(\mathrm{j}\omega_k)]^{-1}$, with

$$\hat{\mathbf{Y}}_k = R_0^{-1}[\mathbf{I} - \hat{\mathbf{S}}_k][\mathbf{I} + \hat{\mathbf{S}}_k]^{-1} \tag{4}$$

Each element (i,j) of matrix $\hat{\mathbf{Z}}_{a,k}$ represents the voltage response at the active block location i, excited by the switching current of the active block j, with all other ports terminated by decoupling capacitors and VRM models.

Suppose now that we compute a rational macromodel in scattering pole-residue form

$$\mathbf{S}(s) = \sum_{n=1}^{N} \frac{\mathbf{R}_n}{s - p_n} + \mathbf{R}_0 \tag{5}$$

by minimizing the cumulative least squares fitting error with respect to the original data (1)

$$E^2 = \sum_{k=1}^{K} E_k^2 = \sum_{k=1}^{K} \|\mathbf{S}(\mathrm{j}\omega_k) - \hat{\mathbf{S}}_k\|^2. \tag{6}$$

This process is standard in Vector Fitting (VF) applications; the global fitting error E can be reduced below a prescribed tolerance δ with a suitable model order N. Once this accurate scattering model is computed, we combine (5) with the terminations (2) to obtain the model of the PDN impedance $\mathbf{Z}_a(s)$, following the same procedure of (2)-(5). We are interested in the characterization of the frequency-dependent error of this impedance model

$$\Delta_k = \|\mathbf{Z}_a(\mathrm{j}\omega_k) - \hat{\mathbf{Z}}_{a,k}\|. \tag{7}$$

We remark that enforcing E_k to be small during the fitting process does not guarantee that Δ_k will be small. In fact, error amplification may occur due to the transformation expressed by (2)-(4), which is termination-dependent. Using a first-order approximation, one may write

$$\Delta_k \approx \mathcal{S}_k E_k, \tag{8}$$

where $\mathcal{S}_k$ can be interpreted as a sensitivity of the PDN impedance with respect to perturbations in the scattering PDN responses under nominal termination conditions (2).

This sensitivity is illustrated on a simple canonical structure. Consider a template PDN formed by a 10×10 mesh of impedances $sL + Z_i(s)$, where $L = 0.25$ nH and $Z_i(s)$ is an equivalent lumped network for the internal conductor impedance and representing both DC and skin effect losses, with identical capacitances $C = 4.43$ pF connecting each node to a reference ground. At one corner (port 3) we connect a 1 mΩ resistor to represent a VRM, in parallel with a 470 μF large bulk capacitor. A single decoupling capacitor ($C = 1$ μF, $R_s = 25$ Ω, $L_s = 0.23$ nH) is connected at the center node (port 2) of the mesh. We are interested in the input impedance $Z_a(s)$ at the opposite corner (port 1) with respect to the VRM. A reduced-order macromodel is computed via standard VF from the 3×3 scattering matrix of the PDN up to 5 GHz. The top panel of Fig. 1 demonstrates the very good accuracy of this macromodel. However, due to the large sensitivity of $\hat{Z}_a(s)$ at low frequency to the model fitting error in the scattering domain (middle panel), the impedance $Z_a(s)$ computed from the macromodel (bottom panel of Fig. 1, thick solid black line) is very different from the nominal $\hat{Z}_a(s)$. Even enforcing an exact DC value in the fit results in a loss of accuracy around 10 MHz (dashed red line). This canonical example shows that even for simple structures the sensitivity $\mathcal{S}_k$ may be large, leading to a complete loss of accuracy when simulating an accurate scattering-based macromodel under typical PDN termination conditions.

III. Formulation

The error amplification issues that were observed in Sec. II can be significantly alleviated by a reweighting process during the macromodel generation. Consider the frequency-dependent sensitivity $\mathcal{S}_k$ depicted in Fig. 1. This plot tells us that the difference between the fitted macromodel $\mathbf{S}(\mathrm{j}\omega_k)$ and the original data $\hat{\mathbf{S}}_k$ is amplified at the frequencies where $\mathcal{S}_k$ is large. A compensation for this effect is achieved by making the macromodel more accurate at these frequencies, i.e., by emphasizing their contribution in the overall fitting error (6). More precisely, we define a modified cost function

$$E_w^2 = \sum_{k=1}^{K} E_{w,k}^2 = \sum_{k=1}^{K} w_k^2 \|\mathbf{S}(\mathrm{j}\omega_k) - \hat{\mathbf{S}}_k\|^2 \tag{9}$$

