Improving the efficiency of passivity compensation schemes via adaptive sampling

Improving the efficiency of passivity compensation schemes via adaptive sampling

S. Grivet-Talocia

Dip. Elettronica, Politecnico di Torino, C. Duca degli Abruzzi 24, 10129 Torino, Italy

Ph. +39 011 5644104, Fax +39 011 5644099 (e-mail grivet@polito.it)

Abstract: A new algorithm for the enforcement of passivity for large lumped macromodels is presented. An accuracy-controlled adaptive frequency sampling combined with an iterative multishift Arnoldi scheme leads to a fast computation of the imaginary Hamiltonian eigenvalues, which are perturbed until passivity is reached. Up to two orders of magnitude speedup factors are achieved with respect to previous formulations.

1 Preliminaries and motivations

Macromodeling of interconnect structures is a common practice for the assessment of Signal Integrity (SI) issues in early stages of the design of high-speed communication and information systems. Macromodels are often derived from tabulated responses coming from direct measurement or full-wave simulations. Excellent fitting algorithms [1] are available for this task. However, when passivity is not explicitly enforced during this process, possible difficulties may arise, since non-passive models may lead to unstable transient simulations due to their ability to generate more energy than they are fed with. This is indeed a serious drawback that in some cases makes the macromodels practically useless. Consequently, there is a strong motivation underlying several research efforts towards efficient schemes for the enforcement of the macromodel passivity (see, e.g., [2, 3, 4]).

Here we concentrate on a class of methods based on the iterative perturbation of Hamiltonian eigenvalues. In fact, there is still significant potential for improving the efficiency of these schemes when the number of internal states, or equivalently, the number of poles of the macromodel is large. In such cases, all schemes suffer from an excessive computational cost that significantly reduces their applicability. In this work, we show that an accuracy-controlled adaptive frequency sampling process, combined with a customized scheme for the determination of few Hamiltonian eigenvalues restricted to small frequency intervals, may result in very significant speedup factors with respect to standard implementations. In some cases speedup factors up to two orders of magnitude have been observed without affecting the quality of the results.

We consider linear macromodels having N poles and P ports, defined by a real, minimal, and strictly stable state-space realization

$$\boldsymbol{H}(s) = \boldsymbol{D} + \boldsymbol{C}\,(s\boldsymbol{I} - \boldsymbol{A})^{-1}\,\boldsymbol{B}\,. \tag{1}$$

Only the scattering representation will be considered for $\boldsymbol{H}(s)$, but all the results in this work will be applicable with obvious modifications to other representations, e.g. impedance or admittance [4]. In the scattering case, $\boldsymbol{H}(s)$ is passive when all singular values of $\boldsymbol{H}(j\omega)$ are uniformly bounded by one at any frequency

$$\sigma_i \leq 1\,, \quad \forall \sigma_i \in \sigma(\boldsymbol{H}(j\omega))\,, \quad \forall \omega\,. \tag{2}$$

We will assume passivity at $s = \infty$ (i.e., the singular values of $\boldsymbol{D}$ are strictly less than one), since this condition can be easily enforced during the fitting stage. We target our analysis to macromodels characterized by sparse state-space matrices, as discussed in [5, 6]. The basic condition underlying all developments requires that $(j\omega\boldsymbol{I} - \boldsymbol{A})$ can be inverted at a small computational cost. Therefore, we will assume $\boldsymbol{A}$ to be block-diagonal, with blocks of size 1 for the synthesis of real poles and of size 2 for complex pole pairs. Note that this choice is always possible in the construction of the state-space realization and leads to a $O(N)$ computational cost for the evaluation of the transfer matrix (1) at a given frequency.