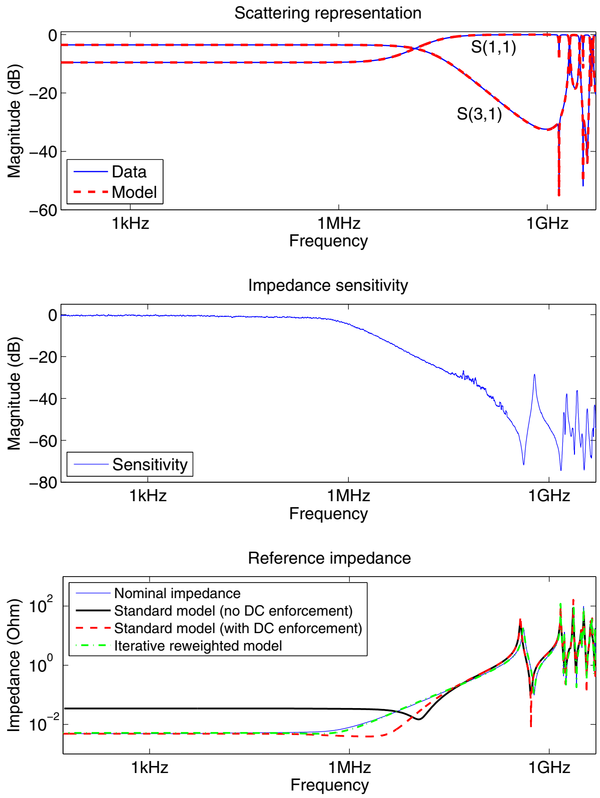

Fig. 1. Illustration of macromodel sensitivity on a canonical PDN structure. Top panel: comparison between macromodel and original responses in the scattering representation. Middle panel: frequency-dependent impedance sensitivity. Bottom panel: comparison between nominal impedance and the corresponding macromodels (standard and iteratively reweighted).

by multiplying the error contribution of each individual frequency by a frequency-dependent weight w_k. This modified cost function is then used as a target accuracy within standard VF iterations (both pole relocation and residue identification stages). In practice, any row k in the linear least squares systems to be solved during VF iterations is multiplied by the corresponding weight w_k, see [6] for more details on practical implementations.

A reasonable choice for the weight w_k is the frequency-dependent sensitivity $\mathcal{S}_k$. However, it may be the case that even this weighted fitting process is not able to attain the desired accuracy. Therefore, we setup the following iterative scheme for $\mu = 0, 1, \ldots$ until the desired accuracy is met.

1) As initialization for $\mu = 0$ we define an initial weight $w_k^0 = \mathcal{S}_k$. The sensitivity can be computed at each frequency ω_k in a closed form by a first-order expansion of Δ_k in (7). Alternatively, before starting the iterations, the raw data $\hat{\mathbf{S}}_k$ are perturbed through the multiplication of all matrix elements by $(1+\varepsilon_k)$, where ε_k is a zero-mean Gaussian variable with prescribed standard deviation σ, and the resulting perturbation $\partial Z_{a,k}$ induced on the target impedance $\hat{\mathbf{Z}}_{a,k}$ is computed. The ratio between the standard deviation of $\partial Z_{a,k}$ and σ in a Monte-Carlo run provides a good estimate of the sensitivity $\mathcal{S}_k$.
2) At each iteration $\mu = 0, 1, \ldots$, a macromodel $\mathbf{S}^\mu(s)$ is computed by weighted VF using the cost function (9) with weight w_k^μ.
3) When macromodel $\mathbf{S}^\mu(s)$ is available, the corresponding PDN impedance $\mathbf{Z}_a^\mu(s)$ is derived and its frequency-dependent deviation Δ_k^μ from the nominal impedance is computed as in (7). If $\Delta_k^\mu < \delta$ at all frequencies, where δ is the desired target accuracy, the iteration is stopped.
4) Otherwise, a new frequency-dependent weight for next iteration is defined as
$$w_k^{\mu+1} = w_k^\mu \cdot \mathcal{F}(\Delta_k^\mu), \tag{10}$$
where $\mathcal{F} : \mathbb{R}^+ \mapsto \mathbb{R}^+$ denotes a non-decreasing function such that $\mathcal{F}(\delta) = 1$. This choice guarantees that the weight of next iteration will further emphasize those frequencies for which the impedance error is significant, and will not modify the weight for those frequencies that are already accurate. The iteration count is then increased $\mu \leftarrow \mu + 1$ and the scheme is restarted from step 2).