The well-known theory of Hamiltonian matrices identifies the frequencies ω_k at which one of the singular values σ_i reaches the threshold $\gamma = 1$ with the purely imaginary eigenvalues $\mu_k = j\omega_k$ of the Hamiltonian matrix

$$\mathcal{M} = \begin{pmatrix} \boldsymbol{A} - \boldsymbol{B}(\boldsymbol{D}^T\boldsymbol{D} - \boldsymbol{I})^{-1}\boldsymbol{D}^T\boldsymbol{C} & -\boldsymbol{B}(\boldsymbol{D}^T\boldsymbol{D} - \boldsymbol{I})^{-1}\boldsymbol{B}^T \\ \boldsymbol{C}^T(\boldsymbol{D}\boldsymbol{D}^T - \boldsymbol{I})^{-1}\boldsymbol{C} & -\boldsymbol{A}^T + \boldsymbol{C}^T\boldsymbol{D}(\boldsymbol{D}^T\boldsymbol{D} - \boldsymbol{I})^{-1}\boldsymbol{B}^T \end{pmatrix}. \tag{3}$$

If some imaginary eigenvalues are found, the macromodel is not passive and compensation must be performed. To this end, an iterative perturbation scheme can be used to displace these eigenvalues from the imaginary axis and to achieve passivity. No details about the main perturbation algorithm will be provided here, since they are available in [4]. We recall that the main numerical tools required by this scheme are eigenvalue determinations of $\mathcal{M}$ and least squares solutions of small linear systems. Unfortunately, the need of an eigensolution for the Hamiltonian matrix becomes a weak point when the dynamic order N is large, since the associated cost scales as $O(N^3)$. As a consequence, when a standard full eigensolver is employed, the standard Hamiltonian-based scheme of [4] can only be applied to moderate-size macromodels with a reasonable computational cost. We remark that the same scaling law applies to the computational cost per iteration of state-of-the art convex optimization approaches [3]. Here we are trying to reduce this cost.

0-7803-9220-5/05/$20.00©2005 IEEE.

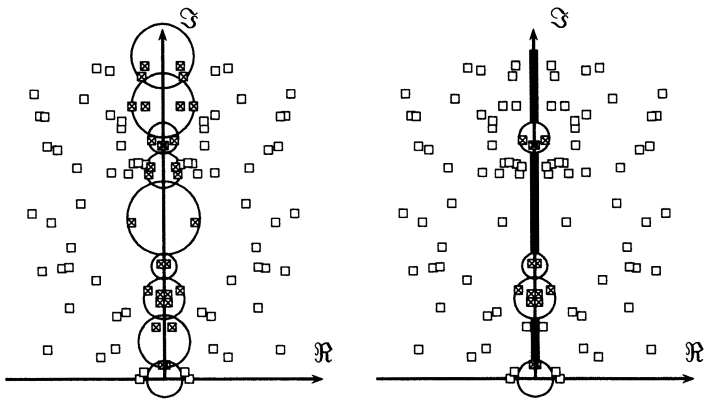

Figure 1: Use of multiple shifts along the imaginary axis for the determination of few eigenvalues of the Hamiltonian matrix close to the imaginary axis (crosses), neglecting the other part of the eigenspectrum (squares). Circles indicate the convergence regions for the restarted Arnoldi process at each shift. The entire bandwidth is processed in left panel, thus requiring a possibly large number of shifts. Right panel shows that a significant reduction in the number of shifts may be achieved by skipping some carefully selected frequency intervals (thick lines).

Some progress in speeding up the algorithm has been documented in [5, 6], where a dedicated eigensolver for the selective determination of the imaginary eigenvalues of Hamiltonian matrices admitting sparse decompositions was presented. This eigensolver is based on a multishift restarted Arnoldi process [7], similar to the Complex Frequency Hopping (CFH) scheme [8]. A bisection process is performed on the full bandwidth encompassing all imaginary eigenvalues (if present), and multiple shifts are placed along the imaginary axis. A dedicated Arnoldi process aimed at the determination of few closest eigenvalues to each shift is run, and the results are collected to gather the full set of imaginary eigenvalues. A schematic illustration is provided in the left panel of Fig. 1. The advantage of this scheme in terms of required operations is significant, since the computational cost scales with the number of states only linearly, as $O(cN)$. Unfortunately, the constant c may be very large, since it is roughly proportional to the number of shifts and to the squared size of the Krylov subspace being used at each shift, the main bottleneck being the unavoidable orthogonalization of the Krylov vectors.