This simple weight updating process is able to converge to a good solution even if the initialization step is skipped and the initial weight is set to $w_k^0 = 1$, due to the automatic detection of the most sensitive frequency points in step 4).

IV. Numerical Examples

We first illustrate the performance of the proposed technique on the canonical PDN example of Fig. 1. The bottom panel reports with a green dash-dot line the PDN impedance $Z_a(s)$ computed from the iteratively reweighted macromodel, showing that the loss of accuracy that was observed for the standard macromodels has been successfully compensated.

The second example is part of a real product (courtesy of Yan Shen Fen, Intel). The structure of interest is a cut out of a PDN with a total of $P = 13$ ports. The first $P_a = 12$ ports correspond to C4 bumps and the last port represents the reference board. The nominal termination network includes a lumped 1 mΩ resistor connected to the board port and suitable models for the on-die capacitance connected to the C4 ports. The actual C4 terminations are all series RC branches with individual values of resistance and capacitance in the range 1.5–3 mΩ and 10–15 nF, respectively. Excitation of the PDN is provided by a total normalized current of $J_{\text{tot}} = 1$ A, uniformly spread through all C4 ports, i.e., $J_{a,k} = J_{\text{tot}}/P_a$ for $k = 1, \ldots, P_a$. The voltage at the C4 ports is considered as the output to define the target impedances $Z_{a,k} = V_k/J_{\text{tot}}$.

The macromodeling results depicted in Fig. 2 confirm that, although the scattering-based macromodel matches closely the scattering responses of the PDN (top panel), the corresponding target impedance is very inaccurate (bottom panel) due to the large sensitivity, especially at low and medium frequencies (middle panel). Conversely, the iteratively reweighted macromodel provides a very accurate impedance response (bottom panel). The middle panel in Fig. 2 also reports the final weight w_k resulting from the proposed iteration. This weight was

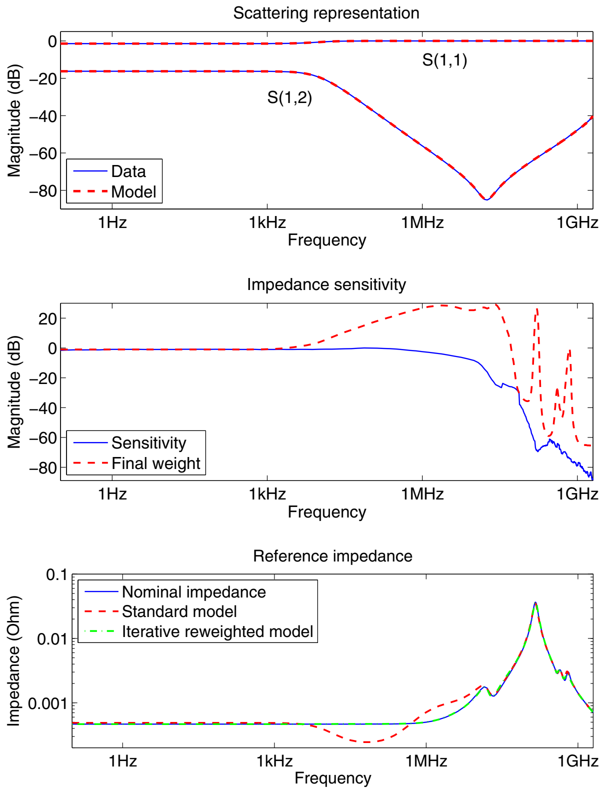

Fig. 2. As in Fig. 1, but for a real PDN structure.

obtained by starting the iterations with $w_k^0 = 1$, i.e., without any prior knowledge of the actual sensitivity response. The figure shows that the algorithm can optimize the weight even with a poor initialization, thus proving quite robust.

V. Conclusions

This paper presented a very simple yet effective iterative process to optimize the accuracy of macromodels when a nominal termination scheme is known. Although special emphasis was put here on PDN structures, the approach is general and in principle applicable to more general structures. The main idea is to exploit knowledge of the nominal terminations by enhancing the fitting accuracy at those frequencies that happen to amplify the inevitable approximation errors that occur during macromodel extraction. The numerical results on both canonical and real structures demonstrate the validity of the approach.