In this work, we show that an adaptive sampling process can be applied to identify some strictly passive frequency bands, where no imaginary eigenvalues can be located, by checking (2). Although it is widely recognized that this procedure may be cumbersome and computationally expensive, this is not the case when a fast inversion of $(j\omega \boldsymbol{I} - \boldsymbol{A})$ is possible and when $N \gg P$. It is obvious that this identification is only approximate, since only a finite number of frequency samples are tested. Indeed, we will use this frequency sweep test only as a preliminary step, aimed at the exclusion of a set of certainly passive frequency bands from the more accurate but expensive check based on the imaginary Hamiltonian eigenvalues. The multishift eigensolver, applied to these reduced-size frequency bands, will only need a reduced number of shifts with respect to its full-bandwidth application, as depicted in the right panel of Fig. 1. Consequently, the computational cost will be significantly reduced, since no effort will be devoted to search for eigenvalues where they cannot be located.

2 Adaptive sampling and imaginary Hamiltonian eigenvalues determination

We highlight here the procedure aimed at the exclusion of certainly passive frequency bands from the imaginary eigenvalue search. As a first step, we estimate a frequency $\omega_{\max}$ providing an upper bound for all imaginary eigenvalues $j\omega_k$ of the Hamiltonian matrix. The existence of this upper frequency is guaranteed by the asymptotic passivity of the macromodel. This upper frequency is quite easy to estimate, being provided by the eigenvalue of maximum magnitude of the Hamiltonian matrix $\mathcal{M}$,

$$\omega_k \leq \omega_{\max} = |\mu_{\max}| = \arg\max_k |\mu_k| \,. \tag{4}$$

The evaluation of $|\mu_{\max}|$ is a standard problem in numerical analysis, since the sequence $\{\boldsymbol{x}, \mathcal{M}\boldsymbol{x}, \mathcal{M}^2\boldsymbol{x}, \dots\}$ with proper normalization converges to the associated eigenvector [7].

Next step is the generation of a set of frequency samples providing a rough representation of the transfer matrix. This set can be determined by the distribution of the macromodel poles $p_n = \alpha_n + j\beta_n$, since each pole contributes to the frequency variations mostly in a bandwidth centered at β_n with size approximately $|\alpha_n|$. Here we consider a total number of $2R + 1$ frequency samples for each pole (only the poles with nonnegative imaginary part are considered), providing uniform phase sampling

$$\omega_{n,r} = \beta_n + \alpha_n \tan \frac{r\pi}{2(R+1)} \quad r = -R, \dots, R \,. \tag{5}$$

We used $R = 3$ in all numerical tests of this work, which proved to be sufficient. The samples generated for each pole are accumulated and sorted. Then, only the frequencies enclosed in $[0, \omega_{\max}]$ are retained, including the boundary samples. A last scan is performed to remove those samples that are closer than a prescribed threshold $\Delta\omega$, which is the initial frequency resolution. In this work we used $\Delta\omega = \omega_{\max}/10N$.

The above procedure leads to a set of frequency samples $\{\omega_m,\ m = 0, \ldots M\}$ with $\omega_0 = 0$ and $\omega_M = \omega_{\max}$. At each of these samples we compute the Hermitian matrix $\boldsymbol{\Theta}_m = \boldsymbol{H}^H(j\omega_m)\boldsymbol{H}(j\omega_m)$ and the associated eigenvalues $\boldsymbol{\Lambda}_m = \mathrm{diag}\{\lambda_{i,m}\}$ and (orthonormal) eigenvectors $\boldsymbol{V}_m$. The eigenvectors at sample $m+1$ will be some perturbation of those at sample m,

$$\boldsymbol{V}_{m+1} = \boldsymbol{V}_m \boldsymbol{\Psi}_{m,m+1} \boldsymbol{P}_{m,m+1} + \delta \boldsymbol{V}_{m,m+1}\,, \tag{6}$$

where $\boldsymbol{\Psi}_{m,m+1} = \mathrm{diag}\{e^{j\psi_i}\}$ takes into account arbitrary phase terms (depending on the eigensolver) and $\boldsymbol{P}_{m,m+1}$ is a permutation matrix. An estimate of this permutation matrix is provided by

$$\widehat{\boldsymbol{P}}_{m,m+1} = \mathcal{I}(|\widetilde{\boldsymbol{P}}_{m,m+1}|)\,, \quad \text{with} \quad \widetilde{\boldsymbol{P}}_{m,m+1} = \boldsymbol{V}_m^H \boldsymbol{V}_{m+1} = \boldsymbol{\Psi}_{m,m+1}\boldsymbol{P}_{m,m+1} + \boldsymbol{V}_m^H \delta \boldsymbol{V}_{m,m+1}\,, \tag{7}$$

where the operator $\mathcal{I}(\cdot)$ rounds towards the nearest integer its matrix argument and the absolute value is taken componentwise to get rid of the phase terms. We attempt reordering of the eigenvectors via inverse permutation by computing

$$\boldsymbol{J}_{m,m+1} = \widehat{\boldsymbol{P}}_{m,m+1}^T \widetilde{\boldsymbol{P}}_{m,m+1} = \widehat{\boldsymbol{P}}_{m,m+1}^T \boldsymbol{\Psi}_{m,m+1}\boldsymbol{P}_{m,m+1} + (\boldsymbol{V}_m \widehat{\boldsymbol{P}}_{m,m+1})^H \delta \boldsymbol{V}_{m,m+1}\,. \tag{8}$$

If the estimate $\widehat{\boldsymbol{P}}_{m,m+1}$ is correct, matrix $\boldsymbol{J}_{m,m+1}$ is a small perturbation of a phase-shifted identity matrix. This condition is verified by checking whether

$$\max_{ij}\left\{|(|\boldsymbol{J}_{m,m+1}| - \boldsymbol{I})|_{ij}\right\} < \varepsilon \tag{9}$$

holds for a sufficiently small threshold ε (the absolute values always being taken componentwise). We use in this work $\varepsilon = 0.2$. If this is not the case, either the estimated $\widehat{\boldsymbol{P}}_{m,m+1}$ is wrong, or the eigenvectors $\boldsymbol{V}_{m+1}$ differ significantly from $\boldsymbol{V}_m$. The interval (ω_m, ω_{m+1}) is therefore flagged as inaccurate due to excessively coarse sampling and it will be refined at the next iteration. Iterative refinement is achieved by inserting an additional frequency sample at the midpoint of each flagged interval, and by repeating the above checks only on the newly generated intervals. The above procedure is very simple and straightforward, yet it allows to track the frequency dependence of individual eigenvectors and eigenvalues λ_i, even if they cross each other between any pair of frequency samples.

Once eigenvector (hence eigenvalue) tracking between any two adjacent samples is achieved, we compute for each eigenvalue and for each sample a worst-case prediction error

$$\Delta_{m+1} = \max_i\left\{\left|\left(\boldsymbol{\Lambda}'_{m+1}\right)_{ii} - \left(\boldsymbol{\Lambda}_{m+1}\right)_{ii}\right|\right\}\,, \quad \text{with} \quad \boldsymbol{\Lambda}'_{m+1} = \boldsymbol{V}_m^H\, \boldsymbol{\Theta}_{m+1} \boldsymbol{V}_m\,. \tag{10}$$