As a final remark, we note that the proposed reweighting scheme was applied here only for the macromodel generation by rational VF, and no discussion on passivity verification and enforcement was carried out. In fact, passivity is a fundamental requirement to guarantee stable and reliable time-domain verifications, which should never be omitted. It turns out that, once the final weighting factor w_k is obtained at the end of the reweighting iterations, the same weighted cost function (9) can be used in any perturbation-based passivity enforcement schemes [11]– [17], resulting in implicit accuracy preservation at the most sensitive frequency points. This application will be documented in a forthcoming report.

References

[1] M. Swaminathan, A.E. Engin, *Power Integrity Modeling and Design for Semiconductors and Systems*, Prentice Hall, 2007.

[2] Swaminathan, M.; Joungho Kim; Novak, I.; Libous, J.P., "Power distribution networks for system-on-package: status and challenges," *IEEE Trans. on Advanced Packaging*, vol. 27, no. 2, pp. 286–300, May 2004.

[3] Swaminathan, M.; Daehyun Chung; Grivet-Talocia, S.; Bharath, K.; Laddha, V.; Jianyong Xie, "Designing and Modeling for Power Integrity," *IEEE Trans. on Electromagnetic Compatibility*, vol. 52, no. 2, pp. 288–310, May 2010.

[4] Haihua Su; Sapatnekar, S.S.; Nassif, S.R., "Optimal decoupling capacitor sizing and placement for standard-cell layout designs," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, no. 4, pp. 428–436, Apr 2003.

[5] Kose, S.; Friedman, E.G., "Distributed On-Chip Power Delivery," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 4, pp. 704–713, Dec. 2012

[6] B. Gustavsen, A. Semlyen, "Rational approximation of frequency responses by vector fitting", *IEEE Trans. Power Delivery*, Vol. 14, N. 3, pp. 1052–1061, July 1999.

[7] D. Deschrijver, B. Haegeman, T. Dhaene, "Orthonormal Vector Fitting: A Robust Macromodeling Tool for Rational Approximation of Frequency Domain Responses", *IEEE Transactions on Advanced Packaging*, Vol. 30, No. 2, pp. 216–225, May 2007.

[8] D. Deschrijver, M. Mrozowski, T. Dhaene, D. De Zutter, "Macromodeling of Multiport Systems Using a Fast Implementation of the Vector Fitting Method," *IEEE Microwave and Wireless Components Letters*, Vol. 18, N. 6, June 2008, pp.383–385.

[9] A. Chinea, S. Grivet-Talocia, "On the Parallelization of Vector Fitting Algorithms," *IEEE Trans. on Components, Packaging, and Manufacturing Technology*, Vol. 1, n. 11, Nov. 2011, pp. 1761–1773.

[10] S. Grivet-Talocia, S.B. Olivadese, P. Triverio, "A compression strategy for rational macromodeling of large interconnect structures," EPEPS 2011, San Jose, CA (USA), October 23-26, 2011, pp. 53–56.

[11] C. P. Coelho, J. Phillips, L. M. Silveira, "A Convex Programming Approach for Generating Guaranteed Passive Approximations to Tabulated Frequency-Data", *IEEE Trans. CAD*, Vol. 23, No. 2, pp. 293–301, Feb. 2004.

[12] S. Grivet-Talocia, "Passivity enforcement via perturbation of Hamiltonian matrices", *IEEE Trans. CAS-I*, Vol. 51, No. 9, pp. 1755-1769, Sept. 2004.

[13] D. Saraswat, R. Achar and M. Nakhla, "Global Passivity Enforcement Algorithm for Macromodels of Interconnect Subnetworks Characterized by Tabulated Data", *IEEE Trans. VLSI Systems*, Vol. 13, No. 7, pp. 819–832, July 2005.

[14] S. Grivet-Talocia, A. Ubolli "On the Generation of Large Passive Macromodels for Complex Interconnect Structures", *IEEE Trans. Adv. Packaging*, vol. 29, No. 1, pp. 39–54, Feb. 2006.

[15] B. Gustavsen, A. Semlyen, "Enforcing passivity for admittance matrices approximated by rational functions", *IEEE Trans. Power Systems*, Vol. 16, N. 1, pp. 97–104, Feb. 2001.

[16] L. De Tommasi, M. de Magistris, D. Deschrijver, T. Dhaene, "An algorithm for direct identification of passive transfer matrices with positive real fractions via convex programming,", *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, Vol. 24, N. 4, pp. 375–386, 2011.

[17] D. Deschrijver, T. Dhaene, "Fast Passivity Enforcement of S-Parameter Macromodels by Pole Perturbation," *IEEE Trans. MTT*, Vol. 57, no. 3, pp. 620–626, 2009.