The interval (ω_m, ω_{m+1}) is flagged as passive, with all eigenvalues λ_i less than the threshold $\gamma = 1$, when

$$\gamma - \max_i\left\{\lambda_i(j\omega_{m+1})\right\} > \beta\Delta_{m+1}\,, \tag{11}$$

where β is a suitable safety factor (we used $\beta = 5$). It is important to note that condition (11) parameterizes how close can an eigenvalue $\lambda_i(j\omega_{m+1})$ is to the threshold in terms of its maximum prediction error Δ_{m+1}. When condition (11) is not satisfied, i.e., when some eigenvalue is too close or larger than the threshold, no conclusions can be drawn on the actual passivity or non-passivity within the corresponding interval. Collecting all these intervals we get the following decomposition

$$[0, \omega_{\max}] = \Omega^{\text{passive}} \cup \Omega^{\text{suspect}}\,, \quad \Omega^{\text{passive}} = \bigcup_{q=1}^{Q}[\omega_{q,0}, \omega_{q,1}]\,, \quad \Omega^{\text{suspect}} = \bigcup_{q=1}^{Q'}[\omega'_{q,0}, \omega'_{q,1}] \tag{12}$$

where Ω^{passive} cannot include any imaginary eigenvalues $j\omega_k$ of $\mathcal{M}$, which (if present) are necessarily located in Ω^{suspect}. The same multishift bisection process already presented in [5, 6] is therefore applied to each of the Q' subintervals $[\omega'_{q,0}, \omega'_{q,1}]$ (see Fig. 1) to identify the complete set of imaginary Hamiltonian eigenvalues. As a result, due to the reduced span of each investigated interval, the number of shifts is very small compared to a full scan of $[0, \omega_{\max}]$, and the computational cost for the determination of all imaginary eigenvalues is drastically reduced. Once these are known, the same perturbation scheme of [4] is applied to displace them from the imaginary axis and to enforce passivity.

3 Numerical results

Table 1 shows numerical results for six different macromodels of various size. All models were obtained by various implementations of the well-known Vector Fitting algorithm [1], applied to frequency-dependent scattering matrices. In particular, cases I and II are high-speed packaging structures (courtesy of Sigrity, Inc.). These two cases were already thoroughly analyzed in [5]. Case III is also a model of a 6-port interconnected system including two power/ground conductors and two signal conductors. Cases IV and V are two macromodels of a 20-port via field under an LGA connector (courtesy of IBM). These two models differ only for the number of poles that were used in the rational approximation. Case VI represents a 3×3 section of a high-speed card-board connector over a bandwidth of 20 GHz (courtesy of IBM).

Results from the computation of the imaginary Hamiltonian eigenvalues are reported in columns 3–6 of Table 1. The column labeled with "Full" reports the CPU time required by the multishift process sweep over the full bandwidth. This is the algorithm that was documented in [5, 6]. The column labeled with "Adaptive" reports the CPU time required by the new proposed algorithm. The

Model specification			Eigenvalue computation			Passivity compensation				
Model	Poles N	Ports P	Full	Adaptive	Speedup	(a)	(b)	(c)	(c) vs (a)	(c) vs (b)
Case I	820	10	48 sec	3.7 sec	13X	119 min	9.1 min	67 sec	106X	8.1X
Case II	1488	12	26 sec	17 sec	1.5X	21 hrs	15 min	6.5 min	194X	2.3X
Case III	450	6	30 sec	8.1 sec	3.7X	16 min	6.6 min	1.4 min	11X	4.7X
Case IV	600	20	46 sec	18 sec	2.5X	36 min	14 min	4.6 min	7.8X	3.0X
Case V	1000	20	160 sec	26 sec	6X	77 min	35 min	5.3 min	15X	6.6X
Case VI	1368	18	420 sec	41 sec	10X	96 min	58 min	4.5 min	21X	13X

Table 1: CPU time required by the computation of the imaginary Hamiltonian eigenvalues (columns 3–6) and for the passivity compensation (columns 7–11) for various test cases. See text for details. All computations were performed with a Pentium IV-based PC.

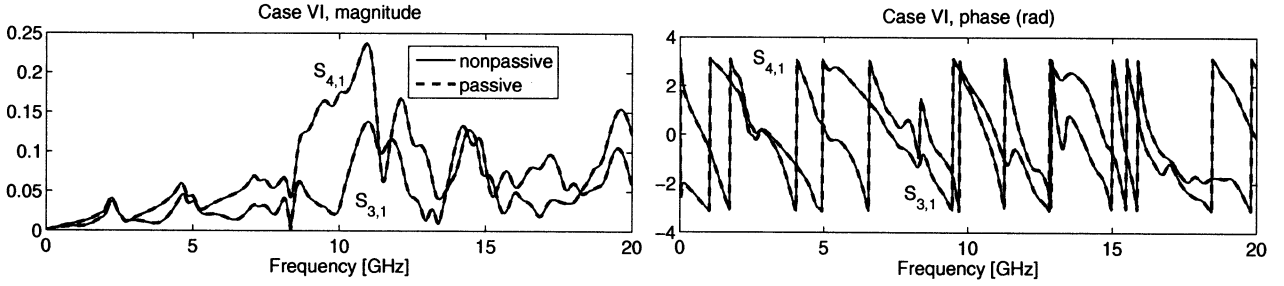

Figure 2: Comparison of Case VI macromodels before and after passivity compensation. Some significant scattering responses before and after passivity compensation are reported in both magnitude and phase.

speedup factor is also reported. We remark that, due to the stringent and conservative conditions on the adaptive sampling process, the same imaginary eigenvalues were detected by the two algorithms for all cases with a maximum relative deviation of 2.2×10^{-14}.

Results from the application of the complete passivity enforcement scheme are presented in the last block of columns. Three implementations are compared: the standard scheme based on a full eigensolver, first documented in [4], in column (a); the scheme based on the multishift iterations applied to the entire bandwidth, first documented in [5, 6], in column labeled (b); and the new scheme based on mixed adaptive sampling/multishift application, in column labeled (c). The speedup factors of the new algorithm with respect to the other two implementations are also reported in the last two columns. These results confirm that the proposed scheme leads to a quite efficient passivity compensation. The compensation is achieved in few minutes for all cases, thus showing significant potential for its automated application to complex multiport interconnect models characterized by a large dynamic order. Figure 2 reports a few representative scattering responses of the passive model to the corresponding ones of the non-passive model before applying the compensation scheme for case VI. The deviation between passive and non-passive models is hardly visible. This confirms that the passivity compensation is performed without compromising the accuracy, and that no overtreatment occurs.

References

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

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

[3] C.P.Coelho, J.Phillips, L.M.Silveira, "A Convex Programming Approach for Generating Guaranteed Passive Approximations to Tabulated Frequency-Data", *IEEE Trans. Computed-Aided Design of Integrated Circuits and Systems*, Vol. 23, No. 2, February 2004.

[4] S. Grivet-Talocia, "Passivity enforcement via perturbation of Hamiltonian matrices", *IEEE Trans. CAS-I*, pp. 1755-1769, vol. 51, n. 9, September, 2004

[5] S. Grivet-Talocia, "Fast passivity enforcement for large and sparse macromodels", in *13th IEEE Topical Meeting on Electrical Performance of Electronic Packaging*, Portland, OR (USA), pp. 247-250, October 25–27, 2004

[6] S. Grivet-Talocia, A. Ubolli "On the Generation of Large Passive Macromodels for Complex Interconnect Structures", to appear in *IEEE Trans. Adv. Packaging*.

[7] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide.* SIAM, Philadelphia, 2000.

[8] E. Chiprout, M.S.Nakhla, "Analysis of Interconnects Networks using Complex Frequency Hopping (CFH)", *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 186-200, Feb. 1995.